الاستدلال الإحصائي (1)

نظرية التقدير

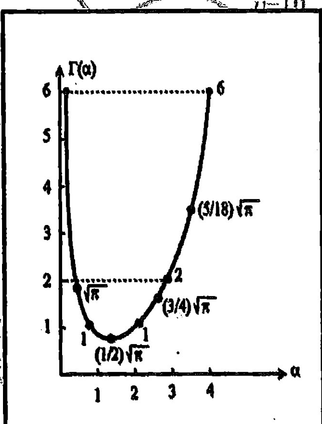
دكتور

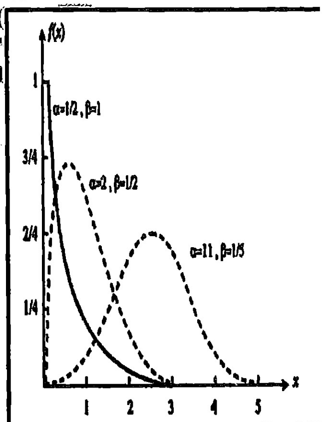
عبد الحفيظ محمد فوزي مصطفى

أستاذ الإحصاء بكلية العلوم

جامعة ناصر

الجماهيرية العظمى





المحتويات

تمهيد

الفصل الأول: مبرهنات النهاية

الفصل الثاني: بعض النماذج الإحصائية الخاصة

الفصل الثالث: المعاينة وتوزيعات المعاينة

الفصل الرابع: الإحصاءات الكافية والمعلومات في العينة

الفصل الخامس: خواص مقدِّرات النقطة

الفصل السادس: طرق إيجاد المقدِّرات النقطية

الفصل السابع: التقدير بفترة

الفصل الثامن: طرق تقدير بييز

مُقدِّمة

يعزى تقدم العلوم للتجريب. حيث ينجز الباحث تجربته الخاصة، فيحصل على بعض المعلومات، ويتوصل بناءً عليها إلى نتائج تتجاوز عادة مواد وعمليات التجربة الخاصة بحيث تشمل كل التجارب المشابهة. أي يقوم بتطبيق منهج الاستدلال الاستقرائي للحصول على معرفة جديدة من الجزء حول الكل.

إن منهج الاستدلال الاستقرائي لا يعطي تعميمات مؤكدة أو حتمية، لكن يمكن قياس درجة الثقة بها فيما إذا ضممت التجربة الخاصة وفق أسس ومبادئ علمية معينة. إن تصميم التجربة، ومن ثم القيام باستدلالات (تعميمات) بناءً على نتائجها يعد من المسائل الأساسية في الإحصاء الرياضي. وعندئذٍ نكون أمام تطبيق منهج الاستدلال الاستقرائي الإحصائي أو باختصار الاستدلال الإحصائي.

يعالج الاستدلال الإحصائي موضوعين هامين، هما:

1. نظرية التقدير.
2. نظرية اختبار الفرضيات.

وهذا الكتاب هو الجزء الأول من كتاب الاستدلال الإحصائي من جزئين، والذي يهدف إلى تقديم مفاهيم وقواعد ومبرهنات نظرية التقدير بشكل متدرج مترابط وشامل بحيث يمكن الطلاب والباحثين -ليس فقط- من استيعاب الأسس الرياضية لهذه النظرية بل من ترجمتها إلى الواقع العملي، وذلك في مختلف الأبحاث التطبيقية التي يمكن أن تعترضهم.

ولتحقيق الهدف، وزعت موضوعات هذا الكتاب على ثمانية فصول، يتقدمها تمهيد يبين مدلول النموذج الإحصائي ومهام الإحصاء الرياضي بالإضافة إلى إيضاح بعض المصطلحات ومدلولات بعض الرموز التي سترد في الفصول المختلفة. وتعتبر الفصول الثلاث الأولى كمقدمة لابد منها لدراسة نظريتي التقدير واختبار الفرضيات.

يتناول الفصل الأول مبرهنات النهاية الأساسية، وذلك بعرض موجز لتقارب متتاليات متغيرات عشوائية، بالإضافة إلى بعض الصيغ الأساسية لقانون الأعداد الكبيرة والنهاية المركزية على هيئة مبرهنات.

يقدم الفصل الثاني النماذج الإحصائية الأكثر استخداماً في التطبيقات الإحصائية مرفقة ببعض الخصائص المميزة ومجالات تطبيقها.

يعالج الفصل الثالث أساسيات المعاينة وتوزيعات المعاينة التي لابد منها لدراسة نظرية التقدير، حيث تم التركيز على المعاينة العشوائية البسيطة لاعتبارها تشكل الأساس لمختلف المعاينات. هذا بالإضافة إلى دراسة بعض الصيغ الخطية والتربيعية الهامة في عناصر العينة وتوزيعاتها الاحتمالية. كما يستعرض إلى جانب ذلك، أهم نتائج نظرية المعاينة للإحصاءات المرتبة والمعاينة في حالة عينة كبيرة الحجم.

يتعرض الفصل الرابع للمعلومات المتوفرة في العينة العشوائية وكيفية تكثيفها في إحصاء واحد أو عدد محدود جداً من الإحصاءات حيث التعامل معها أسهل بكثير من التعامل مع العينة نفسها. وفي هذا الإطار أعطي الإحصاء الكافي والكافي التام أهمية خاصة.

يناقش الفصل الخامس مفهوم التقدير بنقطة وبفترة لمعلمة أو معالم توزيع المجتمع الذي سحبت منه العينة العشوائية، هذا بالإضافة إلى الصفات التي نرغب توفرها في المقدِّر بنقطة. وتم التركيز على المقدِّرات غير المتحيزة ومعايير الكفاءة لها.

يعرض الفصل السادس الطرق المختلفة لإيجاد المقدِّرات النقطية للمعلمة أو المعالم غير المعلومة لتوزيع المجتمع بناءً على معطيات عينة عشوائية. وأعطيت أهمية خاصة لطريقة المعقولية العظمى لاعتبارها أهم وأكثر الطرق استخداماً في الإحصاء لتقدير المعالم لما يتمتع به من خواص وخواص مقاربة.

ويناقش الفصل السابع طرق التقدير بفترة للمعالم غير المعلومة لتوزيع المجتمع الذي سحبت منه العينة العشوائية. وأعطيت الأهمية لبناء فترات الثقة في حالة المعاينة من مجتمع طبيعي.

يتناول الفصل الثامن، والأخير، طرق تقدير بيــيز بنقطــة وبفــترة للمعــالم غــير المعلومة في صيغة توزيع المجتمع الذي سحبت منه عينـــة عشـــوائية، والـــتي ينظـــر إليـــها كـــقيم لمتغيرات عشـــوائية.

إضافة إلى ما سبق، يعالج الكتـــاب، باهتمـــام خـــاص، أمثلـــة كثـــيرة معروضـــة بطريقة موسعة بغية إعطاء القارئ غير المتـــآلف مـــع تقنيـــات نظريـــة التقديـــر تمثيـــلاً محسوساً للأسس الرياضية لهذه النظرية ودليلاً للتطبيق على حـــالات مـــن الواقـــع، كمـــا يشتمل أيضاً على عدد لا بأس به من التمارين غـــير المحلولـــة.

ونرجو أن يكون في الأسلوب الذي قدمت به المـــادة والنـــهج الـــذي اســـتخدم في عرض الأمثلة الكثيرة والمتنوعة ما يمكن القارئ من اكتســـاب المعرفـــة بنظريـــة التقديـــر، ولعل في تقديم المصطلحات باللغتين العربية والإنجليزية ضمن ســـياق النـــص مـــا ييســـر للقارئ فهم المادة ومتابعة المراجع والبحوث الأجنبية الـــتي لا غـــنى عنـــها، وخاصـــةً أن لبساً يحدث أحياناً في بعض المؤلفات عند اختيار المصطلح العلمي العـــربي المناســـب لمـــا يقابله من المصطلحات الإنجليزيـــة.

أخيراً، أود أن أعبر عن شكري الجزيل لكل الذيـــن ســـاعدوني بصفـــة أو بـــأخرى على إخراج هذا الكتاب، وأخص بالذكر الأخ عبدالرحمـــن الفيتـــوري الـــذي لم يدخـــر جهداً في تنسيق وإخراج الكتـــاب.

والله من وراء القصد، فهو نعم المولى ونعم النصير

أ. د. عبدالحفيظ مصطفى
مصراتة في 1999/3/22

تمهيد

1. النموذج الإحصائي الاحتمالي ومهام الإحصاء الرياضي

يفترض عند دراسة نظرية الاحتمالات، أن احتمالات الحوادث الأساسية معلومة. وانطلاقاً من هذه الفرضيات يبنى النموذج الاحتمالي أو الرياضي $(\Omega, \mathcal{A}, P)$ للتجربة المفروضة، حيث Ω فضاء الحوادث الأولية، $\mathcal{A}$ صف من الفئات الجزئية (الحوادث) لـ Ω والتي تحقق شروط σ-الجبرا، P القياس الاحتمالي المعرف على الحوادث $A \in \mathcal{A}$. وعلى أساس هذا النموذج يتم حساب احتمالات الحوادث الأخرى (غير الأساسية) المختلفة.

وبمعرفة النموذج الاحتمالي $(\Omega, \mathcal{A}, P)$، يمكننا بناء النموذج الاحتمالي $\left(\Omega_\xi, \mathcal{A}_\xi, P_\xi\right)$ الذي يولده المتغير العشوائي ξ، الذي يصف التجربة المفروضة. هكذا ببناء النموذج الرياضي للتجربة المفروضة، الموصوفة بالمتغير العشوائي ξ نكون قد افترضنا معرفة توزيعه الاحتمالي، أي معرفة بنية المجتمع المدروس. وبناءً على ذلك تتم دراسة كيفية حساب احتمالات الحوادث التي يمكن أن تقع نتيجة التجربة المفروضة، بالإضافة إلى حساب المميزات العددية المختلفة للمتغير العشوائي ξ. لنوضح ما سبق بمثال يتعلق بمتغير عشوائي ξ خاضع لتوزيع ذي الحدين.

لتكن E تجربة احتمالية تتمثل بإلقاء قطعة نقود معدنية متناظرة (متجانسة) n مرة، تحت نفس الشروط. يبرهن في نظرية الاحتمالات، بافتراض قطعة النقود متناظرة والتكرارات مستقلة، بأن تكرار ظهور الصورة (المتغير العشوائي ξ) يخضع لتوزيع ذي الحدين $B(n,P)$، حيث $P = 0.5$. وعلى أساس هذه الافتراضات

يتم تشكيل النموذج الرياضي $\left(\Omega_\xi, \mathcal{A}_\xi, P_\xi\right)$، حيث $\Omega_\xi = \{1,2,\dots,n\}$ و $\mathcal{A}_\xi$ مجموعة الفئات الجزئية لـ Ω_ξ والتي تحقق شروط σ-الجبرا، أما القياس الاحتمالي P_ξ فيعرف بالعلاقة التالية:

$$P_\xi = P(\xi = m) = C_n^m p^m q^{n-m} \quad ; \quad p + q = 1 \quad , \quad m = 0,1,\dots,n$$

$$C_n^m = \binom{n}{m} = \frac{n!}{m!(n-m)!}$$

وباستخدام هذا النموذج، يمكننا حل المسائل التالية، التي تعتبر أساسية في نظرية الاحتمالات:

1. حساب احتمال ظهور الصورة m مرة بنتيجة التجربة E (إلقاء قطعة نقود معدنية متناظرة n مرة تحت نفس الشروط).
2. حساب احتمال ظهور الصورة أكثر من m_1 مرة وأقل من m_2 مرة بنتيجة التجربة E، حيث $m_1 < m_2$.
3. حساب $P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right)$، حيث $\frac{m}{n}$ التكرار النسبي لظهور الصورة، p احتمال ظهور الصورة في كل تكرار، و ε عدد حقيقي موجب.
4. إيجاد المميزات (الخواص) العددية للمتغير العشوائي ξ، وحل مسائل أخرى متعلقة بحساب احتمالات الحوادث المركبة.

هكذا، تدرس نظرية الاحتمالات احتمالات الحوادث التي يمكن أن تقع بنتيجة التجربة، عند تحقق جملة من الشروط (فمثلاً، عند تحقق الشرطين: قطعة النقود متناظرة، والتكرارات مستقلة). وعندها نعتبر الحوادث الأساسية (في حالة متغير عشوائي منقطع، تؤخذ الحوادث الأولية كحوادث أساسية) معلومة مسبقاً (قبل إجراء التجربة)، وباستخدام طرق نظرية الاحتمالات يتم حساب احتمالات الحوادث المطلوبة.

لكن، نادراً ما تكون بنية المجتمع معلومة، وبالتالي عند إجراء التجربة، لا نعلم

بالضرورة، هل جملة الشروط المفروضة محققة. فمثلاً، هل قطعة النقود متناظرة تماماً؟، هل التكرارات مستقلة فعلاً؟.

هكذا، عند دراسة تجربة مفروضة، قلما يكون الاحتمال P معلوماً تماماً، وغالباً يمكن فقط معرفة أن الاحتمال P عنصر في عائلة معينة من الاحتمالات $\boldsymbol{P}$ وهذه العائلة يمكن أن تشتمل على كل الاحتمالات، التي يمكن تعريفها على $\boldsymbol{A}$ (الحالة غير معينة تماماً). وفي حالات أخرى يمكن تمثيله بعائلة أصغر من الاحتمالات المعطاة في هذه أو تلك الصيغة الواضحة (الحالة، عند توفر معلومات محددة مسبقاً). وبشكل عام إن $\boldsymbol{P}$ عبارة عن عائلة الاحتمالات P المقترحة في الحالة المعينة (من أجل وصف التجربة المفروضة). إذا كانت العائلة $\boldsymbol{P}$ معينة، فيقال إننا نملك نموذج إحصائي احتمالي (أو اختصاراً نموذج إحصائي)، ويفهم منه الثلاثي $(\Omega,\boldsymbol{A},P)$ أو $(\Omega_\xi,\boldsymbol{A}_\xi,P_\xi)$، حيث $P\in\boldsymbol{P}$ و $P_\xi\in\boldsymbol{P}_\xi$.

فمثلاً، في تجربة إلقاء قطعة نقود معدنية n مرة، تحت نفس الشروط، من الواضح أن احتمال ظهور الصورة في الرمية الواحدة غير معلوم مسبقاً، والذي نرمز له بـ θ، وعندئذٍ يمكننا القول فقط بأن $\theta\in\Theta=[0,1]$، وبالتالي عائلة الاحتمالات المقترحة $\boldsymbol{P}$ تأخذ الشكل:

$$\boldsymbol{P}=\{P_\theta;\theta\in\Theta\}$$

حيث:

$$P_\theta(m)=C_n^m\theta^m(1-\theta)^{n-m}\qquad;\qquad m=0,1,\dots,n$$

وبناءً على النتائج العددية الملاحظة $x_1,x_2,\dots,x_n$ (في نظرية الاحتمالات لا يطلب إجراء التجربة فعلياً)، يمكن حل المسائل التالية، التي تعتبر أساسية في الإحصاء الرياضي:

1. اختبار الفرضية (testing hypotheses): إن تكرار ظهور الصورة ع له توزيع ذي الحدين بالمعلمتين المعينتين p,n، أي أن الفرضية الصفرية

$H_0{:}\xi \in B(n,p)$، هذه المسألة في الإحصاء الرياضي تدعى باختبار الفرضية الخاصة بنموذج التوزيع الاحتمالي، ومنها بدالة التوزيع $F(x)$ وتدعى مثل هذه المسألة باختبار الفرضية اللامعلمية (nonparametric).

2. اختبار الفرضية: أن قطعة النقود المعدنية متناظرة، أي أن الفرضية الصفرية خاصة بمعلمة التوزيع $H_0{:}\theta = 0.5$، وتدعى هذه المسألة باختبار الفرضية الخاصة بمعلمة توزيع المتغير العشوائي ξ، أو اختصاراً الفرضية المعلمية (parametric).

3. بناءً على نتائج التجربة المحققة تقدير الاحتمال المجهول θ (احتمال ظهور الصورة)، أي إيجاد قيمة تقريبية للمعلمة θ في التوزيع الثنائي $B(n,\theta)$، وهذه القيمة التقريبية يمكن إيجادها بشكل عدد، الذي يدعى بالتقدير النقطي، (التقدير بنقطة). وإما بشكل فترة (θ_1,θ_2) التي تغطي الاحتمال المجهول θ لظهور الصورة باحتمال كبير، وتدعى هذه المسألة في الإحصاء الرياضي بإيجاد التقديرات بنقطة وبفترة للمعلمة غير المعلومة في النموذج الإحصائي المقترح على الترتيب.

بالإضافة إلى المسائل الواردة أعلاه للإحصاء الرياضي، هناك مسائل أخرى عديدة تواجه الباحث. فمثلاً، هل قطعة النقود ألقيت بنزاهة؟ هل التكرارات مستقلة؟.

مما سبق نخلص إلى القول بأن الإحصاء الرياضي يُقيِّم بنية النموذج الإحصائي المقترح لدراسة وتحليل التجربة الاحتمالية: اختبار صحة اختيار نموذج التوزيع المقترح (ثنائي، بواسون، طبيعي، ... الخ)، تقدير معالم التوزيع المقترح (بعد التأكد من صحة اختيار النموذج)، اختبار الفرضيات الخاصة بمعالم التوزيع المقترح، وذلك كله يتم بناءً على النتائج الملاحظة للتجربة المفروضة. ولإجراء ذلك التقييم لبنية النموذج الإحصائي يجب معرفة كيفية إجراء وتخطيط (تصميم) التجربة، كيفية وصف نتائجها، كيفية تحليل النتائج التجريبية والتنبؤ بنتائج التجربة عند معرفة جملة من الشروط.

إن المسائل الواردة أعلاه (تخطيط، وصف، تحليل وتنبؤ) تعتبر المسائل الأساسية في الإحصاء الرياضي، وفي كتابنا هذا سوف ننطلق من اعتبار أن التجربة الاحتمالية أجريت ويطلب وصف نتائجها وتحليلها للوصول إلى التنبؤات واتخاذ القرارات.

هكذا، النموذج الإحصائي يصف تلك الحالات، التي توجد في النموذج الاحتمالي للتجربة المدروسة، عدم تعيين في الاحتمال P، ومهمة الإحصاء الرياضي تكمن في تقليص عدم التعيين هذا باستخدام المعلومات التي يتم الحصول عليها بملاحظة النتائج التجريبية (البيانات الإحصائية).

وبشكل محدد يعالج الإحصاء الرياضي المسائل التي تعتبر معاكسة لمسائل نظرية الاحتمالات، فهو يستكشف بنية النموذج الإحصائي بناءً على نتائج الملاحظات، وفي هذا الكتاب ينظر إلى أكثر المسائل الإحصائية أساسية والطرق العامة لحلها.

إن نشوء وتطور الإحصاء الرياضي، كبقية فروع الرياضيات، ارتبط بمتطلبات التطبيق، وفي وقتنا الحاضر طرقه تستخدم بشكل واسع في مختلف الفروع التقنية، ويلعب دورا هاماً في الأبحاث الاقتصادية، الزراعية، البيولوجية، الطب، العلوم الفيزيائية، الجغرافيا، علم النفس والأبحاث الاجتماعية، وغيرها، التي كان ينظر إليها سابقاً على أنها أكثر العلوم بعداً عن الرياضيات.

2. المصطلحات والرموز

تمثل المعطيات الإحصائية الأولية، في أغلب الحالات، نتائج ملاحظة فئة منتهية من المتغيرات العشوائية $\boldsymbol{X}=\left(X_1,X_2,...,X_n\right)$، التي تصف التجربة المدروسة. وفي مثل هذه الحالة يقال عادة بأن التجربة تتألف من n تكراراً، ونتيجة التكرار i توصف بالمتغير العشوائي $X_i \;\; ; \;\; i=1,2,....,n$، وتدعى فئة المتغيرات العشوائية الملاحظة بالعينة (sample)، والمتغيرات العشوائية $X_i \;\; ; \;\; i=1,2,....,n$ بعناصر أو مركبات العينة وعددها n بحجم العينة (sample size). ونرمز لقيم $\boldsymbol{X}$ بـ:

$$x = (x_1, x_2, \ldots, x_n)$$

كما ونرمز بـ $x(k)$ للقيمة ذات الترتيب k (من حيث الكبر) ضمن القيم $x_1, \ldots, x_n$.

لتكن $G = \{x\}$ مجموعة القيم الممكنة للمتغير العشوائي $\boldsymbol{X}$. تدعى الفئة G في الإحصاء الرياضي بفضاء العينة *(sample space)*، وعندئذٍ يعرف من G الصف $\boldsymbol{A}$ من الفئات الجزئية التي تشكل σ-الجبرا، ويمكن أن يكون فضاء العينة G مساوياً لـ R^n أو جزءاً منه (في حالة $\boldsymbol{X}$ متغير عشوائي مستمر)، وإما أن يتألف من عدد منتهي أو غير منتهي قابل للعد من النقاط في R^n (في حالة $\boldsymbol{X}$ متغير عشوائي منقطع). وفي هذه الحالة يفهم النموذج الإحصائي على أنه الثنائي $(G, \boldsymbol{P})$، حيث $\boldsymbol{P}$ عائلة التوزيعات المحتملة (المقترحة) للمتغير العشوائي $\boldsymbol{X}$ المعرفة على G. وكما نعلم من نظرية الاحتمالات بأن التوزيع الاحتمالي لأي متغير عشوائي يتعين تماماً بدالة توزيعه، لهذا عادة يعطى النموذج الإحصائي بدلالة عائلة دوال التوزيع المقترحة للعينة $\boldsymbol{X}$، أي بالثنائي $(G, \boldsymbol{F})$، حيث $\boldsymbol{F}$ عائلة دوال التوزيع، التي تنتمي إليها دالة التوزيع المجهولة:

$$F_X = P(X_1 < x_1, \ldots, X_n < x_n) \quad ; \quad -\infty < x_1 < \ldots. < x_n < \infty$$

للعينة $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$

يفترض غالباً أن المركبات X_i , $i = 1,2,\ldots,n$ مستقلة ولها توزيع واحد وهو توزيع متغير عشوائي ما ξ، وهذه الحالة توافق التجربة، التي تتألف من الملاحظات المتكررة والمستقلة على متغير عشوائي ξ، أي أن:

$$F_i = F_{X_i}(x_i) = F_\xi(x_i) \quad ; \quad i = 1,2,\ldots.,n$$

و

$$F_X(x) = F_\xi(x_1) \ldots . F_\xi(x_n)$$

وعندئذٍ يقال بأن العينة $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ مأخوذة من توزيع المتغير العشوائي ξ، وتدعى فئة القيم الممكنة للمتغير العشوائي ξ ذو التوزيع $F_\xi(x)$

بمجتمع الصفة المدروسة أو المجتمع ذو التوزيع $F_\xi(x)$ و $\boldsymbol{X}$ بالعينة المأخوذة من هذا المجتمع. سنرمز لتوزيع المتغير العشوائي ξ بـ $\mathcal{L}(\xi)$، وبالتالي يمكن أن نكتب "$\boldsymbol{X}$ عينة من $\mathcal{L}(\xi)$".

سنقتصر في هذا الكتاب، وللأهمية الخاصة، على دراسة مثل تلك النماذج، على الرغم من أنه في مسائل عدة شرط الاستقلالية والتوزيع الواحد للمركبات $X_i \ ; \ i=1,2,\ldots,n$ غير محقق. ونرمز للنموذج الإحصائي الموافق للملاحظات المتكررة المستقلة باختصار $\boldsymbol{F}=\{F_\xi\}$، أي بالإشارة فقط إلى صف دوال التوزيع المحتملة للمتغير العشوائي الملاحظ ξ، وعندما لا يوجد التباس يمكن إسقاط الدليل ξ ونكتب $\boldsymbol{F}=\{F\}$. وإذا كانت دوال التوزيع $F_\xi \in \boldsymbol{F}$ معرفة بدلالة معلمة θ فئة قيمها الممكنة Θ، فمثل هذا النموذج نرمز له بـ $\boldsymbol{F}=\{F(x,\theta) \ ; \ \theta \in \Theta\}$، ويدعى بالنموذج المعلمي. ويقال في هذه الحالة إن نوع توزيع المتغير العشوائي الملاحظ ξ معلوم، ويمكن أن يكون أحد التوزيعات الاحتمالية الشهيرة (ثنائي، بواسون، طبيعي، .... الخ)، لكنه يعتمد على المعلمة θ غير المعلومة، التي يتبع لها التوزيع، ويمكن أن تكون المعلمة θ سلمية (عددية) أو متجه، وتدعى فئة قيمها الممكنة Θ بفضاء المعلمة (parametric space).
فمثلاً، لنفترض أن $\mathcal{L}(\xi)$ توزيع طبيعي بتباين معلوم σ^2 ومتوسط مجهول θ، أي أن $\mathcal{L}(\xi) \in N(\theta,\sigma^2)$، عندها النموذج الإحصائي يكتب:

$$\boldsymbol{F}=\{F(x,\theta) \ ; \ \theta \in \Theta=(-\infty,+\infty)\}$$

أي أن فضاء المعلمة يمثل بمحور الأعداد الحقيقية، وصيغة كثافة التوزيع الموافقة لـ ξ تكون:

$$f(x,\theta)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\theta)}{2\sigma^2}} \quad ; \quad -\infty < x < +\infty$$

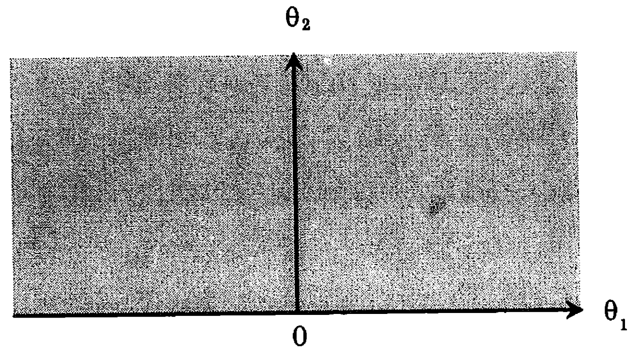
وإذا كان التباين مجهولاً أيضاً: فإن النموذج الإحصائي يأخذ الشكل:

$$\boldsymbol{F} = \left\{F(x,\theta)\ ;\ \boldsymbol{\theta} = (\theta_1,\theta_2) \in \Theta\right\}$$

حيث:

$$\Theta = \left\{\boldsymbol{\theta} = (\theta_1,\theta_2)\ ;\ -\infty < \theta_1 < +\infty\ ,\ 0 < \theta_2 < \infty)\right\}$$

وتمثل بالمنطقة المظللة على الشكل (1).



شكل (1)

وصيغة كثافة التوزيع الاحتمالي الموافقة:

$$f(x,\theta) = \frac{1}{\theta_2\sqrt{2\pi}}\, e^{-(x-\theta)^2/2\theta_2^2} \quad ; \quad -\infty < x < \infty$$

يدعى النموذج $\boldsymbol{F} = \left\{F_\xi\right\}$ بالمستمر (المنقطع)، إذا كانت كل عناصر العائلة $\boldsymbol{F}$ من الدوال المستمرة (المنقطعة).

وفيما يلي سنستخدم رمزاً موحداً $f_\xi(\theta) = f(x;\theta)$ (في حالة النماذج الإحصائية المعلمية $F(x,\theta)$) للدلالة على صيغة كثافة التوزيع الاحتمالي للمتغير العشوائي ξ في حالة النموذج المستمر، وللدلالة على صيغة دالة الاحتمال $P_\theta(\xi = x)$ للمتغير العشوائي ξ في حالة نموذج منقطع.

لنورد المثال التالي الخاص بنموذج معلمي منقطع:

لنفترض أن $\mathcal{L}(\xi)$ ينتمي إلى عائلة توزيعات بواسون، عندها يأخذ النموذج الإحصائي الشكل:

$$\boldsymbol{F} = \{F(x,\theta) \ ; \ \theta \in \Theta = (0,\infty)\}$$

حيث $F(x,\theta)$ يتعين بناءً على الاحتمالات:

$$f(x,\theta) = P(\xi = x) = \frac{\theta^x}{x!}e^{-\theta} \quad ; \quad x = 0,1,2,....$$

إن توزيع الاحتمالات على فضاء العينة G الموافق للمعلمة θ (في حالة نموذج معلمي) نرمز له بـ P_θ، وبشكل مشابه نضيف الدليل θ إلى رمز التوقع والتباين ...الخ، وهذا يعني أن المقدار الموافق يحسب للتوزيع P_θ. فمثلاً: $E_\theta T(X)$ و $\operatorname{var}_\theta(T(X)) = V_\theta T(X)$ يرمزان على الترتيب لتوقع وتباين الدالة $T(X)$ في العينة X، في حالة كون دالة توزيع العينة $F(x,\theta)$، وعندئذٍ لوحدة الرمز نستخدم الكتابة المختصرة بدلالة تكامل ستيلتج:

$$E_\theta T(X) = \int T(x) \, dF_X(x,\theta)$$

وهنا يفهم التكامل على أنه تكامل ذو n بعد على كامل فضاء العينة G، والذي في حالة نموذج مستمر يحسب كتكامل ريمان العادي:

$$\int.....\int T(x_1,.....,x_n) \, f(x_1,\theta).....f(x_n,\theta)dx_1...dx_n$$

وفي حالة نموج منقطع يحسب كمجموع:

$$\sum_{x_1}.....\sum_{x_n} T(x_1,.....,x_n) \, f(x_1,\theta).....f(x_n,\theta)$$

نحتاج في مسائل عدة، بحث متتالية من المتغيرات العشوائية $\{\eta_n\}$ المتقاربة من نهاية محددة η (متغير عشوائي أو ثابت)، عندما $n \to \infty$، وفي عرضنا القادم

سنتعرض لنوعين من التقارب: التقارب بالاحتمال، والتقارب بالتوزيع، ونشير إلى أن المتتالية $\{\eta_n\}$ تدعى متقاربة بالاحتمال من η إذا حققت الشرط التالي:

$$\lim_{n\to\infty} P(|\eta_n - \eta| > \varepsilon) = 0 \quad ; \quad \varepsilon > 0$$

ونكتب ذلك باختصار:

$$\eta_n \xrightarrow{P} \eta$$

ويفهم التقارب بالتوزيع أو اختصاراً $\mathcal{L}(\eta_n) \longrightarrow \mathcal{L}(\eta)$ على أنه تقارب دوال التوزيع الموافقة في كل نقطة باستمرار من دالة النهاية، عندما $n \to \infty$.

أما بقية الرموز والمفاهيم فستعرض في سياق النص.

3. رموز بعض النماذج الإحصائية الأساسية

إن التوزيعات التي نصادفها غالباً في نظرية الاحتمالات لها تسميات ورموز عامة. فمثلاً، التوزيع الطبيعي بمتوسط معلوم μ وتباين معلوم σ^2 يرمز له بالرمز $N(\mu,\sigma^2)$. توزيع ذي الحدين الموافق لـ n و P يرمز له بـ $B(n,P)$، توزيع بواسون بمتوسط λ يرمز له بـ $\Pi(\lambda)$ ... الخ. لهذا إذا كان نوع توزيع المتغير العشوائي الملاحظ ξ معلوم (طبيعي، ثنائي، بواسون، ... الخ)، فإن النموذج الإحصائي الموافق يحمل أيضاً ذلك الاسم. فمثلاً: يقال نموذج طبيعي، نموذج ذي الحدين، نموذج بواسون، ... الخ.

يتضمن الجدول (1) أكثر النماذج الإحصائية استخداماً في التطبيقات (الدالة f تساوي الصفر من أجل قيم x غير الواردة في الجدول 1). فمثلاً، $\mathcal{L}(\xi) \in N(\theta_1,\theta_2^2)$ تعني أن النموذج معطى بفئة دوال الكثافة الطبيعية المحتملة $\{f(x;\theta)\ ;\ \theta \in \Theta\}$، حيث $f(x;\theta)$ كثافة التوزيع الطبيعي و θ المعلمة المجهولة من فضاء المعلمة Θ. ويبين الجدول التالي أهم النماذج الإحصائية.

جــــــدول (1)

بعض النماذج الإحصائية الأساسية

اسم النموذج	رمــز النموذج	دالة الكثافة الاحتمالي أو دالة الاحتمال $f(x,\theta)$	فضاء المعلمة θ
الطبيعي I	$N(\theta,\sigma^2)$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2}$; $-\infty < x < +\infty$	$\{\theta : -\infty < \theta < +\infty\}$
الطبيعي II	$N(\mu,\theta^2)$	$\frac{1}{\theta\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\theta}\right)^2}$; $-\infty < x < +\infty$	$\{\theta : 0 < \theta < +\infty\}$
الطبيعي العام	$N(\theta_1,\theta_2^2)$	$\frac{1}{\theta_2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\theta_1}{\theta_2}\right)^2}$; $-\infty < x < +\infty$	$\left\{\theta = (\theta_1,\theta_2) , -\infty < \theta_1 < +\infty , 0 < \theta_2 < +\infty\right\}$
جــــاما	$\Gamma(\alpha,\theta)$	$\frac{x^{\alpha-1}e^{-x/\theta}}{\Gamma(\alpha)\theta^{\alpha}}$, $x > 0$	$\{\theta : 0 < \theta < +\infty\}$
المنتظم I	$R(0,\theta)$	$\frac{1}{\theta}$, $0 \le x \le \theta$	$\{\theta : 0 < \theta < +\infty\}$
المنتظم العام	$R(\theta_1,\theta_2)$	$\frac{1}{\theta_2-\theta_1}$, $\theta_1 \le x \le \theta_2$	$\{\theta = (\theta_1,\theta_2) , -\infty < \theta_1 < \theta_2 < +\infty\}$
كوشي	$K(\theta)$	$\frac{1}{\pi}\frac{1}{1+(x-\theta)^2}$, $-\infty < x < +\infty$	$\{\theta : -\infty < \theta < +\infty\}$
ذو الحدين	$B(n,\theta)$	$C_n^x\theta^x(1-\theta)^{n-x}$, $x = 0,1,\ldots,n$	$\{\theta : 0 < \theta < 1\}$
البواسوني	$\Pi(\theta)$	$\frac{\theta^x}{x!}e^{-\theta}$, $x = 0,1,\ldots$	$\{\theta : 0 < \theta < +\infty\}$
الثنائي السالب	$\overline{B}(r,\theta)$	$C_{r+x-1}^r\,\theta^x(1-\theta)^x$, $x = 0,1,2,\ldots$	$\{\theta : 0 < \theta < 1\}$
كاي مربع	χ_v^2	$\frac{1}{2^{v/2}\Gamma(v/2)} x^{\frac{v}{2}-1} e^{-\frac{x}{2}}$; $x > 0$	$\{v ; v = 1,2,\ldots\}$

تابع جدول (1)

اسم النموذج	رمــز النموذج	دالة الكثافة الاحتمالي أو دالة الاحتمال $f(x,\theta)$	فضاء المعلمة θ
بيتا	$Be(\upsilon_1,\upsilon_2)$	$\frac{1}{\beta(\upsilon_1,\upsilon_2)}x^{\upsilon_1-1}(1-x)^{\upsilon_2-1}\ ;\ 0<x<1$	$\{(\upsilon_1,\upsilon_2)\ ;\ \upsilon_1,\upsilon_2>0\}$
إفْ	F	$\frac{1}{\beta\left(\frac{\upsilon_1}{2},\frac{\upsilon_2}{2}\right)}x^{\frac{\upsilon_1}{2}-1}(1-x)^{\frac{\upsilon_2}{2}-1}\ ;\ 0<x<1$	$\{(\upsilon_1,\upsilon_2)\ ;\ \upsilon_1,\upsilon_2=1,2,\ldots\}$
تِ	$S(\upsilon)$	$\frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\sqrt{2\pi\upsilon}}\left(1+\frac{t^2}{\upsilon}\right)^{\frac{\upsilon+1}{2}}\ ;\ -\infty<t<\infty$	$\{\upsilon\ ;\ \upsilon=1,2,\ldots\}$

هنـا ولاحقـاً، عندمـا نكتـب $N(\theta,\sigma^2)$ ، نعـني بذلـك توزيـع طبيعـي بمتوسـط مجهـول وتبـاين معلـوم، وكذلـك نعـني بـالرمز $N(\mu,\theta^2)$ توزيـع طبيعـي بمتوسـط معلـوم وتبـاين مجهـول. وبعبـارة أخـرى فـالرمزين μ ، σ^2 ثوابـت معينـة، طالمـا لم نشـر إلى خلاف ذلـك.

وأخـيراً لا بـد لنـا مـن الإشـارة إلى أن هنـاك بعـض الحـالات الخاصـة للنمـاذج الـواردة في الجـدول رقـم (1) تحمـل تسـميات خاصـة، فمثـلاً النمـوذج $B(1,\theta)$ يدعـى بنمـوذج برنـولي، والنمـوذج $\overline{B}(1,\theta)$ بالهندسـي، ... الخ.

الفصل الأول:

مبرهنات النهاية

1.1 مقدمة

تدرس نظرية الاحتمالات القوانين التي تظهر عند تكرار التجارب الاحتمالية مرات عديدة تحت نفس الشروط. وتسمح دراسة مثل هذه القوانين بالتنبؤ بنتائج تلك التجارب، وعندما يكون عدد التكرارات كبيراً بكفاية، فإن بعض مميزات المتغيرات والحوادث العشوائية تصبح تقريباً غير عشوائية، ويقال عادة في مثل هذه الحالة أن المميزات تتمتع بصفة الاستقرارية (invariant). فمثلاً، التكرار النسبي يتمتع بهذه الصفة، أي إذا كانت شروط التجربة ثابتة أو ثابتة تقريباً، فإن التكرار النسبي يتأرجح حول قيمة ثابتة. وبعبارة أخرى، عندما تكون $n \to \infty$، فإن التكرار النسبي يفقد صفة العشوائية، وكذلك الوسط الحسابي لنتائج التكرارات $\overline{X} = \frac{1}{n}\sum X_i$ يتمتع أيضاً بصفة الاستقرارية عندما تكون n كبيرة.

يتطلب صف هام من المسائل في الإحصاء الرياضي، تعيين نهاية دوال معينة في n متغير عشوائي، وكذلك نهاية توزيعاتها عندما $n \to \infty$، وبعبارة أخرى، إذا كان $X = (X_1, \ldots, X_n)$ متغير عشوائي ذو n بعد و $g_n(X)$ دالة ما، التي تعتبر أيضاً متغير عشوائي، ويطلب تعيين نهاية $g_n(X)$ ونهاية توزيعها (التوزيع المقارب لها) عندما $n \to \infty$. إن مجموعة المبرهنات التي تعطي الحل لمثل هذه المسائل تدعى بمبرهنات النهاية (limit theorems). وهذه المبرهنات تصنف إلى نوعين:

1. قانون الأعداد الكبيرة: وهي مجموعة مبرهنات النهاية، التي تعين نهاية متوسط متغيرات عشوائية، عند شروط معينة، وهذه تندرج تحت اسم "قانون الأعداد الكبيرة (law of large numbers)".

2. مبرهنات النهاية المركزية: وهي مجموعة مبرهنات النهاية التي تعين نهاية توزيعات دوال في متغيرات عشوائية، تحت شروط معينة، وتندرج تحت اسم "مبرهنات النهاية المركزية (central limit theorems)".

يهدف هذا الفصل إلى تقديم بعض الصيغ الأساسية لقانون الأعداد الكبيرة على هيئة مبرهنات، بالإضافة إلى مبرهنات أساسية في النهاية المركزية.

سنتعرض في البداية، وبصورة موجزة لتقارب متتاليات متغيرات عشوائية.

2.1 تقارب متتاليات متغيرات عشوائية

CONVERGENCE OF SEQUENCES OF RANDOM VARIABLES

لتكن $X_1, X_2, \dots, X_n, \dots$ متتالية متغيرات عشوائية نرمز لها بـ $\{X_{(n)}\}$، المعرفة على الفضاء الاحتمالي $(\Omega, \mathcal{A}, P)$. بما أن المتغيرات العشوائية المشكلة لهذه المتتالية عبارة عن دوال في الحوادث الأولية $w \in \Omega$، باختيار $w_0 \in \Omega$ نحصل على المتتالية العددية $\{X_n(w_0)\}$، التي يمكن أن تكون متقاربة أو متباعدة.

تعريف 1.2.1

نقول عن متتالية المتغيرات العشوائية $\{X_n\}$ إنها متقاربة عند الحادث الأولي $w_0 \in \Omega$ إذا كانت المتتالية العددية $\{X_n(w_0)\}$ متقاربة.

تعريف 2.2.1

نقول إن متتالية المتغيرات العشوائية $\{X_n\}$ متقاربة عند الحادث $A \subset \Omega$، إذا كانت هذه المتتالية متقاربة عند كل حادث أولي $w \in A$.

تعريف 3.2.1

يقال عن متتالية المتغيرات العشوائية $\{X_n\}$ إنها متقاربة من المتغير العشوائي X عند الحادث A، إذا تحقق الشرط الآتي:

$$\lim_{n\to\infty} X_n(w) = X(w) \quad ; \quad \forall\, w \in A \tag{1.2.1}$$

مثال 1.2.1

لتكن متتالية المتغيرات العشوائية $\{X_n\}$، حيث $X_n(w) = w^n \;;\; n = 1,2,\ldots$ المعرفة على الفضاء $\Omega = [-1,1]$.

بما أن:

$$X(w) = \lim_{n\to\infty} X_n(w) = \lim_{n\to\infty} w^n = \begin{cases} 1 & ; \quad w = 1 \\ \\ 0 & ; \quad 0 \le w < 1 \end{cases}$$

وغير معرف عندما $-1 < w < 0$، فإن متتالية المتغيرات العشوائية $\{X_{(n)}\}$ تتقارب من $X = 1$ عندما $w = 1$ وتتقارب من $X = 0$ عند الحادث $A = [0,1)$.

وبصورة خاصة، إذا كان $A = \Omega$ فيقال بأن متتالية المتغيرات العشوائية $\{X_n\}$ متقاربة من X باحتمال يساوي الواحد أو تتقارب تقارباً شبه أكيد (almost certain convergence).

تعريف 4.2.1: التقارب شبه الأكيد Almost Certain Convergence

نقول عن متتالية المتغيرات العشوائية $\{X_n\}$ إنها متقاربة من X باحتمال يساوي الواحد (تقارب شبه أكيد)، إذا تحقق الشرط الآتي:

$$P\left(\left\{w \in \Omega \;;\; \lim_{n\to\infty} X_n(w) = X(w)\right\}\right) = 1 \tag{2.2.1}$$

ويرمز لهذا التقارب بـ:

$$P\left(X_n \longrightarrow X\right) = 1 \quad \text{أو} \quad P\left(X_n \not\longrightarrow X\right) = 0$$

أو في بعض الأدبيات الإحصائية بـ $X_n \xrightarrow{CP} X$

تعريف 5.2.1: التقارب بالاحتمال Convergence in Probability

نقول عن متتالية من المتغيرات العشوائية $\{X_n\}$ بأنها متقاربة بالاحتمال من المتغير العشوائي (أو غير العشوائي) X، إذا تحقق الشرط:

$$\lim_{n\to\infty} P\left(\left|X_n - X\right| \le \varepsilon\right) = 1 \quad ; \quad \varepsilon > 0 \tag{3.2.1}$$

ويرمز للتقارب بالاحتمال عادة بـ:

$$X_n \xrightarrow{P} X$$

نلاحظ عندما $X_n \xrightarrow{P} X$، فهذا يعني أن المتتالية $\{X_n\}$ قد تكون متقاربة من X عندما تقع بعض الحوادث الأولية، وقد تكون غير متقاربة عند وقوع بعض الحوادث الأولية الأخرى المرتبطة بالتجربة نفسها، فتقارب هذه المتتالية إذن هو حادث مرتبط بالتجربة المفروضة، وأن وقوع هذا الحادث يكافئ وقوع جميع الحوادث $\{\left|X_n - X\right| \le \varepsilon\}$ باستثناء عدد منته منها وذلك مهما يكن العدد الموجب ε. وبناءً على ذلك، وتعريف التقارب في التحليل الرياضي، نجد أن هنالك اختلافاً بين التقارب بالاحتمال والتقارب العادي، ويتمثل هذا الاختلاف بما يلي:

إذا كانت $\{X_n\}$ متقاربة، بمفهوم التحليل الرياضي، من X عندما $n \to \infty$، فهذا يعني بدءاً من قيمة معينة $n_0 \in N$:

$$\left|X_n - X\right| \le \varepsilon \quad ; \quad n > n_0 \tag{4.2.1}$$

أما إذا كانت المتتالية $\{X_n\}$ متقاربة بالاحتمال من X عندما $n \to \infty$، فإنه من أجل بعض القيم لـ n يمكن أن تكون المتباينة $\left|X_n - X\right| \le \varepsilon$ غير محققة.

ملاحظة 1.2.1

إذا كانت متتالية المتغيرات العشوائية $\{X_n\}$ متقاربة من X باحتمال يساوي الواحد، فهي بالضرورة متقاربة بالاحتمال من X. ويمكن إثبات ذلك بسهولة، ونتركه للقارئ كتمرين.

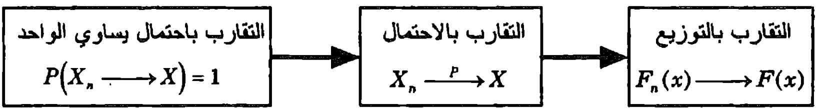
تعريف 6.2.1

نقول عن متتالية دوال توزيع $\{F_n(x) = P(X < x)\}$ إنها متقاربة بالتوزيع (convergence in distribution) من الدالة $F(x)$ ، إذا تحقق الشرط:

$$\lim_{n\to\infty} F_n(x) = F(x) \quad ; \quad \forall\, x \in R \qquad (5.2.1)$$

حيث إن $F(x)$ دالة مستمرة عند النقطة x. وكمثال على التقارب بالتوزيع نذكر تقارب توزيع ذي الحدين $B(n,P)$ من التوزيع الطبيعي $N(np,npq)$.

نلاحظ مما سبق أن هناك علاقة بين أنواع التقارب (التقارب بالاحتمال، التقارب باحتمال يساوي الواحد، التقارب بالتوزيع)، وهذه العلاقة ممثلة بالشكل (1.2.1).



شكل (1.1.1)

3.1 متباينة تشيبيشف TCHEBICHEV INEQUALITY

سنتطرق في هذا البند لمتباينة شهيرة تملك قيمة تطبيقية محدودة، لكنها ذات قيمة نظرية كبيرة لأنها تستخدم أساساً لإثبات مجموعة من المبرهنات المنتمية لقانون الأعداد الكبيرة. وهذه المتباينة تدعى بـ "متباينة تشيبيشف".

مبرهنــة 1.3.1

إذا كـان X متغـير عشـوائي حقيقـي و $g(X)$ دالـة حقيقـة غـير سـالبة، فـإن:

$$P\left(g(X) \geq k\right) \leq \frac{E[g(X)]}{k} \quad ; \quad k > 0 \tag{1.3.1}$$

الإثبـــــات

لنفـترض X متغـير عشـوائي مسـتمر دالـة كثافتـه $f(x)$، عندئـذٍ:

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_x(x) dx = \int_{\{x:g(x)\geq k\}} g(x) f_x(x) dx + \int_{\{x:g(x)<k\}} g(x) f_x(x) dx$$

$$\geq \int_{\{x:g(x)\geq k\}} g(x) f_x(x) dx \geq \int_{\{x:g(x)\geq k\}} k f_x(x) dx = kP\left(g(X) \geq k\right)$$

وبالقســمة علـى k نجـد:

$$P\left(g(X) \geq k\right) \leq \frac{E[g(X)]}{k}$$

وهـو المطلـوب.

وبشــكل مشــابه يمكــن الإثبــات في حالــة X متغــير عشــوائي منقطــع، وذلــك باسـتبدال التكـامل بـالمجموع.

متباينـة تشيبيشـف

إذا كان X متغيراً عشوائياً بمتوسط $EX = \mu$ وتباين منتهي $\operatorname{var} X = \sigma^2$، فإن:

$$P\left(|X - \mu| \geq \varepsilon\right) = P\left((X - \mu)^2 \geq \varepsilon^2\right) \leq \frac{\sigma^2}{\varepsilon^2} \quad ; \quad \varepsilon > 0 \tag{2.3.1}$$

الإثبـــــات

بوضع $g(X) = (X - \mu)^2$ و $k = \varepsilon^2$ في العلاقة (1.3.1) نحصل على العلاقة (2.3.1)

بما أن:

$$\Omega_X = \{|X-\mu| \geq \varepsilon\} \cup \{|X-\mu| < \varepsilon\}$$

والحادثان $\{|X-\mu| \geq \varepsilon\}$ و $\{|X-\mu| < \varepsilon\}$ متنافيان، فإن:

$$P(\Omega_X) = P(\{|X-\mu| \geq \varepsilon\}) + P(\{|X-\mu| < \varepsilon\}) = 1$$

ومنه:

$$P(\{|X-\mu| < \varepsilon\}) = 1 - P(\{|X-\mu| \geq \varepsilon\}) \geq 1 - \frac{\sigma^2}{\varepsilon^2} \qquad (3.3.1)$$

وهي صيغة مكافئة لـ (2.3.1).

بوضع $\varepsilon = r\sigma$ حيث $r > 0$، في العلاقة (3.3.1) نجد:

$$P(\{|X-\mu| < r\sigma\}) = P(\mu - r\sigma < X < \mu + r\sigma) \geq 1 - \frac{1}{r^2}$$

أي أن احتمال أن يأخذ المتغير العشوائي قيماً في الفترة $(\mu - r\sigma, \mu + r\sigma)$ لا يقل عن $1 - \frac{1}{r^2}$، فمثلاً، عندما $r = \frac{1}{2}$:

$$P(\mu - 2\sigma < X < \mu + 2\sigma) \geq \frac{3}{4}$$

وهذا يعني من أجل متغير عشوائي X، تباينه منتهي، فإن $\frac{3}{4}$ قيمه الممكنة على الأقل تقع في إطار إنحرافين معياريين عن متوسطه.

هكذا نجد أن متباينة تشيبيشف تطبق على أي متغير عشوائي X شريطة أن يكون تباينه منتهي، ولا نحتاج لمعرفة توزيع X. وهذه المتباينة لا تعطي طريقة عملية لتقدير الاحتمالات، لأن الفرق بين طرفي المتباينة (3.3.1) عادة كبير. وعندما $r \leq 1$ أي $\left(\frac{1}{r} \geq 1\right)$، فإن المتباينة (3.3.1) لا تفيدنا شيئاً.

مثـــــــال 1.3.1

لتكـن لدينـا شـحنة كبـيرة مـن بضاعـة مصنعـة، نسـبة العطـب فيهـا p، ونريـد تعيـين n، بحيـث إذا أخذنـا عينـة عشـوائية $(X_1, X_2, \dots, X_n)$ حجمهـا n أو أكـــثر مـــن هـــذه البضاعـة كنـا واثقـين باحتمـال لا يقـل عـن 0.95 بـأن نســبة العطـــب في العينـــة يختلـف عن P بـأقل مـن 0.1.

لـنرمز بـــ $i = 1,2,\dots,n$; X_i لنتيجـة السـحب رقـم i عنـد تشـكيل العينـة العشـوائية، حيـث $X_i = 0$ إذا كـانت الوحـدة المسـحوبة سـليمة، و $X_i = 1$ إذا كـانت معطوبـة. ولتكـن $X = \sum_1^n X_i$ عـدد الوحـدات المعطوبـة في العينـة العشـوائية المؤلفـة مـن n وحـدة. ويصبح المطلوب تعيين n بحيـث:

$$P\left(\left|\frac{X}{n} - p\right| < 0.1\right) \geq 0.95 \quad ; \quad EX = p$$

$$P\left(\left|\frac{X}{n} - p\right| < 0.1\right) = P(|X - np| < 0.1n) \Rightarrow$$

$$P(|X - np| \geq 0.1n) \leq 1 - 0.95 = 0.05 \qquad .....(1)$$

وحسـب متباينـة تشيبيشـف:

$$P(|X - np| \geq 0.1n) \leq \frac{\operatorname{var} X}{(0.1n)^2} = \frac{npq}{(0.1n)^2} \leq \frac{n/4}{(0.1n)^2} = \frac{25}{n} \qquad .....(2)$$

لأن: $pq \leq \frac{1}{4}$، وبمقارنـة (1)، (2) نجـد:

$$\frac{25}{n} \leq 0.05 \Rightarrow n \geq 500$$

مثـــــــال 3.3.1

لتكـن لدينـا تجربـة عشـوائية، احتمـال ظهـور الحـادث A يسـاوي $\frac{1}{2}$، أثبـت أن

عدد مرات ظهور الحادث A في 1000 تكراراً مستقلاً للتجربة المفروضة يقع بين 400 و 600 باحتمال لا يقل عن 0.97.

لنرمز بـ X لعدد مرات ظهور الحادث A في n تكراراً مستقلاً للتجربة المفروضة.

بما أن $n = 1000$, $\operatorname{var} X = npq = 250$ و $EX = np = 500$، فحسب متباينة تشيبيشف:

$$P\left(|X - EX| < \varepsilon\right) \geq 1 - \frac{\operatorname{var} X}{\varepsilon^2}$$

$$P\left(|X - 500| < 100\right) \geq 1 - \frac{250}{(100)^2} = 0.975$$

وهو المطلوب.

4.1 قانون الأعداد الكبيرة الضعيف

THE WEAK LAW OF LARGE NUMBERS

إن مجموعة المفاهيم المتعلقة بالتقارب بالإضافة إلى متباينة تشيبيشف الواردة أعلاه، تمكن من إيضاح محتوى قانون الأعداد الكبيرة.

يصوغ قانون الأعداد الكبيرة العلاقة بين المميزات النظرية والتجريبية للتجارب العشوائية. وهذا القانون له صيغ مختلفة تتجلى بمجموعة من المبرهنات (تشيبيشف، برنولي، بواسون، ...).

تعريف 1.4.1: قانون الأعداد الكبيرة الضعيف

نقول عن متتالية من المتغيرات العشوائية $\{X_n\}$ إنها تخضع لقانون الأعداد الكبيرة الضعيف إذا حققت الشرط:

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{1}^{n} X_i - \frac{1}{n}\sum_{1}^{n} EX_i\right| < \varepsilon\right) = 1 \quad ; \quad \varepsilon > 0 \qquad (1.4.1)$$

1.4.1 صيغة تشيبيشف لقانون الأعداد الكبيرة الضعيف

مبرهنة 1.4.1: مبرهنة تشيبيشف **Tchebichev Theorem**

إذا كانت متتالية المتغيرات العشوائية $X_1, X_2, \ldots, X_n, \ldots$ مستقلة مثنى مثنى، وذات تباينات $\operatorname{var} X_i = \sigma_i^2\ ;\ i = 1,2,\ldots$ منتهية ومحدودة بعدد ثابت c، فإن:

$$\lim_{n\to\infty} P\left(\left\{\left|\frac{1}{n}\sum_1^n X_i - \frac{1}{n}\sum_1^n EX_i\right| < \varepsilon\right\}\right) = 1 \quad ; \quad \varepsilon > 0 \tag{2.4.1}$$

الإثبات

حسب الفرض، المتغيرات $X_i\ ;\ i = 1,2,\ldots,n$ مستقلة وبالتالي:

$$\operatorname{var}\left(\frac{1}{n}\sum_1^n X_i\right) = \frac{1}{n^2}\sum_1^n \operatorname{var} X_i = \frac{1}{n^2}\sum_1^n \sigma_i^2$$

$$\leq \frac{nc}{n^2} = \frac{c}{n}$$

وبناءاً على متباينة تشيبيشف (العلاقة 3.3.1)

$$P\left(\left\{\left|\frac{1}{n}\sum_1^n X_i - \frac{1}{n}\sum_1^n EX_i\right| < \varepsilon\right\}\right) \geq 1 - \left[\frac{\operatorname{var}\left(\frac{1}{n}\sum_1^n X_i\right)}{\varepsilon^2}\right] \geq 1 - \frac{c}{n\varepsilon^2}$$

وبأخذ نهاية الطرفين، عندما $n \to \infty$، نجد:

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_1^n X_i - \frac{1}{n}\sum_1^n EX_i\right| < \varepsilon\right\} \geq 1$$

وبما أن الاحتمال لا يمكن أن يتجاوز الواحد، فنحصل على العلاقة المطلوبة.

ونحصل بالانتقال إلى الحادث المعاكس على صيغة أخرى مكافئة لـ (2.4.1):

$$\lim_{n\to\infty} P\left(\left\{\left|\frac{1}{n}\sum_{1}^{n} X_i - \frac{1}{n}\sum_{1}^{n} EX_i\right| \geq \varepsilon\right\}\right) = 0 \quad ; \quad \varepsilon > 0 \tag{3.4.1}$$

نلاحظ أن المتوسط $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ عبارة عن متغير عشوائي، بينما $\frac{1}{n}\sum_{1}^{n} EX_i$ مقدار ثابت.

وكحالة خاصة، إذا كان $EX_i = \mu \; ; \; i = 1,2,...$ ، فإن (2.4.1) تكتب:

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{1}^{n} X_i - \mu\right| < \varepsilon\right) = 1 \quad ; \quad \varepsilon > 0 \tag{4.4.1}$$

وهي تعطي الأساس لتقدير المتوسط الحسابي لمجتمع. فمثلا، لنفترض إجراء عملية قياس لمقدار فيزيائي a. بتكرار عملية القياس n مرة بشكل مستقل، تحت نفس الشروط، يحصل الملاحظ على النتائج $x_1,...,x_n$ غير المتطابقة تماماً، عندها يؤخذ المتوسط الحسابي للقيم الملاحظة كتقريب لـ a.

مثال 1.4.1

إذا كانت $X_1, X_2,..., X_n$ القياسات المستقلة لمقدار فيزيائي a، وكانت:

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - a)^2$$

هل يمكن اعتبار:

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - a)^2$$

قيمة تقريبية لتباين أخطاء جهاز القياس وليكن σ^2؟

يمكن اعتبار S_n^2 قيمة تقريبية لـ σ^2 ، إذا كان:

$$\lim_{n\to\infty} P\left(\left|S_n^2 - \sigma^2\right| < \varepsilon\right) = 1 \quad ; \quad \varepsilon > 0 \qquad ..... (1)$$

بما أن المتغيرات العشوائية $X_1, X_2, \ldots, X_n$ مستقلة ولها نفس التوزيع (حسب الفرض)، فإن المتغيرات العشوائية $Y_i = (X_i - a)^2\ ,\ i = 1, \ldots, n$ مستقلة ولها توزيع واحد.

$$EY_i = E(X_i - a)^2 = EX_i^2 - 2aEX_i + a^2 \quad ; \quad i = 1, \ldots, n$$
$$= \sigma^2 + \mu^2 - 2a\mu + a = \sigma^2 + (\mu - a)^2$$

حيث إن:

$$\mu = EX_i$$

ولكي تتحقق المساواة $EY_i = \sigma^2$ يجب أن يكون $\mu = a$، وهذا يعني عدم وجود أخطاء نظامية، وبالتالي حسب المبرهنة (2.4.1) فالعلاقة (1) محققة.

تبين مبرهنة تشيبيشف أن المتوسط الحسابي لعدد كبير من المتغيرات العشوائية المستقلة ذات تباينات منتهية ومحدودة بعدد ثابت يفقد صفة العشوائية، أي يقترب بالاحتمال من مقدار ثابت.

2.4.1 صيغة برنولي لقانون الأعداد الكبيرة الضعيف

مبرهنة 2.4.1: مبرهنة برنولي **Bernoulli's Theorem**

إذا كان X متغيراً عشوائياً يساوي عدد مرات ظهور حادث A في n تكراراً مستقلاً لتجربة عشوائية و p احتمال ظهور الحادث A في كل تكرار، فإن:

$$\lim_{n \to \infty} P\left(\left|\frac{X}{n} - p\right| < \varepsilon\right) = 1 \quad ; \quad \varepsilon > 0 \qquad (5.4.1)$$

الإثبات

يمكن اعتبار X متغير عشوائي مؤلف من مجموع n متغيراً عشوائياً.

$$X = \sum_1^n X_i$$

حيث X_i متغير عشوائي يساوي عدد مرات وقوع الحادث A في التكرار i، أي يفترض القيمتين 0,1 باحتمال p و $q = 1 - p$ على الترتيب.

بما أن:

$$EX_i = p \quad , \quad \operatorname{var} X_i = pq \leq \frac{1}{4} \;\; ; \;\; i = 1, \ldots, n$$

فبتطبيق مبرهنة (1.4.1) نجد:

$$\lim_{n \to \infty} P\left(\left|\frac{1}{n}\sum_{n}^{1} X_i - \frac{1}{n}\sum_{1}^{n} EX_i\right| < \varepsilon\right) = \lim_{n \to \infty} P\left(\left|\frac{X}{n} - p\right| < \varepsilon\right) = 1$$

وهو المطلوب.

عندما تكون n كبيرة بكفاية، فإن التكرار النسبي $W(A) = \frac{X}{n}$ يكون قريباً جداً من القيمة p (يفقد صفة العشوائية). وتكمن الأهمية التطبيقية لقانون الأعداد الكبيرة الضعيف بصيغة برنولي في أنه يعتبر أساس الطريقة الإحصائية لتقدير الاحتمالات.

تجدر الإشارة، إلى أن مبرهنة بيرنولي أهم صيغة، وأسبق تاريخياً، لقانون الأعداد الكبيرة، وهي تصوغ العلاقة بين التكرار النسبي لوقوع حادث واحتماله. وبرهانها كان صعباً، إلى أن جاء العالم تشيبيشف وقدم الإثبات الوارد أعلاه.

3.4.1 صيغة بواسون لقانون الأعداد الكبيرة الضعيف

مبرهنة 3.4.1: مبرهنة بواسون **Poisson's Theorem**

إذا كان في متتالية من التكرارات المستقلة، احتمال وقوع حادث A في التكرار رقم i يساوي P_i، فإن:

$$\lim_{n \to \infty} P\left(\left|\frac{X}{n} - \frac{1}{n}\sum_{1}^{n} P_i\right| < \varepsilon\right) = 1 \quad ; \quad \varepsilon > 0 \tag{6.4.1}$$

حيث إن X عدد مرات ظهور الحادث A في التكرارات الـ n الأولى.

الإثبات

لدينا:

$$X = \sum_1^n X_i \quad , \quad EX_i = P_i \quad , \quad \operatorname{var} X_i = p_i q_i \le \frac{1}{4}$$

حيث $i = 1,2,\ldots\ ;\ X_i$ عدد مرات ظهور الحادث A في التكرار رقم i.

بتطبيق مبرهنة تشيبيشف (1.4.1) نحصل على المطلوب

4.4.1 صيغة ماركوف لقانون الأعداد الكبيرة الضعيف

مبرهنة 4.4.1: مبرهنة ماركوف **Markov's Theorem**

إذا كانت متتالية المتغيرات العشوائية $\{X_n\}$، وكان:

$$\lim_{n\to\infty} \operatorname{var}\left(\frac{1}{n}\sum_n^1 X_i\right) = 0$$

فإن:

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_1^n X_i - \frac{1}{n}\sum_1^n EX_i\right| < \varepsilon\right) = 1 \quad ; \quad \varepsilon > 0 \qquad (7.4.1)$$

يمكن إثبات صحة هذه المبرهنة بسهولة بتطبيق مبرهنة تشيبيشف. وكحالة خاصة، إذا كانت المتغيرات العشوائية $X_1, X_2, \ldots, X_n, \ldots$ مستقلة مثنى مثنى، فإن شرط ماركوف يكتب على الصورة:

$$\lim_{n\to\infty}\left(\frac{1}{n^2}\sum_{i=1}^n \operatorname{var} X_i\right) = 0$$

نلاحظ أن مبرهنة تشيبيشف (1.4.1) حالة خاصة من مبرهنة ماركوف.

5.1 قانون الأعداد الكبيرة القوي

THE STRONG LAW OF LARGE NUMBERS

رأينا في البند (4.1) أن قانون الأعداد الكبيرة الضعيف، تحت شروط معينة، يبين أن متوسط متتالية المتغيرات العشوائية المستقلة تتقارب بالاحتمال من متوسط القيم المتوقعة لها. والسؤال الذي يطرح نفسه الآن: هل يمكن تقديم عبارة احتمالية حول متتالية المتوسطات $\left\{\overline{X}_n \; ; \; \overline{X}_n = \frac{1}{n}\sum_1^n X_i \; , \; n = 1,2,\ldots\right\}$؟ والجواب يعطى بقانون الأعداد الكبيرة القوي.

تعريف 1.5.1

نقول عن متتالية المتغيرات العشوائية $\{X_n\}$ بأنها تخضع لقانون الأعداد الكبيرة القوي إذا كان مهما يكن $\varepsilon > 0$ و $\delta > 0$ يمكن إيجاد عدد صحيح موجب n_0، بحيث يتحقق الشرط:

$$P\left(\left|\overline{X}_n - \sum_{i=1}^{n} EX_i\right| < \varepsilon \; ; \; n = n_0, n_0+1, \ldots, n_0+k\right) \geq 1-\delta \qquad (1.5.1)$$

من أجل كل قيم $n \geq n_0$ وكل قيم k.

متباينة كالماغوروف ***Kalmagorov Inequality***

إذا كانت $X_1, X_2, \ldots, X_n$ متغيرات عشوائية مستقلة وذات تباينات منتهية $\sigma_1^2, \sigma_2^2, \ldots, \sigma_v^2$، ولنفترض $\sum_1^n \sigma_i^2 = c_n$ و $EX_i = \mu_i \; ; \; i = 1,\ldots,n$ فإن:

$$P\left(\left|\sum_1^n (X_i - \mu_i)\right| < \varepsilon\right) \geq 1 - \frac{c_n}{\varepsilon^2} \quad ; \quad \varepsilon > 0 \qquad (2.5.1)$$

مبرهنة 1.5.1: مبرهنة كالماغوروف **Kalmagorov Theorem**

إذا كانت متتالية المتغيرات العشوائية $\{X_n\}$ مستقلة مثنى مثنى وتحقق الشرط:

$$\sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} < +\infty \qquad (3.5.1)$$

فإنها تخضع لقانون الأعداد الكبيرة القوي.

وبناءً على ذلك، إذا كانت متتالية المتغيرات العشوائية $\{X_n\}$ مستقلة مثنى مثنى وذات تباينات محدودة بعدد ثابت c، فإنها تخضع لقانون الأعداد الكبيرة القوي.

إحدى النتائج الهامة المتعلقة بقانون الأعداد الكبيرة القوي توصل إليها العالم كالماغوروف وهي:

"إذا كانت $\{X_n\}$ متتالية من المتغيرات العشوائية المستقلة مثنى مثنى ولها نفس التوزيع، فإن الشرط اللازم والكافي لكي تخضع هذه المتتالية لقانون الأعداد الكبيرة القوي يتمثل في وجود القيمة المتوقعة $EX_i = \mu$".

6.1 مبرهنات النهاية المركزية *CENTRAL LIMIT THEOREMS*

تطرقنا في البندين السابقين لمسائل تقارب بعض المتغيرات العشوائية إلى نهاية معينة، بغض النظر عن توزيعاتها الاحتمالية. ورأينا أن المعرفة المطلوبة للمتغيرات العشوائية موضوع الدراسة لا تتعدى قيمها المتوقعة وتبايناتها.

توجد مجموعة أخرى من المبرهنات في نظرية الاحتمالات تخص التوزيع الاحتمالي الذي ينتهي إليه توزيع المتغير العشوائي $Y_n = \sum_{i=1}^{n} X_i$، عندما $n \to \infty$.

وهذه المجموعة من المبرهنات هامة جداً في الإحصاء، وتحمل اسماً عاماً "مبرهنات النهاية المركزية"، وهي تبين الشروط العامة المفروضة على المتغيرات العشوائية $X_1,\ldots,X_n$ لكي يؤول مجموعها Y_n إلى التوزيع الطبيعي عندما $n \to \infty$.

لتكن $\{X_n\}$ متتالية من المتغيرات العشوائية المستقلة، ولنرمز بـ $EX_i = m_i$ ، $\operatorname{var} X_i = \sigma_i^2 \; ; \; i = 1,2,\ldots$ للقيمة المتوقعة والتباين لـ X_i على الترتيب، وكذلك بـ $\operatorname{var} Y_n = \sigma_{Y_n}^2 = \sum_{i=1}^{n} \sigma_i^2$ ، $EY_n = m_{Y_n} = \sum_{i=1}^{n} m_i$ للقيمة المتوقعة والتباين لـ Y_n. وعندئذٍ يمكن التعبير عن المحتوى العام لمبرهنات النهاية المركزية كالآتي:

بغض النظر عن توزيعات المتغيرات العشوائية $X_1,\ldots,X_n$ فإن توزيع مجموعها Y_n، تحت شروط معينة مفروضة على المتغيرات العشوائية X_i، ينتهي إلى التوزيع الطبيعي $N\left(m_{Y_n}, \sigma_{Y_n}^2\right)$، عندما $n \to \infty$.

1.6.1 مبرهنة النهاية المركزية بصيغة ليابونوف

Liapounov Central Limit Theorem

إذا كانت $\{X_n\}$ متتالية من المتغيرات العشوائية تحقق الشروط الآتية:

1. القيم المتوقعة $EX_i = m_i \, , \, i = 1,2,\ldots$ موجودة.
2. التباينات $\operatorname{var} X_i = \sigma_i^2 \; ; \; i = 1,2,\ldots$ موجودة.
3. العزوم المركزية المطلقة من المرتبة الثالثة $E\left(\left|X_i - m_i\right|^3\right)$ موجودة.
4.

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} E\left(\left|X_i - m_i\right|^3\right)}{\left(\sum_{i=1}^{n} \sigma_i^2\right)^{3/2}} = 0 \tag{1.6.1}$$

فـإن توزيــع المجمــوع $Y_n = \sum_{i=1}^{n} X_i$ ينتهــي إلى التوزيــع الطبيعــي $N\left(m_{Y_n}, \sigma_{Y_n}^2\right)$. عندمـا $n \to \infty$. وهـذا يعـني عندمـا تكـون n كبـيرة بكفايــة فـإن التوزيــع التقريــبي لــ Y_n هـــو $N\left(m_{Y_n}, \sigma_{Y_n}^2\right)$.

نلاحـظ مـن شـرط ليــابونوف (1.6.1) أن:

$$\lim_{n\to\infty} \frac{\sigma_i^2}{\sum_{i=1}^{n} \sigma_i^2} = 0 \tag{2.6.1}$$

وهـذا يعـني أن تبـاين كـل متغـير عشـوائي X_i يدخـل في تشـكيل المتغــير العشــوائي $Y_n = \sum_{i=1}^{n} X_i$ يمثـل نسـبة صغـيرة مـن تبــاين Y_n. وبعبــارة أخــرى، بغــض النظــر عــن اختـلاف تباينـات المتغــيرات العشــوائية X_i، فـإن هــذه التباينــات يجــب أن لا تختلــف عـن بعضهـا بشـكل كبـير، مـن حيـث تأثيرهـا علـى تبـاين المتغـير العشـوائي Y_n، وعندئـذٍ يقـال: لا يسـود في المجمـوع Y_n أي متغـير مـن المتغــيرات العشــوائية X_i.

وإذا رمزنــا بـــ $Z_n = \dfrac{Y_n - m_{Y_n}}{\sigma_{Y_n}}$، فـإن Z_n يخضــع تقريبــاً للتوزيــع الطبيعــي المعيـاري، عندمـا تكـون n كبـيرة.

إثبـات مبرهنـة ليـابونوف معقـد، ولا مجـال لذكـره هنـا، لــذا سـنقبل هــذه المبرهنـة دون إثبـات. وسـنكتفي بإثباتهـا في حالـة خاصـة كافيـة مـن أجــل معظـم القضايـا الــتي نحتاجهـا، بالإضافـة إلى أنهـا كافيـة لمعظـم التطبيقـات الإحصائيــة.

إذا كــانت $\{X_n\}$ متتاليــة مــن المتغــيرات العشــوائية المســتقلة محققــة للشــروط الآتيـة:

1. تخضـع لتوزيــع احتمـالي واحـد.
2. القيــم المتوقعــة $EX_i = m \, , \, i = 1,2,\ldots$ موجــودة.
3. التباينـــات $\operatorname{var} X_i = \sigma^2 \, , \, i = 1,2,\ldots$ موجـــودة.

فإن توزيـع المتغـير العشـوائي:

$$Z_n = \frac{Y_n - nm}{\sigma\sqrt{n}} \quad ; \quad Y_n = \sum_{i=1}^{n} X_i$$

ينتهـي إلى التوزيـع الطبيعـي المعيـاري $N(0,1)$ عندمـا $n \to \infty$.

الإثبـــات

بمـا أن المتغـيرات العشـوائية $X_i \; ; \; i = 1,2,\ldots$ مسـتقلة (حسـب الفـرض) فـإن المتغـيرات العشـوائية $(X_i - m_i) \; ; \; i = 1,2,\ldots$ مسـتقلة بدورهـا أيضـاً.

لـنرمز بـ $\varphi_i(t) \; ; \; i = 1,2,\ldots$ للدالـة المميـزة للمتغـير العشـوائي $(X_i - m_i)$، وبـ $\varphi_n(t)$ للدالة المميزة للمتغير العشـوائية Z_n، ولتكـن $F_n(x)$ دالـة توزيـع Z_n.

بنـاءً على خـواص الدالـة المميزة، فـإن الدالـة المميزة للمتغـير العشـوائي $\frac{X_i - m_i}{\sigma\sqrt{n}}$ ولنرمز لهـا بـ $\varphi_i^*(t)$

$$\varphi_i^*(t) = \varphi_i\left(\frac{t}{\sigma\sqrt{n}}\right) \quad ; \quad i = 1,2,\ldots$$

ومن ثـم:

$$\varphi_n(t) = \left[\varphi_i\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n$$

وحيـث إن العزمين المركزيـين الأول والثـاني همـا 0 و σ^2 علـى الـترتيب، فـإن:

$$\varphi_i(t) = 1 - \frac{1}{2}\sigma^2 t^2 + o(t^2)$$

وباستبدال t بـ $\frac{t}{\sigma/\sqrt{n}}$ نجـد:

$$\varphi_i\left(\frac{t}{\sigma/\sqrt{n}}\right)=1-\frac{t^2}{2n}+\frac{R(n,t)}{n}$$

حيث إن $\lim\limits_{n\to\infty} R(n,t)=0$ وذلك من أجل t ثابتة. وبالتالي:

$$\varphi_n(t)=\left[1-\frac{t^2}{2n}+\frac{R(n,t)}{n}\right]^n$$

وبأخذ نهاية الطرفين، عندما $n\to\infty$ نجد:

$$\lim_{n\to\infty}\varphi_n(t)=e^{-\frac{1}{2}t^2}$$

وهذه ما هي إلا الدالة المميزة للتوزيع الطبيعي المعياري $N(0,1)$، أي أن:

$$\lim_{n\to\infty}F_n(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-\frac{1}{2}z^2}dz=\Phi(z)$$

وهذا يعني عندما تكون n كبيرة فإن:

$$F_n(z)\approx\Phi(z)\ ;\ -\infty<z<+\infty$$

ومن ثم فإن التوزيع التقريبي لـ $Y_n=\sum\limits_{i=1}^{n}X_i$ عندما تكون n كبيرة، هو التوزيع الطبيعي $N(nm,n\sigma^2)$.

فمثلاً، عند قياس مقدار فيزيائي ما a، نتائج القياس تمثل متغيراً عشوائياً ع، وبتكرار قياس المقدار a، تحت نفس الشروط، نحصل على متتالية $\{X_n\}$ من المتغيرات العشوائية المستقلة. إذا كانت القياسات لا تتضمن أخطاء نظامية، فإن القيمة المتوقعة لـ $X_i\ ;\ i=1,2,\ldots$ هي عبارة عن القيمة الحقيقية لقياس المقدار a، أي أن $EX_i=a$. وعندئذٍ تباينات المتغيرات X_i تصف دقة القياسات، وبما أننا لا نميز بين القياسات (لا نعتبر أي قياس هو الأساس)، فإن شروط مبرهنة

النهاية المركزية محققة، لذا فالوسط الحسابي للقياسات المأخوذة للمقدار a هو متغير عشوائي، وتوزيعه يقترب من التوزيع الطبيعي بازدياد عدد القياسات.

2.6.1 مبرهنة النهاية المركزية بصيغة موافر ولابلاس

أشرنا عند دراسة مبرهنة ليابونوف إلى أن المتغيرات العشوائية $X_i \; ; \; i = 1,\dots,n$، الداخلة في تشكيل المجموع $Y_n = \sum_{i=1}^{n} X_i$، تخضع لتوزيع احتمالي واحد، لكن لم نحدد نوع التوزيع (طبيعي، ذي الحدين، بواسون، ....)، أي يمكن أن تخضع لأي توزيع احتمالي. إذا افترضنا أنها تتبع توزيع برنولي بمعلمة p، فما هو التوزيع المقارب للمجموع $Y_n = \sum_{i=1}^{n} X_i$، وبعبارة أخرى ما هو التوزيع المقارب لتوزيع ذي الحدين؟ هذا ما تجيب عليه مبرهنة موافر ولابلاس.

مبرهنة موافر ولابلاس **Laplace-Moiver Theorem**

إذا كانت متتالية المتغيرات العشوائية $\{X_n\}$ مستقلة ولكل منها توزيع بيرنولي $B(1,p)$، فإن توزيع المجموع $Y_n = \sum_{i=1}^{n} X_i$ يتقارب من التوزيع الطبيعي $N(np,npq)$ بازدياد n، أي أن:

$$\lim_{n\to\infty} B(n,p) = N(np,npq)$$

الإثبات

كما نعلم

$$P(X_i = x_i) = p^{x_i} q^{1-x_i} \quad ; \quad x_i = 0,1 \quad , \quad q = 1-p \; ; \; 0 \le p \le 1$$

$$EX_i = \sum_{i=0}^{1} x_i p^{x_i} q^{1-x_i} = p$$

$$\operatorname{var} X_i = EX_i - (EX_i)^2 = pq + p^2 - p^2 = pq \le \frac{1}{4} \quad ; \quad i = 1,2,\dots$$

أي أن شروط مبرهنة ليابونوف (في صيغتها المبسطة) محققة، ومن ثم توزيع

$Y_n = \sum_{i=1}^{n} X_i$ ينتهـي إلى التوزيـع الطبيعـي $N(np,npq)$.

لكـن Y_n يتبع توزيع ذي الحديـن $B(n,p)$ ، وبالتـالي:

$$\lim_{n\to\infty} B(n,p) = N(np,npq)$$

وعندما تكـون n كبـيرة فـإن:

$$B(n,p) \approx N(np,npq)$$

تمـــارين

1. كم مرة يجب رمـي قطعـة نقـود متجانسـة حتـى يكـون:

$$P\left(0.4 < \frac{m}{n} < 0.6\right) \geq 0.90$$

2. يرغـب بـاحث علمـي تقدير متوسـط مجتمـع مسـتخدماً عينـة عشـوائية حجمهـا كـاف للقـول بـأن احتمـال ألا يتجـاوز الفـرق بـين متوسـط العينـة ومتوسـط المجتمـع 25٪ مـن الانحـراف المعيـاري هـو 0.59، فمـا هـو حجـم العينـة المطلـوب؟

3. X متغـير عشـوائي توقعـه الريـاضي $EX = 1$ وانحرافـه المعيـاري $\sigma = 0.2$ ، قـدر الحـد الأدنـى لاحتمـالات الحـوادث:

$$A = \{0.5 \leq X < 1.5\} \quad , \quad B = \{0.75 \leq X < 1.35\} \quad , \quad C = \{X < 2\}$$

4. X متغير عشوائي يساوي عدد الأيام المشمسة من السنة في منطقـة معينـة. إذا علمـت بأن $EX = 100$ يوم و $\sigma = 20$ يوم، قدر احتمال كل من الحادثين:

$$A = \{X \geq 150\} \quad , \quad B = \{X \geq 200\}$$

5. كم مرة يجب رمي قطعة نقود متجانسة للحصول على تكرار نسبي لظهور الصورة يقع في المجال [0.4,0.6]، وذلك باحتمال لا يقل عن 0.975؟

6. في استفتاء للرأي العام حول قضية ما، نريد أن يكون حجم العينة بحيث يكون الاحتمال 1٪ فقط أن نحصل على نسبة تأييد للقضية المطروحة على الاستفتاء أقل من 50٪، في حين أن حقيقة هذه النسبة 52٪، فكم يجب أن يكون حجم العينة؟

7. هل يمكن تطبيق قانون الأعداد الكبيرة على متتالية المتغيرات العشوائية المستقلة مثنى مثنى $X_1, X_2, ...$ في كل من الحالات الآتية:

$$a)\quad P\left(X_k = \pm 2^k\right) = 2^{-(2k+1)} \quad , \quad P\left(X_k = 0\right) = 1 - 2^{-2k}$$

$$b)\quad P\left(X_k = \pm ak\right) = \frac{1}{2k^2} \quad , \quad P\left(X_k = 0\right) = 1 - \frac{1}{k^2}$$

$$c)\quad P\left(X_k = \pm 2^k\right) = \frac{1}{2^k} \quad , \quad P\left(X_k = 0\right) = 1 - \frac{1}{2^{2k-1}}$$

$$d)\quad P\left(X_k = a\right) = \frac{k}{2k+1} \quad , \quad P\left(X_k = -a\right) = \frac{k+1}{2k+1}$$

8. في إحدى المدارس الابتدائية، يجب استقبال 200 طفل وطفلة في السنة الأولى، أوجد احتمال أن يكون من بينهم 100 طفلة إذا علمت بأن احتمال ولادة ذكر يساوي 0.515.

9. ليكن لدينا 625 تكراراً مستقلاً لتجربة مفروضة، إذا علمت بأن احتمال ظهور الحادث A في كل تكرار ثابت ويساوي 0.8، فأوجد احتمال أن يكون انحراف التكرار النسبي لظهور الحادث A عن احتماله الثابت، بالقيمة المطلقة، لا يزيد عن 0.06.

10. نسبة العطب في إنتاج صناعي 61٪، سحبت عينة عشوائية بحجم

$n = 1100$، فما هو احتمال أن تحتوي هذه العينة على 17 وحدة معطوبة على الأكثر؟

11. نعلم أن دواءً معيناً فعال في 80٪ من الحالات التي يستخدم فيها للعلاج، وقد وُجد أن نوعاً جديداً من هذا الدواء أثبت فعالية في 85 من الحالات المائة الأولى التي استخدم فيها، هل يمكن القول أن هذا الدواء الجديد أفضل من القديم؟

12. لتكن $X_1, X_2, \ldots$ متتالية من المتغيرات العشوائية ذات توقعات رياضية متساوية وتباينات محدودة، إذا علمت بأن التغاير المشترك $C_{ij} = E\left[(X_i - m_i)(X_j - m_j)\right]$ سالب فهل يمكن تطبيق قانون الأعداد الكبيرة على هذه المتتالية؟

13. عين الدالة المميزة للمتغير العشوائي:

$$Y_n = \frac{\sum_{k=1}^{n}(X_k - EX_k)}{\sqrt{\sum_{k=1}^{n} VX_k}}$$

ثم أوجد نهايتها عندما $n \to \infty$، إذا علمت بأن المتغيرات العشوائية $X_1, X_2, \ldots$ مستقلة ولها توزع احتمالي واحد:

$$P(X_k = m) = \frac{\lambda^m}{m!}e^{-\lambda} \quad ; \quad x = 0,1,\ldots$$

14. احتمال ظهور الحادث A في كل تكرار لتجربة مفروضة 0.25، فما هو احتمال ظهور الحادث A عند 300 تكرار للتجربة المفروضة:

1. 75 مرة؟
2. أكثر من 85 وأقل من 100 مرة؟

الفصل الثاني:

بعض النماذج الإحصائية الخاصة

1.2 مقدمة

يهدف صف هام من المسائل في الإحصاء الرياضي إلى إيجاد توزيعات متغيرات عشوائية تصف مجتمعات مختلفة. وإحدى المسائل المركزية: دراسة مجتمعات أشكال توزيعاتها معلومة، أي أن صيغ دوال التوزيع $F(x;\theta)$ أو دوال الكثافة(دوال الاحتمال في حالة المنقطع) $f(x;\theta)$ معلومة، لكنها تعتمد على معلمة غير معلومة θ (وحيدة البعد أو متعددة الأبعاد)، وبمعرفة القيمة الحقيقية (أو التقريبية) لـ θ يتعين تماماً (أو تقريباً) توزيع المجتمع. ويعني هذا أن لدينا عائلة من التوزيعات الاحتمالية المختلفة فيما بينها قيمة المعلمة θ، وتوزيع مجتمع الصفة المدروسة هو أحد هذه التوزيعات الاحتمالية (عضو في هذه العائلة).

يدعى التوزيع الاحتمالي الذي يعتمد على معلمة غير معلومة θ بالتوزيع الاحتمالي الإحصائي أو اختصاراً بالنموذج الإحصائي (statistical model).

ويهدف هذا الفصل إلى تقديم النماذج الإحصائية الأكثر استخداماً في الإحصاء الرياضي والتي لها تسميات ورموز عامة. فمثلاً، نموذج ذي الحدين، ويرمز له بـ $B(n,\theta)$. هذا بالإضافة إلى تقديم بعض الخصائص لكل نموذج إحصائي كالمتوسط، التباين، الدالة المميزة ...

نميز بين نوعين من النماذج الإحصائية: المنقطعة والمستمرة، وذلك حسب طبيعة المتغير العشوائي ع الذي يصف خاصة من خواص المجتمع قيد الدراسة.

2.2 بعـض النمـاذج الإحصائيـة المنقطعـة الخاصـة

SOME SPECIAL STATISTICAL DISCRETE MODELS

نقـدم في هـذا البنـد أهـم النمـاذج الإحصائيـة المنقطعـة في الإحصـاء الريـاضي، ونرفق كـل نمـوذج ببعـض خواصـه المميـزة ومجـالات تطبيقـه.

1.2.2 نمـــوذج بـــيرنولي *Bernoulli Model*

يقـال عـن متغير عشـوائي ξ إنـه يخضـع لنمـوذج بـيرنولي، إذا كـانت صيغـة دالـة كثافته مـن الشـكل:

$$f(x;\theta)=\theta^{x}(1-\theta)^{1-x} \quad ; \quad x=0,1 \quad ; \quad 0\le\theta\le1 \tag{1.2.2}$$

حيـث إن θ غـير معلومـة، ويمكـن أن تـأخذ أيـة قيمـة في الفـترة $[0,1]$. وعندئـذٍ عائلـة توزيعـات بـيرنولي:

$$\{f(x;\theta)\ ;\ 0\le\theta\le1\}$$

ويرمـز عـادة لنمـوذج بـيرنولي بـ $B(1,\theta)$. وعندئـذٍ نكتـب $\mathcal{L}(\xi)\in B(1,\theta)$ وتقـرأ أن توزيـع ξ هـو أحـد توزيعـات $B(1,\theta)$، أو ξ يخضـع لنمـوذج بـيرنولي بمعلمة غير معلومـة θ. نلاحـظ عنـد كـل قيمـة معينـة لـ θ نحصـل على توزيـع احتمـالي (توزيعاً بيرنولياً معينـاً تمامـاً). وإذا علمنـا القيمـة الحقيقـة أو التقريبيـة لـ θ الموافقـة لتوزيـع ξ ولتكـن θ_0 عندئـذٍ نكتـب $\mathcal{L}(\xi)=B(1,\theta_0)$.

إذا رمزنـا للقيمـة المتوقعـة والتبـاين والدالـة المميـزة للنمـوذج $B(1,\theta)$ بـ $E_\theta\xi$، $V_\theta\xi$ و $\varphi(t;\theta)$ علـى الـترتيب، فـإن:

$$E_\theta\xi=\theta \quad , \quad V_\theta\xi=\theta(1-\theta) \quad ; \quad \varphi(t;\theta)=\left[1+\theta\left(e^{it}-1\right)\right] \tag{2.2.2}$$

الإثبـــات

$$\varphi(t;\theta)=E_\theta e^{it\xi}=\sum_{x=0}^{1}e^{itx}\theta^x(1-\theta)^{1-x}$$

$$=\left[1+\theta\left(e^{it}-1\right)\right]=\left[(1-\theta)+\theta e^{it}\right]$$

وكما نعلـم مـن نظرية الاحتمـالات:

$$E_\theta\xi=\frac{1}{i}\frac{\partial\varphi(t;\theta)}{\partial t}\bigg|_{t=0}=\frac{1}{i}i\theta e^{it}\bigg|_{t=0}=\theta$$

$$E_\theta\xi^2=\frac{1}{i^2}\frac{\partial^2\varphi(t;\theta)}{\partial t^2}\bigg|_{t=0}=\frac{1}{i^2}i^2\theta e^{it}\bigg|_{t=0}=\theta$$

$$V_\theta\xi=E_\theta\xi^2-\left(E_\theta\xi\right)^2=\theta-\theta^2=\theta(1-\theta)$$

يستخدم هـذا النمـوذج إذا كنـا أمـام تجربـة أو محاولـة عشـوائية لهـا ناتجـان فقـط هما ظهـور الحـدث A أو عـدم ظهـوره (ظهـور $\overline{A}$)، وأن θ احتمـال ظهـور الحـدث A، ويكـون $\xi(\overline{A})=0$ ، $\xi(A)=1$. ومثـل هـذه المحاولـة تدعـى بمحاولـة بـيرنولي (Bernoulli trial)، ويدعى هـذا النموذج أحياناً بنموذج ذي الحديـن النقطـي.

2.2.2 نمـوذج ذي الحديـن *Binomial Model*

يقـال عـن متغير عشـوائي ξ إنـه يخضـع لنمـوذج ذي الحديـن إذا كـانت صيغـة دالة احتمالـه كـالآتي:

$$f(x;n,\theta)=C_n^x\theta^x(1-\theta)^{n-x}\quad;\quad x=0,1,\ldots..,n\quad;\quad 0\le\theta\le1\qquad(3.2.2)$$

ويرمـز عـادة لهـذا النمـوذج بــ $B(n,\theta)$ ، وهـو يعتمـد علـى معلمتـين θ، n غـير معلومتـين بشـكل عـام، وغالبـاً تكـون n معلومـة.

ويسـتخدم هـذا النمـوذج إذا كـان ξ متغـير عشـوائي يصـف نتـائج n تكـراراً مسـتقلاً لمحاولـة بـيرنولي (n تكـراراً مسـتقلاً لتجربـة مفروضـة ذات نـاتجين A و $\overline{A}$

واحتمـــال ظهـــور الحـــادث A ثابت في كل تكـــرار ويســـاوي θ). أي أن نمـــوذج برنولي حالة خاصة من نموذج ذي الحدين (موافق لـــ $n = 1$).

$$E_\theta \xi = n\theta \quad ; \quad V_\theta \xi = n\theta(1-\theta) \quad ; \quad \varphi(t;n,\theta) = \left[(1-\theta) + \theta e^{it}\right]^n \qquad (4.2.2)$$

الإثبـــات

$$\varphi(t;n,\theta) = E_\theta e^{it\xi} = \sum_{x=0}^{n} C_n^x e^{itx} \theta^x (1-\theta)^{n-x}$$

$$= \left[(1-\theta) + \theta e^{it}\right]^n$$

$$E_\theta \xi = \frac{1}{i} \left. \frac{\partial \varphi(t;n,\theta)}{\partial t} \right|_{t=0} = n\theta$$

$$E_\theta \xi^2 = \frac{1}{i^2} \left. \frac{\partial^2 \varphi(t;n,\theta)}{\partial t^2} \right|_{t=0} = n\theta\left[(1-\theta) + n\theta\right]$$

$$V_\theta \xi = E_\theta \xi^2 - \left(E_\theta \xi\right)^2 = n\theta(1-\theta)$$

ولأهمية نموذج ذي الحديـــن في التطبيقـــات الإحصائيـــة، ولتســـهيل اســـتخدامه، أعدت له جداول خاصة حسب قيم n و θ تعطي مباشرة احتمـــالات القيـــم الممكنـــة لــ ξ [الملحق جـــدول (1)].

3.2.2 نمـــوذج بواســـون Possion Model

يقال عن توزيع متغير عشـــوائي ξ إنـــه يخضـــع لنمـــوذج بواســـون، إذا كـــانت صيغة دالة احتمالـــه:

$$f(x;\theta) = \frac{\theta^x}{x!} e^{-\theta} \qquad ; \qquad x = 0,1,\ldots. \quad , \quad \theta > 0$$

ويـــرمز عـــادة لهـــذا النموذج بـــ $\Pi(\theta)$ وهـــو يعتمـــد علـــى معلمـــة مجهولـــة θ، وعندئذٍ نكتب $\mathcal{L}(\xi) \in \Pi(\theta)$.

إن الدالة المميزة لـ $\Pi(\theta)$

$$\varphi(t;\theta)=\sum_{x=0}^{\infty}e^{itx}\frac{\theta^{x}}{x!}e^{-\theta}=e^{-\theta\left(1-e^{it}\right)} \tag{6.2.2}$$

ومنها:

$$\begin{aligned} E_{\theta}\xi&=\frac{1}{i}\frac{\partial\varphi(t,\theta)}{\partial t}\bigg|_{t=0}=\theta \\ V_{\theta}\xi&=\mathrm{E}_{\theta}\xi^{2}-\left(E_{\theta}\xi\right)^{2}=\frac{1}{i^{2}}\frac{\partial^{2}\varphi(t,\theta)}{\partial t^{2}}\bigg|_{t=0}-\theta^{2}=\theta^{2}+\theta-\theta^{2}=\theta \end{aligned} \tag{7.2.2}$$

نلاحظ أن:

$$E_{\theta}\xi=V_{\theta}\xi=\theta \tag{8.2.2}$$

وهذه صفة مميزة لنموذج بواسون.

يدعى نموذج بواسون بنموذج الحوادث النادرة، حيث يصف الكثير من الظواهر التي تحدث في واحدة الزمن أو الفضاء. هذا بالإضافة إلى أن نموذج ذي الحدين ينتهي ضمن شروط معينة (عندما $n\to\infty$ و $\theta\to 0$ و $n\theta=c$ حيث c ثابت) إلى نموذج بواسون $\Pi(n\theta)$، وبالتالي عندما تكون n كبيرة و θ صغيرة، بحيث تكون $n\theta$ ثابتة أو ثابتة تقريباً، فيمكن تقريب نموذج ذي الحدين $B(n,\theta)$ بنموذج بواسون $\Pi(n\theta)$. ولتسهيل عملية استخدامه في التطبيقات الكثيرة أعدت له جداول خاصة [الملحق جدول (2)].

4.2.2 النموذج فوق الهندسي *The Hypergeometric Model*

يقال عن متغير عشوائي ξ أن توزيعه ينتمي لعائلة التوزيعات فوق الهندسية إذا كانت دالة احتماله معرفة كالآتي:

$$f(x;N,k,n)=\frac{C_{k}^{x}C_{N-k}^{n-x}}{C_{N}^{n}}\quad;\quad x=0,1,\ldots,n \tag{9.2.2}$$

حيث إن n ، k ، N أعداد صحيحة موجبة و $n \le k \le N$. ويرمز لهذا النموذج عادة بـ $H(N,k,n)$ وهو يعتمد على ثلاث معالم n ، k ، N، أي أن $\theta = (N,k,n)$، وأن:

$$E_\theta \xi = n\frac{k}{N} \quad , \quad V_\theta \xi = n\frac{k}{N}\,\frac{N-k}{N}\,\frac{N-n}{N-1} \tag{10.2.2}$$

نترك إثبات ذلك للقارئ على سبيل المثال.

يستخدم هذا النموذج عندما نكون أمام حالة من النوع الآتي:

ليكن ξ متغيراً عشوائياً يساوي عدد الوحدات المعطوبة في عينة بحجم n مشكلة بطريقة المعاينة بدون إرجاع (without replacement) من صندوق يحتوي على N وحدة منتوج منها k وحدة معطوبة. عندئذٍ ξ يتبع التوزيع فوق الهندسي.

5.2.2 نموذج ذي الحدين السالب
Negative Binomial Distribution

يقال عن متغير عشوائي ξ أن توزيعه عضو في عائلة توزيعات ذي الحدين السالب إذا كانت صيغة دالته الاحتمالية كالآتي:

$$f(x;r,\theta) = C^{x}_{r+x-1}\theta^{r}(1-\theta)^{x} \quad ; \quad x = 0,1,2,\ldots \tag{11.2.2}$$

حيث: $0 \le \theta \le 1$ و r عدد صحيح موجب، ونرمز لهذا النموذج بـ $\overline{B}(r,\theta)$.

يستخدم هذا النموذج في الحالة التي نكون فيها أمام تكرار لتجربة بيرنولي (محاولة بيرنولي) تحت نفس الشروط حتى الحصول على r نجاحاً. فيكون ξ عندئذٍ يساوي عدد مرات الفشل حتى الحصول على r نجاحاً، أي أن $r+\xi$ يساوي عدد التكرارات اللازمة حتى الحصول على r نجاحاً، وهنالك ظواهر عديدة في الحياة العملية تخضع لمثل هذا التوزيع.

ويمكن إثبات أن:

$$E_\theta\xi=\frac{r(1-\theta)}{\theta}\quad,\quad V_\theta\xi=\frac{r(1-\theta)}{\theta^2}\quad,\quad \varphi(t;r,\theta)=\theta^r\left[1-(1-\theta)e^{it}\right]^{-r} \quad (12.2.2)$$

كحالة خاصة، عندما $r=1$، فنموذج ذي الحدين السالب يؤول إلى نموذج يدعى بالنموذج الهندسي (geometric distribution) أو نموذج باسكال (Pascal) ويمكن إثبات أن توزيع مجموع متغيرات عشوائية مستقلة ولكل منها نفس التوزيع الهندسي عبارة عن توزيع ذي الحدين السالب.

6.2.2 النموذج المتعدد الحدود *The Mulinomial Model*

لنفترض لدينا تجربة أو محاولة تنتهي إلى واحدة فقط من الحوادث المتنافية $C_1,\dots,C_k$، ولتكن احتمالاتها الموافقة على الترتيب $\theta_1,\theta_2,\dots,\theta_k$، حيث $i=1,\dots,k$; $\theta_i>0$ و $\sum_1^k\theta_i=1$. لنفترض الآن أننا كررنا التجربة أو المحاولة n مرة بشكل مستقل، وكان $\xi=(X_1,\dots,X_k)$ متغير عشوائي ذي k بعد يصف عدد مرات ظهور الحوادث $C_1,\dots,C_k$ ضمن الـ n تكراراً مستقلاً للتجربة المفروضة، وهذا يعني أن X_1 يساوي عدد مرات ظهور الحادث C_1، X_2 عدد مرات ظهور الحادث C_2، وهكذا X_k عدد مرات ظهور الحادث C_k، بحيث أن $\sum_1^k X_i=n$. عندئذٍ المتغير العشوائي ξ يخضع لنموذج يدعى بالنموذج المتعدد الحدود، ودالة احتماله من الشكل:

$$f(x;\theta_1,\dots,\theta_k)=\frac{n!}{x_1!\dots x_k!}\theta_1^{x_1}\dots\theta_k^{x_k} \quad (13.2.2)$$

حيث إن $x=(x_1,\dots,x_n)$. ويرمز لهذا النموذج عادة بـ $M(n;\theta_1,\dots,\theta_k)$.

إن الدالة المميزة للتوزيع $M(n;\theta_1,\dots,\theta_k)$ هي:

$$\varphi(t_1,\dots,t_k;\theta_1,\dots,\theta_k)=\left(\theta_1e^{it_1}+\dots+\theta_ke^{it_k}\right)^n \quad (14.2.2)$$

ومن ثم:

$$E_\theta X_i = n\theta_i \quad , \quad V_\theta X_i = n\theta_i(1-\theta_i)$$
$$\text{cov}_\theta(X_i, X_j) = -n\theta_i\theta_j \tag{15.2.2}$$

حيث إن: $\theta = (\theta_1, \dots, \theta_k)$

3.2 بعض النماذج الإحصائية المستمرة الخاصة
SOME SPECIAL CONTINUOS STATISTICAL MODELS

يهدف هذا البند إلى تقديم بعض أهم النماذج الإحصائية المستمرة في الإحصاء الرياضي مرفقة ببعض خواصها.

1.3.2 النموذج المنتظم *The Unoform Rectangular Model*

إذا كان ξ متغير عشوائي مستمر كثافة توزيعه من الشكل:

$$f(x;\theta_1,\theta_2) = \begin{cases} \dfrac{1}{\theta_2-\theta_1} & ; \quad \theta_1 \le x \le \theta_2 \\ 0 & ; \quad x < \theta \text{ أو } x > \theta_2 \end{cases} \tag{1.3.2}$$

حيث إن $\theta_1 < \theta_2$ معلمتين غير معلومتين، فيقال إن توزيع ξ عضو في عائلة التوزيعات المنتظمة (أو المستطيلة)، التي تختلف عن بعضها بقيمة أحد المعلمتين $\theta_1 > \theta_2$ أو الاثنين معاً. ويرمز لهذا النموذج عادة بـ $R(\theta_1,\theta_2)$، وهو يعتمد على المعلمتين θ_1, θ_2 غير المعلومتين. ونكتب $\mathcal{L}(\xi) \in R(\theta_1,\theta_2)$. وفي حالات عدة تكون $\theta_1 = 0$ فنحصل على النموذج $R(0,\theta_2)$.

إن دالة توزيع النموذج $R(\theta_1,\theta_2)$ هي:

$$F(x;\theta_1,\theta_2) = \begin{cases} 0 & ; \quad x \le \theta_1 \\ \dfrac{x-\theta_1}{\theta_2-\theta_1} & ; \quad \theta_1 < x \le \theta_2 \\ 1 & ; \quad x > \theta_2 \end{cases} \tag{2.3.2}$$

كما أن:

$$E_\theta \xi = \frac{\theta_1 + \theta_2}{2} \quad , \quad V_\theta \xi = \frac{(\theta_2 - \theta_1)^2}{12}$$

$$\varphi(t,\theta) = \frac{e^{i\theta_2 t} - e^{i\theta_1 t}}{i(\theta_2 - \theta_1)t} \quad ; \quad \theta = (\theta_1, \theta_2) \tag{3.3.2}$$

يعتبر النموذج $R(\theta_1,\theta_2)$ من أبسط النماذج الإحصائية المستمرة، لكنه مفيد جداً في الإحصاء الرياضي، وتبدو فائدته بشكل خاص بما يلي:

إذا كان X متغيراً عشوائياً مستمراً دالة توزيعه $F(x)$ ، فإن المتغير العشوائي $Y = F(X)$ يتبع التوزيع المنتظم $R(0,1)$. وهذه الحقيقة يمكن إثباتها على النحو التالي.

بما أن:

$$H(y) = P\big(F(X) < y\big) = \begin{cases} 0 & ; \quad y \le 0 \\ y & ; \quad 0 < y \le 1 \\ 1 & ; \quad y > 1 \end{cases}$$

وهذا يعني أن الدالة $H(y)$ [قارن بالعلاقة (2.3.2)، حيث: $\theta_1 = 0, \theta_2 = 1$] ما هي إلا دالة التوزيع المنتظم $R(0,1)$ ، وهو المطلوب. ونلاحظ بسهولة أن متوسط وتباين المتغير العشوائي $Y = F(X)$ هما:

$$EY = \frac{1}{2} \quad , \quad VY = \frac{1}{12}$$

2.3.2 النموذج الطبيعي العام *The General Normal Model*

يعتبر من أهم النماذج المستمرة، حيث إن توزيعات أو نهايات توزيعات العديد من الظواهر إن لم نقل أغلبيتها تنتمي لعائلة التوزيعات الطبيعية بالإضافة إلى

أن غالبية الطرق والأساليب في الإحصاء الرياضي تعتمد على التوزيعات الطبيعية، وهذا ما يبدو بوضوح في الموضوعات اللاحقة.

يقال عن توزيع متغير عشوائي ξ إنه عضو في عائلة التوزيعات الطبيعية إذا كانت دالة توزيعه من الشكل:

$$f(x;\theta_1,\theta_2)=\frac{1}{\theta_2\sqrt{2\pi}}e^{-\frac{1}{2}(x-\theta_1)^2/\theta_2^2}\quad ;\quad -\infty<x<+\infty \tag{4.3.2}$$

حيث إن θ_1,θ_2^2 معلمتين غير معلومتين. تدعى هذه العائلة من التوزيعات الطبيعية التي تتميز عن بعضها بقيمة إحدى المعلمتين θ_1,θ_2^2 (أو الاثنين معا) بعائلة التوزيعات الطبيعية العامة، يرمز للنموذج (4.3.2) بـ $N(\theta_1,\theta_2^2)$ ويدعى بالنموذج الطبيعي العام، وهو يعتمد على معلمتين مجهولتين.

ومن الواضح بإجراء التحويل $Z=\dfrac{\xi-\theta_1}{\theta_2}$ نحصل على التوزيع الطبيعي المعياري $N(0,1)$ (توزيع معين تماماً).

إن الدالة المميزة للنموذج $N(\theta_1,\theta_2^2)$ هي:

$$\varphi(t;\theta_1,\theta_2^2)=e^{i\theta_1 t-\frac{1}{2}\theta_2^2t^2} \tag{5.3.2}$$

وباستخدام طريقة الاشتقاق نجد متوسط وتباين $N(\theta_1,\theta_2^2)$

$$\begin{aligned}E_\theta\xi&=\frac{1}{i}\frac{\partial\varphi(t;\theta_1,\theta_2)}{\partial t}\bigg|_{t=0}=\theta_1\\ V_\theta\xi&=\frac{1}{i^2}\frac{\partial^2\varphi(t;\theta_1,\theta_2)}{\partial t^2}\bigg|_{t=0}=\theta_2^2\end{aligned} \tag{6.3.2}$$

إذا رمزنا بـ $\theta=(\theta_1,\theta_2^2)$. فإن النموذج الطبيعي العام يعتمد على معلمة متجه (معلمة ذات بعدين).

مبـــــرهنة 1.3.2

إذا كـان X_1,X_2 متغـيرين عشـوائيين مسـتقلين، وكـان X_1 يتبـع $N\left(\mu_1,\sigma_1^2\right)$ و X_2 يتبـع $N\left(\mu_2,\sigma_2^2\right)$، فـإن $\eta = c_1X_1 + c_2X_2$، حيـث c_1,c_2 ثابتـان حقيقيـان (ليس كلاهمـا صفـراً)، يتبـع التوزيـع الطبيعـي:

$$N\left(\sum_{i=1}^{2} c_i\mu_i , \sum_{i=1}^{2} c_i^2\sigma_i^2\right)$$

الإثبـــات

إذا رمزنـا بـ $\varphi(t)$ للدالـة المميزة للمتغـير العشـوائي η و بـ $\varphi_i(t)\ ;\ i = 1,2$ للدالة المميزة لـ X_i، فـإن:

$$\varphi(t) = \varphi_1\left(c_1t\right)\,\varphi_2\left(c_2t\right) \tag{7.3.2}$$

لأن X_1 و X_2 مستقلان (حسب الفـرض)، وحسـب العلاقـة (5.3.2) نجـد:

$$\varphi_i\left(c_it\right) = e^{ic_i\mu_it - \frac{1}{2}c_i^2\sigma_i^2t^2} \quad ; \quad i = 1,2$$

ومن ثـم:

$$\varphi(t) = e^{i\left(c_1\mu_1+c_2\mu_2\right)t - \frac{1}{2}\left(c_1^2\sigma_1^2+c_2^2\sigma_2^2\right)t} \tag{8.3.2}$$

وهذه ما هي إلا الدالة المميزة للتوزيع الطبيعي $N\left(\sum_{i=1}^{2} c_i\mu_i , \sum_{i=1}^{2} c_i^2\sigma_i^2\right)$ وهو المطلوب.

يمكن تعميم تلك الخاصة على n من المتغيرات العشوائية المستقلة $X_1,\ldots,X_n$ حيث:

$$\mathcal{L}\left(X_i\right) = N\left(\mu_i,\sigma_i^2\right) \quad ; \quad i = 1,2,\ldots,n$$

عندئـذ توزيـع المتغير العشـوائي $\eta = \sum_{i=1}^{n} c_iX_i$ هـو:

$$N\left(\sum_{i=1}^{n} c_i\mu_i , \sum_{i=1}^{n} c_i^2\sigma_i^2\right)$$

كحالة خاصة، إذا كانت المعلمة θ_2^2 معلومة ولنرمز لها بـ σ^2، بينما θ_1 غير معلومة، فنحصل على النموذج الطبيعي $N\left(\theta,\sigma^2\right)$، وللتمييز سنطلق عليه اسم النموذج الطبيعي *I*. وبشكل مشابه، إذا كانت المعلمة θ_1 معلومة ولنرمز لها بـ μ، بينما θ_2^2 غير معلومة، فنحصل على النموذج الطبيعي $N\left(\mu,\theta^2\right)$ ونطلق عليه اسم النموذج الطبيعي *II*.

3.3.2 النموذج الطبيعي الثنائي العام

The General Bivariate Normal Model

تطرقنا في الفقرة السابقة للنموذج الطبيعي للمتغيرات العشوائية وحيدة البعد، وللأهمية الاستثنائية لهذا النموذج في الإحصاء، من المفيد دراسة النموذج الطبيعي المتعدد. وقبل ذلك نرى الفائدة في دراسة النموذج الطبيعي الثائي بشيءٍ من التفصيل.

يقال إن المتغير العشوائي ذو البعدين $Y=\left(X_1,X_2\right)$ يتبع النموذج الطبيعي الثنائي العام، إذا كانت صيغة دالة كثافته من الشكل:

$$f\left(x_1,x_2;\mu,\Sigma\right)=\frac{\sqrt{\left|\Sigma^{-1}\right|}}{2\pi}e^{-\frac{1}{2}Q\left(x_1,x_2\right)}\quad ;\quad \left(x_1,x_2\right)\in R^2 \qquad (9.3.2)$$

حيث إن:

$$Q=\sigma^{11}\left(x_1-\mu_1\right)^2+2\sigma^{12}\left(x_1-\mu_1\right)\left(x_2-\mu_2\right)+\sigma^{22}\left(x_2-\mu_2\right)^2 \qquad (10.3.2)$$

و

$$\Sigma=\left\|\sigma_{ij}\right\|\quad ;\quad \sigma_{ij}=\operatorname{cov}\left(X_i,X_j\right)\quad ,\quad i,j=1,2$$

مصفوفة التغاير (covariance matrix)، $\Sigma^{-1}=\left\|\sigma^{ij}\right\|$ معكوس المصفوفة Σ، $\left|\Sigma^{-1}\right|$ محددة المصفوفة Σ^{-1} وأن:

$$\mu=\left(\mu_1,\mu_2\right)\quad ;\quad EX_i=\mu_i\quad ;\quad i=1,2$$

نلاحظ بسهولة أن المصفوفة Σ محددة موجبة (أي أن $|\Sigma| \neq 0$)، ومن ثم فالصيغة التربيعية Q محددة موجبة.

وإذا رمزنا بـ ρ لمعامل الارتباط بين X_1 و X_2، يمكن كتابة العلاقة (9.3.2) على النحو الآتي:

$$f(x_1,x_2;\mu,\Sigma) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho}} e^{-\frac{1}{2}Q(x_1,x_2)} \quad ; \quad (x_1,x_2) \in R^2 \qquad (11.3.2)$$

حيث أن:

$$Q = \frac{1}{(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}\right] \quad ;$$

$$\sigma_i^2 = \text{var}\, X_i \;;\; i = 1,2$$

ونرمز للنموذج الطبيعي الثنائي العام بـ $N(\mu,\Sigma)$ حيث إن μ و Σ غير معلومتين.

إن الدالة المميزة للنموذج $N(\mu,\Sigma)$ هي:

$$\varphi(t_1,t_2) = \exp\left[i(\mu_1 t_1 + \mu_2 t_2) - \frac{1}{2}\left(\sigma_{11}t_1^2 + \sigma_{22}t_2^2 + 2\sigma_{12}t_1 t_2\right)\right] \qquad (12.3.2)$$

مبرهنة 2.3.2

إذا كان (X_1,X_2) متغيراً عشوائياً يتبع النموذج الطبيعي الثنائي العام $N(\mu,\Sigma)$، فأن المتغير العشوائي $\eta = c_1X_1 + c_2X_2$ يتبع النموذج:

$$N\left(\sum_{i=1}^{2} c_i\mu_i , \sum_{i,j=1}^{2} \sigma_{ij}c_i c_j\right)$$

ويمكن إثبات صحة هذه المبرهنة بسهولة اعتماداً على العلاقة (12.3.2) وخواص الدالة المميزة.

لندرس الآن المتغير العشوائي الشرطي $X_2|X_1$، حيث (X_1,X_2) يتبع النموذج الثنائي $N(\mu,\Sigma)$.

إن النمـوذج الهامشـي لـ X_1:

$$f_1(x_1;\mu,\Sigma)=\int_{-\infty}^{+\infty} f(x;\mu,\Sigma)dx_2$$

$$=\frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{1}{2}(x_1-\mu_1)^2/\sigma_1^2} \quad ; \quad x_1\in R$$

حيث إن: $\sigma_1^2=\sigma_{11}$، وباستخدام العلاقة (10.3.2) نجد:

$$f(x_2|x_1;\mu,\Sigma)=\frac{f(x_1,x_2)}{f_1(x_1)}=\frac{1}{\sigma_2\sqrt{2\pi(1-\rho)}}e^{-\frac{1}{2}Q(x_1,x_2)+\frac{1}{2}(x_1-\mu_1)^2/\sigma_1^2} \qquad (13.3.2)$$

وبإجراء بعض الإصلاحات الجبرية نجد:

$$f(x_2|x_1;\mu,\Sigma)=\frac{1}{\sigma_2\sqrt{2\pi(1-\rho^2)}}\exp\left\{\frac{1}{-2\sigma_2^2(1-\rho^2)}\left[(x_2-\mu_2)-\rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1)\right]^2\right\}$$

. . . . . (14.3.2)

أي أن $(X_2|X_1)$ يتبع النمـوذج الطبيعـي:

$$N\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1)\ ,\ \sigma_2^2(1-\rho^2)\right)$$

وهذا يعـني:

$$E(X_2|X_1)=\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1) \quad , \quad \text{var}(X_2|X_1)=\sigma_2^2(1-\rho^2) \qquad (15.3.2)$$

وهي عبارة عن معادلة خط انحدار X_2 على X_1. وبشكل مشابه نجد توزيع $X_1|X_2$.

4.3.2 النمـوذج الطبيعي المتعـدد العـام

The General k-Negative Normal Model

إن النمـوذج الطبيعـي ذي $k>2$ بعـد وخواصـه عبـارة عـن تعميـم مباشـر لحالـة $k=2$، الـواردة في الفقـرة السـابقة.

دالة كثافة النموذج الطبيعي المتعدد العام (ذي k بعد)

$$f(x_1,\dots,x_k;\mu,\Sigma)=\frac{\sqrt{|\Sigma^{-1}|}}{(2\pi)^{k/2}}e^{-\frac{1}{2}Q(x_1,\dots,x_k)} \quad ; \quad (x_1,\dots,x_k)\in R^k \qquad (16.3.2)$$

حيث إن:

$$\mu=(\mu_1,\dots,\mu_k) \quad , \quad Q(x_1,\dots,x_k)=\sum_{i,j=1}^{k}\sigma^{ij}(x_i-\mu_\iota)(x_j-\mu_\varphi) \qquad (17.3.2)$$

و $\Sigma^{-1}=\|\sigma^{ij}\|$ معكوس المصفوفة $i,j=1,\dots,k$; $\Sigma=\|\sigma_{ij}\|$.

نفترض أن Σ محددة موجبة $(|\sigma_{ij}|\neq 0)$.

إذا رمزنا بـ μ لمتجه القيم المتوقعة $(\mu_1,\dots,\mu_k)$، فنرمز للنموذج الطبيعي المتعدد العام بـ $N(\mu,\Sigma)$ ودالته المميزة تكون:

$$\varphi(t_1,\dots,t_k;\mu,\Sigma)=\exp\left[i\sum_{i=1}^{k}\mu_i t_i-\frac{1}{2}\sum_{i=1}^{k}\sigma_{ij}t_i t_j\right] \qquad (18.3.2)$$

مبرهنة 3.3.2

إذا كان $(X_1,\dots,X_k)$ متغيراً عشوائياً يتبع النموذج الطبيعي المتعدد $N(\mu,\Sigma)$، فإن المتغير العشوائي $\eta=c_1X_1+\dots+c_kX_k$ يتبع النموذج الطبيعي العام:

$$N\left(\sum_{i=1}^{k}c_i\mu_i\ ,\ \sum_{i,j=1}^{k}\sigma_{ij}c_ic_j\right)$$

يمكن الإثبات بالاعتماد على الدالة المميزة وخواصها.

5.3.2 نموذج جاما *Gamma Model*

تدعى الدالة $\Gamma(\alpha)$ المعرفة كالآتي:

$$\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^{-x}dx \quad ; \quad x>0\ ,\ \alpha>0 \qquad (19.3.2)$$

بدالـة جاما (gamma function) أو بتكـامل آيلر مـن النـوع الثـاني.

تحقـق دالـة جاما $\Gamma(\alpha)$ الخـواص الآتيـة:

.1 $$\Gamma(1) = \Gamma(2) = 1$$

وهـذه تنتـج مباشـرة مـن العلاقـة (19.3.2) بإعطــاء $\alpha = 1,2$.

.2 $$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$$

يمكـن إثبـات صحـة هـذه الخاصـة باستخدام العلاقـة (19.3.2) وتطبيـق التكـامل بالتجزئـة:

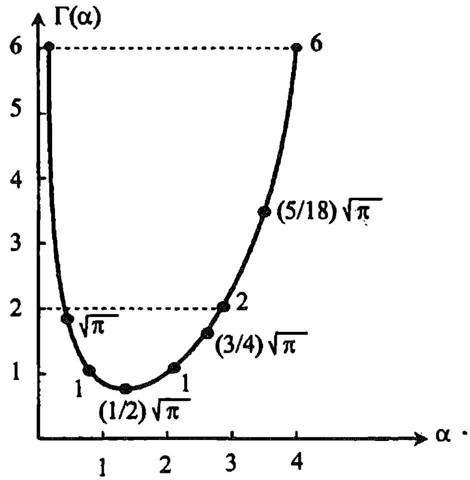
$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx = \left(-x^{\alpha-1}e^{-x}\Big|_0^\infty + (\alpha-1)\int_0^\infty x^{\alpha-2}e^{-x}dx\right.$$
$$= (\alpha - 1)\Gamma(\alpha - 1)$$

وكحالـة خاصـة، إذا كـان α عـدد صحيـح، فـإن:

$$\Gamma(\alpha) = (\alpha - 1)!$$

.3 $$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

ويمكـن تمثيـل الدالـة $\Gamma(\alpha)$ بيانياً كمـا هـو مبـين علـى الشـكل (1.3.2).



شكل 1.3.2

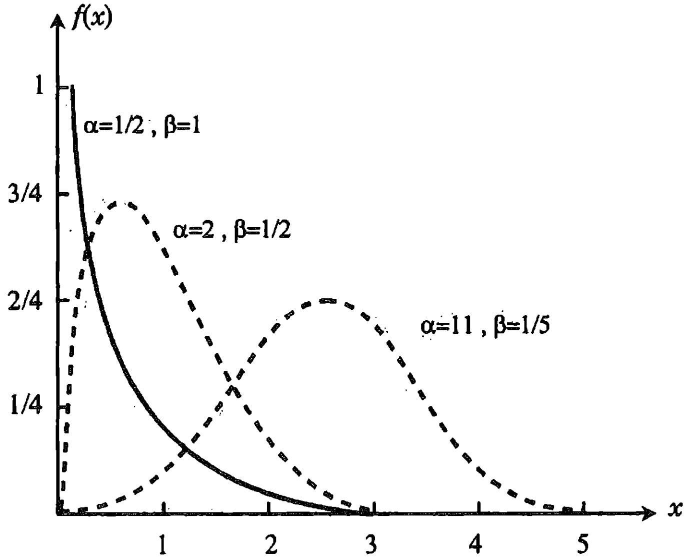
يمكن من الدالة $\Gamma(\alpha)$ تعريف نموذج له دورٌ هامٌ في الإحصاء، يدعى بنموذج جاما (gamma model) نرمز له بـ $\Gamma(\alpha,\beta)$ وهو يعتمد على معلمتين α,β غير معلومتين.

يقال عن توزيع متغير عشوائي ξ إنه عضو في عائلة التوزيعات الجماوية إذا كانت صيغة دالة كثافته من الشكل:

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta} & ; \quad x>0\\ 0 & ; \quad x\le 0\end{cases} \tag{20.3.2}$$

حيث إن $\Gamma(\alpha)$ دالة جاما المعرفة بالعلاقة (19.3.2) و $\alpha>0\ ,\ \beta>0$ ونرمز لنموذج Γ بـ $\Gamma(\alpha,\beta)$.

نلاحظ أن نموذج جاما يتعلق بالمعلمتين α,β، ويبين الشكل (2.3.2) منحنيات توزيعات جاما الموافقة لبعض القيم الخاصة لـ α,β.



شكل (2.3.2)

الدالة المميزة للنموذج $\Gamma(\alpha,\beta)$ هي:

$$\varphi(t;\alpha,\beta) = (1 - i\beta t)^{-\alpha} \tag{21.3.2}$$

الإثبات

$$\varphi(t;\alpha,\beta) = Ee^{it\xi} = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty} e^{itx}x^{\alpha-1}e^{-x/\beta}dx$$

$$= \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty} x^{\alpha-1}e^{-\left(\frac{1}{\beta}-it\right)x}\,dx$$

وبإجراء التحويل $y = \left(\frac{1}{\beta} - it\right)x$ نجد:

$$\varphi(t;\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\frac{1}{\left(1/\beta - it\right)^{\alpha}}\int_0^{\infty} y^{\alpha-1}e^{-y}dy$$

$$\varphi(t;\alpha,\beta) = (1 - i\beta t)^{-\alpha}$$

وهو المطلوب.

العزم الابتدائي من المرتبة r لنموذج جاما (20.3.2)

$$E_{\theta}\xi^{r} = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty} x^{r+\alpha-1}\, e^{-x/\beta}\, dx \quad ; \quad \theta = (\alpha,\beta)$$

وبوضع $y = \frac{x}{\beta}$ نجد:

$$E_{\theta}\xi^{r} = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\int_0^{\infty} y^{r+\alpha-1}\, e^{-y}\, dy = \frac{\beta^{r}\Gamma(\alpha + r)}{\Gamma(\alpha)} \quad ; \quad \theta = (\alpha,\beta) \tag{22.3.2}$$

وبإعطاء $r = 1,2$ نحصل على:

$$E_{\theta}\xi = \alpha\beta \;,\; E_{\theta}\xi^{2} = \alpha(\alpha + 1)\beta^{2} \;,\; \mathrm{var}_{\theta}\,\xi = E_{\theta}\xi^{2} - \left(E_{\theta}\xi\right)^{2} = \alpha\beta^{2} \tag{23.3.2}$$

ويمكن الوصول إلى نفس النتيجة باستخدام الدالة المميزة $\varphi(t;\alpha,\beta)$ وتطبيق العلاقة:

$$E_\theta\xi^r = \frac{1}{i^r}\frac{\partial^r\varphi(t;\alpha,\beta)}{\partial t^r}\bigg|_{t=0}$$

نلاحظ من العلاقتين (20.3.2) و (21.3.2) أن المعلمة β تلعب دور المقياس، بحيث إنه إذا كان $\mathcal{L}(\xi)\in\Gamma(\alpha,1)$ فإن $\mathcal{L}(\beta\xi)\in\Gamma(\alpha,\beta)$. وبناءً على هذه الخاصة، لدراسة المميزات العددية لنموذج جاما يكفي دراستها من أجل قيم معينة لـ β. فمثلاً عند $\beta = 1$ أو $\beta = 2$، والقيمة الثانية أكثر فائدة، لأن النموذج $\Gamma(\alpha,2)$ يلعب دوراً هاماً في الإحصاء ويدعى بنموذج كاي مربع.

وهناك حالة خاصة من نموذج جاما تلعب دوراً هاماً في الإحصاء، وهي الحالة الموافقة لـ $\alpha = 1 , \beta = \theta$ ، وتدعى بالنموذج الأسي (exponential model) ونرمز له بـ $\Gamma(1,\theta)$. أي أن:

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}e^{-x/\theta} & ; \quad x > 0 \\ 0 & ; \quad x \le 0 \end{cases} \tag{24.3.2}$$

ودالته المميزة:

$$\varphi(t;\theta) = (1 - i\theta t)^{-1} = \frac{1}{(1 - i\theta t)} \tag{25.3.2}$$

وإذا كان $\mathcal{L}(\xi)\in\Gamma(1,\theta)$ ، فإن:

$$E_\theta\xi^r = \frac{\theta^r\Gamma(1+r)}{\Gamma(1)} = \theta^r\Gamma(r+1) \tag{26.3.2}$$

وبإعطاء $r = 1,2$ نجد:

$$E_\theta \xi = \theta \quad , \quad E_\theta \xi^2 = 2\theta^2 \quad , \quad \underset{\theta}{\mathrm{var}}\, \xi = \theta^2$$

مــــبرهنة 4.3.2

إذا كانت $X_1,\dots,X_n$ متغيرات عشوائية مستقلة من توزيع $\mathcal{L}(\xi) \in \Gamma(1,\theta)$، فإن المتغير العشوائي $\sum_{j=1}^{n} X_j = Y$ يخضع لتوزيع $\Gamma(n,\theta)$.

يمكن إثبات ذلك باستخدام الدالة المميزة على النحو الآتي:

$$\varphi_Y(t,\theta) = \prod_{j=1}^{n} \varphi_{X_j}(t,\theta) = (1 - i\theta t)^{-n}$$

وبالمقارنة مع العلاقة (21.3.2) نجد أن $\varphi_Y(t)$ ما هي إلا الدالة المميزة للنموذج $\Gamma(n,\theta)$، حيث $\alpha = n\ ,\ \beta = \theta$. وحسب العلاقة (23.3.2)، فإن:

$$E_\theta Y = n\theta \quad , \quad E_\theta Y^2 = n(n+1)\theta^2 \quad , \quad V_\theta Y = n\theta^2 \tag{28.3.2}$$

$$E_\theta\left(\frac{1}{Y}\right) = \frac{1}{\theta^n \Gamma(n)} \int_0^\infty \frac{1}{y} y^{n-1} e^{-y/\theta} dy = \frac{1}{\theta^n \Gamma(n)} \int_0^\infty y^{n-2} e^{-y/\theta} dy$$

وبوضع $z = \dfrac{y}{\theta}$ نجد:

$$E_\theta\left(\frac{1}{Y}\right) = \frac{\theta^{n-1}}{\theta^n \Gamma(n)} \int_0^\infty z^{n-2} e^{-z} dz = \frac{1}{\theta \Gamma(n)} \Gamma(n-1) = \frac{1}{n-1}\,\frac{1}{\theta} \tag{27.3.2}$$

وهنالك صيغة أخرى مكافئة للنموذج الأسي نرمز لها بـ $\Gamma\left(1, \dfrac{1}{\theta}\right)$:

$$f(x;\theta) = \begin{cases} \theta e^{-\theta x} & ; \quad x > 0 \\ 0 & ; \quad x \le 0 \end{cases}$$

وعندئذٍ تبقى كل النتائج أعلاه صحيحة على أن نستبدل θ بـ $\frac{1}{\theta}$.

مبرهنة 5.3.2

إذا كان X متغيراً عشوائياً له التوزيع المنتظم $R(0,1)$ ، فإن توزيع المتغير العشوائي $Y = -\ln X$ له توزيع جاما $\Gamma(1,1)$.

نترك إثبات صحة هذه المبرهنة كتمرين للقارئ.

للنموذج $\Gamma(\alpha,\beta)$ دورٌ هامٌ عند حل المسائل الإحصائية المتعلقة بصيغ تربيعية في متغيرات عشوائية لها توزيعات طبيعية.

6.3.2 نموذج كاي مربع (نموذج χ^2) Chi-Square Model

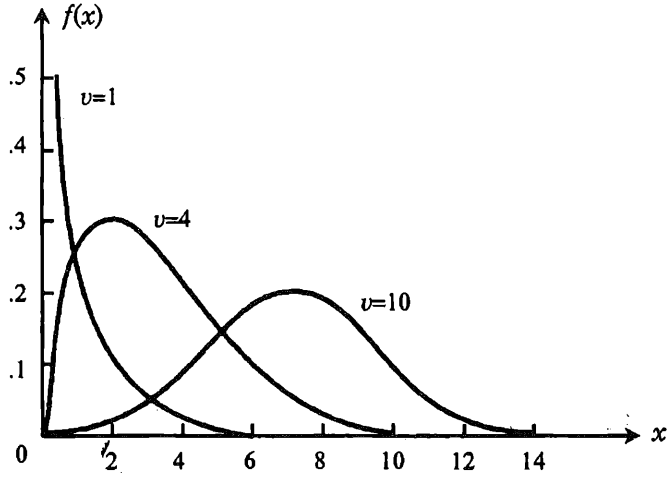
نميز ضمن عائلة التوزيعات الجماوية عائلة أصغر تدعى بعائلة توزيعات χ^2.

نقول عن توزيع متغير عشوائي ξ إنه عضو في عائلة توزيعات χ^2 إذا كانت صيغة دالة كثافته الاحتمالية من الشكل:

$$f(x;\upsilon) = \frac{1}{2^{\upsilon/2}\Gamma\left(\upsilon/2\right)} x^{\frac{\upsilon}{2}-1} e^{-\frac{1}{2}x} \quad ; \quad x > 0 \tag{29.3.2}$$

ويرمز لنموذج χ^2 بـ χ^2_υ (أو $\chi^2_{(\upsilon)}$)، الذي يعتمد على معلمة غير معلومة υ تدعى بدرجات الحرية، وهو حالة خاصة من نموذج $\Gamma(\alpha,\beta)$ الموافق لـ $\alpha = \frac{\upsilon}{2}$ و $\beta = 2$، أي أن $\chi^2_\upsilon = \Gamma\left(\frac{\upsilon}{2},2\right)$.

هكذا، فإن نموذج χ^2_υ يشتمل على معلمة غير معلومة واحدة υ ويبين الشكل (4.3.2) منحني توزيع χ^2 الموافق بعض درجات حرية مختلفة.



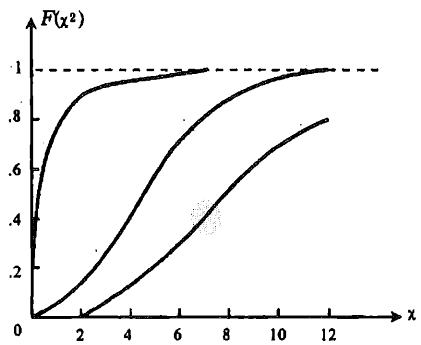
شكل (4.3.2)

دالة توزيع χ^2 تأخذ الشـــكل:

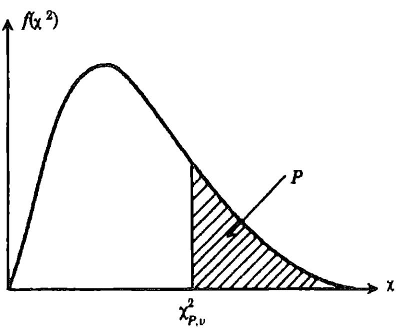
$$F(x;\upsilon) = P_\upsilon(\xi < x) = \begin{cases} \dfrac{1}{2^{\upsilon/2}\,\Gamma(\upsilon/2)} \displaystyle\int_0^x x^{\upsilon/2-1} e^{-x/2} dx & ; \ x > 0 \\ 0 & ; \ x \le 0 \end{cases} \qquad (30.3.2)$$

ويرمز عادة للمتغير العشوائي الذي توزيعه يخضع للنموذج χ_υ^2 بـــ χ^2.

ويبين الشكل (5.3.2) منحني دالة توزيع χ^2 الموافق لدرجـــات حريـــة مختلفـــة.



شكل (5.3.2) دالة توزيع χ_υ^2



شكل (6.3.2) توزيع χ_υ^2

تستخدم غالباً في التطبيقات القيمة الحرجة لتوزيع χ_{υ}^{2} والتي نرمز لها بـ $\chi_{p,\upsilon}^{2}$. القيمة الحرجة $\chi_{p,\upsilon}^{2}$ الموافقة لاحتمال مفروض p هي عبارة عن قيمة المتغير العشوائي χ^{2} التي تحقق الشرط الآتي:

$$P\left(\chi^{2}>\chi_{p,\upsilon}^{2}\right)=\int_{\chi_{p,\upsilon}^{2}}^{\infty} f\left(\chi^{2}\right) d \chi^{2}=p$$

ومن وجهة نظر هندسية، إيجاد القيمة الحرجة $\chi_{p,\upsilon}^{2}$ يتمثل في اختيار قيمة لـ χ^{2} بحيث تكون مساحة المنطقة المخططة على الشكل (6.3.2) مساوية لـ p.
ويتضمن الجدول (11) في الملحق قيم $\chi_{p,\upsilon}^{2}$ الموافقة لدرجات حرية $\upsilon \leq 35$ واحتمال P.

بوضع $\alpha=\frac{\upsilon}{2}$ و $\beta=2$ في العلاقة (21.3.2) نحصل على الدالة المميزة للنموذج χ_{υ}^{2}، أي أن:

$$\varphi(t, \upsilon)=(1-2 i t)^{-\frac{\upsilon}{2}} \tag{31.3.2}$$

وأيضاً بوضع $\beta=2$ وإعطاء $r=1,2$ في العلاقة (22.3.2) نجد:

$$E_{\upsilon} \chi^{2}=2 \alpha=\upsilon \quad, \quad V_{\upsilon} \chi^{2}=4 \alpha=2 \upsilon \tag{32.3.2}$$

مبرهنة 6.3.2

إذا كانت المتغيرات العشوائية $\xi_{1}, \ldots, \xi_{n}$ مستقلة و $\mathcal{L}\left(\xi_{i}\right) \in \chi_{\upsilon_{i}}^{2} ; i=1, \ldots, n$، فإن $\xi=\sum_{i=1}^{n} \xi_{i}$ فإن $\mathcal{L}(\xi) \in \chi_{\upsilon}^{2}$ حيث $\upsilon=\sum_{i=1}^{n} \upsilon_{i}$.

الإثبات

بناءً على خاصة الدالة المميزة:

$$\varphi_{\xi}(t)=\varphi_{\sum_{1}^{n} \xi_{i}}(t)=\prod_{i=1}^{n} \varphi_{\xi_{i}}\left(t, \upsilon_{i}\right)=\prod_{i=1}^{n}(1-2 i t)^{-\frac{\upsilon_{i}}{2}}=(1-2 i t)^{-\frac{1}{2} \sum_{1}^{n} \upsilon_{i}}=(1-2 i t)^{-\frac{\upsilon}{2}}$$

أي أن:

$$\mathcal{L}(\xi) \in \chi^2_{\upsilon} \quad ; \quad \upsilon = \sum_{i=1}^{n} \upsilon_i$$

وهو المطلــوب.

إذا كان $\mathcal{L}(\eta_i^2) = \chi^2_{(1)}\ ; i = \overline{1,n}$ فـــإن:

$$E\eta_i^2 = 1 \quad , \quad V\eta_i^2 = 2$$

وحسب مبرهنة النهاية المركزية، عندما $n \to \infty$، نجــد:

$$\mathcal{L}(\eta^*) = N(n,2n) \quad ; \quad \eta^* = \sum_{i=1}^{n} \eta_i^2$$

وعلى ذلــك:

$$\mathcal{L}\left(\frac{\eta^* - n}{\sqrt{2n}}\right) = N(0,1)$$

وهـــذا يعني عندما تكون n كبيرة $(n \geq 30)$ يمكن اســـتخدام التوزيـــع الطبيعـــي $N(n,2n)$ كتقـــريب جيـــد للتـــوزيع $\chi^2_{(n)}$. وبالتـــالي عنـــدما $n \geq 30$ يمكـــن استخدام العلاقــة:

$$P(\chi^2 < x) \approx \Phi\left(\frac{x-n}{\sqrt{2n}}\right)$$

لحساب $P(\chi^2 < x)$، أو لإيجاد القيم الحرجة الموافقة لـــــ n درجـــة حريـــة واحتمـــال معطى p، ويمكن استخدام الصيغة التقريبيــــة:

$$P(\chi^2 < x) \approx \Phi\left(\sqrt{2x} - \sqrt{2n-1}\right)$$

عندما تكون n و x كبـــيرتين. لـــذا لا يشـــتمل الملحـــق جـــدول (11) علـــى قيـــم $\upsilon > 35$.

مبرهنة 7.3.2

إذا كان $\mathcal{L}(\eta) = N(0,1)$ ، فإن $\mathcal{L}(\eta^2) = \chi_1^2$ ، أي أن توزيع مربع المتغير العشوائي ξ الخاضع للتوزيع الطبيعي المعياري هو χ^2 بدرجة حرية واحدة.

الإثبات

يمكن إثبات ذلك على النحو الآتي:

$$\varphi_{\eta^2}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itz^2} e^{-\frac{1}{2}z^2} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(1-2it)z^2} dz$$

وبإجراء التحويل $u = (1-2it)^{\frac{1}{2}} z$ نجد:

$$\varphi_{\eta^2}(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{(1-2it)^{1/2}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}u^2} du = (1-2it)^{-\frac{1}{2}}$$

وهذه ما هي إلا الدالة المميزة للتوزيع χ_1^2 .

نتيجة

بناءً على المبرهنتين (6.3.2) و (7.3.2) نجد: إذا كانت المتغيرات العشوائية $\eta_1, \dots, \eta_n$ مستقلة ولكل منها توزيع طبيعي معياري $N(0,1)$ ، فإن:

$$\mathcal{L}\left(\sum_{1}^{n} \eta_i^2\right) = \chi_n^2$$

لأن: $\mathcal{L}(\eta_i^2) = \chi_1^2 \; ; i = 1, \dots, n$.

هكذا، إذا كان $\mathcal{L}(\eta^*) = \chi_n^2$ ، فإن المتغير العشوائي η^* يمكن كتابته على الشكل: $\eta^* = \sum_{i=1}^{n} \eta_i^2$ ، حيث المتغيرات العشوائية $\eta_i^2 \; ; i = 1, \dots, n$ مستقلة و $\mathcal{L}(\eta_i^2) = \chi_{(1)}^2 \; ; i = 1, \dots, n$. ومصطلح "درجات الحرية" مرتبط أساساً هذا التصور.

مـــبرهنة 8.3.2

إذا كـــان X_1^2 و X_2^2 متغـــيرين عشوائيين مستقلين، حيث X_1^2 يتبع توزيـــع $\chi_{v_1}^2$ و $X_1^2 + X_2^2$ يتبـــع تـــوزيع $\chi_v^2 (v > v_1)$، فـــإن X_2^2 يخضـــع أيضاً لتوزيـــع χ^2 بـــ $v_2 = v - v_1$ درجة حرية.

يمكن إثبات هذه الخاصة بسهولة باســـتخدام العلاقـــة (31.3.2).

سنحتاج لاحقاً إلى إحدى التعميمات الهامة لتوزيع χ^2. لتكن $\eta_1, \ldots, \eta_n$ متغـــيرات عشوائية مستقلة و $\mathcal{L}(\eta_i) = N(\mu_i, 1)$، وبالتالي فإن:

$$\mathcal{L}\left(\sum_{i=1}^{n} (\eta_i - \mu_i)^2\right) = \chi_n^2$$

ونحتاج أحياناً لمعرفة توزيع مجموع مربعات المتغيرات العشوائية η_i. ويدعى التـــوزيع الاحتمالي للمتغير العشوائي $\eta^* = \sum_1^n \eta_i^2$ بتوزيع χ^2 اللامـــركزي بـــ n درجة حريـــة بالمعلمة $\lambda^2 = \sum_{i=1}^{n} \mu_i^2$ ونرمز له بـــ $\chi_{n,\lambda}^2$ (أو $\chi_{(n,\lambda)}^2$). أي أن:

$$\mathcal{L}\left(\sum_1^n \eta_i^2\right) = \chi_{(n,\lambda)}^2$$

ودالة كثافة هذا التوزيـــع:

$$f_n(\chi^2, \lambda^2) = \frac{(\chi^2)^{\frac{n}{2}-1}}{2^{\frac{n}{2}}} e^{-(\chi^2+\lambda^2)/2} \sum_{j=0}^{\infty} \frac{(\chi^2\lambda^2)^j}{j!\,2^{2j}\,\Gamma\left(\frac{n}{2} + j\right)} \tag{33.3.2}$$

ونلاحظ بســـهولة أن:

$$E_\lambda\left(\sum_1^n \eta_i^2\right) = n + \lambda^2 \quad , \quad V_\lambda\left(\sum_{i=1}^{n} \eta_i^2\right) = 2(n + 2\lambda^2)$$

وعندما $\lambda = 0$ فإن توزيع $\chi_{n,\lambda}^2$ اللامركزي يؤول إلى توزيع χ_n^2 المركزي (العادي).

7.3.2 نمــوذج بيتــا (نمــوذج β) Beta Model

يقـال عـن توزيـع متغير عشـوائي مسـتمر ξ إنـه عضـو في عائلـة توزيعـات β، إذا كانت صيغـة كثافة توزيعه الاحتماليـة مـن الشـكل:

$$f(x;\upsilon_1,\upsilon_2)=\frac{\Gamma(\upsilon_1+\upsilon_2)}{\Gamma(\upsilon_1)\Gamma(\upsilon_2)}x^{\upsilon_1-1}(1-x)^{\upsilon_2-1} \quad ; \quad 0<x<1 \tag{34.3.2}$$

حيــث إن υ_1,υ_2 أعــداد حقيقيـــة موجبـــة، ويرمـــز لهـــذا النمـــوذج عـــادة بـــ $Be(\upsilon_1,\upsilon_2)$، وهـو يعتمـد على معلمتين غـــير معلومتـــين υ_1 و υ_2.

دالــة توزيـــع $Be(\upsilon_1,\upsilon_2)$، ونرمـــز لهـــا بـــ $I_x(\upsilon_1,\upsilon_2)$. وتدعـــى بدالـــة بيتـــا غـــير التامـة (الناقصـة):

$$I_x(\upsilon_1,\upsilon_2)=\frac{\Gamma(\upsilon_1+\upsilon_2)}{\Gamma(\upsilon_1)\Gamma(\upsilon_2)}\int_{0}^{x}x^{\upsilon_1-1}(1-x)^{\upsilon_2-1}dx \quad ; \quad 0<x<1 \tag{35.3.2}$$

وهنــالك جــداول خاصــة تعطــي قيــم $I_x(\upsilon_1,\upsilon_2)$ مــن أجــل $x=0.01,\ldots,1$ و $\upsilon_1,\upsilon_2=0.05,\ldots,50$ ويدعــى المقــدار $\frac{\Gamma(\upsilon_1)\Gamma(\upsilon_2)}{\Gamma(\upsilon_1+\upsilon_2)}$ بدالـــة بيتـــا (beta function) ونرمز لهـا بــ $\beta(\upsilon_1,\upsilon_2)$. ومـن خـواص دالـة الكثافـة نجـد أن:

$$\beta(\upsilon_1,\upsilon_2)=\frac{\Gamma(\upsilon_1)\Gamma(\upsilon_2)}{\Gamma(\upsilon_1+\upsilon_2)}=\int_{0}^{1}x^{\upsilon_1-1}(1-x)^{\upsilon_2-1}dx \tag{36.3.2}$$

العـزم الابتدائـي مـن المرتبـة r للتوزيـــع (34.3.2)

$$E_\theta\xi^r=\frac{1}{\beta(\upsilon_1,\upsilon_2)}\int_{0}^{1}x^{\upsilon_1+r-1}(1-x)^{\upsilon_2-1}dx \quad ; \quad \theta=(\upsilon_1,\upsilon_2)$$

وبنـاءً علـى العلاقـة (36.3.2) نجـد:

$$E_\theta\xi=\frac{\Gamma(\upsilon_1+\upsilon_2)}{\Gamma(\upsilon_1)\Gamma(\upsilon_2)}\ \frac{\Gamma(\upsilon_1+r)\Gamma(\upsilon_2)}{\Gamma(\upsilon_1+\upsilon_2+r)}=\frac{\Gamma(\upsilon_1+\upsilon_2)\Gamma(\upsilon_1+r)}{\Gamma(\upsilon_1+\upsilon_2+r)\Gamma(\upsilon_1)} \tag{37.3.2}$$

ومن ثـم:

$$E_\theta \xi = \frac{\Gamma(\upsilon_1+\upsilon_2)\Gamma(\upsilon_1+r)}{\Gamma(\upsilon_1+\upsilon_2+r)\Gamma(\upsilon_1)} = \frac{\upsilon_1}{\upsilon_1+\upsilon_2}$$

$$V_\theta \xi = E_\theta \xi^2 - (E_\theta \xi)^2 = \frac{\upsilon_1\upsilon_2}{(\upsilon_1+\upsilon_2)^2(\upsilon_1+\upsilon_2+1)} \tag{38.3.2}$$

مـــــبرهنة 8.3.2

إذا كـــان X_1, X_2 متغـــيرين عشـــوائيين مســـتقلين لهمـــا توزيـــع $\Gamma(\upsilon_2,1)$ و $\Gamma(\upsilon_1,1)$ علـى الـترتيب، فـإن المتغير العشـوائي

$$U = \frac{X_1}{X_1+X_2}$$

لـه توزيـع بيتـا $Be(\upsilon_1,\upsilon_2)$.

الإثبـــــات

دالـة كثافـة المتغـير العشـوائي (X_1, X_2)

$$\frac{1}{\Gamma(\upsilon_1)\Gamma(\upsilon_2)} x_1^{\upsilon_1-1} x_2^{\upsilon_2-1} e^{-x_1-x_2}$$

وبالانتقـال إلى المتغـيرين Z, U، حيـث إن:

$$u = \frac{x_1}{x_1+x_2} \quad , \quad z = x_2$$

نجد:

$$\frac{1}{\Gamma(\upsilon_1)\Gamma(\upsilon_2)} u^{\upsilon_1-1} z^{\upsilon_1+\upsilon_2-1} (1-u)^{-(\upsilon_1+1)} e^{-z/(1-u)} \, dudz$$

وبإجراء التكـامل بالنسـبة لـ z مـن 0 إلى ∞ نجـد:

$$f(u;\upsilon_1,\upsilon_2)=\frac{\Gamma(\upsilon_1+\upsilon_2)}{\Gamma(\upsilon_1)\Gamma(\upsilon_2)}u^{\upsilon_1-1}(1-u)^{\upsilon_2-1}\quad;\quad 0<u<1$$

وهو عبارة عن توزيع $Be(\upsilon_1,\upsilon_2)$.

8.3.2 نموذج F — *F Model*

هناك نموذج آخر يلعب دوراً هاماً يتعلق بالمعاينة من مجتمع طبيعي، يدعى نموذج F (نموذج فيشر)، الذي يستنتج من نموذج β على النحو الآتي:

باستبدال υ_1,υ_2 بـ $\frac{\upsilon_1}{2},\frac{\upsilon_2}{2}$ على الترتيب في نموذج بيتا (34.3.2) نجد:

$$f(x;\upsilon_1,\upsilon_2)=\begin{cases}kx^{\frac{\upsilon_1}{2}-1}(1-x)^{\frac{\upsilon_2}{2}-1} & ;\quad 0<x<1\\ 0 & ;\quad x\le 0 \text{ or } x\ge 1\end{cases}\tag{39.3.2}$$

حيث إن $k=\Gamma\left(\frac{\upsilon_1+\upsilon_2}{2}\right)\Big/\Gamma\left(\frac{\upsilon_1}{2}\right)\Gamma\left(\frac{\upsilon_2}{2}\right)$.

لنبحث الآن عن النموذج الإحصائي الاحتمالي للمتغير $Y=\frac{\upsilon_2\xi}{\upsilon_1(1-\xi)}$

حيث إن ξ يخضع للنموذج (39.3.2).

نلاحظ أن:

$$\xi=\frac{\upsilon_1 Y}{\upsilon_2+\upsilon_1 Y}\quad\Rightarrow\quad\frac{dx}{dy}=\frac{\upsilon_1\upsilon_2}{(\upsilon_2+\upsilon_1 y)^2}$$

وبما أن:

$$g(y;\upsilon_1,\upsilon_2)=f(x;\upsilon_1,\upsilon_2)\left|\frac{dx}{dy}\right|$$

فإن:

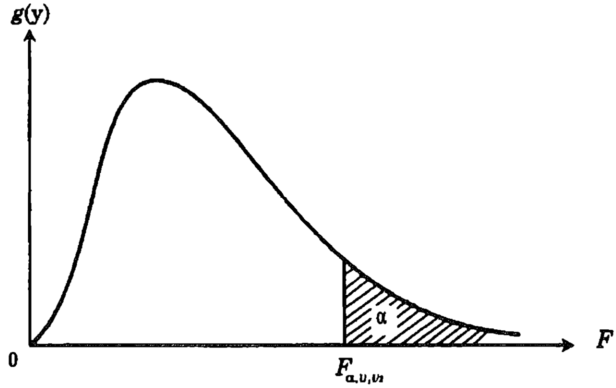
$$g(y)=g(y;\upsilon_1,\upsilon_2)=\frac{\Gamma\left(\frac{\upsilon_1+\upsilon_2}{2}\right)}{\Gamma\left(\upsilon_1\!/_2\right)\Gamma\left(\upsilon_2\!/_2\right)}\left(\frac{\upsilon_1}{\upsilon_2}\right)^{\frac{\upsilon_1}{2}}y^{\frac{\upsilon_1}{2}-1}\left(1+\frac{\upsilon_1}{\upsilon_2}y\right)^{-\frac{\upsilon_1+\upsilon_2}{2}}\quad;\quad y>0$$

(40.3.2)

يدعى النموذج المعرف بدالة الكثافة الاحتمالية (40.3.2) بنموذج F بـ υ_1 و υ_2 درجة حرية ونرمز له بـ $F_{\upsilon_1,\upsilon_2}$. ونظراً لأهمية هذا النموذج أعدت له جداول خاصة. والجدول (12) في الملحق يتضمن قيم $F_{\alpha,\upsilon_1,\upsilon_2}$، حيث إن:

$$\int_{F_{\alpha,\upsilon_1,\upsilon_2}}^{\infty} g(y)dy = \alpha$$

وذلك من أجل $\alpha = 0.01$ و $\alpha = 0.05$ والقيم المختلفة لـ υ_1، υ_2. والشكل (7.3.2) يبين منحني نموذج F والقيمة الحرجة $F_{\alpha,\upsilon_1,\upsilon_2}$.



شكل (7.3.2)

مبرهنة 9.3.2

إذا كان X_1, X_2 متغيرين عشوائيين مستقلين لكل منهما توزيع χ^2 بـ υ_1, υ_2 درجة حرية على الترتيب، فإن:

$$Y = \frac{X_1 / \upsilon_1}{X_2 / \upsilon_2}$$

له توزيع F بـ υ_1، υ_2 درجة حرية. نترك إثبات صحة ذلك للقارئ على سبيل المثال.

نتيجة:

إذا كان X متغيراً عشوائياً يخضع لتوزيع F بـ m و n درجة حرية، فإن المتغير العشوائي $Y = 1/X$ يخضع بدوره لتوزيع F، لكن بـ n و m درجة حرية. وبناءً على هذه النتيجة فإن:

$$F_{1-\alpha;n,m} = \frac{1}{F_{\alpha;m,n}} = F^{-1}_{\alpha;m,n}$$

لأن:

$$\alpha = P\left(X > F_{\alpha;m,n}\right) = P\left(Y < \frac{1}{F_{\alpha;m,n}}\right) = 1 - P\left(Y > \frac{1}{F_{\alpha;m,n}}\right)$$

$$1-\alpha = P\left(Y > \frac{1}{F_{\alpha;m,n}}\right) \Rightarrow F_{1-\alpha;n,m} = \frac{1}{F_{\alpha;m,n}}$$

9.3.2 نموذج t *t-Model*

لتوزيع t دوراً هاماً عند إيجاد تقديرات معالم نموذج مقترح لمتغير عشوائي ملاحظ ξ. وحسب تعريف نموج t بـ υ درجة حرية، ونرمز له بـ $S(\upsilon)$ هو عبارة عن نموذج المتغير العشوائي:

$$t = \frac{\xi}{\sqrt{\eta/\upsilon}} \qquad (41.3.2)$$

حيث إن المتغيرين العشوائيين ξ و η مستقلان و $\mathcal{L}(\xi) = N(0,1)$ و $\mathcal{L}(\eta) \in \chi^2_\upsilon$.

يمكن إيجاد دالة الكثافة $s_\upsilon(x)$ للنموذج $S(\upsilon)$ باستخدام الطريقة العامة لحساب كثافة توزيع حاصل قسمة متغيرين عشوائيين مستقلين. إن دالة الكثافة المشتركة لـ ξ و η تعطى كما يلي:

$$f(x,y) = f_\xi(x).f_\eta(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \dfrac{1}{\Gamma(\upsilon/2)2^{\upsilon/2}} y^{\frac{\upsilon}{2}-1} e^{-\frac{y}{2}} & ; \ -\infty < x < +\infty, y > 0 \\ 0 & ; \ y \le 0 \end{cases}$$

ومنهـا دالـة الكثافـة المشـتركة لـ t و η :

$$g(t,y) = f(t\sqrt{y/\upsilon},y)\left|\frac{dx}{dt}\right| = \begin{cases} \dfrac{1}{\sqrt{2\pi\upsilon}\,\Gamma(\upsilon/2)2^{\upsilon/2}}\, y^{\frac{\upsilon-1}{2}}\, e^{-\frac{y}{2}\left(1+\frac{t^2}{\upsilon}\right)} & ; \ -\infty < x < +\infty, y > 0 \\ 0 & ; \ y \leq 0 \end{cases}$$

حيـث إن $t = \dfrac{x}{\sqrt{y/\upsilon}}$ و $\dfrac{dx}{dt} = \sqrt{y/\upsilon}$ (بافـتراض y ثـابت).

بـإجراء تكـامل الدالـة $g(t,y)$ بالنسـبة لــ y، وبافـتراض $z = \dfrac{y}{2}\left(1+\dfrac{t^2}{\upsilon}\right)$ نجـد:

$$S_\upsilon(t) = \frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\sqrt{2\pi\upsilon}\,\Gamma\left(\frac{\upsilon}{2}\right)}\left(1+\frac{t^2}{\upsilon}\right)^{-\frac{\upsilon+1}{2}} \quad ; \quad -\infty < t < +\infty \qquad (42.3.2)$$

وهـذه هـي دالـة كثافـة نمـوذج t بـ υ درجـة حريـة. وكـان أول مـن حصـل علـى هـذا النمـوذج هـو غوسـيت (W. S. Gosset)، الـذي نشـر بحثـه تحـت اسـم مسـتعار "ستيودنت *Student*"، لـذا يدعـى هـذا النمـوذج أيضـاً بنمـوذج سـتيودنت.

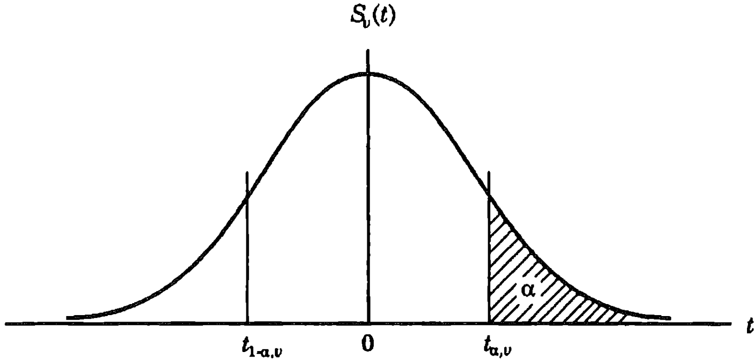
يـدل الرمـز υ في الصيغـة (42.3.2) علـى عـدد درجـات الحريـة للمتغـير العشـوائي η في العلاقـة (41.3.2).

ولأهمية نموذج t أعدت له جداول خاصة تعطي القيم الحرجة $t_{\alpha,\upsilon}$، حيث إن:

$$\int_{t_{\alpha,\upsilon}}^{\infty} s_\upsilon(t)dt = \alpha$$

ويبـدو بوضـوح α و $t_{\alpha,\upsilon}$ علـى الشـكل (8.3.2)، ويتضمـن الجـدول (10) في الملحق على قيـم $t_{\alpha,\upsilon}$ مـن أجـل:

$$\alpha = 0.005\ ,\ 0.01\ ,\ 0.025\ ,\ 0.05\ ,\ 0.10$$
$$\upsilon = 1,2,\ldots,29$$



شكل (8.3.2)

لا تشتمل جداول t على قيم $t_{\alpha,\upsilon}$ السالبة (أي عندما $\alpha > 0.50$) لأن نموذج t متناظر بالنسبة لـ $t = 0$، وبالتالي $t_{1-\alpha,\upsilon} = -t_{\alpha,\upsilon}$. وعندما تكون $\upsilon \geq 30$ يستخدم التوزيع الطبيعي المعياري $N(0,1)$ كتقريب جيد لتوزيع t.

يمكن إثبات أن القيمة المتوقعة والتباين للمتغير العشوائي t يساويان على الترتيب.

$$E_\upsilon t = 0 \quad , \quad V_\upsilon t = \operatorname{var} t = \frac{\upsilon}{\upsilon - 2} \quad , \quad \upsilon > 2$$

نلاحظ من الصيغة (42.3.2) أن نموذج t يعتمد على معلمة واحدة فقط هي عدد درجات الحرية υ. وعندما $\upsilon = 1$ يؤول نموذج t إلى توزيع كوشي، الذي يعرف بدالة الكثافة:

$$f(t) = \frac{1}{\sqrt{2}\pi\left(1+t^2\right)} \qquad ; \qquad -\infty < t < \infty$$

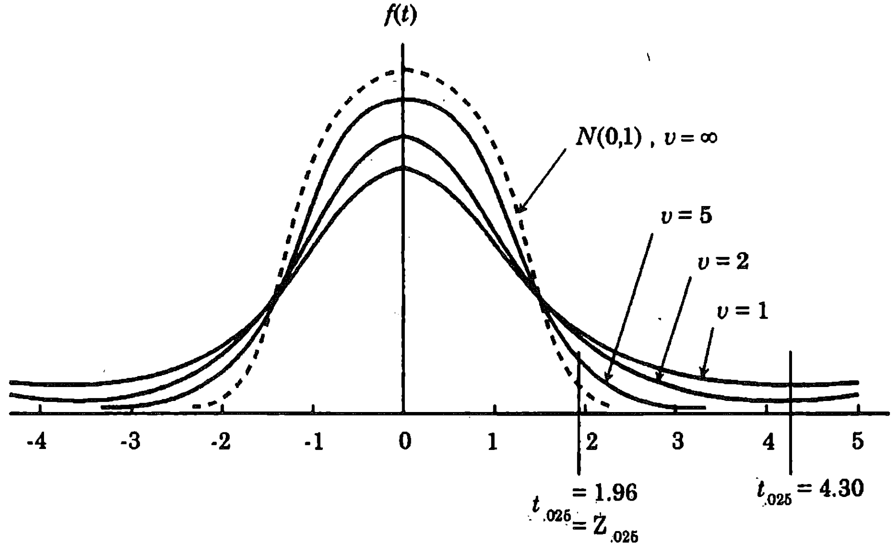
وعندما $\upsilon \to \infty$، فإن:

$$\left(1+\frac{t^2}{\upsilon}\right)^{-\frac{\upsilon+1}{2}} \longrightarrow e^{-\frac{t^2}{2}} \qquad , \qquad \frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\Gamma\left(\frac{\upsilon}{2}\right)\sqrt{\pi\upsilon}} \longrightarrow \frac{1}{\sqrt{2\pi}}$$

وهـذا يعـني أن متتاليـة كثافـات الاحتمـال f_n للمتغـيرات العشـوائية الـتي لهـا نمـوذج t تتقـارب إلى التوزيـع الطبيعـي المعيـاري:

$$S_v(t) \longrightarrow N(0,1)$$

بازديـاد عـدد درجـات الحريـة v بـدون حـدود. ويبـين الشـكل (9.3.2) منحـني توزيـع t عنـد درجـات حريـة v مختلفـة. ونلاحـظ بوضـوح أنـه بازديـاد عـدد درجـات الحريـة v فـإن منحـني توزيـع t يقـترب أكـثر فـأكثر مـن منحـني التوزيـع الطبيعـي المعيـاري. وإذا كـانت درجـة الحريـة v صغـيرة، فـإن احتمـالات الإنحرافـات الكبـيرة إلى حـد مـا أكـبر بالمقارنـة مـع التوزيـع الطبيعـي المعيـاري، وعندمـا $t > 2$ فمنحـني توزيـع t يتوضـع فـوق المنحـني الطبيعـي المعيـاري.



شكل (9.3.2): توزيع t

نلاحـظ ممـا سبق أن هنـالك نوعـين مـن النمـاذج: النمـاذج الاحتماليـة (التوزيـع الاحتمـالي معـين تمامـاً، أي لدينـا توزيـع واحـد)، والنمـاذج الاحتماليـة الإحصائيـة (نوع التوزيع الاحتمـالي معلـوم لكنـه يعتمـد على معلمـة غير معلومـة).

الفصل الثالث:

المعاينة وتوزيعات المعاينة

1.3 مقدمة

لن نعمد في هذا الفصل إلى بحث كامل لنظرية المعاينة، بل سنقتصر على ما هو لازم وضروري لدراسة نظرية التقدير . لذا سنتعرض فقط للحالات التي توافق نموذجاً لملاحظات متكررة ومستقلة على متغير عشوائي حقيقي ع، ونقدم بعض المفاهيم الأساسية في نظرية المعاينة كالاستدلال الإحصائي، العينة العشوائية، التوزيع التجريبي، عزوم العينة، الإحصاءات المرتبة ...الخ. ومن ثم نبحث في توزيعات المعاينة لبعض الصيغ الخطية والتربيعية في متغيرات العينة، بالإضافة إلى دراسة موجزة للإحصاءات المرتبة وتوزيعاتها الاحتمالية، حيث إن هذه الأخيرة مهمة جداً عند دراسة الاستدلال الإحصائي اللامعلمي، وأخيراً، دراسة تقارب بعض مميزات العينة وتوزيعات المعاينة في حالة عينات كبيرة.

2.3 الاستدلال الإحصائي STATISTICAL INFERENCE

يعزى دائماً تقدم العلوم للتجريب، وإذا أنجز الباحث تجربة وتحصل على بعض المعلومات، وبناءً على هذه المعلومات توصل إلى نتائج معينة، وهذه النتائج عادة تتجاوز مواد وعمليات التجربة الخاصة. وبعبارة أخرى، يستطيع الباحث التعميم من التجربة الخاصة إلى صف كل التجارب المشابهة.

يدعى هذا النوع من التمديد من الخاص إلى العام بـ "الاستدلال الاستقرائي inductive inference"، ويعتبر إحدى طرق الوصول إلى معرفة جديدة.

لا تعطي طريقة الاستدلال الاستقرائي تعميمات مؤكدة، لكن يمكن قياس درجة الشك (أو درجة الثقة) بتلك التعميمات فيما إذا أنجزت التجربة وفق مبادئ معينة. وإحدى مهام الإحصاء تتمثل في تقديم الأساليب لإجراء استدلالات استقرائية وقياس درجة الثقة في كل استدلال وذلك بعبارة احتمالية.

وتجدر الإشارة هنا إلى نوع آخر من الاستدلال، الذي يدعى "الاستدلال الاستنتاجي deductive inference"، في حين أن النتائج التي نتوصل إليها بواسطة الاستدلال الاستقرائي احتمالية، إلا أن النتائج التي نتوصل إليها بواسطة الاستدلال الاستنتاجي قطعية (مؤكدة).

ولإيضاح الاستدلال الاستنتاجي، لنفترض لدينا العبارتين الآتيتين:

1. إحدى الزوايا الداخلية في كل مثلث قائم تساوي 90 درجة.
2. المثلث A قائم.

إذا قبلنا بالعبارتين السابقتين فنصل قطعاً إلى النتيجة:

3. إحدى الزوايا الداخلية في المثلث A تساوي 90 درجة.

تدعى العبارة (1) بالمقدمة الكبرى، والعبارة (2) بالمقدمة الصغرى، بينما العبارة (3) فتدعى بالنتيجة.

نلاحظ أن طريقة الاستدلال الاستنتاجي هي انتقال من العام إلى الخاص، أي عكس الاستدلال الاستقرائي، بالإضافة إلى أنها تعطي نتائج مؤكدة.

بالرغم من أن الاستدلال الاستنتاجي مهم جداً، وخاصة في الرياضيات، للوصول إلى معارف جديدة، إلا أن الكثير من المعارف الجديدة في مختلف الميادين يتم الوصول إليها باتباع طريقة الاستدلال الاستقرائي.

لنوضح الآن الاستدلال الاستقرائي بمثال بسيط. لنفترض لدينا صندوق تخزين يحتوي على 10 مليون بذرة ورد، ونعلم أن كل بذرة تنتج وردة بيضاء أو حمراء

فيما لو زرعت. ونريد معرفة ما يلي: كم عدد (أو نسبة) البذور ضمن الـ 10 مليون بذرة التي ستنتج ورود بيضاء؟ الطريقة الوحيدة التي تمكننا الإجابة على هذا السؤال وبدقة تامة هي زراعة كل البذور الـ 10 مليون وملاحظة عدد الورود البيضاء. لكن هذه الطريقة غير عملية لأننا نريد بيع البذور، وحتى إذا لم نرد بيع البذور، فنرغب في الإجابة على هذا السؤال بأقل تكلفة وجهد ممكنين وأحياناً في أسرع وقت. طبعاً، بدون زراعة كل البذور، ومن ثم ملاحظة عدد الورود البيضاء لا يمكننا بشكل مؤكد معرفة عدد البذور في صندوق التخزين التي ستنتج بذوراً بيضاء.

الطريقة المنطقية الأخرى هي: نستطيع زراعة جزء من البذور، وعلى أساس ألوان الورود الناتجة نحصل على إفادة (تقدير) حول عدد البذور ضمن الـ 10 مليون بذرة التي سوف تنتج وروداً بيضاء. وهذه إجابة غير مؤكدة عن عدد البذور التي ستنتج وروداً بيضاء، لكن يمكن صياغة عبارة احتمالية (درجة الثقة) بالإجابة فيما إذا كان ذلك الجزء من البذور مأخوذاً بطريقة معينة.

هكذا، فالاستدلال الاستقرائي: نأخذ جزءاً من الـ 10 مليون بذرة، نزرعها، نلاحظ عدد الورود البيضاء، وعلى أساسها نقدر عدد البذور التي ستنتج وروداً بيضاء ضمن الـ 10 مليون بذرة. أي من معرفة ألوان جزء نعمم على الكل (10 مليون بذرة) وهذا التعميم غير مؤكد. وإذا كان هذا الجزء من الكل مأخوذ بطريقة علمية (محققة لشروط معينة، وهذا ما سنتطرق إليه لاحقاً)، فيمكننا قياس درجة الثقة في التعميم باستخدام نظرية الاحتمالات، وعندئذٍ يطلق على الاستدلال الاستقرائي بـ "الاستدلال الإحصائي statistical inference".

يعالج الاستدلال الإحصائي مسألتين هامتين هما:

1. التقدير.
2. اختبار الفرضيات.

3.3 العينة العشوائية وتوزيعها

RANDOM SAMPLE AND ITS DISTRIBUTION

ليكن ξ متغيراً عشوائياً وحيد البعد (one-dimensional) دالة توزيعه $F(x)$ معرفة على R (فئة الأعداد الحقيقية)، ولنرمز لتوزيع ξ بشكل عام بـ $\mathcal{L}(\xi)$. تعرف العينة العشوائية بحجم n مأخوذة من مجتمع دالة توزيعه $F(x)$ على أنها متغير عشوائي ذي n بعد $X=(X_1,..,X_n)$ دالة توزيعه:

$$F_X(x)=F_X(x_1,\dots,x_n)=\prod_{i=1}^{n}F(x_i) \qquad (1.3.3)$$

معرفة على R^n.

نلاحظ من هذا التعريف أن العناصر أو المركبات $X_1,\dots,X_2$ للعينة العشوائية مستقلة ولكل منها نفس التوزيع، وهو توزيع المجتمع الذي سحبت منه، أي $F(x)$. ومن أجل الاختصار نقول عادة أن X عينة عشوائية بحجم n من $F(x)$ (أو من $\mathcal{L}(\xi)$). وبملاحظة أن العينة بحجم n من $F(x)$ عبارة عن مثال بسيط للعمليات العشوائية المنتهية، لذا تدعى أحياناً بالعينة العشوائية البسيطة (simple random sample) أو بالمعاينة العشوائية البسيطة (simple random sampling)، وفيما يلي حيث نذكر كلمة عينة نقصد بها عينة عشوائية.

إذا كان المتغير العشوائي ξ من النوع المنقطع (المتقطع) دالة احتماله $f(x)$، فإن دالة احتمال ودالة توزيع العينة X

$$f_X(x)=f_X(x_1,\dots,x_n)=\prod_{i=1}^{n}f(x_i)$$

$$F_X(x)=\prod_{i=1}^{n}\left(\sum_{x_{ij}<x_i}f(x_{ij})\right) \qquad (2.3.3)$$

حيث إن $x_{ij}\ ;\ j=1,2,\dots$ القيم الممكنة للمركبة $X_i\ ;\ i=1,\dots,n$.

وإذا كان المتغير ξ من النوع المستمر (المتصل)، فإن دالة كثافة ودالة توزيع العينة X:

$$f_X(x) = f(x_1,\ldots,x_n) = \prod_{i=1}^{n} f(x_i)$$
$$F_X(x) = \prod_{i=1}^{n}\left(\int_{-\infty}^{x_i} f(x)dx\right) \tag{3.3.3}$$

وهذا بناءً على مفهوم استقلال المتغيرات العشوائية في نظرية الاحتمالات.

بعد ملاحظة العينة $X = (X_1,\ldots,X_n)$، تكون القيم الفعلية لـ $X_1,\ldots,X_n$ معلومة، أي لدينا عينة عشوائية مسحوبة فعلاً من المجتمع $F(x)$، نرمز لها عادة بـ $x = (x_1,\ldots,x_n)$ وتدعى بالعينة العشوائية الملاحظة (المشكلة فعلاً).

ملاحظة 1.3.3

إن تعريفنا للمعاينة العشوائية يقتصر على المعاينة من مجتمع غير منتهي أو منتهي لكن السحب مع الإعادة، لأن المركبات $X_i\ ;\ i = 1,\ldots,n$ مستقلة.

4.3 التوزيع التجريبي *EXPERIMENTAL DISTRIBUTION*

إن مسألة استدلال التوزيع الاحتمالي للمتغير العشوائي الملاحظ ξ [توزيع المجتمع $\mathcal{L}(\xi)$] أو بعض خصائصه، في حالة كون $\mathcal{L}(\xi)$ غير معلوم، تتطلب في حالات عدة تبويب المعطيات الإحصائية التي تقدمها عينة عشوائية ملاحظة $x = (x_1,\ldots,x_n)$ من هذا المجتمع.

إذا كان المتغير العشوائي ξ من النوع المنقطع وحجم العينة n صغيراً، فنأخذ القيم المختلفة في العينة الملاحظة x ونرتبها تصاعدياً ومن ثم نعين التكرارات النسبية الموافقة لها، فمثلاً، إذا كان القيم الملاحظة $x_i\ ;\ i = 1,\ldots,n$ تشتمل على $k \le n$ قيمة مختلفة ولنرمز لها بـ $z_j\ ;\ j = 1,\ldots,k$، حيث $z_1 < z_2 < \ldots < z_k$، وكانت $\frac{n_j}{n}\ ;\ j = 1,\ldots,k$ التكرارات النسبية الموافقة لها، فتدعى فئة الثنائيات

بالتوزيع التجريبي للمتغير العشوائي الملاحظ ع الموافق للعينة $x = (x_1,\ldots,x_n)$، حيث إن $\{(z_j, n_j/n)\ ;\ j = 1,\ldots,k\}$، أي التوزيع التكراري النسبي لمعطيات العينة المشاهدة، $\sum_{j=1}^{k} \frac{n_j}{n} = 1$. وللتمييز نستخدم أحياناً مصطلح التوزيع النظري لـ ع للدلالة على التوزيع الاحتمالي له (توزيع المجتمع) ويتم عرض التوزيع التجريبي لـ ع جدولياً، كالجدول (1.4.3).

جدول (1.4.3): التوزيع التجريبي للمتغير العشوائي المنقطع ع

القيم الملاحظة المختلفة z_j	z_1	z_2		z_k
التكرارات النسبية الموافقة $\frac{n_j}{n}$	$\frac{n_1}{n}$	$\frac{n_2}{n}$		$\frac{n_k}{n}$

أما إذا كنا ندرس متغيراً عشوائياً مستمراً ع، أو كان منقطعاً لكن حجم العينة n كبير، فإن تبويب المعطيات الإحصائية يتطلب تقسيم المدى $r = x_{(n)} - x_{(1)}$ إلى k من الفئات الجزئية غير المتقاطعة، وتبسط الحسابات إلى حد كبير إذا كانت الفئات متساوية المدى $\frac{r}{k} \approx 1$، وهذا ما سوف نعتمده دوماً. وبعدئذٍ يتم حساب التكرارات النسبية الموافقة لتلك الفئات n_j/n. ونشير هنا إلى أن عدد الفئات k اختياري وعادة لا يقل عن 5 ولا يزيد عن 20، ويعتمد هذا الاختيار على طبيعة الظاهرة أو الخاصة المدروسة، بحيث يمكن اعتبار عناصر الفئة متجانسة بالنسبة للصفة المدروسة ع. ونحصل بالنتيجة على التوزيع التجريبي للمتغير العشوائي الملاحظ ع، كما هو مبين في الجدول (2.4.3).

جدول (2.4.3): التوزيع التجريبي للمتغير العشوائي ع

الفئات $(c_{j-1}, c_j]$	$(c_0, c_1]$	$(c_1, c_2]$	. . . .	$(c_{k-1}, c_k]$
التكرار النسبي $\frac{n_j}{n}$	$\frac{n_1}{n}$	$\frac{n_2}{n}$		$\frac{n_k}{n}$

حيث $\left(c_{j-1}, c_j\right]$ الفئة رقم $j = 1,2,\ldots,k$ و $\sum_{1}^{k} \frac{n_j}{n} = 1$.

يعتبر التكرار النسبي $\frac{n_j}{n}$ للحادث $B_j = \left(c_{j-1}, c_j\right]$ متغيراً عشوائياً (لأنه يتغير بشكل عام من عينة ملحوظة لأخرى) بقيمة متوقعة تساوي احتمال هذا الحادث، أي أن:

$$E\left(\frac{n_j}{n}\right) = P\left(c_{j-1} < \xi \le c_j\right) = P\left(B_j\right) = P_j \tag{1.4.3}$$

لأن المتغير العشوائي $Y = n_j$ يتبع توزيع ذي الحدين $B\left(n, p_j\right)$.

وكما نعلم من قانون الأعداد الكبيرة [مبرهنة بيرنولي (2.3.1)] أنه إذا كررت التجربة المفروضة n مرة تحت نفس الشروط، فإن التكرار النسبي لظهور الحادث B_j ينتهي إلى احتمال هذا الحادث، وذلك عندما $n \to \infty$، أي أن:

$$\lim_{n\to\infty} \frac{n_j}{n} = p_j$$

وهذا يعني، بملاحظة الجدولين (1.4.3) و (2.4.3)، أن السطر الثاني في كليهما يحتوي على القيم التقديرية للاحتمالات $p_j \; ; \; j = 1,2,\ldots,k$. أي أن:

$$\frac{n_j}{n} \approx p_j \quad ; \quad j = 1,\ldots,k$$

1.4.3 دالة التوزيع التجريبي
Experimental Distribution Function

لنفترض من أجل أي عدد حقيقي x فالمتغير Z_i معرف على النحو الآتي:

$$Z_i = \begin{cases} 1 & ; \quad x_i < x \\ 0 & ; \quad x_i \ge x \end{cases} \quad , \quad i = 1,2,\ldots,n \tag{2.4.3}$$

وعليه $Z=\sum_{i=1}^{n} Z_i$ متغـــير عشـــوائي يمثل عدد عناصر العينة الملاحظـة $x=(x_1,\dots,x_n)$ الأصغر من x.

إذا رمـــزنا بـــ $F_n^*(x)=\dfrac{Z}{n}$، فتدعى الدالـــة $F_n^*(x)$ بدالـــة التوزيـــع التجريـــبي لــــ ξ (experimental distribution function)، الموافقـــة للعينـــة الملاحظـــة x.

وللتمييـــز نسمي دالـــة التوزيع $F(x)$ للمتغير العشـــوائي ξ بدالـــة التوزيـــع النظـــري (توزيع المجتمع). وفي حالات عدة حيث لا يوجد التباس نســـقط الدليـــل n، أي نرمـــز بـــ $F^*(x)$ لدالة التوزيع التجريبي لــــ ξ.

مثـــال 1.4.3

إذا كان التوزيع التجريبي للمتغير العشوائي ξ معطى بالجدول (3.4.3). فأوجد دالـــة التوزيع التجريبي $F^*(x)$.

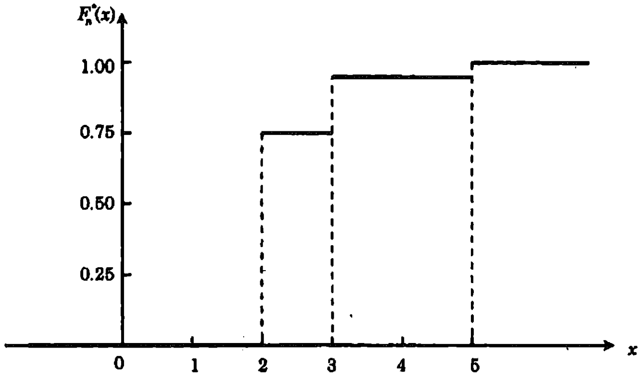
جدول (3.4.3): التوزيع التجريبي لـ ξ

القيم الملاحظة لـ ξ	2	3	5
التكرار النسبي $\frac{n_j}{n}$	0.75	0.20	0.05

من تعريف دالة التوزيع التجريـــبي $F_n^*(x)=\dfrac{Z}{n}$، نجـــد أن:

$$F_n^*(x)=\begin{cases} 0 & ; \quad x\le 2 \\ 0.75 & ; \quad 2<x\le 3 \\ 0.95 & ; \quad 3<x\le 5 \\ 1 & ; \quad x>5 \end{cases}$$

والتمثيل البياني لدالة التوزيع التجريبي $F_n^*(x)$ يبينـــه الشـــكل (1.4.3).

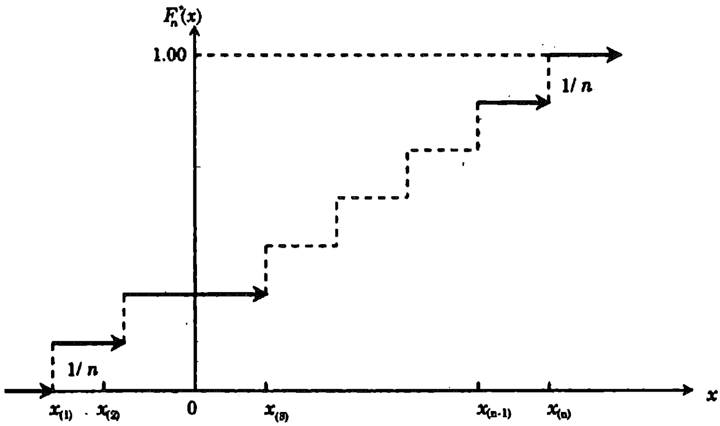


شكل (1.4.3)

إذا كانت عناصر المتجه x مختلفة فإن دالة التوزيع التجريبي $F_n^*(x)$ تعطى كما يلي:

$$F_n^*(x) = \begin{cases} 0 & ; \quad x \le x_{(1)} \\ k/n & ; \quad x_{(k)} < x \le x_{(k+1)} \quad , \quad k = 1,2,\dots,n-1 \\ 1 & ; \quad x > x_{(n)} \end{cases}$$

وهذا يعني، في هذه الحالة، أن قيم القفزات متساوية وتساوي $1/n$.
والتمثيل البياني لـ $F_n^*(x)$ يأخذ الشكل (2.4.3).



شكل (2.4.3)

وفي الحالة العامة يمكن كتابة دالة التوزيع التجريبي $F_n^*(x)$ على النحو الآتي:

$$F_n^*(x) = \frac{1}{n}\sum_{k=1}^{n} e\left(x - x_{(k)}\right) \tag{3.4.3}$$

حيث $e(x)$ ما هي إلا دالة الواحدية للقفزات (دالة خوفسيد):

$$e(x) = \begin{cases} 0 & ; \quad x \le 0 \\ 1 & ; \quad x > 0 \end{cases}$$

وفي العلاقة (3.4.3) يبدو بوضوح ارتباط $F_n^*(x)$ بالعينة x.

إن لدالة التوزيع التجريبي دوراً هاماً وأساسياً في الإحصاء الرياضي، وتكمن أهميتها الخاصة بأنه بازدياد عدد الملاحظات للمتغير العشوائي ξ تقترب دالة التوزيع التجريبي $F_n^*(x)$ من دالة التوزيع النظري $F(x)$. وهذا ما سنراه لاحقاً.

2.4.3 التمثيل البياني للتوزيعات التجريبية

نلاحظ مما سبق أن التوزيع التجريبي يعتبر طريقة مفيدة لتمثيل المعطيات الإحصائية (معطيات العينة الملاحظة) والتي تمكننا من استدلال بعض النتائج عن التوزيع الاحتمالي للمتغير العشوائي الملاحظ ξ عندما يكون مجهولاً بشكل كامل أو جزئي. هذا وتوجد طرق أخرى لتمثيل المعطيات الإحصائية، وإحدى هذه الطرق تتمثل في بناء المدرج والمضلع للتوزيع التجريبي لـ ξ. ويستخدم مضلع التوزيع التجريبي عندما يكون المتغير العشوائي الملاحظ ξ مستمراً أو منقطعاً، أما مدرج التوزيع التجريبي فيقتصر استخدامه على حالة كون المتغير العشوائي الملاحظ ξ مستمراً، أو أن العينة العشوائية X ذات حجم كبير (المعطيات الإحصائية مبوبة حسب الفئات).

لنوضح كيفية بناء مدرج ومضلع توزيع تجريبي لمتغير عشوائي ξ من خلال المثال التالي.

مثـــــال 2.4.3

لتكـن لدينـا نتـائج دراسـة متانـة 200 نمـوذج مـن البيتـون المسـلح، كمـا هـي واردة في الجـدول التـالي:

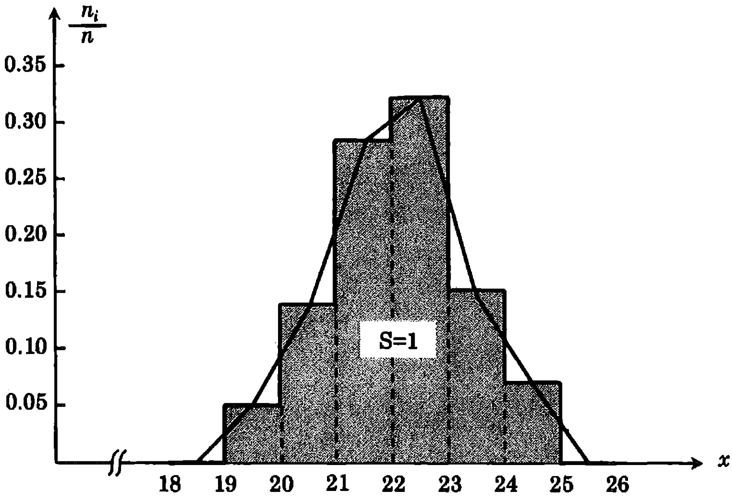
جـدول 4.4.3:

مجال المتانة (ميجا باسكال)	التكرار النسبي $\frac{n_j}{n}$
20–19	0.05
21–20	0.13
22–21	0.28
23–22	0.32
24–23	0.15
25–24	0.07

المطلـوب بنـاء مـدرج ومضلـع التوزيـع التجريبي المعطـى.

لبنـاء المـدرج التكـراري لهـذا التوزيـع التجريـبي نرسـم محوريـن متعـامدين، يمثـل المحـور الأفقـي الفـترات الجزئيـة والمحـور الشـاقولي التكـرارات النسـبية. ثـم نحـدد الفـترات الجزئيـة للقيـم الملاحظـة والتكـرارات النسـبية الموافقـة علـى المحـول الأفقـي والشـاقولي علـى الـترتيب، وبعـد ذلـك نقيـم مسـتطيلات علـى تلـك الفـترات الجزئيـة بارتفاعـات تسـاوي التكـرارات النسبية الموافقـة، فنحصـل بالنتيجـة علـى المـدرج التكـراري المبـين بالشـكل (3.4.3).

لرسـم مضلـع التوزيـع التجريـبي، نحـدد منتصفـات الأضـلاع العليـا للمسـتطيلات المبينـة علـى الشـكل (3.4.3)، ثـم نصـل بينهـا علـى الـترتيب بخطـوط مسـتقيمة فنحصـل علـى المطلـوب. ولإغـلاق المضلـع نضيـف فئـة قبـل الفئـة الأولى وأخـرى بعـد الفئـة الأخـيرة بحيث يكـون التكـرار النسبي لكـل منهـا صفـراً، فنحصـل علـى المضلـع المبـين علـى الشـكل (3.4.3).



شكل (3.4.3)

يستخدم مدرج ومضلع التوزيع التجريبي لاستدلال نوع نموذج التوزيع لـ ع. فمثلاً، النموذج المبين على الشكل (3.4.3) يذكرنا بالنموذج الطبيعي العام، لهذا يمكن افتراض أن نوع توزيع المتغير العشوائي الملاحظ ع هو طبيعي، هكذا، يمكن اعتبار الشكل البياني للتوزيع التجريبي لمتغير عشوائي ع كمناظر إحصائي لشكل التوزيع النظري له.

5.3 الإحصاء وعزوم العينة

STATISTIC AND SAMPLE MOMENT

إن تشكيل العينة العشوائية ليس هدفاً بحد ذاته، بل الهدف هو الوصول إلى استدلالات حول المجتمع الذي سحبت منه، وهذه الاستدلالات قد تكون حول توزيع المجتمع أو حول بعض أو كل مميزاته العددية (معالمه). لهذا في نظرية المعاينة لا نكتفي بمعرفة التوزيع التجريبي أو التوزيع الاحتمالي للعينة العشوائية، بل نهتم بمعرفة دالة أو أكثر في المتغيرات العشوائية المكونة للعينة، حيث تستخدم للقيام

باستدلالات حول معالم المجتمع. ومن أهم تلك الدوال ما تدعى بعزوم العينة (sample moments). سنعرف أولاً الإحصاء أو الإحصاءة (statistic).

تعريف 1.5.3: الإحصاء Statistic

إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من التوزيع $\mathcal{L}(\xi)$، فإن أي دالة في العينة فقط $T=T(X)$ تدعى "إحصاء".

نلاحظ من هذا التعريف أن الإحصاء $T=T(X)$ لا يحتوي على أي معلمة مجهولة، كما أنه يعتبر متغيراً عشوائياً لأن قيمته، بشكل عام، تتغير من عينة ملاحظة لأخرى، وبالتالي له توزيع احتمالي يدعى بتوزيع المعاينة (sampling distribution) للإحصاء T.

فمثلاً، إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من التوزيع $\mathcal{L}(\xi)\in N(\theta_1,\theta_2^2)$، حيث θ_1,θ_2 معلمتين مجهولتين، فإن الدالة $X_1-\theta_1$ وكذلك الدالة $\sum_1^n X_i/\theta_2$ ليست إحصاء (لأن كليهما ليس دالة في العينة فقط، بل يحويان معلمة مجهولة)، لكن $\sum_1^n X_i$، $\frac{X_1+X_2}{4}$ و $\sum_1^n \ln X_i$ إحصاءات.

إحدى المسائل الأساسية في الإحصاء هي إيجاد الإحصاءات المناسبة لتقدير معالم المجتمع.

سنعرف فيما يلي ونناقش بعض الإحصاءات الهامة وهي عزوم العينة.

تعريف 2.5.3: عزوم العينة Sample moments

لتكن $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع المتغير العشوائي الملاحظ ξ ولنرمز له بـ $f(x)$. عندئذٍ فالعزم الابتدائي (حول 0) للعينة X من المرتبة r نرمز له بـ $A_r=A_r(X)$، ويعرف كالآتي:

$$A_r(X)=\frac{1}{n}\sum_{i=1}^{n}X_i^r \quad ; \quad r=1,2,\ldots \tag{1.5.3}$$

وكحالة خاصة، عندما $r = 1$ ، نحصل على متوسط العينة، ويرمز له عادة بـ $\overline{X}$ أو $\overline{X}_n$ ، أي أن:

$$A_1(X) = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{2.5.3}$$

وبشكل مشابه، العزم المركزي (حول $\overline{X}$) للعينة من المرتبة r، نرمز له بـ $M_r = M_r(X)$ ، ويعرف كالآتي:

$$M_r(X) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^r \quad ; \quad r = 1,2,\ldots \tag{3.5.3}$$

وكحالة خاصة، عندما $r = 2$ ، نحصل على تباين العينة، ويرمز له عادة بـ $S^2 = S^2(X)$ أو $S_n^2 = S_n^2(X)$ ؛ أي أن:

$$M_2(X) = S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \tag{4.5.3}$$

وكما نعلم من نظرية الاحتمالات، إذا رمزنا بـ α_r و μ_r للعزم الابتدائي والمركزي من المرتبة r للمتغير العشوائي ξ (العزوم النظرية) على الترتيب، فإن:

$$\begin{aligned} &\alpha_r = E\xi^r \quad ; \quad r = 1,2,\ldots \\ &\mu_r = E(\xi - \mu)^r \quad ; \quad r = 1,2,\ldots \quad , \quad \mu = \alpha_1 \end{aligned} \tag{5.5.3}$$

وأن هنالك علاقة بينهما وهي:

$$\mu_r = \sum_{j=0}^{r}(-1)^j C_r^j \mu^j \alpha_{r-j} \tag{6.5.3}$$

وأن العلاقة (6.5.3) موجودة أيضاً بين العزوم الابتدائية والعزوم المركزية للعينة:

$$M_r = \sum_{j=0}^{r}(-1)^j C_r^j \overline{X}^j A_{r-j} \quad ; \quad C_r^j = \frac{r!}{j!(r-j)!} \tag{7.5.3}$$

وبإعطاء $r = 2,3,4$ نجد:

$$\begin{aligned} &S^2 = A_2 - \overline{X}^2 \quad , \quad M_3 = A_3 - 3\overline{X}A_2 + 2\overline{X}^3 \\ &M_4 = A_4 - 4\overline{X}A_3 + 6A_2\overline{X}^2 - 3\overline{X}^4 \end{aligned} \tag{8.5.3}$$

6.3 متوسط وتباين متوسط العينة

MEAN AND VARIANCE OF SAMPLE MEAN

نفترض $X=(X_1,\dots,X_n)$ عينة عشوائية من توزيع $f(x)$ بمتوسط μ وتباين σ^2، ونريد تعيين القيمة المتوقعة والتباين لمتوسط العينة $\overline{X}$.

لدينا حسب التعريف (العلاقة 2.5.3):

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i=\frac{1}{n}\left(X_1+\dots+X_n\right)$$

بأخذ القيمة المتوقعة للطرفين، نجد:

$$E\overline{X}=\frac{1}{n}\left(EX_1+\dots+EX_n\right)$$

لكن $X_1,\dots,X_n$ متغيرات عشوائية مستقلة ولكل منها التوزيع $f(x)$، إذن:

$$EX_1=EX_2=\dots=EX_n=\mu$$

وبالتالي:

$$E\overline{X}=\mu \qquad (1.6.3)$$

إيجاد $\operatorname{var}\overline{X}$:

بما أن:

$$\operatorname{var}\overline{X}=E\left(\overline{X}-E\overline{X}\right)^2=E\left(\overline{X}-\mu\right)^2$$

و

$$\overline{X}-\mu=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\mu\right)$$

فإن:

$$\begin{aligned}\operatorname{var}\overline{X}&=E\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\mu\right)\right]^2\\&=\frac{1}{n^2}\sum_{i=1}^{n}E\left(X_i-\mu\right)^2+\frac{1}{n^2}\sum_{i\neq j}E\left[\left(X_i-\mu\right)\left(X_j-\mu\right)\right]\end{aligned} \qquad (2.6.3)$$

لكن حسب تعريف العينـــة العشـــوائية، فالمتغـــيرات $X_i \; ; \; i = 1,2,\ldots,n$ مســـتقلة ولها نفس التوزيـــع، إذن:

$$E(X_i - \mu)^2 = \sigma^2 \quad ; \quad i = 1,\ldots,n$$
$$\operatorname{cov}(X_i, X_j) = E[(X_i - \mu)(X_j - \mu)] = 0 \quad ; \quad i \neq j \tag{3.6.3}$$

وبالتعويض في (2.6.3)، نجـــد:

$$\operatorname{var} \overline{X} = V\overline{X} = \frac{\sigma^2}{n} \tag{4.6.3}$$

7.3 متوسط وتباين العينة

Mean and Variance of Sample Variance

لنتأمل الآن تباين العينة، المعـــرف بالعلاقـــة (4.5.3):

$$\begin{aligned}
S^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[(X_i - \mu) - \frac{1}{n}\sum_{j=1}^{n}(X_j - \mu)\right]^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{(X_i - \mu)^2 - \frac{2}{n}(X_i - \mu)\sum_{j=1}^{n}(X_j - \mu) + \frac{1}{n^2}\left[\sum_{j=1}^{n}(X_j - \mu)\right]^2\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 - \frac{2}{n^2}\left[\sum_{j=1}^{n}(X_j - \mu)\right]^2 + \frac{1}{n^2}\left[\sum_{j=1}^{n}(X_j - \mu)\right]^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 - \frac{1}{n^2}\left[\sum_{j=1}^{n}(X_j - \mu)\right]^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 - \frac{1}{n^2}\left[\sum_{j=1}^{n}(X_j - \mu)^2 + \sum_{i \neq j}(X_i - \mu)(X_j - \mu)\right] \\
&= \frac{n-1}{n^2}\sum_{i=1}^{n}(X_i - \mu)^2 - \frac{1}{n^2}\sum_{i \neq j}(X_i - \mu)(X_j - \mu)
\end{aligned}$$

بأخذ القيمة المتوقعة والاستفادة مـــن (3.6.3)، نجـــد:

$$ES^2 = \frac{n-1}{n}\sigma^2 \tag{1.7.3}$$

إيجـــاد $\operatorname{var} S^2$:

كما نعلـم:

$$\operatorname{var} S^2 = ES^4 - \left(ES^2\right)^2$$

ولحساب ES^4 نضع $Y_i = X_i - \mu$ في عبارة S^2 ثم نربعـــها، فنجـــد:

$$S^2 = \frac{n-1}{n^2}\sum_{i=1}^{n} Y_i^2 - \frac{1}{n^2}\sum_{i\neq j}^{n} Y_i Y_j$$

$$S^4 = \frac{(n-1)^2}{n^4}\left(\sum_{i=1}^{n} Y_i^2\right)^2 + \frac{1}{n^4}\left(\sum_{i\neq j}^{n} Y_i Y_j\right)^2 - \frac{2(n-1)}{n^4}\sum_{i=1}^{n} Y_i^2 \sum_{i\neq j}^{n} Y_i Y_j$$

وبما أن $EY_i = EX_i - \mu = 0$، فبأخذ توقـــع S^4 نجـــد:

$$ES^4 = \frac{(n-1)^2}{n^4}\sum_{i=1}^{n} EY_i^4 + \frac{(n-1)^2}{n^4}\sum_{i\neq j}^{n} EY_i^2 EY_j^2 + \frac{2}{n^4}\sum_{i\neq j}^{n} EY_i^2 EY_j^2$$
$$= \frac{(n-1)^2}{n^3}\mu_4 + \frac{(n-1)^2+2}{n^3}(n-1)\mu_2^2$$

وبالتـالي:

$$\operatorname{var} S^4 = ES^4 - \left(ES^2\right)^2$$
$$= \frac{(n-1)^2}{n^3}\mu_4 + \frac{(n-1)^2+2}{n^3}(n-1)\mu_2^2 - \frac{(n-1)^2}{n^2}\mu_2^2 \tag{2.7.3}$$
$$= \frac{(n-1)^2}{n^3}\left(\mu_4 - \frac{n-3}{n-1}\mu_2^2\right) \quad ; \quad \mu_2 = \sigma^2$$

يمكن تلخيص ما سبق على النحـــو الآتي:

إذا كانت $X=(X_1,\ldots,X_n)$ عينة من توزيع $f(x)$ بمتوسط μ وتباين σ^2، فإن متوسط وتباين متوسط العينة $\overline{X}$ هما:

$$E\overline{X}=\mu \quad , \quad \operatorname{var}\overline{X}=\frac{\sigma^2}{n}$$

بالإضافة إلى ذلك، إذا كان العزم المركزي من المرتبة الرابعة لتوزيع المجتمع منتهي، عندئذٍ تباين العينة S^2 له متوسط وتباين هما:

$$ES^2=\frac{n-1}{n}\sigma^2$$
$$\operatorname{var}S^2=\frac{(n-1)^2}{n^3}\left(\mu_4-\frac{n-3}{n-1}\sigma^4\right)$$

ملاحظة

يستخدم عادة في التطبيقات الإحصائية تباين العينة المعدل S^{*2}.

$$S^{*2}(X)=\frac{n}{n-1}S^2(X) \qquad (3.7.3)$$

وبالتالي:

$$ES^{*2}=E\left(\frac{n}{n-1}S^2\right)=\frac{n}{n-1}ES^2=\sigma^2 \qquad (4.7.3)$$

$$\begin{aligned}\operatorname{var}S^{*2}&=\operatorname{var}\left(\frac{n}{n-1}S^2\right)=\left(\frac{n}{n-1}\right)^2\operatorname{var}S^2\\&=\frac{1}{n}\left(\mu_4-\frac{n-3}{n-1}\sigma^2\right)\end{aligned} \qquad (5.7.3)$$

ومبرر استخدام S^{*2} في التطبيقات الإحصائية هو أن $ES^{*2}=\sigma^2$، وهذا يعني S^{*2} مقدّر غير متحيز للتباين، وتجاوزاً يدعى S^{*2} بتباين العينة $X=(X_1,\ldots,X_n)$.

8.3 توزيعات المعاينة لمجموع ومتوسط عينة

SAMPLING DISTRIBUTIONS OF SAMPLE SUM AND MEAN

سنحتاج في موضوعات قادمة لمعرفة توزيع مجموع ومتوسط عينة عشوائية، ولإيجاد ذلك هناك طريقتان سنتطرق إليها في الفقرتين التاليتين.

1.8.3 طريقة التكرار

لدراسة طريقة التكرار لإيجاد توزيع مجموع ومتوسط عينة، وللتبسيط نفترض أولاً أن العينة بحجم $n = 2$، إذا رمزنا بـ $G_2(z)$ لدالة توزيع المجموع $Z = X_1 + X_2$، فإن:

$$G_2(z) = P(X_1 + X_2 < z) = \int_D d\big(F(x_1)F(x_2)\big)$$

حيث إن $D = \{(x_1, x_2) \in R^2 \ ; \ x_1 + x_2 < z\}$، وبما أن X_1, X_2 متغيرين عشوائيين مستقلين، يمكننا كتابة:

$$\begin{aligned} G_2(z) &= \int_D d\big(F(x_1)F(x_2)\big) = \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{Z-x_1} dF(x_2)\right] dF(x_1) \\ &= \int_{-\infty}^{+\infty} F(z - x_1)\,dF(x_1) \end{aligned} \qquad (1.8.3)$$

ويمكن تعميم ذلك على مجموع عينة حجمها $n > 2$ فنجد:

$$G_n(z) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{Z-(x_1+\ldots+x_{n-3})} \int_{-\infty}^{Z-(x_1+\ldots+x_{n-2})} F(z - x_1 - \ldots - x_{n-1})\,dF(x_{n-1})\,dF(x_{n-2}) \ldots dF(x_1)$$

. . . . (2.8.3)

وإذا رمزنا بـ $H_n(\bar{x})$ لدالة توزيع متوسط العينة $\overline{X}$، فإن:

$$H_n(\bar{x}) \equiv G_n(n\bar{x})$$

هكذا، إذا كانت $X=(X_1,\dots,X_n)$ عينة عشوائية من توزيع $F(x)$، فإن دالة توزيع مجموع العينة $Z=\sum_{i=1}^{n}X_i$ تعطى بالعلاقة (2.8.3)، ودالة توزيع متوسط العينة $\overline{X}$ بالعلاقة:

$$H_n(\bar{x}) \equiv G_n(n\bar{x}) \tag{3.8.3}$$

نلاحظ أن هذه الطريقة صعبة وتتطلب إجراءات متتالية من التكاملات. في بعض الحالات، وخاصة عندما يكون الإحصاء الذي نرغب في معرفة توزيعه الاحتمالي عبارة عن دالة خطية في متغيرات عشوائية مستقلة، يمكن دائماً الحصول على التوزيع باستخدام طريقة سهلة تدعى بطريقة الدالة المميزة.

2.8.3 طريقة الدالة المميزة

تعتبر طريقة الدالة المميزة طريقة بسيطة ومفيدة لتعيين توزيع المعاينة للمجاميع والمتوسطات في حالات عدة.

في البداية سنُذكّر بتعريف الدالة المميزة، ومن ثم نقدم بعض خواصها الأساسية بصورة مبرهنات بدون إثبات، حيث إن إثباتها بسيط نتركه للقارئ.

تعريف 1.8.3: الدالة المميزة The Characteristic Function

ليكن ξ متغير عشوائي حقيقي دالة توزيعه $F(x;\theta)$. تدعى الدالة $\varphi(t;\theta)$ المعرفة كالآتي:

$$\varphi(t;\theta) = E_\theta\left(e^{it\xi}\right) = \int_{-\infty}^{+\infty} e^{itx}\,dF(x;\theta) \tag{4.8.3}$$

بالدالة المميزة للنموذج $F(x;\theta)$ أو الدالة المميزة لـ ξ بالمعلمة θ، حيث إن $i=\sqrt{-1}$ و t عدد حقيقي.

وبشكل مشابه، بفرض $\eta=(X_1,\dots,X_n)$ متغير عشوائي ذي n بعد، دالة توزيعه

$F(x_1,\ldots,x_n;\theta)$ ، فــإن الــدالـــة المميـــزة $\varphi(t_1,\ldots,t_n;\theta)$ للمتغــير العشــوائي η تعرف كــالآتي:

$$\begin{aligned}\varphi_h(t_1,\ldots,t_n;\theta) &= E_\theta\left[\exp\left(i\sum_{j=1}^{n} t_j X_j\right)\right] \\ &= \int_{R^n}\exp\left(i\sum_{j=1}^{n} t_j X_j\right)dF(x_1,\ldots,x_n;\theta)\end{aligned} \tag{5.8.3}$$

مـــبرهنة 1.8.3

إذا كـــان X_1,X_2 متغيرين عشوائيين دالتي توزيعهما $F_1(x),F_2(x)$ على الــترتيب، وكانت $\varphi_1(t),\varphi_2(t)$ الدالتين المميزتين لــ X_1,X_2 علــى الــترتيب، فــإن الشــرط اللازم والكافي لتساوي التوزيعين $F_1(x),F_2(x)$ ، أي $F_1(x)\equiv F_2(x)$ هـــو تســاوي دالتيهما المميزتين $\varphi_1(t)\equiv\varphi_2(t)$.

هذا التوافق الأحادي (تقابل واحد لواحد) بـــين دوال التوزيـــع والـــدوال المميـــزة له أهمية كبيرة في نظرية الاحتمالات والإحصــاء الريــاضي، حيـــث يعطـــي الأســاس لإيجاد التوزيعات الاحتمالية والنماذج الإحصائية لمتغيرات عشوائية مـــن خـــلال دوالهـــا المميزة.

مـــبرهنة 2.8.3

إذا كـــانت Y دالة خطيـــة في k متغـــير عشـــوائي مســـتقل $X_1,\ldots,X_k$ معرفـــة كـالآتي:

$$Y=\sum_{j=1}^{n} c_j X_j$$

وكانت $\varphi(t)$ الدالة المميزة لــ Y و $\varphi_i(t_i)\ ;\ i=1,\ldots,k$ الدالة المـــيزة للمتغـــير العشوائي X_i ، فإن:

$$\varphi(t)=\prod_{j=1}^{k}\varphi_j(c_j t_j) \tag{6.8.3}$$

مبرهنة 3.8.3

إذا كانت $X=(X_1,\dots,X_2)$ عينة عشوائية من توزيع $F(x;\theta)$، وكانت $\varphi(t;\theta)$ الدالة المميزة للتوزيع $F(x;\theta)$، فإن الدالة المميزة لمجموع العينة $Z=X_1+\dots+X_n$:

$$\varphi_Z(t;\theta)=\left[\varphi(t;\theta)\right]^n \tag{7.8.3}$$

والدالة المميزة لمتوسط العينة $\overline{X}$:

$$\varphi_{\overline{X}}(t;\theta)=\left[\varphi\left(\frac{t}{n};\theta\right)\right]^n \tag{8.8.3}$$

مبرهنة 4.8.3

بفرض $X=(X_1,\dots,X_n)$ عينة عشوائية من توزيع $F(x;\theta)$ دالته المميزة $\varphi(t;\theta)$. عندئذٍ إذا كان:

$$\left[\varphi(t;\theta)\right]^n=\varphi(t;n\theta)$$

فإن توزيع مجموع العينة Z ومتوسط العينة $\overline{X}$ هو من الشكل $F(z;n\theta)$ و $F(n\overline{x};n\theta)$ على الترتيب.

مبرهنة 5.8.3

إذا كان العزم الابتدائي من المرتبة r للمتغير العشوائي ξ موجوداً، فإن:

$$\varphi(t)=1+\sum_{j=1}^{r}\frac{(it)^j}{j!}\alpha_k+0\left(t^r\right)$$

حيث إن:

$$\lim_{t\to\infty}\frac{0\left(t^r\right)}{t^r}=0$$

3.8.3 توزيعات المعاينة لمجموع ومتوسط عينة في بعض الحالات الخاصة

يمكننا الآن باستخدام الدالة المميزة وبعض خواصها الأساسية معرفة توزيع المعاينة لمجموع ومتوسط عينة في بعض الحالات الخاصة والهامة، التي سنحتاجها في فصول لاحقة.

.1 إذا كانت العينة العشوائية $X = (X_1,\dots,X_n)$ من توزيع بيرنولي $B(1,\theta)$، فإن توزيع المعاينة للمجموع $Z = \sum_{i=1}^{n} X_i$ هو توزيع ذي الحدين $B(n,\theta)$.

إن الدالة المميزة للتوزيع $B(1,\theta)$ هي:

$$\varphi(t) = \left[(1-\theta) + \theta e^{it}\right]$$

وحسب المبرهنة (3.8.3) [العلاقة (7.8.3)]، فإن الدالة المميزة لـ $Z = \sum_{i=1}^{n} X_i$:

$$\varphi_Z(t) = \left[(1-\theta) + \theta e^{it}\right]^n \tag{9.8.3}$$

وهذه ما هي إلا الدالة المميزة للتوزيع $B(n,\theta)$، أي أن توزيع Z هو $B(n,\theta)$.

وحسب المبرهنة (3.8.3) [العلاقة (8.8.3)] فالدالة المميزة لـ $\overline{X}$ هي:

$$\varphi_{\overline{X}}(t;\theta) = \varphi_Z\left(\frac{t}{n};\theta\right) = \left[(1-\theta) + \theta e^{\frac{it}{n}}\right]^n \tag{10.8.3}$$

وهذه أيضاً الدالة المميزة لنموذج ذي الحدين $B(n,\theta)$، أي أن توزيع المعاينة لـ $\overline{X}$ هو توزيع ذي الحدين $B(n,\theta)$.

.2 إذا كانت $X = (X_1,\dots,X_n)$ عينة عشوائية من توزيع بواسون $\Pi(\theta)$،

فإن توزيع المعاينة لـ $Z = \sum_{i=1}^{n} X_i$ هو نموذج بواسون $\Pi(n\theta)$.

إن الدالة المميزة للنموذج $\Pi(\theta)$ هي:

$$\varphi(t;\theta) = e^{-\theta\left(1-e^{it}\right)}$$

ومنها حسب المبرهنة (3.8.3) [العلاقة (7.8.3)] فالدالة المميزة لـ Z هي:

$$\varphi_Z(t;\theta) = e^{-n\theta\left(1-e^{it}\right)} \tag{11.8.3}$$

وهذه ما هي إلا الدالة المميزة لنموذج بواسون $\Pi(n\theta)$. ومن ثم توزيع المعاينة لـ $\overline{X}$ هو أيضاً نموذج بواسون بمتوسط $n\theta$.

3. إذا كانت $X = (X_1,\dots,X_n)$ عينة عشوائية من نموذج طبيعي عام $N(\theta_1,\theta_2^2)$، فإن توزيع المعاينة لـ $Z = \sum_{1}^{n} X_i$ هو $N(n\theta_1, n\theta_2^2)$، وتوزيع المعاينة لـ $\overline{X}$ يكون $N(\theta_1,\theta_2^2/n)$.

إن الدالة المميزة للتوزيع $N(\theta_1,\theta_2^2)$ هي:

$$\varphi(t;\theta) = e^{i\theta_1 t - \frac{1}{2}\theta_2^2 t^2} \quad ; \quad \theta = (\theta_1,\theta_2^2)$$

ومن ثم فالدالة المميزة لـ Z هي:

$$\varphi_Z(t;\theta) = \left[\varphi(t;\theta)\right]^n = e^{i(n\theta_1)t - \frac{1}{2}(n\theta_2^2)t^2} \tag{12.8.3}$$

وهذه الأخيرة ما هي إلا الدالة المميزة للنموذج $N(n\theta_1, n\theta_2^2)$، ومن ثم فهو توزيع المعاينة لـ Z. وللحصول على الدالة المميزة لـ $\overline{X}$ نستبدل t بـ $\frac{t}{n}$ في $\varphi_Z(t;\theta)$:

$$\varphi_{\overline{X}}(t;\theta) = \varphi_Z\left(\frac{t}{n};\theta\right) = e^{i\theta_1 t - \frac{1}{2}\left(\frac{\theta_2^2}{n}\right)t^2} \tag{13.8.3}$$

وهذه الدالة المميزة للنموذج $N\left(\theta_1,\theta_2^2/n\right)$، أي أن توزيع المعاينة لـ $\overline{X}$ هو من الشكل $N\left(\theta_1,\theta_2^2/n\right)$.

4. إذا كانت $\left(Y_1,\ldots,Y_n\right)$، حيث $Y_i=\left(X_{1i},\ldots,X_{ki}\right)\ ;\ i=1,2,\ldots,n$، عينة عشوائية من توزيع طبيعي ذي بعد k

$$N(\mu,\Sigma)\quad;\quad \mu=\left(\mu_1,\ldots,\mu_k\right)\quad,\quad \Sigma=\left\|\sigma_{ij}\right\|;i,j=1,\ldots,k$$

فإن توزيع المعاينة لمتجه المجاميع $\left(Z_1,\ldots,Z_k\right)\ ;\ Z_r=\sum_{i=1}^{n}X_{ri},r=1,2,\ldots,k$ هو التوزيع الطبيعي $N(n\mu;n\Sigma)$، وتوزيع المعاينة لمتجه المتوسطات $\left(\overline{X}_1,\ldots,\overline{X}_n\right);\overline{X}_r=\frac{1}{n}\sum_{i=1}^{n}X_{ri}$ يكون $N\left(\mu,\frac{1}{n}\Sigma\right)$.

الإثبات هنا، مشابه للحالة (3)، ولذا نترك ذلك للقارئ على سبيل المثال.

5. إذا كانت $\left(X_1,\ldots,X_n\right)$ عينة عشوائية من توزيع $\Gamma(\alpha,1)$، فإن توزيع المعاينة لـ $Z=\sum_{i=1}^{n}X_i$ هو توزيع جاما $\Gamma(n\alpha,1)$.

الدالة المميزة لنموذج $\Gamma(\alpha,1)$ هي:

$$\varphi(t;\alpha)=(1-it)^{-\alpha}$$

ومن ثم فالدالة المميزة لـ Z هي:

$$\varphi_Z(t;\alpha)=(1-it)^{-n\alpha}\qquad(14.8.3)$$

وهي الدالة المميزة لتوزيع $\Gamma(n\alpha,1)$. وللحصول على الدالة المميزة لمتوسط المعاينة $\overline{X}$، نستبدل t بـ $\frac{t}{n}$ في عبارة $\varphi_Z(t;\alpha)$، فنجد:

$$\varphi_{\overline{X}}(t;\alpha)=\varphi_Z\left(\frac{t}{n};\alpha\right)=\left(1-i\frac{1}{n}t\right)^{-n\alpha}\qquad(15.8.3)$$

وهي عبارة عن الدالة المميزة لتوزيع جاما $\Gamma(n\alpha,1/n)$.

6. إذا كانت $(X_{k1},\ldots,X_{kn_k}),\ldots\ldots,(X_{11},\ldots,X_{1n_1})$ عبارة عن k عينة مستقلة من التوزيعات الطبيعية $N(\theta_{1k},\theta_{2k}^2),\ldots,N(\theta_{12},\theta_{22}^2),N(\theta_{11},\theta_{21}^2)$ على الترتيب، وكان $\overline{X}_1,\ldots,\overline{X}_n$ متوسطات هذه العينات على الترتيب، و $c_1,\ldots,c_k$ ثوابت (ليست كلها أصفاراً)، فعندئذٍ توزيع المعاينة لـ:

$$Y = c_1\overline{X}_1+\ldots+c_k\overline{X}_k$$

هو التوزيع الطبيعي:

$$N\left(\sum_{i=1}^{k}c_i\theta_{1i}\ ,\ \sum_{i=1}^{k}\frac{c_i^2\theta_{2i}^2}{n_i}\right)$$

باستخدام الدوال المميزة نحصل على المطلوب مباشرة.

نلاحظ، بوضع $c_1=1\ ,\ c_2=-1\ ,\ c_3=\ldots\ldots=c_k=0$، نحصل على النتيجة التالية:

توزيع المعاينة للفرق $\overline{X}_1-\overline{X}_2$ بين متوسطي عينتين مستقلتين مأخوذتين من توزيعين طبيعيين $(X_{11},\ldots,X_{1n_1})$ ، $(X_{21},\ldots,X_{2n_2})$ $N(\theta_{12},\theta_{22}^2)$ و $N(\theta_{11}^2,\theta_{21}^2)$ على الترتيب هو التوزيع الطبيعي:

$$N\left((\theta_{11}-\theta_{12})\ ,\ \frac{\theta_{21}^2}{n_1}+\frac{\theta_{22}^2}{n_2}\right)$$

9.3 توزيعات المعاينة لصيغ تربيعية معينة في عينات من توزيع طبيعي

بحثنا في الفقرة السابقة (3.8.3) توزيعات المعاينة الأساسية لمجاميع ومتوسطات عينات عشوائية، وبشكل خاص مجاميع ومتوسطات عينات عشوائية (صيغ خطية في عناصر العينة) من توزيعات طبيعية (الحالات 3،4،6). سنعالج في هذا البند

توزيعات المعاينة لمجموع مربعات (صيغ تربيعية) في عناصر العينة العشوائية كالتباين وصيغ تربيعية أخرى للعينات من توزيع طبيعي وحيد البعد.

سنعرض في البداية بعض التعاريف والمبرهنات الضرورية لدراسة الصيغ التربيعية في متغيرات عشوائية طبيعية وتوزيعات المعاينة لها.

1.9.3 الصيغ الخطية والتربيعية في متغيرات عشوائية طبيعية

Linear A Quadratic Forms in Normal Random Variables

نفترض أن $X=(X_1,\ldots,X_n)$ متغيراً عشوائياً ذا n بعد، و $A=\|a_{ij}\|_1^n$ مصفوفة حقيقية متناظرة من المرتبة $n\times n$. يدعى المقدار $G=X'AX$ بالشكل التربيعي في المتغيرات $X_i\,;\,i=1,\ldots,n$، وعندئذٍ تدعى A بمصفوفة الشكل التربيعي، وإذا كانت المصفوفة A محددة موجبة $(|A|\neq 0)$، فإن الشكل التربيعي G محدد موجب. نلاحظ أن:

$$G=X'AX=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}X_iX_j \tag{1.9.3}$$

و X' منقول X.

ويدعى المقدار $t_k=\sum_{i=1}^{n}b_{ki}X_i$ بالشكل الخطي في المتغيرات X_i، حيث b_{ki} أعداد ثابتة. وإذا كان لدينا m شكل خطي $t_k\,;\,k=1,\ldots,m$، فيمكننا كتابته بصيغة المصفوفات على النحو $t=BX$، حيث $B=\|b_{ki}\|$ مصفوفة مستطيلة من المرتبة $m\times n$. وفيما يلي سنرمز بـ O للمصفوفة الصفرية (zero matrix or null matrix) وبـ E_n للمصفوفة المحايدة أو الواحدية (identity) والدليل n يشير إلى مرتبة المصفوفة.

مبرهنة 1.9.3

بفرض $X=(X_1,\ldots,X_n)$ عينة عشوائية من التوزيع $N(0,1)$، ولدينا الشكل التربيعي $G=X'AX$ و m صيغة خطية $t=BX$. عندئذٍ إذا كان

$BA = O$، فإن المتغيرين G و $t_k \;;\; k = 1,\ldots,m$ مستقلان.

الإثبات

بما أن A مصفوفة حقيقة متناظرة، فهي قابلة التقاطر تعامدياً (orthogonally diagonalizable)، أي يمكن إيجاد مصفوفة متعامدة U (orthogonal matrix) $\left(U^tU = E_n\right)$، بحيث إن $U^tAU = D$ مصفوفة قطرية عناصرها $\lambda_i \;;\; i = 1,\ldots,n$ عبارة عن القيم الذاتية للمصفوفة A (حلول المعادلة المميزة $\left|A - \lambda E_n\right| = 0$). وتعتبر الأعمدة u_j للمصفوفة $U = \|u_1 u_2 \ldots u_n\|$ المتجهات الذاتية للمصفوفة A الموافقة للقيم الذاتية $\lambda_1,\ldots,\lambda_n$ على الترتيب، أي أن:

$$Au_j = \lambda_j u_j \quad ; \quad j = 1,\ldots,n$$

يفهم في لغة المصفوفات المتجه كمصفوفة عمود و U^t أو U' منقول المصفوفة U.

لتكن r رتبة المصفوفة A (rank A) و $\lambda_1,\ldots,\lambda_r$ القيم الذاتية لها المختلفة عن الصفر، وبما أن $U^tAU = D$، فيمكن كتابة:

$$A = UDU^t = \sum_{j=1}^{r} \lambda_j u_j u_j^t \tag{2.9.3}$$

وحسب الفرض:

$$BA = O \;\Rightarrow\; \sum_{j=1}^{r} \lambda_j \left(Bu_j\right) u_j^t = O$$

وبضرب طرفي المساواة الأخيرة من اليمين بالمتجه u_s، وباعتبار أن المتجهات u_j متعامدة (لأن المصفوفة U متعامدة)، أي أن $u_j^t u_s = 0 \;;\; j \neq s$، نجد:

$$Bu_s = O \quad ; \quad s = 1,2,\ldots,r \tag{3.9.3}$$

لنأخذ الآن المتجه العشوائي $\left(t_1,\ldots,t_m,u_1^tX,\ldots,u_r^tX\right)$. يخضع هذا المتجه للتوزيع الطبيعي ذي $m + r$ بعد، لأن مركباته (متغيرات عشوائية) عبارة عن

أشكال خطية في المتجه الطبيعـــي X.

يمكن بناءً على العلاقة (2.9.3) كتابـــة:

$$G = X^t AX = \sum_{j=1}^{r} \lambda_j \left(X^t u_j\right)\left(u_j^t X\right) = \sum_{j=1}^{r} \lambda_j \left(u_j^t X\right)^2 \qquad (4.9.3)$$

وبالتـــالي، نصـــل إلى المطلـــوب، إذا أمكـــن إثبـــات أن المتغـــيرات العشـــوائية $t_1,\ldots,t_m, u_1^t X,\ldots,u_r^t X$ غير مرتبطة، وذلك لأنهـــا خاضعـــة لتوزيعـــات طبيعيـــة، أي يكفي إثبـــات أن:

$$\operatorname{cov}\left(t_i, u_j^t X\right) = 0 \quad ; \quad i = 1,\ldots,m \ , \ j = 1,\ldots,r$$

وهذا يعني أن $t_i, u_j^t X$ مســـتقلان.

لإثبات ذلك، لـــنرمز بـــ $b_i^t \ , \ i = 1,\ldots,m$ لأســـطر المصفوفـــة B، وحســـب العلاقة (23.3.2)، نجـــد:

$$\begin{aligned}\operatorname{cov}\left(t_i, u_j^t X\right) &= \operatorname{cov}\left(b_i^t X, u_j^t X\right)\\ &= E\left(b_i^t X u_j^t X\right) - E\left(b_i^t X\right)E\left(u_j^t X\right)\\ &= E\left(b_i^t X u_j^t X\right) = b_i^t E\left(XX^t\right)u_j\\ &= b_i^t E_n u_j = b_i^t u_j = 0\end{aligned}$$

حيث إن $E\left(b_i^t X\right) = E\left(u_j^t X\right) = 0$، لأن توزيع X_i هـــو $N(0,1)$، و $b_i^t u_j = 0$ لأن $Bu_j = 0 \ ; \ j = 1,\ldots,r$ [العلاقة (3.9.3)]. وهو المطلوب.

لنرى الآن الشكلين الـــتربيعيين:

$$G_1 = X^t AX \qquad , \qquad G_2 = X^t BX$$

حيث كل من A و B مصفوفة حقيقة متناظرة مـــن المرتبـــة n.

مـــبرهنة 2.4.3

إذا كان $AB = BA = O$، فإن المتغيرين العشوائيين $G_1 ; G_2$ مستقلان.

الإثبـــات

لنفترض من أجـــل المصفوفـة A العلاقـة (2.9.3) صحيحـة. ويمكننـا كتابـة المصفوفة B على النحـو الآتي:

$$B=\sum_{j=1}^{S}\upsilon_j\mathbf{v}_j\mathbf{v}_j' \quad , \quad S=RankB$$

وحسب الفرض $AB=O$، إذن:

$$AB=\sum_{i=1}^{r}\sum_{j=1}^{S}\lambda_i\upsilon_j\boldsymbol{u}_i\left(\boldsymbol{u}_i'\mathbf{v}_j\right)\mathbf{v}_j'=O$$

وبضرب هذه المساواة من اليسار بـ $\boldsymbol{u}_i'$ ومن اليمين بـ $\mathbf{v}_j$، نجد:

$$\boldsymbol{u}_i'\mathbf{v}_j=0 \quad , \quad i=1,\dots,r \quad , \quad j=1,\dots,s$$

وهذا يعني أن المتجهات $\boldsymbol{u}_i$ متعامدة مع المتجهات $\mathbf{v}_j$. مـن هنـا، كمـا أشـرنا سابقاً، فالمتغيرات العشوائي $\boldsymbol{u}_i'X$ و $\mathbf{v}_j'X$ غير مرتبطـة. وبمـا أن توزيعـها المشـترك طبيعي فإنها مستقلة. وبالتالي، بمـا أن:

$$G_1=\sum_{i=1}^{r}\lambda_i\left(\boldsymbol{u}_i'X\right)^2 \quad , \quad G_2=\sum_{j=1}^{S}\lambda_j\left(\mathbf{v}_j'X\right)^2$$

فإن المتغيرين G_1،G_2 مستقلان.

2.9.3 توزيعات المعاينة لصيغ تربيعية في متغيرات عشوائية طبيعية

لنرمز بـ trA لأثر المصفوفة المربعة A (مجموع عناصرهـا القطريـة).

مـبرهنة 3.9.3

لتكن $X=\left(X_1,\dots,X_n\right)$ عينة عشـوائية مـن التوزيـع $N(0,1)$، $G=X'AX$ و $r\le n$ رتبة المصفوفة A. إذا كانت $A^2=A$ (A مصفوفة متسـاوية القـوى)، فـإن $\mathcal{L}(G)=\chi^2_{(r)} \quad ; \quad r=trA$.

الإثبـــات

لنفترض من أجل A العلاقة (2.9.3) صحيحـــة، عندهـــا مـــن خاصـــة التنـــاظر ومتســـاوية القـــوى نجـــد أن $\lambda_1 = \lambda_2 =....= \lambda_r = 1$ ، وبنـــاءً علـــى العلاقـــة (4.9.3) نجـــد:

$$G = \sum_{k=1}^{r}\left(u_k^t X\right)^2$$

وبما أن المتجهات u_k متعامدة ومستقلة، فإن المتغيرات العشـوائية $u_k^t X\ ,\ k = 1,...,r$ مستقلة ولكل منها التوزيع الطبيعي $N(0,1)$. وبالتالي، بناءً علـــى المـــبرهنتين (6.3.2) و (7.3.2) لتوزيع χ^2 فإن $\mathcal{L}(G) = \chi^2_{(r)}$ ، وأن:

$$t_r A = t_r\left(U^2 UD\right) = t_r D = \lambda_1 +....+ \lambda_r = r$$

مـــبرهنة 4.9.3

إذا كان المتجه $Y = \left(Y_1,...,Y_n\right)$ خاضعـــاً للتوزيـــع $N(\mu,\Sigma)$ ، فـــإن الشـــكل الـــتربيعي $G = (Y-\mu)^t\Sigma(Y-\mu)$ يخضـــع لتوزيـــع $\chi^2_{(n)}$. حيـــث إن Σ مصفوفـــة التغاير و μ مصفوفة القيم المتوقعة لـــ $Y_i\ ;\ i = 1,...,n$.

الإثبـــات

بما أن Σ مصفوفة حقيقيـــة ومتنـــاظرة، فـــهي قابلـــة للتقـــاطر. وبالتـــالي يمكـــن إيجاد المصفوفة U التي تحول المصفوفـــة Σ إلى شـــكل قطـــري $U^t\Sigma U = D$. وبمـــا أن العناصر القطريـــة λ_i للمصفوفـــة D موجبـــة، فـــإن المصفوفـــة $D^{-1/2}$ عبـــارة عـــن المصفوفـــة القطريـــة بالعنـــاصر $\lambda_i^{-1/2}$. لنبحـــــث الآن عـــــن توزيـــــع المتجـــــه $Z = D^{-1/2}U^t(Y-\mu)$.

بما أن $\mathcal{L}(Y) = N(\mu,\Sigma)$ ، وبـــأخذ $V = LY$ ، حيـــث L مصفوفـــة التحويـــل الخطي، فإن $\mathcal{L}(V) = N\left(L\mu, L\Sigma L^t\right)$ ، وبالتالي:

$$\mathcal{L}(Z) = N(O, E_n)$$

لكن $Y - \mu = UD^{1/2}Z$، إذن:

$$G = ZD^{-1/2}U'\Sigma^{-1}UD^{1/2}Z = Z'E_nZ = Z'Z$$

وبناءً على المبرهنتين (6.3.2) و (7.3.2) لتوزيع χ^2، نجد:

$$\mathcal{L}(G) = \chi^2_{(n)}$$

مبرهنة 5.9.3

إذا كانت $X = (X_1, \ldots, X_n)$ عينة عشوائية من التوزيع $N(\mu, \sigma^2)$ وكان $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ و $S^2(X) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ متوسط وتباين العينة X على الترتيب، فإنهما مستقلان وتوزيع المتغير العشوائي $\frac{nS^2}{\sigma^2}$ هو $\chi^2_{(n-1)}$.

الإثبات

بما أن (حسب الفرض):

$$\mathcal{L}(X_i) = N(\mu, \sigma^2) \quad , \quad i = 1,2,\ldots,n$$

فإن:

$$\mathcal{L}\left(Y_i = \frac{X_i - \mu}{\sigma}\right) = N(0,1) \quad , \quad i = 1,2,\ldots,n$$

أي أن $Y = (Y_1, \ldots, Y_n)$ عينة عشوائية من التوزيع $N(0,1)$. ونلاحظ أن:

$$\bar{Y} = \frac{\bar{X} - \mu}{\sigma} \quad , \quad S^2(Y) = \frac{S^2(X)}{\sigma^2}$$

ولإثبات استقلال المتغيرين $\bar{X}$، $S^2(X)$ يكفي إثبات استقلال المتغيرين $\bar{Y}$، $S^2(Y)$.

لنأخذ مصفوفة السطر $b' = \left[\frac{1}{n} \ldots\ldots \frac{1}{n}\right]$ من المرتبة $1 \times n$، عندئذٍ المصفوفة $B = \|b \ldots\ldots b\|$ من المرتبة n حقيقة ومتناظرة.

نلاحظ $\overline{Y} = b'Y$ و $nS^2(Y) = (Y - BY)'(Y - BY)$، وبالتالي يمكننا أن نكتب $nS^2(Y) = Y'AY$، حيث إن $A = E_n - B$ مصفوفة متساوية القوى.

بما أن $b'A = b' - b'B = b' - b' = O$، فإن $BA = O$، وحسب المبرهنة (1.9.3) فإن المتغيرين $\overline{Y}$، $S^2(Y)$ مستقلان، وبالتالي $\overline{X}$، $S^2(X)$ مستقلان.

من الواضح أن توزيع $\sqrt{n}\,\overline{Y} = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ هو $N(0,1)$. وبما أن:

$$\frac{nS^2(X)}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left[\left(X_i - \mu\right) - \left(\overline{X} - \mu\right)\right]^2$$

$$= \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 - \left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}\right)^2$$

وحسب المبرهنتين (6.3.2) و (7.3.2) لتوزيع χ^2 فإن:

$$\mathcal{L}\left(\sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2\right) = \chi^2_{(n)}$$

$$\mathcal{L}\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}\right)^2 = \chi^2_{(1)}$$

وبالتالي، بناءً على المبرهنة (8.3.2) لتوزيع χ^2 نجد:

$$\mathcal{L}\left(\frac{nS^2(X)}{\sigma^2}\right) = \chi^2_{(n-1)}$$

مـــــبرهنة 6.9.3

إذا كانت $X=(X_1,\dots,X_n)$ عينة عشـــوائية مـــن التوزيـــع $N(\mu,\sigma^2)$ وكـــان S^2،$\overline{X}$ متوسط وتباين العينة X على الترتيب، فإن:

$$\mathcal{L}\left(\frac{\overline{X}-\mu}{S/\sqrt{n-1}}\right)=S_{(n-1)}$$

أي أن المتغير العشوائي $\dfrac{\overline{X}-\mu}{S/\sqrt{n-1}}$ يخضع لتوزيع t بـــ $n-1$ درجة حرية.

الإثبـــــات

لنفترض في العلاقـــة (*41.3.2*) أن:

$$\xi=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\qquad,\qquad \eta=\frac{nS^2(X)}{\sigma^2}$$

وعلى ذلـــك:

$$\frac{\overline{X}-\mu}{S/\sqrt{n-1}}=\frac{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{nS^2}{\sigma^2}\Big/(n-1)}}=\frac{\xi}{\sqrt{\dfrac{\eta}{(n-1)}}}$$

وبما أن:

$$\mathcal{L}\left(\xi=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)=N(0,1)$$

$$\mathcal{L}\left(\eta=\frac{nS^2}{S/\sqrt{n}}\right)=\chi^2_{(n-1)}$$

إذن حسب تعريف توزيع t، فـــإن:

$$\mathcal{L}\left(\frac{nS^2}{S/\sqrt{n}}\right)=S_{(n-1)}$$

مـــبرهنة 7.9.3

إذا كانت $X=(X_1,\dots,X_m)$ و $Y=(Y_1,\dots,Y_n)$ عينتين عشوائيتين مستقلتين مـــن التوزيع $N(\mu,\sigma^2)$، وكان $\bar{X}$،$S^2(X)$ و $\bar{Y}$،$S^2(Y)$ متوسط وتباين كل منـــهما علـــى الترتيب، فإن المتغير العشوائي:

$$t=\sqrt{\frac{mn(m+n-2)}{m+n}}\;\frac{\bar{X}-\bar{Y}}{\sqrt{mS^2(X)+nS^2(Y)}} \qquad (5.9.3)$$

يخضع لتوزيع t بـــ $(m+n-2)$ درجة حرية.

الإثبـــات

يمكن كتابة المتغير العشوائي t علـــى النحـــو الآتي:

$$t=\frac{\left[\left(\bar{X}-\bar{Y}\right)/\sigma\right]\sqrt{\dfrac{mn}{m+n}}}{\sqrt{\dfrac{mS^2(X)+nS^2(Y)}{\sigma^2}\Big/(m+n-2)}}$$

حيـــث إن:

$$E\left(\bar{X}-\bar{Y}\right)=0 \qquad , \qquad \operatorname{var}\left(\bar{X}-\bar{Y}\right)=\frac{m+n}{mn}\sigma^2$$

بما أن:

$$\mathcal{L}\left(\frac{mS^2(X)}{\sigma^2}\right)=\chi^2_{(m-1)} \qquad , \qquad \mathcal{L}\left(\frac{nS^2(Y)}{\sigma^2}\right)=\chi^2_{(n-1)}$$

فحسب الخاصة (2) لتوزيع χ^2 نجـــد:

$$\mathcal{L}\left(\frac{mS^2(X)+nS^2(Y)}{\sigma^2}\right)=\chi^2_{(m+n-2)}$$

وبسهولة يمكن إثبـــات أن:

$$\mathcal{L}\left(\frac{\overline{X}-\overline{Y}}{\sigma}\sqrt{\frac{mn}{m+n}}\right)=N(0,1)$$

وبالتالي، بناءً على تعريف توزيع t (41.3.2) نجـــد أن توزيـــع المتغـــير العشـــوائي t المعرف بالعلاقـــة (5.9.3) هـــو $S(m+n-2)$، أي أن:

$$\mathcal{L}\left(\frac{\left[\left(\overline{X}-\overline{Y}\right)/\sigma\right]\sqrt{\frac{mn}{m+n}}}{\sqrt{\frac{mS^2(X)+nS^2(Y)}{\sigma^2}\Big/(m+n-2)}}\right)=S(m+n-2)$$

وبشكل مشابه، يمكن إثبات أن المتغـــير العشـــوائي:

$$t=\frac{\left(\overline{X}-\overline{Y}\right)-(\mu_1-\mu_2)}{\sqrt{mS^2(X)+nS^2(Y)}}\sqrt{\frac{mn(m+n-2)}{m+n}}$$

يخضع لتوزيع t بـــ $(m+n-2)$ درجة حرية، حيث إن $X=(X_1,\ldots,X_m)$ من التوزيع $N(\mu_1,\sigma^2)$ و $Y=(Y_1,\ldots,Y_n)$ من التوزيع $N(\mu_2,\sigma^2)$.

10.3 الإحصاءات المرتبة *THE ORDER STATISTICS*

تلعب نظرية المعاينة للإحصاءات المرتبة دوراً هامـــاً في مســـائل مختلفـــة للاســـتدلال الإحصائي، التي سنتطرق إليها فيما بعد. لذا سنبحث في هذا البنـــد بعـــض أهـــم نتـــائج نظرية المعاينة للإحصاءات المرتبـــة.

نفـــترض $X=(X_1,\ldots,X_n)$ عينـــة عشـــوائية مـــن توزيـــع مســـتمر $F(x)$ و $x=(x_1,\ldots,x_n)$ قيمـــة ملاحظـــة لـــ X. إن القيـــم الملاحظـــة $x_1,\ldots,x_n$ يمكـــن ترتيبها تصاعديـــاً حســـب قيمـــها (مـــن الأصغـــر إلى الأكـــبر) ولـــنرمز لهـــا بـــ $x_{(1)},\ldots,x_{(n)}$، حيـــث إن:

$$x_{(1)}\leq x_{(2)}\leq\ldots\leq x_{(n)} \qquad (1.10.3)$$

ولنرمز بـ $X_{(k)}$ للمتغير العشوائي الذي يأخذ القيمة $x_{(k)}\ ;\ k=1,\dots,n$ عند أي ملاحظة x للعينة X، وبذلك نعرف من العينة X متتالية جديدة من المتغيرات العشوائية $X_{(1)},X_{(2)},\dots,X_{(n)}$، حيث إن:

$$X_{(1)} \le X_{(2)} \le \dots \le X_{(n)} \tag{2.10.3}$$

تدعى المتغيرات الجديدة $X_{(k)}\ ;\ k=1,\dots,n$ بالإحصاءات المرتبة للعينة، ويدعى المتغير $X_{(k)}$ بالإحصاء ذو الترتيب k، ومن ثم تدعى $\left(X_{(1)},\dots,X_{(k)}\right)$ بالعينة المرتبة. كما تدعى المتسلسلة المعرفة بالعلاقة (2.10.3) بالمتسلسلة المتغيرة (variable sery) للعينة X. وهذا يعني أن المتسلسلة المتغيرة لـ X ما هي إلا عينة عناصرها مرتبة تصاعدياً حسب قيمها. ومن أجل ملاحظة x لـ X، فإن المتسلسلة المعرفة بالعلاقة (2.10.3) تأخذ شكل المتسلسلة العددية المعرفة بالعلاقة (1.10.3). وهذه الأخيرة تعتبر الشكل الأولي لتمثيل المعطيات الإحصائية (العينة الملاحظة).

1.10.3 التوزيع الاحتمالي الهامشي لإحصاء مرتب

The Marginal Probability distribution of An Individual Order Statistic

نقدم في هذه الفقرة التوزيعات الاحتمالية الهامشية للإحصاءات المرتبة $X_{(k)}\ ;\ k=1,\dots,n$ لعينة عشوائية X مأخوذة من توزيع مستمر $F(x)$.

مبرهنة 1.10.3

إذا كانت $X=\left(X_{(1)},\dots,X_{(n)}\right)$ عينة عشوائية من توزيع مستمر $\mathcal{L}(\xi)$ دالة كثافته الاحتمالية $f(x)$ ودالة توزيعه هي:

$$F(x)=P(\xi<x)$$

وكانت $Y_1=X_{(1)},\dots,Y_n=X_{(n)}$ الإحصاءات المرتبة للعينة X، فإن التوزيع الاحتمالي للإحصاء المرتب $Y_i=X_{(i)}$ يعطى بدالة الكثافة:

$$h_i(y) = h(y) = \frac{n!}{(i-1)!(n-i)!} f(y)[F(y)]^{i-1}[1 - F(y)]^{n-i} \qquad (3.10.3)$$

الإثبــات

حسب تعريف دالة الكثافة في نظريــة الاحتمــالات:

$$h(y) = \lim_{\Delta y \to 0} \frac{H(y+\Delta y) - H(y)}{\Delta y} = \lim_{\Delta y \to 0} \frac{P(y \le Y_i < y + \Delta y)}{\Delta y}$$

حيث $h(y)$ و $H(y)$ دالة الكثافة الاحتمالية ودالة التوزيــع للمتغــير العشــوائي Y_i على الــترتيب.

إن تحقــق الحــادث $y \le Y_i < y + \Delta y$ يعــني أن $(i-1)$ مــن المركبــات $X_i\ ;\ i = 1,\dots,n$ للعينــة X أقــل مــن y ومركبــة واحــدة تنتمــي للفــترة $[y, y + \Delta y)$ وأن $(n-i)$ مركبة مــن X أكــبر أو تســاوي $y + \Delta y$. وباســتخدام التوزيع المتعدد الحدود نجــد:

$$P(y \le Y_i < y + \Delta y) =$$
$$= \frac{n!}{(i-1)!1!(n-i)!}[F(y)]^{i-1}[F(y+\Delta y) - F(y)][1 - F(y+\Delta y)]^{n-i}$$

وبتقسيم الطرفين على Δy، وبأخذ النهاية عندمــا $\Delta y \to 0$، نجــد:

$$\lim_{\Delta y \to 0} \frac{P(y \le Y_i < y + \Delta y)}{\Delta y} =$$
$$= \frac{n!}{(i-1)!1!(n-i)!}[F(y)]^{i-1}[1 - F(y)]^{n-i} \lim_{\Delta y \to 0} \frac{F(y+\Delta y) - F(y)}{\Delta y}$$

وبالتالي:

$$h_i(y) = \frac{n!}{(i-1)!1!(n-i)!} f(y)[F(y)]^{i-1}[1 - F(y)]^{n-i}$$

وهو المطلــوب.

بوضع $i=1$ في العلاقة (3.10.3) نحصل على دالة الكثافة الاحتمالية لتوزيع $Y_1 = X_{(1)}$:

$$h_1(y) = \frac{n!}{(n-1)!} f(y)[1-F(y)]^{n-1} = nf(y)[1-F(y)]^{n-1} \tag{4.10.3}$$

وبوضع $i=n$ نحصل على دالة كثافة توزيع $Y_n = X_{(n)}$:

$$h_n(y) = nf(y)[F(y)]^{n-1} \tag{5.10.3}$$

وبشكل مشابه نحصل على توزيع بقية الإحصاءات المرتبة.

مثال 1.10.3

بفرض $X = (X_1,\dots,X_n)$ عينة عشوائية من التوزيع المنتظم $R(0,1)$، أوجد توزيع كل من $Y_1 = X_{(1)}$, $Y_n = X_{(n)}$.

بما أن دالة كثافة ودالة توزيع $R(0,1)$ هي على الترتيب:

$$f(x) = \begin{cases} 1 & ; \quad 0 \le x \le 1 \\ 0 & ; \quad x < 0 \text{ or } x > 1 \end{cases}$$

$$F(x) = \begin{cases} 0 & ; \quad x \le 0 \\ x & ; \quad 0 < x \le 1 \\ 1 & ; \quad x > 1 \end{cases}$$

فحسب العلاقة (3.10.3) نجد دالة كثافة توزيع $Y_i = X_{(i)}$

$$h_i(y) = \frac{n!}{(i-1)!(n-i)!} y^{i-1}(1-y)^{n-i} \quad ; \quad 0 < y \le 1$$

بوضع $i=1$ نحصل على توزيع $Y_1 = X_{(1)}$

$$h_1(y) = n(1-y)^{n-1} \quad ; \quad 0 < y \le 1$$

وبوضع $i = n$ نحصل على توزيع $Y_n = X_{(n)}$

$$h_n(y) = ny^{n-1} \quad ; \quad 0 < y \le 1$$

مثـــال 2.10.3

إذا كانت $X = (X_1,\dots,X_n)$ عينة عشوائية من توزيع أسي:

$$f(x;\theta) = \theta e^{-\theta x} \quad ; \quad x > 0 \ , \ \theta > 0$$

فأوجد توزيع كل من $Y_2 = X_{(2)}$ و $Y_{n-1} = X_{(n-1)}$

بما أن دالة التوزيع الأسي:

$$F(x;\theta) = \begin{cases} 0 & ; \ x \le 0 \\ 1 - e^{-\theta x} & ; \ x > 0 \end{cases}$$

فبالتعويض في العلاقة (3.10.3) نجد توزيع $Y_i = X_{(i)}$

$$h_i(y) = \frac{n!}{(i-1)!(n-i)!}\theta e^{-(n+1-i)\theta x}\left(1 - e^{-\theta x}\right)^{i-1} \quad ; \quad x > 0$$

بوضع $i = 2, n-1$ نحصل على توزيع Y_2 و Y_{n-1} على الترتيب.

$$h_2(y) = n(n-1)\theta e^{-(n-1)\theta x}\left(1 - e^{-\theta x}\right) \quad ; \quad x > 0$$

$$h_{n-1}(y) = n(n-1)\theta e^{-2\theta x}\left(1 - e^{-\theta x}\right)^{n-2} \quad ; \quad x > 0$$

2.10.3 التوزيع المشترك لإحصائين مرتبين

The Joint Distribution of Two Order Statistics

مـــبرهنة 2.10.3

إذا كانت $X = (X_1,\dots,X_n)$ عينة عشوائية من مجتمع دالة كثافته الاحتمالية $f(x)$ ودالة توزيعه $F(x)$، فإن دالة الكثافة الاحتمالية المشتركة للإحصائين

المرتبين $1 \le i < j \le n$; $Y_i = X_{(i)} , Y_j = X_{(j)}$ هي:

$$h_{Y_iY_j}(x,y) = h_{ij}(x,y) =$$

$$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} f(x)f(y)\left[F(x)\right]^{i-1}\left[F(y)-F(x)\right]^{j-i-1}\left[1-F(y)\right]^{n-j} \tag{6.10.3}$$

الإثبات

نعلم أن:

$$h_{ij}(x,y) = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{P\left(x \le Y_i < x+\Delta x \,,\, y \le Y_j < y+\Delta y\right)}{\Delta x \Delta y}$$

وتحقق الحادث $\left(x \le Y_i < x+\Delta x \,,\, y \le Y_j < y+\Delta y\right)$ يعني تحقق الحوادث:

$$\begin{array}{lll}
Y_k < x & ; & k \le i-1 \\
x \le Y_k < x+\Delta x & ; & k = i \\
x+\Delta x \le Y_k < y & ; & i < k \le j-1 \\
y \le Y_k < y+\Delta y & ; & k = j \\
Y_k \ge y+\Delta y & ; & k > j
\end{array}$$

نلاحظ أننا أمام n تكرار مستقل لتجربة مفروضة، فضاء الحوادث الأولية الموافق لها مجزأ إلى خمس حوادث متنافية مثنى مثنى، وهي الواردة أعلاه، وأن احتمالات هذه الحوادث على الترتيب:

$$\begin{aligned}
P_1 &= F(x) \\
P_2 &= F(x+\Delta x) - F(x) \\
P_3 &= F(y) - F(x+\Delta x) \\
P_4 &= F(y+\Delta y) - F(y) \\
P_5 &= 1 - F(y+\Delta y)
\end{aligned}$$

ومن ثم احتمال ظهور الحادث الأول $(i-1)$ مرة، الحادث الثاني مرة واحدة، الحادث الثالث $(j-i-1)$ مرة، الحادث الرابع مرة واحدة والحادث الخامس $(n-j)$ مرة ضمن الـ n تكرار مستقل للتجربة المفروضة يعطى بقانون التوزيع المتعدد الحدود:

$$P\left(x \le Y_i < x+\Delta x \,,\, y \le Y_j < y+\Delta y\right) =$$

$$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!}\left[F(x)\right]^{i-1}\left[F(x+\Delta x)-F(x)\right]$$

$$\times\left[F(y)-F(x+\Delta x)\right]^{j-i-1}\left[F(y+\Delta y)-F(y)\right]\left[1-F(y+\Delta y)\right]^{n-j}$$

بتقسيم الطرفين على $\Delta x\Delta y$، وأخذ النهاية عندما $\Delta x \to 0 \,,\, \Delta y \to 0$ نحصل على دالة الكثافة المشتركة $h_{ij}(x,y)$ المعرفة بالعلاقة (6.10.3).

بوضع $i=1 \,,\, j=n$ في العلاقة (6.10.3) نحصل على دالة الكثافة الاحتمالية المشتركة للإحصائين المرتبين $Y_1 = X_{(1)} \,,\, Y_n = X_{(n)}$:

$$h_{1n}(x,y) = n(n-1)f(x)f(y)\left[F(y)-F(x)\right]^{n-2} \quad ; \quad x<y \qquad (7.10.3)$$

مثال 3.10.3

بفرض $X = \left(X_1,\ldots,X_n\right)$ عينة من التوزيع المنتظم $R(0,1)$، أوجد التوزيع المشترك للإحصائين $Y_1 = X_{(1)} \,,\, Y_n = X_{(n)}$.

حيث أن:

$$f(x) = 1 \quad ; \quad 0 \le x \le 1$$
$$F(x) = x \quad ; \quad 0 < x \le 1$$

وبالتعويض في العلاقة (6.10.3):

$$h_{ij}(x,y) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!}x^{i-1}(y-x)^{j-i-1}(1-y)^{n-j} \;;\; 0<x<y\le 1$$

وبإعطاء $i=1 \,,\, j=n$ نحصل على كثافة التوزيع الاحتمالي لـ $\left(Y_1,Y_n\right)$

$$h_{1n}(x,y) = n(n-1)(y-x)^{n-2} \qquad ; \qquad 0 < x < y \le 1$$

ويمكن الحصول على نفس النتيجة بـــالتطبيق في العلاقـــة (7.10.3).

مثـــال 4.10.3

بفرض $X = (X_1,\ldots,X_n)$ عينة عشوائية من مجتمع كثافـــة توزيعـــه:

$$f(x) = \frac{1}{2}e^{-\frac{x}{2}} \qquad ; \qquad x > 0$$

أوجد التوزيع المشترك للإحصائين $Y_1 = X_{(1)}$ ، $Y_n = X_{(n)}$.

بما أن:

$$F(x) = \frac{1}{2}\int_0^x e^{-\frac{x}{2}}dx = 1 - e^{-\frac{x}{2}} \qquad ; \qquad x > 0$$

فحسب العلاقة (7.10.3)

$$h_{1n}(x,y) = \frac{n(n-1)}{4}e^{-(x+y)/2}\left[\left(1-e^{-y/2}\right)-\left(1-e^{-x/2}\right)\right]^{n-2} \qquad ; \qquad 0 < x < y$$

مـــبرهنة 3.10.3

إذا كـــانت $X = (X_1,\ldots,X_n)$ عينة عشوائية من مجمع كثـــافة توزيعه $f(x)$ ودالة توزيعه $F(x)$ ، وكانت $Y_1,\ldots,Y_n$ الإحصاءات المرتبة للعينة X، فإن دالة الكثافة المشـتركة لها هي:

$$h(y_1,\ldots,y_n) = n!\prod_{i=1}^{n} f(y_i) \qquad ; \qquad y_1 < y_2 < \ldots. < y_n \tag{8.10.3}$$

الإثبـــات

نعلـم أن:

$$h(y_1,y_2,\dots,y_n)=\lim_{\Delta y_i\to 0}\frac{1}{\prod_{i=1}^{n}\Delta y_i}P(y_1\le Y_1<y_1+\Delta y_1,\dots,y_n\le Y_n<y_n+\Delta y_n)$$

وتحقق الحادث $(y_1\le Y_1<y_1+\Delta y_1,\dots,y_n\le Y_n<y_n+\Delta y_n)$ يعني تحقيق الحوادث $(y_i\le Y_i<y_i+\Delta y_i)\ ;\ i=1,\dots,n$ المتنافية مثنى مثنى في آن واحد، حيث توجد ملاحظة واحدة من العينة X في كل فترة $(y_i\le Y_i<y_i+\Delta y_i)$ باحتمال $F(y_i+\Delta y_i)-F(y_i)$ وحسب قانون التوزيع المتعدد الحدود:

$$P(y_1\le Y_1<y_1+\Delta y_1,\dots,y_n\le Y_n<y_n+\Delta y_n)=n!\prod_{i=1}^{n}\left[F(y_i+\Delta y_i)-F(y_i)\right]$$

وبتقسيم الطرفين على $\prod_{i=1}^{n}\Delta y_i$ وجعل $\Delta y_i\to 0$ نجد:

$$h(y_1,\dots,y_n)=n!\prod_{i=1}^{n}\lim_{\Delta y_i\to 0}\frac{F(y_i+\Delta y_i)-F(y_i)}{\Delta y_i}=$$

$$=n!\prod_{i=1}^{n}f(y_i)\quad ;\quad y_1<y_2<\dots<y_n$$

وهو المطلوب.

فمثلاً، إذا كانت لدينا معطيات المثال (4.10.3)، نجد:

$$h(y_1,\dots,y_n)=\left(\frac{1}{2}\right)^n n!e^{-\frac{1}{2}\sum_{1}^{n}y_i}\quad ;\quad 0<y_1<y_2<\dots<y_n$$

وإذا كانت لدينا معطيات المثال (3.10.3)، فإن:

$$h(y_1,\dots,y_n)=n!\quad ;\quad 0<y_1<y_2<\dots<y_n\le 1$$

ملاحظة 1.10.3

يوجد فرق بين توزيع العينة العشوائية $X=(X_1,\dots,X_n)$ وتوزيع العينة العشوائية

المرتبة $Y_i = X_{(i)}$, $i = 1,\dots,n$; $Y = (Y_1,\dots,Y_n)$. فمثلاً، في حالة المعاينة من توزيع $R(0,1)$، فإن:

$$f(x_1,\dots,x_n) = \prod_{i=1}^{n} f(x_i) = 1 \quad ; \quad 0 \le x_i \le 1 \,,\ i = 1,\dots,n$$

بينما:

$$h(y_1,\dots,y_n) = n! \quad ; \quad 0 < y_1 < y_2 < \dots < y_n \le 1$$

وفي حالة المعاينة من توزيع $\Gamma(1,2)$ فإن:

$$f(x_1,\dots,x_n) = \left(\frac{1}{2}\right)^n e^{-\frac{1}{2}\sum_1^n x_i}$$

بينما:

$$h(y_1,\dots,y_n) = \left(\frac{1}{2}\right)^n n!\, e^{-\frac{1}{2}\sum_1^n y_i}$$

11.3 توزيع دوال في الإحصاءات المرتبة

DISTRIBUTION OF FUNCTIONS OF ORDER STATISTICS

استخلصنا في الفقرتين السابقتين التوزيعات الهامشية والمشتركة للإحصاء المرتبة ذاتها. وسنبحث في هذه الفقرة توزيع دوال معينة في الإحصاءات المرتبة. إحدى أبسط الدوال في الإحصاءات المرتبة هو المتوسط الحسابي لها:

$$\frac{1}{n}\sum_{i=1}^{n} Y_i \quad ; \quad Y_i = X_{(i)}$$

وبما أن:

$$\frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

فتوزيع المعاينة لمتوسط الإحصاءات المرتبة هو نفسه توزيع المعاينة لمتوسط العينة $\boldsymbol{X}$، وهذا الأخير تم بحثه في البند 8.3.

سنناقش في الفقرات التالية توزيع وسيط العينة، مدى العينة ونصف مدى العينة. لذا لابد في البداية من تعريف وسيط، مدى ونصف مدى العينة.

نفترض $Y_1,\ldots,Y_n$ الإحصاءات المرتبة لعينة عشوائية $X_1,\ldots,X_n$ من توزيع مستمر $F(x)$.

تعريف 1.11.3: وسيط العينة Sample Median

يعرف وسيط العينة $\boldsymbol{X}$ على أنه الإحصاء المرتب الأوسط إذا كانت n فردية ومتوسط الإحصائين المرتبين الأوسطين إذا كانت n زوجية. وإذا رمزنا بـ $\tilde{X}$ لوسيط العينة العشوائية $\boldsymbol{X}$، فإن:

$$\tilde{X}=\begin{cases} Y_{m+1} & ; \quad n=2m+1 \\ \dfrac{Y_m+Y_{m+1}}{2} & ; \quad n=2m \end{cases} \tag{1.11.3}$$

حيث إن m عدد صحيح موجب.

ونعلم من نظرية الاحتمالات أن وسيط المجتمع $F(x)$، ولنرمز له بـ $\tilde{\mu}$، هو تلك القيمة التي تحقق الشرط:

$$F(\tilde{\mu})=\frac{1}{2}$$

تعريف 2.11.3: مدى العينة Sample Range

مدى العينة $\boldsymbol{X}=\left(X_1,\ldots,X_n\right)$ نرمز له بـ R ويعرف على النحو الآتي:

$$R=Y_n-Y_1=X_{(n)}-X_{(1)} \tag{2.11.3}$$

تعريف 3.11.3: نصف مدى العينة Sample Midrange

نصف مدى العينة $X=(X_1,\dots,X_n)$ نرمز له بـ T ويعرف على النحو التالي:

$$T=\frac{Y_n+Y_1}{2}=\frac{R}{2}+Y_1 \tag{3.11.3}$$

1.11.3 توزيع وسيط العينة *Distribution of Sample Median*

بما أن وسيط العينة $\tilde{X}$ عبارة عن إحصاء، أي متغير عشوائي فله توزيع احتمالي يدعى بتوزيع المعاينة لـ $\tilde{X}$. ولإيجاد هذا التوزيع نميز حالتين: n فردية، n زوجية.

إذا كان حجم العينة X فردياً، فإن $\tilde{X}=Y_{m+1}$، ومن ثم توزيع $\tilde{X}$ هو عبارة عن توزيع الإحصاء المرتب Y_{m+1}، وهذا الأخير نحصل عليه بوضع $n=2m+1\ ,\ i=m+1$ في العلاقة (3.10.3).

$$h(\tilde{x})=\frac{(2m+1)!}{m!m!}f(\tilde{x})\left[F(\tilde{x})\right]^m\left[1-F(\tilde{x})\right]^m \tag{4.11.3}$$

أما إذا كان حجم العينة X زوجياً $(n=2m)$، فإن $\tilde{X}=\dfrac{Y_m+Y_{m+1}}{2}$

وعندئذٍ نشتق توزيع $\tilde{X}$ على النحو الآتي:

نعين توزيع (Y_m,Y_{m+1})، وذلك بوضع $i=m\ ,\ j=m+1$ في العلاقة (6.10.3):

$$h(x,y)=\frac{2m!}{(m-1)!(m-1)!}f(x)f(y)\left[F(x)\right]^{m-1}\left[1-F(y)\right]^{m-1}$$

وبإجراء التحويل من (Y_m,Y_{m+1}) إلى $(\tilde{X},Z)$، حيث إن $Z=Y_{m+1}$. وبملاحظة أن المحدد الجاكوبي $J=2$ نجد:

$$g(\tilde{x},z)=\frac{2(2m)!}{(m-1)!(m-1)!}f(z)f(2\tilde{x}-z)\left[F(2\tilde{x}-z)\right]^{m-1}\left[1-F(z)\right]^{m-1}\ ;\ \tilde{x}\le z$$

(5.11.3)

ومن ثم دالة الكثافة الاحتمالية لوسيط العينة هي:

$$g(\tilde{x})=\int_{\tilde{x}}^{\infty}g(\tilde{x},z)dz \qquad (6.11.3)$$

مثال 1.11.3

بفرض $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع:

$$f(x)=2e^{-2x}\quad ;\quad x>0$$

أوجد توزيع $\tilde{X}$، إذا علمت أن حجم العينة فردي $(n=2m+1)$.

بما أن:

$$f(x)=2e^{-2x}\quad ;\quad x>0$$

فإن:

$$F(x)=1-e^{-2x}\quad ;\quad x>0$$

وبما أن $n=2m+1$ فردي، عندئذٍ توزيع $\tilde{X}$ نحصل عليه باستخدام العلاقة (4.11.3)

$$\begin{aligned}h(\tilde{x})&=\frac{(2m+1)!}{m!m!}2e^{-2\tilde{x}}\left(1-e^{-2\tilde{x}}\right)^m\left(e^{-2\tilde{x}}\right)^m\\&=\frac{2(2m+1)!}{m!m!}e^{-2(1+m)\tilde{x}}\left(1-e^{-2\tilde{x}}\right)^m\qquad ;\qquad \tilde{x}>0\end{aligned}$$

فمثلاً، إذا كانت $n=11$، أي $m=5$:

$$h(\tilde{x})=\frac{2\times 11!}{5!5!}e^{-12\tilde{x}}\left(1-e^{-2\tilde{x}}\right)^5\quad ;\quad \tilde{x}>0$$

مثـــال 2.11.3

بفـــرض Y_1,Y_2,Y_3 الإحصـــاءات المرتبـــة لعينـــة عشـــوائية $X=(X_1,X_2,X_3)$ مـن التوزيـع المنتظـم $R(0,1)$، أوجـد توزيـع الوسـيط $\widetilde{X}$ واحسـب متوسـطه وتباينـه.

بمـا أن:

$$f(x)=1 \quad ; \quad 0\le x\le 1$$
$$F(x)=x \quad ; \quad 0<x\le 1$$

فبتطبيـق العلاقـة (4.11.3)، نجـد:

$$g(\tilde{x})=\frac{3!}{1!1!}\tilde{x}(1-\tilde{x})=6\tilde{x}(1-\tilde{x}) \quad ; \quad 0<\tilde{x}\le 1$$

$$E\widetilde{X}=6\int_0^1\tilde{x}^2(1-\tilde{x})d\tilde{x}=6\left(\frac{\tilde{x}^3}{3}-\frac{\tilde{x}^4}{4}\right|_0^1=\frac{1}{2} \ , \ E\widetilde{X}^2=6\int_0^1\tilde{x}^3(1-\tilde{x})d\tilde{x}=\frac{3}{10}$$

$$\operatorname{var}\widetilde{X}=E\widetilde{X}^2-\left(E\widetilde{X}\right)^2=\frac{3}{10}-\frac{1}{4}=\frac{1}{20}$$

2.11.3 توزيـع مـدى ونصـف مـدى العينـة

Distribution of Sample Range and Midrange

لاشـتقاق توزيـع كـل مـن مـدى العينـة $R=Y_n-Y_1$ ونصـف مـدى العينـة $T=\dfrac{Y_1+Y_n}{2}$، نبحـث أولاً عـن توزيـع (Y_1,Y_n) وهـو:

$$h_{in}(x,y)=n(n-1)f(x)f(y)\left[F(y)-F(x)\right]^{n-2} \quad ; \quad x<y$$

وبـالتحويل مـن (Y_1,Y_n) إلى (R,T)، حيـث إن:

$$\left.\begin{aligned}R&=Y_n-Y_1\\ T&=\frac{Y_1+Y_n}{2}\end{aligned}\right\} \Rightarrow \left\{\begin{aligned}x&=\frac{2t-r}{2}\\ y&=\frac{2t+r}{2}\end{aligned}\right.$$

ومن ثـم جـاكوبي التحويـل:

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial t} \end{vmatrix} = \begin{vmatrix} -\frac{1}{2} & 1 \\ \frac{1}{2} & 1 \end{vmatrix} = -1$$

وبالتالي، فـإن دالـة الكثافـة لـ (R,T)

$$g(r,t) = n(n-1)f\left(t-\frac{r}{2}\right)f\left(t+\frac{r}{2}\right)\left[F\left(t+\frac{r}{2}\right)-F\left(t-\frac{r}{2}\right)\right]^{n-2} \; ; \; r > 0 \quad (7.11.3)$$

وبناءً علـى ذلـك فدالـة الكثافـة الاحتماليـة للمـدى R

$$g(r) = \int_{-\infty}^{+\infty} g(r,t)dt \quad (8.11.3)$$

ودالـة الكثافـة الاحتماليـة لنصـف المـدى:

$$g(t) = \int_{0}^{\infty} g(r,t)dr \quad (9.11.3)$$

مثـــال 3.11.3

بفـــرض $X = (X_1,\dots,X_n)$ عينـــة عشـــوائية مـــن التوزيـــع المنتظـــم $R(0,1)$، أوجـد:

1. التوزيـع المشـترك للمـدى R ونصـف المـدى T.
2. التوزيـع الهامشـي للمـدى R.
3. التوزيـع الهامشـي لنصـف المـدى T.

بمـا أن:

$$\begin{aligned} f(x) &= 1 \quad ; \quad 0 \le x \le 1 \\ F(x) &= x \quad ; \quad 0 < x \le 1 \end{aligned}$$

فإن دالـة الكثافـة المشـتركة لـ R و T هـي:

$$g(r,t) = n(n-1)r^{n-2} \quad ; \quad \begin{matrix} 0 < t - \dfrac{r}{2} < 1 \\ 0 < t + \dfrac{r}{2} < 1 \end{matrix} \qquad (10.11.3)$$

ومن ثم دالة الكثافة الهامشية لـ R:

$$g(r) = \int_{\frac{r}{2}}^{1-\frac{r}{2}} n(n-1)r^{n-2}dt = n(n-1)(1-r)r^{n-2} \quad ; \quad 0 < r < 1 \qquad (11.11.3)$$

ودالة الكثافة الهامشية لـ T:

$$g(t) = n(n-1)\int_{0}^{\min[2t,2(1-t)]} r^{n-2}dr = \begin{cases} n(2t)^{n-1} & ; \quad -1 < t < 0 \\ n[2(1-t)]^{n-1} & ; \quad 0 < t < 1 \end{cases} \qquad (12.11.3)$$

والقيمة المتوقعة لـ R

$$ER = \frac{n-1}{n+1}$$

12.3 توزيع المعاينة لـ $F^*(x)$ SAMPLE DISTRIBUTION OF

نلاحظ من تعريف دالة التوزيع التجريبي $F_n^*(x)$ (البند 5.3) أنها عبارة عن إحصاء، عند كل قيمة حقيقية x، لكونها دالة في بيانات العينة، أي متغيراً عشوائياً، وبالتالي لها توزيع احتمالي تحدده المبرهنة الآتية.

مبرهنة 1.12.3

لتكن $F_n^*(x)$ دالة التوزيع التجريبي لمتغير عشوائي ع الموافقة لعينة عشوائية $x = (x_1, \ldots, x_n)$ من توزيع $\mathcal{L}$(ع)، ولنفترض $F(x)$ دالة التوزيع النظري لـ ع. عندئذٍ توزيع $F_n^*(x)$ يعطي بالعلاقة الآتية:

$$P\left(F_n^*(x) = \frac{k}{n}\right) = C_n^k[F(x)]^k[1-F(x)]^{n-k} \quad ; \quad k = 0,1,\ldots,n \qquad (1.12.3)$$

الإثبـــات

نلاحـظ بوضـوح أن المتغير العشـوائي $F_n^*(x)$ يفـترض القيـم:

$$0,\frac{1}{n},\frac{2}{n},\ldots\ldots,\frac{n-1}{n},\frac{n}{n}=1$$

وإذا رمزنا بـ:

$$Z_i=\begin{cases}1 & ; \quad x_i<x\\ 0 & ; \quad x_i\ge x\end{cases}$$

فـإن Z_i متغـير عشـوائي يتبـع توزيـع بـيرنولي $B(1,F(x))$، ومـن ثـم $Z=\sum_{i=1}^{n}Z_i$ يتبـع توزيـع ذي الحدين $B(n,F(x))$، حيـث أن Z يمثـل عـدد القيـم x_i الأقـل مـن x. ولكـن حسـب التعريـف:

$$F_n(x)=\frac{1}{n}\sum_{i=1}^{n}Z_i$$

وعلى ذلـك يكـون لـه نفـس توزيـع المعاينـة لمتوسط عينـة عشـوائية مـأخوذة مـن توزيـع بـيرنولي [راجـع (1) في الفقـرة 3.8.3]، أي أن:

$$P\left(F_n^*(x)=\frac{k}{n}\right)=C_n^kF^k(x)\,[1-F(x)]^{n-k} \qquad ; \qquad k=0,1,\ldots,n$$

وهـو المطلـوب.

بمـا أن:

$$F_n^*(x)=\frac{Z}{n}$$

فإن:

$$E\left[F_n^*(x)\right]=E\left(\frac{Z}{n}\right)=\frac{1}{n}EZ=\frac{1}{n}nF(x)=F(x)$$

$$\operatorname{var}\left[F_n^*(x)\right]=V\left(\frac{Z}{n}\right)=\frac{1}{n}F(x)[1-F(x)] \qquad (2.12.3)$$

13.3 المشمولات أو المغطيات COVERAGES

لنفترض أن $X = (X_1,...,X_n)$ عينة عشوائية من توزيع مستمر $\mathcal{L}(\xi)$ دالة كثافته $f(x)$ ودالة توزيعه $F(x)$، ولتكن $Y = (Y_1,...,Y_n)$ العينة المرتبة الموافقة لـ X.

توجد خاصتان مهمتان للإحصاءات المرتبة، مفيدتان جداً عند دراسة بعض طرق الاستدلال الإحصائي، وهما

1. المساحة المحددة بدالة الكثافة $f(x)$ وأي إحصائين مرتبين مستقلة عن دالة الكثافة هذه.

2. الإحصاءات المرتبة $Y_1,...,Y_n$ تقسم (في المتوسط) المساحة تحت منحني التوزيع $f(x)$ إلى $(n+1)$ جزء متساوٍ، ومساحة كل جزء تساوي $1/(n+1)$.

إثبات الخاصة (1)

إذا رمزنا بـ $U_i = F(Y_i)\ ;\ i = 1,...,n$، فإن U_i متغير عشوائي يخضع للتوزيع المنتظم $R(0,1)$ [راجع فقرة (1.3.2)]، وبناءً على ذلك $U = (U_1,...,U_n)$ عبارة عن الإحصاءات المرتبة لعينة عشوائية مأخوذة من مجتمع يتبع التوزيع المنتظم $R(0,1)$، ومن ثم الفرق:

$$Z_{ij} = U_j - U_i = F(Y_j) - F(Y_i) = P(Y_i < \xi < Y_j)$$

عبارة عن المساحة المحددة بمنحني التوزيع $f(x)$ والإحصائين المرتبين $Y_i, Y_j\ ;\ i < j$، ويصبح المطلوب إثبات أن التوزيع الاحتمالي للفرق Z_{ij} لا يعتمد على التوزيع $\mathcal{L}(\xi)$.

بناءً على نتائج المثال (3.10.3)، فإن كثافة الاحتمال لـ (U_i, U_j) هي:

$$h\left(u_i,u_j\right)=\frac{n!}{(i-1)!(j-i-1)!(n-j)!}u_i^{i-1}\left(u_j-u_i\right)^{j-i-1}\left(1-u_j\right)^{n-j}$$
$$;\quad 0<u_i<u_j<1$$

وبالتحويل من المتغير $\left(U_i,U_j\right)$ إلى المتغير $\left(U_i,Z_{ij}\right)$، وبملاحظة أن المحدد الجاكوبي $J=1$، نجد:

$$K\left(u,z\right)=\frac{n!}{(i-1)!(j-i-1)!(n-j)!}u^{i-1}z^{j-i-1}\left(1-z-u\right)^{n-j}\quad;\quad 0<u<1$$
$$;\quad 0<z+u<1$$

وبإجراء تكامل الدالة $K(u,z)$ بالنسبة لـ u نحصل على دالة الكثافة الاحتمالية لـ Z_{ij}:

$$K\left(z\right)=\frac{n!}{(i-1)!(j-i-1)!(n-j)!}z^{j-i-1}\int_0^{1-Z}u^{i-1}(1-z-u)^{n-j}du$$

ولحساب هذا التكامل نضع $y=\frac{u}{1-z}$:

$$K\left(z\right)=\frac{n!}{(i-1)!(j-i-1)!(n-j)!}z^{j-i-1}(1-z)^{n-j+i}\int_0^1 y^{i-1}(1-y)^{n-j}dy \quad (1.13.3)$$

وبناءً على العلاقة (36.3.2) والخاصة (2) لدالة جاما [راجع الفقرة (5.3.2)]، نجد:

$$\int_0^1 y^{i-1}(1-y)^{n-j}dy=\frac{\Gamma(i)\Gamma(n-j+1)}{\Gamma(n-j+i+1)}=\frac{(i-1)!(n-j)!}{(n-j+i)!}$$

وبالتعويض في (1.13.3) نحصل على دالة كثافة توزيع $Z_{ij}=Z$

$$\begin{aligned}K\left(z\right)&=\frac{n!}{(i-1)!(j-i-1)!(n-j)!}\frac{(i-1)!(n-j)!}{(n-j+1)!}z^{j-i-1}(1-z)^{n-j+i}\\&=\frac{n!}{(j-i-1)!(n-j+i)!}z^{j-i-1}(1-z)^{n-j+i}\quad;\quad 0<z<1\end{aligned}\quad (2.13.3)$$

وهذه دالة كثافة توزيع $Be(j-i,n-j+i+1)$ ، وهو لا يعتمـــد علـــى توزيـــع المجتمع $f(x)$.

إثبـــات الخاصـــة (2)

$$EZ = \int_0^1 zK(z)dz = \frac{n!}{(j-i-1)!(n-j+i)!}\int_0^1 z^{j-i}(1-z)^{n-j+i}dz$$

$$= \frac{n!}{(j-i-1)!(n-j+i)!}\frac{\Gamma(j-i+1)\Gamma(n-j+i+1)}{\Gamma(n+2)}$$

$$= \frac{n!}{(j-i-1)!(n-j+i)!}\frac{(j-i)!(n-j+i)!}{(n+1)n!} = \frac{j-i}{n+1}$$

وبوضع $j=i+1$ نجد $EZ = \dfrac{1}{n+1}$. وهـــو المطلـــوب.

تعريف 1.13.3: المشـــمولات أو المغطيـــات Coverages

تدعـــى الفـــترات $\left(-\infty, X_{(1)}\right], \left(X_{(1)}, X_{(2)}\right], \ldots, \left(X_{(n)}, +\infty\right)$ بخلايـــا العينـــة X، ونرمز لها بـــ $B_1, B_2, \ldots, B_{n+1}$ على الترتيب، كما تدعـــى الـــدوال (الإحصـــاءات):

$$C_1 = F(Y_1) - F(-\infty) = U_1$$

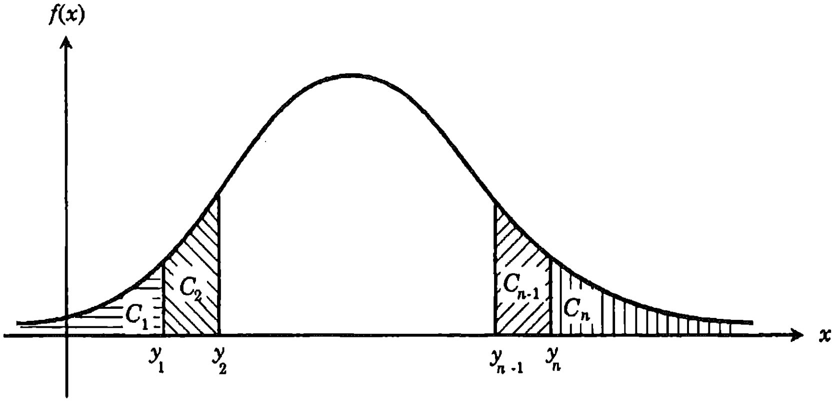
$$C_2 = F(Y_2) - F(Y_1) = U_2 - U_1$$

(3.13.3)

$$C_n = F(Y_n) - F(Y_{n-1}) = U_n - U_{n-1}$$

$$C_{n+1} = F(+\infty) - F(Y_n) = 1 - U_n \qquad ; \quad Y_i = X_{(i)} \ ; \ i = 1, \ldots, n$$

في هذه الخلايا بالمشمـــولات أو المغطيات (coverages)، حيث إن مجموعها $\sum_{i=1}^{n+1} C_i = 1$ ، وهذا يعني أن C_i غطاء الخلية B_i حيث $i = 1, \ldots, n+1$ ، ويبدو ذلك بوضـــوح علـــى الشكل (1.13.3).



شكل 1.13.3

بما أن الغطاء $C_j ; j = 1,\ldots,n$ متغير عشوائي، فله إذن توزيع احتمالي، وهذا التوزيع يمكن إيجاده بسهولة بوضع $i = j-1 , j = j$ في العلاقة (2.13.3):

$$K_{C_j}(z) = n(1-z)^{n-1} \qquad ; \qquad 0 < z < 1 \qquad (4.13.3)$$

وبناءً على ذلك:

$$E(C_i) = n\int_0^1 z(1-z)^{n-1}\,dz = n\frac{\Gamma(2)\Gamma(n)}{\Gamma(n+2)} = \frac{1}{n+1}$$

$$E(C_i\ ) = n\int_0^0 z^n(1-z)^{n-1}\,dz = n\frac{\Gamma(3)\Gamma(n)}{\Gamma(n+3)} = \frac{2}{(n+2)(n+1)}$$

$$V(C_i) = E(C_i^2) - [E(C_i)]^2 = \frac{n}{(n+2)(n+1)^2}$$

14.3 المعاينة في حالة عينات كبيرة

SAMPLING FOR LARGE SAMPLES

بحثنا في البنود السابقة من هذا الفصل أساسيات نظرية المعاينة في حالة عينات بحجم محدود n. ولنتأمل الآن النتائج التي يمكن الحصول عليها من نظرية المعاينة عندما نجعل $n \to \infty$، في هذه الحالة لدينا مجتمع دالة توزيعه $F(x)$ وندرس

متتالية من المتغيرات العشوائية $X_1, X_2, \ldots$ المستقلة مثنى مثنى، وأي فئة من n من تلك المتغيرات تعتبر عينة عشوائية بحجم n مأخوذة من المجتمع $F(x)$.
وسنعتمد في هذا البند على مبرهنات النهاية الواردة في الفصل الأول ونتائجها المفيدة لدراسة السلوك التقاربي لمميزات العينة وإيجاد التوزيعات التقريبية لدوال مختلفة في العينة، في حالة عينة كبيرة الحجم.

1.14.3 تقارب عزوم العينة بالاحتمال
Convergence of Sample Moments in Probability

أحد أبسط النتائج المتعلقة بعزوم العينة من أجل عينات كبيرة معطاة بالمبرهنة الآتية:

مبرهنة 1.14.3

إذا كانت $X = (X_1, \ldots, X_n)$ عينة عشوائية من توزيع $\mathcal{L}(\xi)$، وكان العزم من المرتبة r لـ ξ موجوداً، فإن العزم من المرتبة r للعينة X يتقارب بالاحتمال من العزم من المرتبة r لـ ξ، عندما $n \to \infty$.

الإثبات

لنرمز للعزم الابتدائي من المرتبة r المعرف بالعلاقة (1.5.3) بـ A_{nr} للإشارة إلى علاقة عزم العينة بحجمها n.

كما نعلم:

$$EA_{nr} = \alpha_r$$

$$VA_{nr} = \frac{\alpha_{2r} - \alpha_r^2}{n}$$

وحسب متباينة تشيبيشف (العلاقة 3.2.1)

$$P\left(\left|A_{nr} - \alpha_r\right| \le \varepsilon\right) \ge 1 - \frac{\alpha_{2r} - \alpha_r^2}{n\varepsilon^2}$$

وبأخذ نهاية الطرفين، عندما $n \to \infty$، نجد:

$$\lim_{n\to\infty} P\left(\left|A_{nr} - \alpha_r\right| \le \varepsilon\right) = 1$$

لأن الاحتمال لا يتجاوز الواحد والعزم الابتدائي من المرتبة r لـ ξ موجود (حسب الفرض)، وهو المطلوب.

وبشكل مشابه، يمكن الإثبات من أجل العزم المركزي M_{nr} للعينة $\boldsymbol{X}$ على أنه يتقارب بالاحتمال من العزم المركزي μ_r لـ ξ عندما $n \to \infty$.

تعني المبرهنة (1.14.3) أن عزوم العينة تعتبر قيماً تقريبية جيدة للعزوم النظرية الموافقة لها (إذا كانت هذه الأخيرة موجودة) في حالة عينات كبيرة الحجم.

وكحالة خاصة، عندما $r = 1$ فإن $A_{n1} = \overline{X}$ و $\alpha_1 = \mu$، أي أن متوسط العينة متقارب بالاحتمال من متوسط المجتمع μ. وكذلك، عندما $r = 2$ فإن $M_{n2} = S^2$ و $\mu_2 = \sigma^2$، أي أن تباين العينة متقارب بالاحتمال من تباين المجتمع σ^2.

بصورة عامة، تبقى المبرهنة (1.14.3) صحيحة من أجل أي دالة مستمر لعدد منته من الكميات A_{nr} (فمثلاً، M_{nr} دالة مستمرة في $A_{n1},\ldots,A_{nr}$ على شكل كثير حدود من الدرجة r). وهذه تعتبر نتيجة للمبرهنة العامة الآتية حول التقارب بالاحتمال لدالة مستمرة في متغيرات عشوائية.

مبرهنة 2.14.3

لتكن المتغيرات العشوائية $\eta_1(n), \eta_2(n), \ldots, \eta_r(n)$ متقاربة بالاحتمال، عندما $n \to \infty$، إلى الثوابت $c_1, c_2, \ldots, c_r$ على الترتيب، عندها من أجل أي دالة مستمرة $\varphi(x_1, x_2, \ldots, x_r)$ فالمتغير العشوائي:

$$\eta(n) = \varphi\left(\eta_1(n), \eta_2(n), \ldots, \eta_r(n)\right) \xrightarrow{P} \varphi\left(c_1, c_2, \ldots, c_r\right)$$

الإثبات

بما أن φ دالة مستمرة، فإنه مهما يكن العدد الحقيقي $\varepsilon > 0$ يمكن إيجاد

عدد حقيقي $\delta = \delta(\varepsilon)$، بحيث إن:

$$\left|\varphi(x_1, x_2, \ldots, x_r) - \varphi(c_1, c_2, \ldots, c_r)\right| < \varepsilon$$

من أجل $|x_i - c_i| < \delta \; ; \; i = 1,2,\ldots,r$.

إذا رمزنا بـ $B_i = \left\{|\eta_i(n) - c_i| < \delta\right\} \; ; \; i = 1,2,\ldots,r$، فإن الحادث $B = B_1 \cap B_2 \cap \ldots \cap B_r$ محتوى في الحادث $C = \left\{|\eta(n) - \varphi(c_1, c_2, \ldots, cr)| < \varepsilon\right\}$، ومن ثم:

$$P(C) \geq P(B) = 1 - P(\overline{B}) = 1 - P\left(\overline{B}_1 \cup \overline{B}_2 \cup \ldots \cup \overline{B}_r\right) = 1 - \sum_{i=1}^{r} P\left(\overline{B}_i\right)$$

وبناءً على التقارب بالاحتمال، حسب الفرض، للمتغيرات العشوائية $\eta_i(n)$، فإنه من أجل أي عدد حقيقي δ وأي $\gamma > 0$ يمكن إيجاد عدد طبيعي $n_i = n_i(\gamma)$، بحيث إن:

$$P\left(\overline{B}_i\right) = P\left(|\eta_i(n) - c_i| \geq \delta\right) < \frac{\gamma}{r} \quad ; \quad n \geq n_i \; , \; i = 1,\ldots,r$$

بأخذ $n_0 = \max(n_1, n_2, \ldots, n_r)$ نجد أن المتباينات:

$$P\left(\overline{B}_i\right) < \frac{\gamma}{r} \quad ; \quad i = 1,2,\ldots,r$$

محققة في آن واحد عندما $n \geq n_0$، وبالتالي:

$$P\left(|\eta(n) - \varphi(c_1, c_2, \ldots, c_r)| < \varepsilon\right) \geq 1 - \gamma \quad ; \quad n \geq n_0$$

وهو المطلوب.

نشير أيضاً لاستخدام إحصائي آخر للمبرهنة (2.14.3)، عند دراسة خصوص منحني توزيع متغير عشوائي مستمر، حيث نحتاج في الغالب لإيجاد معامل الإلتواء γ_1 ومعامل التطاول γ_2، المعرفين كما يلي:

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} \quad , \quad \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 \qquad (1.14.3)$$

وإذا كـــانت العينة العشـــوائية $X=(X_1,X_2,\dots,X_n)$ من توزيع مستمر $\mathcal{L}(\xi)$، فإن معامل الإلتواء والتطاول للعينة العشوائية Γ_{n1},Γ_{n2} يعرفان على النحو الآتي:

$$\Gamma_{n1}=\frac{M_{n3}}{S_n^3} \quad , \quad \Gamma_{2n}=\frac{M_{n4}}{S_n^4}-3 \tag{2.14.3}$$

وبمـــلاحظة أن $\Gamma_{ni}\ ,\ i=1,2$ دالتين مستمرتين، عنـــدما $\mu_2>0$، في عزوم العينـــة (راجع العلاقة 8.5.3)، وبالتالي حسب المبرهنـــة (2.14.3) فـــإن Γ_{n1} و Γ_{n2} يتقاربـــان بالاحتمال إلى المميزين النظريين الموافقين لهما γ_1 و γ_2، عندما $n\to\infty$.

2.14.3 التوزيعات التقريبية للعزوم الابتدائيـــة للعينـــة

Approximating Distribution of Sample Elementary Moments

نفترض أن العـــزوم المختلفة من المـــــرتبة r للمجتمـــع $\mathcal{L}(\xi)$ موجـــودة، وهـــذا ما نتخذه دائمـــاً.

إذا كـــــان تـــوزيع المتغـــــير العشـــــوائي η_n يتقـــارب بالتـــــوزيع (convergence in distribution) من توزيع المتغير العشوائي η، عندما $n\to\infty$، فإننـــا نكتب ذلك:

$$\mathcal{L}(\eta_n)\longrightarrow\mathcal{L}(\eta) \quad \text{أو} \quad \mathcal{L}(\eta_n)\xrightarrow[n\to\infty]{}\mathcal{L}(\eta)$$

وفيما يلي، إذا كـــان المتغـــير العشـــوائي η_n يتقـــارب مـــن التوزيـــع الطبيعـــي بالمعلمتين $\sigma_n^{*2}\ ,\ \mu_n^*$ عندما $n\to\infty$. أي لـــه التوزيـــع الحـــــدي $N(\mu_n^*,\sigma_n^{*2})$ فنعـــبر عن ذلـــك:

$$\mathcal{L}(\eta_n)\xrightarrow[n\to\infty]{}N(\mu_n^*,\sigma_n^{*2})$$

وهذا يعـــني:

$$\mathcal{L}\left(\frac{\eta_n-\mu_n^*}{\sigma_n^*}\right)\xrightarrow[n\to\infty]{}N(0,1)$$

يعطى التوزيع الحدي للعزم الابتدائي A_{nr} للعينة بالمبرهنة الآتية.

مبرهنة 3.14.3

إذا كان A_{nr} العزم الابتدائي من المرتبة r لعينة عشوائية بحجم n من مجتمع $\mathcal{L}(\xi)$ عزمه الابتدائي α_r موجود (منتهٍ)، فإن التوزيع الحدي لـ A_{nr} هو $N\left(\alpha_r, \frac{\alpha_{2r} - \alpha_r^2}{n}\right)$، أي أن:

$$\mathcal{L}(A_{nr}) \xrightarrow[n\to\infty]{} N\left(\alpha_r, \frac{\alpha_{2r} - \alpha_r^2}{n}\right) \tag{3.14.3}$$

الإثبات

كما نعلم A_{nr} عبارة عن متوسط n من المتغيرات العشوائية X_i^r , $i = 1,\dots,n$ المستقلة مثنى مثنى ولها نفس التوزيع، وهو توزيع ξ^r، إذن، حسب مبرهنة النهاية المركزية، فإن:

$$\mathcal{L}(A_{nr}) \xrightarrow[n\to\infty]{} N\left(\alpha_r, \frac{\alpha_{2r} - \alpha_r^2}{n}\right)$$

ومن ثم:

$$\mathcal{L}\left(\frac{A_{nr} - \alpha_r}{\sqrt{(\alpha_{2r} - \alpha_r^2)/n}}\right) \xrightarrow[n\to\infty]{} N(0,1) \tag{4.14.3}$$

وهو المطلوب.

وهذا يعني أن A_{nr} يتوزع طبيعياً بالتقارب (asymptatically normally distributed) وفق $N\left(\alpha_r, \frac{\alpha_{2r} - \alpha_r^2}{n}\right)$ من أجل n كبيرة. ومن ثم $nA_{nr} = \sum_{i=1}^{n} X_i^r$ يتوزع طبيعياً بالتقارب وفق $N\left(n\alpha_r, n(\alpha_{2r} - \alpha_r^2)\right)$ من أجل n كبيرة.

وكحالة خاصة، بإعطاء $r=1$ فإن $A_{n1}=\overline{X}$ ، $nA_{nr}=\sum_{i=1}^{n}X_i$ وبالتالي فإن التوزيع التقريبي لـ $\overline{X}$ ، $\sum_{i=1}^{n}X_i$ هما على الترتيب $N\left(\mu,\frac{\sigma^2}{n}\right)$ ، $N\left(n\mu,n\sigma^2\right)$.

نلاحظ أن المبرهنة (3.14.3) لا تشترط توزيعاً معيناً للمجتمع الذي سحبت منه العينة العشوائية $X=\left(X_1,\ldots,X_n\right)$ ، وبعبارة أخرى، بغض النظر عن توزيع المجتمع فإن كلاً من المتغيرين العشوائيين A_{nr} و nA_{nr} يتوزع طبيعياً بالتقارب وفق $N\left(\alpha_r,\frac{\alpha_{2r}-\alpha_r^2}{n}\right)$ و $N\left(n\alpha_r,n\left(\alpha_{2r}-\alpha_r^2\right)\right)$ على الترتيب.

نتيجة 1.14.3

إذا كانت $X=\left(X_1,\ldots,X_n\right)$ عينة عشوائية من توزيع ذي الحدين $B(n,p)$ ، فإن كلاً من $\overline{X}$ و $\sum_{1}^{n}X_i$ يتوزع طبيعياً بالتقارب وفق $N(np,pq)$ و $N\left(n^2p,n^2pq\right)$ على الترتيب.

تمكننا المبرهنة (3.14.3) من تقدير (في حالة عينات كبيرة) احتمال إنحراف عزم العينة A_{nr} عن العزم النظري الموافق α_r.

$$P\left(\sqrt{\frac{n}{\alpha_{2r}-\alpha_r^2}}\left|A_{nr}-\alpha_r\right|<t\right)\approx\frac{1}{\sqrt{2\pi}}\int_{-t}^{t}e^{-\frac{1}{2}z^2}dz=2\Phi(t)-1 \qquad (5.14.3)$$

حيث إن $\Phi(t)$ دالة التوزيع الطبيعي المعياري $N(0,1)$.

3.14.3 التوزيع التقريبي لـ $F_n^*(x)$

رأينا في البند (12.3) أن دالة التوزيع التجريبي $F_n^*(x)$ عبارة عن إحصاء، عند كل قيمة حقيقية معينة x، وتخضع لتوزيع ذي الحدين $B(n,F(x))$. وعندما تكون n كبيرة (نتيجة 1.14.3) فإن:

$$B(n,F(x))\approx N\left\{F(x),\frac{1}{n}F(x)[1-F(x)]\right\} \qquad (6.14.3)$$

أي أن:

$$\mathcal{L}\left\{\sqrt{n}\left[F_n^*(x)-F(x)\right]\right\}\approx N\left[0,F(x)\left(1-F(x)\right)\right]$$

وإذا كنا نرغب في تقدير $F(x)$ لكل x (بدلاً من قيمة حقيقية معينة x)، فيجب دراسة السلوك التقاربي لـ $F_n^*(x)$ من $F(x)$ لجميع قيم x.

مبرهنة 4.14.3

إذا كانت $F_n^*(x)$ و $F(x)$ دالتي التوزيع التجريبي والنظري للمتغير العشوائي ξ، فإن:

$$\lim_{n\to\infty}P\left(\left|F_n^*(x)-F(x)\right|<\varepsilon\right)=1 \tag{7.14.3}$$

مهما تكن $-\infty<x<+\infty$ و $\varepsilon>0$.

الإثبات

بما أن $F_n^*(x)$ عبارة عن التكرار النسبي لظهور الحادث $\{\xi<x\}$ في n تكرار مستقل، واحتمال الظهور في كل تكرار ثابت يساوي $F(x)$، فحسب قانون الأعداد الكبيرة الضعيف بصيغة بيرنولي [مبرهنة (2.4.1)]:

$$F_n^*(x)\xrightarrow{P}F(x)$$

أي أن العلاقة (7.14.3) محققة.

وتعني المبرهنة (7.14.3) أن $F_n^*(x)$ يتقارب بالاحتمال وبانتظام في x من $F(x)$، عندما $n\to\infty$، ومن ثم عندما تكون n كبيرة فإن:

$$F_n^*(x)\approx F(x)$$

لجميع قيم x.

وبناءً على ذلك يمكن اعتبار $F_n^*(x)$، كمقدِّر لـ $F(x)$، يتقارب بالاحتمال من $F(x)$ لجميع قيم x.

مـــبرهنة 5.14.3: مبرهنة كليفينكو Klevenco Theorem

بناءً على شروط المبرهنة السابقة، فـــإن:

$$P\left(\lim_{n\to\infty}\sup_{-\infty<x<\infty}\left|F_n^*(x)-F(x)\right|=0\right)=1 \qquad (8.14.3)$$

وتنص على أنه باحتمال يســـاوي الواحـــد فـــإن $F_n^*(x)$ تتقـــارب بانتظـــام في x من $F(x)$، وهذا يعـــني أن $F_n^*(x)$، كمقـــدِّر لــــ $F(x)$، يتقـــارب مـــن $F(x)$ بانتظام لجميع قيم x وذلك باحتمال يساوي الواحد، وبعبـــارة أخـــرى يكـــون الفـــرق $\left|F_n^*(x)-F(x)\right|$ صغيراً بقدر ما نريد (مـــن أجـــل كـــل قيمـــة لــــ x) باحتمـــال يساوي الواحد، عندما يكون حجم العينـــة كبـــيراً.

وإذا رمزنا بـــ:

$$D_n=D_n(X)=\sup_{-\infty<x<\infty}\left|F_n(x)-F(x)\right| \qquad (9.14.3)$$

فإن D_n متغير عشوائي منقطع يقيس الإنحراف الأعظم للدالة $F_n^*(x)$ عـــن الدالـــة $F(x)$ على كامل محور الأعداد الحقيقية، ومن ثم يمكن كتابة العلاقـــة (8.14.3) علـــى الصورة:

$$P\left(\lim_{n\to\infty}D_n=0\right)=1 \qquad (10.14.3)$$

مـــبرهنة 6.14.3: مبرهنة كالماغوروف Kolmagorov Theorem

إذا كان المتغير العشوائي الملاحظ ξ مستمراً، دالة توزيعه $F(x)$، فمن أجل أي قيمـــة حقيقية معينة $z>0$ تتحقق العلاقة الآتية:

$$\lim_{n\to\infty}P\left(D_n<\frac{z}{\sqrt{n}}\right)=K(z)=1-2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2j^2z^2}\approx 1-2e^{-2z^2} \qquad (11.14.3)$$

وعندما تكون n كبيرة $(n\geq 20)$ فإن:

$$P\left(D_n < \frac{z}{\sqrt{n}}\right) \approx K(z) \approx 1 - 2e^{-2Z^2} \qquad (12.14.3)$$

أو:

$$P\left(D_n \geq \frac{z}{\sqrt{n}}\right) \approx 2e^{-2Z^2}$$

وهـذا يعـني أنـه يمكـن اسـتخدام التوزيـع $K(z)$ كتقريـب جيـد لـ $P\left(D_n < \frac{z}{\sqrt{n}}\right)$ في التطبيقات الإحصائية وذلـك عندمـا تكـون $n \geq 20$.

تستخدم مبرهنة كالماغوروف لإيجاد الحدين (القيمتين) اللذيـن تقـع بينـهما دالـة التوزيع $F(x)$، عندما تكون هذه الأخـيرة غـير معلومـة، وذلـك باحتمـال معـين $\alpha \in (0,1)$، حيـث إن:

$$P\left(D_n < \frac{z_\alpha}{\sqrt{n}}\right) = P\left(F_n^*(n) - \frac{z_\alpha}{\sqrt{n}} < F(x) < F_n^*(n) + \frac{z_\alpha}{\sqrt{n}}\right) \quad ; \quad \forall\, x \in R$$

وعندما تكون $n \geq 20$ فإن z_α تعين مـن المسـاواة:

$$K(z) = \alpha$$

هكذا، عندما تكون $n \geq 20$، فإن المتباينـة:

$$F_n^*(n) - \frac{z_\alpha}{\sqrt{n}} < F(x) < F_n^*(n) + \frac{z_\alpha}{\sqrt{n}} \quad ; \quad \forall\, x \in R \qquad (13.14.3)$$

محققة باحتمال يسـاوي α تقريبـاً، لكـن $0 \leq F(x) \leq 1$، وبالتـالي يمكـن كتابـة المتباينة (13.14.3) على النحـو الآتي:

$$\max\left(0\ ,\ F_n^*(x) - \frac{z_\alpha}{\sqrt{n}}\right) < F(x) < \min\left(1\ ,\ F_n^*(x) + \frac{z_\alpha}{\sqrt{n}}\right)$$

تكمن أهمية الإحصـاء $D_n(X)$ في أن توزيعـه مسـتقل عـن توزيـع المجتمـع الـذي سحبت منه العـينة X، أي أن D_n حـر التوزيـع، ويمكـن إثبـات ذلـك

على النحـو الآتـي:

إذا رمزنـا بـ $u = F(x)$ فـإن $x = F^{-1}(u)$ وبـالتعويض في (9.14.3)، نجـد:

$$D_n = \sup_{0<u<1}\left|F_n^*\left[F^{-1}(u)\right] - u\right|$$

وبالإنتقال إلى المتغيرات العشـوائية:

$$U_i = F(X_i) \qquad ; \qquad i = 1,\ldots,n$$

$$U_{(1)} \le U_{(2)} \le \ldots\ldots \le U_{(n)} \qquad ; \qquad U_{(i)} = F\left(X_{(i)}\right) \ , \ i = 1,\ldots,n$$

الإحصـاءات المرتبـة للعينـة $(U_1,\ldots,U_n)$. وبملاحظـة أن المتباينـة $X_{(i)} \le x$ مكافئـة للمتباينـة $U_{(i)} \le u$ وبـالتعويض في العلاقـة (3.4.3)، نجـد:

$$F_n^*\left[F^{-1}(u)\right] = \frac{1}{n}\sum_{i=1}^{n} e\left(x - x_{(i)}\right) = \frac{1}{n}\sum_{i=1}^{n} e\left(u - u_i\right) = \Phi_n^*(u) \qquad (14.14.3)$$

يعـني هـذا أن الدالـة $F_n^*(x)$ لا تعتمـد علـى الدالـة $F(x)$، وكمـا نعلـم أن توزيـع $U_i = F(X_i)$ هـو التوزيـع المنتظـم $R(0,1)$ و $\Phi_n^*(u)$ دالـة التوزيـع التجريبي الموافقـة للعينـة الملاحظـة $u = (u_1,\ldots,u_n)$ المـأخوذة مـن التوزيـع $R(0,1)$، وعلـى ذلـك فـإن $D_n(X)$ يطـابق $\sup_{0\le u\le 1}\left|\Phi^*(u) - u\right|$، ومـن ثـم توزيـع D_n مسـتقل عـن $F(x)$ (لا يعتمـد علـى توزيـع المجتمـع الأساسـي). وهـذه الخاصـة لهـا أهميـة كبـيرة، حيـث يكفـي حسـاب وجدولـة توزيـع D_n مـرة واحـدة مـن أجـل عينـات عشـوائية مـأخوذة مـن التوزيـع $R(0,1)$. بنـاءً علـى ذلـك وعلـى التوزيـع المقـارب $K(z)$ تم بنـاء جـدول اختبـار كالمـاغوروف، الـذي يعطـي القيـم الحرجـة λ الموافقـة لكـل مـن الاحتمـال α وحجـم العينـة n، بحيـث:

$$P(D_n \ge \lambda) = \alpha$$

وعندما تكـون n كبـيرة، فـإن:

$$P\left(D_n \ge \frac{z}{\sqrt{n}}\right) \approx 2e^{-2Z^2}$$

ومنها:

$$P\left(D_n \geq \frac{z}{\sqrt{n}}\right) = \alpha \quad \Rightarrow \quad \alpha \approx 2e^{-2Z^2}$$

وبأخذ لوغاريتم الطرفين، نجد:

$$z_\alpha \approx \sqrt{-\frac{1}{2}\ln\frac{\alpha}{2}}$$

وهذا يعني، عندما تكون n كبيرة فإن الحد الحرج λ_α يعين من الصيغة التقريبية:

$$\lambda_\alpha = \frac{z_\alpha}{\sqrt{n}} \approx \sqrt{-\frac{1}{2}\ln\frac{\alpha}{2}} \tag{15.14.3}$$

بالإضافة إلى ماسبق، فإن خاصة حر التوزيع لـ D_n، أدت إلى استخدام واسع لهذا الإحصاء في التطبيقات الإحصائية، وخاصة في مجال الاستدلال الإحصائي اللامعلمي.

مبرهنة 7.14.3: مبرهنة سميرنوف Smirnov Theorem

إذا كانت $F_{1n}^*(x)$, $F_{2n}^*(x)$ دالتي التوزيع التجريبي للمتغير العشوائي المستمر ع الموافقتين لعينتين عشوائيتين بحجم n_1 , n_2 على الترتيب، والمأخوذتين من توزيع $F(x)$، وكان:

$$D_{n_1 n_2} = \sup_{-\infty<x<\infty}\left|F_{1n_1}^*(x) - F_{2n_1}^*(x)\right|$$

فإن:

$$\lim_{n_1, n_2 \to \infty} P\left(\sqrt{n_1 n_2/(n_1+n_2)}\, D_{n_1 n_2} < z\right) = K(z) \tag{26.14.3}$$

من أجل أي عدد حقيقي $z > 0$، حيث إن الدالة $K(z)$ معرفة بالعلاقة (12.14.3)، ولهذه المبرهنة استخدام هام في اختبار الفرضيات.

4.14.3 التوزيع المقارب لوسيط العينة
Asymptotic Distribution of Sample Median

وجدنا في الفقرة (1.11.3) التوزيع الاحتمالي لوسيط عينة عشوائية $X = (X_1,\dots,X_n)$ مأخوذة من توزيع مستمر $F(x)$. وذلك عندما يكون حجم العينة n صغيراً أو كبيراً. سنبحث في هذه الفقرة عن التوزيع التقريبي لوسيط عينة عشوائية كبيرة الحجم.

كما نعلم، إذا كانت $X = (X_1,\dots,X_n)$ عينة عشوائية من توزيع مستمر $F(x)$، فإن $U_i = F(X_i)\ ;\ i = 1,\dots,n\ ;\ (U_1,\dots,U_n)$ عينة عشوائية من توزيع منتظم $R(0,1)$، وإذا كانت $Y_i = X_{(i)}\ ;\ i = 1,\dots,n\ ;\ (Y_1,\dots,Y_n)$ العينة المرتبة الموافقة لـ X، فإن $U_{(i)} = F(Y_i)\ ;\ i = 1,\dots,n\ ;\ (U_{(1)},\dots,U_{(n)})$ العينة المرتبة الموافقة لـ $(U_1,\dots,U_n)$ وأن توزيع الإحصاء $U_{(i)}$ هو:

$$h_i(u) = \frac{\Gamma(n+1)}{\Gamma(i)\Gamma(n-i+1)} u^{i-1}(1-u)^{n-i} \quad , \quad 0 < u < 1$$

وقيمته المتوقعة وتباينه:

$$E\left(U_{(i)}\right) = \frac{i}{n+1} \quad , \quad V\left(U_{(i)}\right) = \frac{i(n-i+1)}{(n+2)(n+1)^2}$$

ومن ثم بافتراض حجم العينة فردي $(n = 2m+1)$، فإن القيمة المتوقعة والتباين لوسيط عينة عشوائية $(U_1,\dots,U_n)$ مسحوبة من توزيع منتظم $R(0,1)$ هما (بوضع $i = m+1$):

$$E\left(\tilde{U}\right) = \frac{1}{2} \quad , \quad V\left(\tilde{U}\right) = \frac{1}{4(n+2)} \tag{13.14.3}$$

ومن أجل n كبيرة:

$$V\left(\tilde{U}\right) \approx \frac{1}{4n}$$

وبناءً على المبرهنة الآتية: إذا كان X متغيراً عشوائياً بمتوسط μ وتباين σ^2، وكانت $Y = g(X)$ دالة ما في X، فيمكن الحصول على $E(Y)$ و $V(Y)$ تقريبياً وفق العلاقتين:

$$EY \cong g(\mu) + \frac{1}{2}\sigma^2 g''(\mu)$$
$$VY \cong \sigma^2\left[g'(\mu)\right]^2 \tag{14.14.3}$$

يمكن إيجاد قيمة تقريبية لكل من القيمة المتوقعة والتباين للإحصاء Y_i باستخدام العلاقة بين Y_i و $U_{(i)}$ على النحو الآتي:

$$U_{(i)} = F(Y_i) \;\; \Rightarrow \;\; Y_i = F^{-1}\left(U_{(i)}\right)$$

ومن ثم:

$$E(Y_i) \approx F^{-1}\left[EU_{(i)}\right] = F^{-1}\left(\frac{i}{n+1}\right)$$

$$V(Y_i) \approx \frac{i(n-i+1)}{(n+2)(n+1)^2}\;\frac{1}{\left\{f\left[F^{-1}\left(\frac{i}{n+1}\right)\right]\right\}^2}$$

وبوضع $i = m + 1$ نجد:

$$E\widetilde{X} \approx F^{-1}\left(\frac{1}{2}\right) = \widetilde{\mu}$$

$$V\widetilde{X} \approx \frac{1}{4(n+2)}\;\frac{1}{\left\{f\left[F^{-1}\left(\frac{1}{2}\right)\right]\right\}^2}$$

$$= \frac{1}{4(n+2)}\;\frac{1}{f^2(\widetilde{\mu})} \approx \frac{1}{4nf^2(\widetilde{\mu})}$$

عندما تكون n كبيرة. وهذا يعني أن وسيط العينة تقدير غير متحيز تقريباً لوسيط المجتمع $\widetilde{\mu}$.

مبرهنة

إذا كانت $X=(X_1,\dots,X_n)$ عينة عشوائية من مجتمع كثافة توزيعه الاحتمالية $f(x)$ ودالة توزيعه $F(x)$، فإن وسيط هذه العينة $\tilde{X}$ يتوزع تقريباً وفق التوزيع الطبيعي بمتوسط $E\tilde{X}=\tilde{\mu}$ وتباين $V\tilde{X}=\dfrac{1}{4nf^2(\tilde{\mu})}$، عندما تكون n كبيرة. أي أن:

$$\mathcal{L}(\tilde{X}) \approx N\left(\tilde{\mu}\,,\ \frac{1}{4nf^2(\tilde{\mu})}\right)$$

ومن ثم:

$$\mathcal{L}\left[2\sqrt{n}f(\tilde{\mu})\left(\tilde{X}-\tilde{\mu}\right)\right] \approx N(0,1)$$

تمارين

1. إذا كان ع متغيراً عشوائياً يصف درجات الطلاب في مقرر ما (حيث الدرجة من 100)، وأخذت 18 ملاحظة مستقلة، فكانت على النحو الآتي:

81 ، 57 ، 59 ، 56 ، 72 ، 68 ، 46 ، 54 ، 42 ، 82 ، 87 ، 82 ، 85 ، 82 ، 66 ، 60 ، 82 ، 59

المطلوب:

1. بناء جدول التوزيع التكراري لدرجات الطلاب.

2. إيجاد دالة التوزيع التجريبي لـ ع الموافقة للعينة الملاحظة.

3. رسم المدرج التكراري للتوزيع التجريبي لدرجات الطلاب.

2. ليكن لدينا التوزيع التجريبي لمتغير عشوائي مستمر كما هو مبين في الجدول

الآتـي:

جدول التوزيع التجريبي لـ ع

القيم الملاحظة x	1	2	3	4	5	6	7
التكرار النسبي n_i/n	0.5	0.10	0.15	0.30	0.20	0.15	0.05

المطلـوب:

1. إيجـاد دالـة التوزيـع التجريبي لــ ع.

2. رسـم المـدرج والمضلـع للتوزيـع التجريبي لــ ع.

3. ما هو نوع توزيع ع الذي تستقرئه من شكل المدرج أو المضلع التكراري؟

3. إذا كـانت $X=(X_1,\dots,X_n)$ عينـة عشـوائية مـن مجتمـع يتبـع توزيـع $f(x,\theta)$، فبـين أي مـن الـدوال الآتيـة يمثـل إحصـاء:

$$T_1=\sum X_i+\theta \quad , \quad T_2=\frac{X_1+X_2}{2} \quad , \quad T_3=\prod_{i=1}^{n}\frac{X_i}{\theta^2}$$

$$T_4=\sum_{i=1}^{n-1}\ln X_i \quad , \quad T_5=e^{\theta X_i} \quad , \quad T_6=\frac{X_1X_2}{X_n}$$

4. إذا كـانت $X=(X_1,X_2,X_3)$ عينـة عشـوائية مـن التوزيـع المنتظـم $R(0,1)$، فأوجد توزيـع كـل مـن المتغـيرين:

$$\frac{X_1+X_2}{2} \quad , \quad \bar{X}$$

5. إذا كـانت $X=(X_1,X_2)$ عينـة عشـوائية مـن التوزيـع $N(0,1)$، فـأوجد:

1. توزيـع $\frac{X_2-X_1}{\sqrt{2}}$.

2. توزيع $\frac{(X_1+X_2)^2}{(X_1-X_2)^2}$.

3. توزيع $\frac{(X_1+X_2)}{\sqrt{(X_1-X_2)^2}}$.

4. توزيع $\frac{X_2^2}{X_1^2}$.

6. بفرض $X=(X_1,\ldots,X_n)$ عينة عشوائية من التوزيع $N(0,1)$، أوجد:

1. توزيع $k<n$; $\frac{(\overline{X}_k+\overline{X}_{n-k})}{2}$.

2. توزيع $k<n$; $k\overline{X}_k^2+(n-k)\overline{X}_{n-k}^2$.

3. توزيع $\frac{X_1}{X_n}$.

7. إذا كانت $X=(X_1,X_2)$ و $Y=(Y_1,Y_2)$ عينتين عشوائيتين مستقلتين من التوزيع $N(0,1)$ و $N(1,1)$ على الترتيب، فأوجد:

1. توزيع $\overline{X}+\overline{Y}$.

2. $(Y_1-Y_2)\Big/\sqrt{\left[(X_2-X_1)^2+(Y_1-Y_2)^2\right]\Big/2}$.

3. $(2\overline{X}-2)^2\Big/(X_2-X_1)^2$.

8. لتكن $i=1,2,3$; $\mathcal{L}(X_i)=N(i,i^2)$ و X_1,X_2,X_3 متغيرات عشوائية مستقلة مثنى مثنى:

1. أعطِ مثالاً لإحصاء يتبع توزيع χ^2 بثلاث درجات حرية.

2. أعطِ مثالاً لإحصاء يتبع توزيع $F_{1,2}$.

3. أعط مثالاً لإحصاء يتبع توزيع $S(2)$.

9. إذا كانت $\boldsymbol{X} = \left(X_1,\dots,X_n\right)$ عينة عشوائية من التوزيع المنتظم $R(-2,2)$، فأوجد:

1. توزيع كل من المتغيرين $X_{(1)}$ و $X_{(n)}$.

2. التوزيع المشترك لـ $X_{(1)}$ و $X_{(2)}$.

3. توزيع وسيط العينة $\tilde{X}$ بافتراض أن n عدد فردي.

10. إذا كانت $\boldsymbol{X} = \left(X_1,\dots,X_n\right)$ عينة عشوائية من توزيع $\mathcal{L}(\xi) \in \Gamma(1,\theta)$، فأوجد:

1. التوزيع الاحتمالي لـ $\boldsymbol{X} = \left(X_1,\dots,X_n\right)$.

2. التوزيع الاحتمالي لـ $\left(X_{(1)},\dots,X_{(n)}\right)$.

3. التوزيع الاحتمالي لكل من المتغيرين $X_{(1)}$ و $X_{(n)}$ ثم احسب متوسط وتباين كل منهما.

4. التوزيع الاحتمالي لكل من المدى $R = X_{(n)} - X_{(1)}$ ونصف المدى T.

11. إذا كانت $\boldsymbol{X} = \left(X_1,\dots,X_n\right)$ عينة عشوائية من التوزيع $R(-3,3)$ فأوجد متوسط وتباين كل من المتغيرات الآتية:

$$\frac{X_{(1)} + X_{(3)}}{2} \quad , \quad X_{(n)} - X_{(1)} \quad , \quad X_{(k+1)} \;;\; n = 2k+1$$

12. إذا كانت $\boldsymbol{X} = \left(X_1,\dots,X_n\right)$ عينة عشوائية من التوزيع:

$$f(x) = 2e^{-2(x-3)} \quad ; \quad x \geq 3$$

فأوجد:

1. التوزيع الاحتمالي لـ $X_{(n)} - X_{(1)}$.

2. التوزيع الاحتمالي لوسيط العينة $\tilde{X}$.

3. التوزيع الاحتمالي لـ $R = X_{(n)} - X_{(1)}$.

13. إذا كانت $\boldsymbol{X} = (X_1, X_2)$ عينة عشوائية من التوزيع الأسي:

$$f(x) = \frac{1}{2} e^{-\frac{1}{2}x} \quad ; \quad x > 0$$

فأوجد توزيع X_1/X_2 .

14. إذا كانت $\boldsymbol{X} = (X_1, X_2)$ عينة عشوائية من التوزيع المنتظم $R(0,1)$ ، وكانت $\boldsymbol{Y} = (Y_1, Y_2)$ العينة المرتبة الموافقة لـ $\boldsymbol{X}$، فأوجد التوزيع الشرطي لـ Y_2 بفرض $Y_1 = y_1$.

15. إذا كانت $\boldsymbol{X} = (X_1, \ldots, X_n)$ عينة عشوائية من التوزيع المنتظم $R\left(a - \sqrt{3}b \,,\, a + \sqrt{3}b\right)$ ، حيث a, b ثوابت حقيقية، وكانت $Y_1, \ldots, Y_n$ الإحصاءات المرتبة الموافقة لـ $\boldsymbol{X}$، فأوجد:

1. متوسط وتباين $Y_2 - Y_1$.

2. متوسط وتباين $\dfrac{Y_1 + Y_2}{2}$.

3. توزيع وسيط العينة $\widetilde{X}$.

16. إذا كانت $\boldsymbol{X} = (X_1, \ldots, X_n)$ عينة عشوائية من توزيع $F(x)$ ، وكانت العزوم المركزية (حول المتوسط) من المرتبة $2r$ لهذا التوزيع موجودة، فأثبت أن:

$$VM_{nr} = \frac{1}{n}\left[\mu_{2r} - \mu_r^2 - r(r-1)\mu_2\mu_{r-2} + r\mu_{r-2}\left(r\mu_2\mu_{r-1}\right) - 2\mu_{r+1}\right] + 0\left(\frac{1}{n^2}\right)$$

17. إذا كانت $\boldsymbol{X} = \left(X_1, \ldots, X_{n_1}\right)$ و $\boldsymbol{Y} = \left(Y_1, \ldots, Y_{n_2}\right)$ عينتين عشوائيتين مستقلتين مأخوذتان من التوزيعين $N\left(\mu_1, \sigma^2\right)$ و $N\left(\mu_2, \sigma^2\right)$ على الترتيب، فأثبت أن توزيع

المتغير العشوائي:

$$\frac{\left(\overline{X}-\overline{Y}\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{(n_1-1)S_1^{*2}+(n_2-1)S_2^{*2}}{n_1+n_2-2}\right)}}$$

يتبع توزيع $S(n-2)$، حيث إن

$$S_2^{*2}=\frac{1}{n_2-1}\sum_{i=1}^{n_2}\left(Y_i-\overline{Y}\right)^2 \quad و \quad S_1^{*2}=\frac{1}{n_1-1}\sum_{i=1}^{n_1}\left(X_i-\overline{X}\right)^2$$

18. إذا كان لدينا عينة عشوائية X بحجم $(n=2m+1)$ من توزيع $R(0,1)$، فأثبت أن وسيط هذه العينة له توزيع $Be(m+1,m+1)$.

19. إذا كانت $\left(X_1,\ldots,X_n\right)$ عينة عشوائية من توزيع مستمر $F(x)$، وكانت $D_n=\sup_x\left|F_n^*(x)-F(x)\right|$، فأوجد توزيع المتغير العشوائي المنقطع D_n عندما $n=3$ ، $n=5$.

20. ليكن ξ متغيراً عشوائياً يتبع التوزيع الأسي:

$$f(x)=\frac{1}{3}e^{-\frac{x}{3}} \quad ; \quad x>0$$

سحبت عينة عشوائية بحجم $n=10$ من هذا التوزيع فكانت على النحو الآتي:

1 ، 3 ، 5 ، 0.2 ، 4 ، 0.5 ، 0.6 ، 2.2 ، 0.3 ، 3

أوجد قيمة الإحصاء $D_n=\sup_x\left|F_n^*(x)-F(x)\right|$.

21. إذا كانت $X=\left(X_1,\ldots,X_n\right)$ عينة عشوائية من توزيع مستمر دالة توزيعه:

$$F(x)=1-e^{-0.5x} \quad ; \quad x>0$$

فـأوجد دالـة الكثافـة الاحتماليـة للمتغـيرات العشـوائية الآتيـة:

1. $\overline{X}$ ، $\widetilde{X}$.

2. $1-F(Y_j)$ ، $F(Y_j)$.

3. $F(Y_2)-F(Y_1)$ ، $F(Y_n)-F(Y_1)$.

4. $F(Y_n)-F(Y_1)+F(Y_2)-F(Y_1)$.

حيـــث إن $Y_i = X_{(i)}\ ;\ i=\overline{1,n}$.

22. إذا كـانت $\boldsymbol{X}=(X_1,\ldots,X_{10})$ عينـة عشـوائية مـن التوزيـع $\mathcal{L}(\xi)=N(0,1)$ فأوجد:

1. توزيـــع $Y_{10}=X_{(10)}$.

2. $P(Y_3<\xi<Y_4)$ ، حيــــــث إن $Y_i = X_{(i)}\ ;\ i=\overline{1,10}$.

3. التوقـع الريـاضي للمتغـــير العشـــوائي Y_1-Y_{10}.

23. إذا كانت $\boldsymbol{X}=(X_1,\ldots,X_{10})$ و $\boldsymbol{Y}=(Y_1,\ldots,Y_{25})$ عينتـين عشـوائيتين مسـتقلتين من التوزيع الطبيعي $N(150,28.6)$ ، فعين:

$$E(\overline{X}-\overline{Y})\quad,\quad \operatorname{var}(\overline{X}-\overline{Y})\quad,\quad P(|\overline{X}-\overline{Y}|<4)$$

24. إذا كـانت $\boldsymbol{X}=(X_1,\ldots,X_n)$ عينـة عشـوائية مـــن توزيـــع:

$$f(x;\theta)=\theta e^{-\theta x}\quad ;\quad x>0\ ,\ \theta>0$$

فـأوجد:

1. التوزيـع المقـارب لــ $\overline{X}_n$.

2. التوزيـع المقـارب لــ $\widetilde{X}_n$.

الفصل الرابع:

الإحصاءات الكافية والمعلومات في العينة

1.4 مقدمة

كما نعلم أن المجتمع الإحصائي عبارة عن فئة منتهية أوغير منتهية من المفردات. وغالباً يقتصر الاهتمام على صفة مميزة أو أكثر لمفردات المجتمع الإحصائي، ويعبر عن الصفة المميزة قيد الدراسة بمتغير عشوائي حقيقي ξ. وأن لكل متغير عشوائي ξ توزيع احتمالي، وكل توزيع احتمالي له مميزاته العددية (معالمه)، التي تعبر عن الخصائص الجوهرية لهذا التوزيع، ويمكن تعيين تلك المعالم المختلفة بناءً على معرفة التوزيع الاحتمالي وذلك باستخدام قوانين نظرية الاحتمالات. لكن في مسائل الإحصاء عادة توزيع المتغير العشوائي ξ غير معلوم، وفي حالات عدة يمكن معرفة عائلة التوزيعات التي ينتمي إليها التوزيع الاحتمالي لـ ξ، أي معرفة النموذج الاحتمالي الإحصائي (النموذج الإحصائي) له، وهذا الأخير يعتمد على معلمة θ (وحيدة البعد أو متعددة الأبعاد) غير معلومة. وهدف الإحصائي أن يعرف قيمتها، ومن ثم يتعين توزيع ξ تماماً. ليس من اليسير معرفة القيمة الحقيقية لـ θ الموافقة لتوزيع ξ، حيث هذا يتطلب معرفة قيمة ξ عند كل مفردات المجتمع الإحصائي، ولأسباب عدة (قد تكون: فساد عناصر المجتمع الإحصائي، تعذر الوصول إلى جميع عناصر المجتمع الإحصائي، التكاليف الباهظة، الجهد الكبير اللازم، ...الخ) يتعذر ذلك. وعندئذٍ نلجأ إلى أخذ عينة عشوائية $X=(X_1,\dots,X_n)$ من توزيع ξ (حيث كما أشرنا سابقاً نقتصر على المجتمعات

الإحصائية غير المنتهية أو المنتهية لكن المعاينة بالإرجاع)، وبناءً على معطيات العينة العشوائية X نقوم بتقدير القيمة الحقيقية لـ θ. ويتطلب التقدير عادة إيجاد الدالة المناسبة في عناصر العينة (الإحصاء $T = T(X)$) التي تدعى بمقدِّر θ، وقيمتها عند عينة مشاهدة $x = (x_1,\ldots,x_n)$ تدعى تقديراً للقيمة الحقيقية لـ θ الموافقة لتوزيع ξ. وبعبارة أخرى نبحث عن تكثيف مناسب للمعطيات المتوفرة في العينة X حول المعلمة θ. لهذا نطمح في أن يشتمل الإحصاء $T = T(X)$ على كافة المعلومات المتوفرة في العينة العشوائية X حول المعلمة θ، أي لا تؤدي عملية التكثيف هذه إلى خسارة في المعلومات (loss of information) المتوفرة في العينة العشوائية X حول المعلمة θ. إذا وجد مثل هذا الإحصاء $T = T(X)$، فيدعى بالإحصاء الكافي (sufficient statistic) للمعلمة θ أو التوزيع $f(x;\theta)$ للمتغير العشوائي ξ.

نلاحظ إذا كانت $x = (x_1,\ldots,x_2)$ عينة عشوائية من توزيع ξ، فإن كل المعلومات المتوفرة لدينا حول المعلمة θ تقتصر على معطيات هذه العينة، ولذلك عندما نقول إن $T = T(X)$ إحصاء كافي لـ θ، فهذا يعني أن الإحصاء $T(X)$ امتص أو احتوى على كافة المعلومات المتوفرة في العينة $x = (x_1,\ldots,x_n)$ حول المعلمة θ، وبشكل عام يقال: إن كل المعلومات المتوفرة في العينة العشوائية X حول المعالم المختلفة لـ ξ يمكن تكثيفها بإحصاءات كافية. لكن، بما أن معرفة قيمة θ في $f(x;\theta)$ تعين لنا تماماً التوزيع الاحتمالي الخاص بالمتغير العشوائي الملاحظ ξ، لذلك سنقتصر البحث عن الإحصاء الكافي المناسب للمعلمة θ دون المعالم الأخرى المختلفة لـ ξ غير الواردة في صيغة التوزيع.

هكذا، يمثل الإحصاء الكافي $T = T(X)$ لمعلمة θ تكثيفاً للمتغيرات العشوائية $X_i\ ;\ i = 1,2,\ldots,n$، وبالتالي يعطي أسلوباً جيداً لتقديم معطيات العينة العشوائية X، وخاصة عندما يكون حجم العينة العشوائية X كبيراً. ولذلك من المفيد البحث عن الإحصاء الكافي، الذي يقدم أكبر تكثيف ممكن للمعطيات، وبهذا المعنى يتم الحديث عن الإحصاء الكافي الأصغر (minimal)، حيث

إن المعلمة θ في $f(x;\theta)$ لها غالباً أكثر من إحصاء كافي، وهذا ما سنلاحظه لاحقاً.

2.4 الإحصاءات الكافية SUFFICIENT STATISTICS

تعريف 1.2.4: الإحصاء الكافي The Sufficient Statistic

يدعى الإحصاء $T = T(X)$ بالكافي من أجل التوزيع $f(x;\theta)$ (أو بالإحصاء الكافي من أجل المعلمة θ)، عندما يكون واضحاً عن أي نموذج يدور الحديث) إذا وفقط إذا كان التوزيع الشرطي (المشروط) للعينة العشوائية $X = (X_1, X_2, ..., X_n)$ بافتراض $T(X) = t$ مستقل عن المعلمة θ، أي إذا رمزنا بـ $f(x|t;\theta)$ لدالة الكثافة الشرطية (دالة الاحتمال الشرطي في حالة المنقطع) للمتغير العشوائي $X = (X_1, X_2, ..., X_n)$، بافتراض $T(X) = t$، فإن:

$$f(x|t;\theta) = \frac{f(x;\theta)}{g(t;\theta)} \tag{1.2.4}$$

لا يعتمد على θ، حيث إن:

$f(x;\theta)$ دالة كثافة أو دالة احتمال العينة العشوائية X.

$g(t;\theta)$ دالة كثافة أو دالة احتمال الإحصاء $T(X)$.

يمكن تفسير التوزيع $f(x|t;\theta)$ كتوزيع X على السطح $T(X) = t$، وعندما يكون $T(X)$ إحصاء كافي، فإن هذا التوزيع مستقل (لا يعتمد) عن θ. وهذا يعني أن معرفة أين تقع النقطة العشوائية $(X_1, X_2, ..., X_n)$ على السطح $T(X) = t$ لا تخبرنا أو تزودنا بأية معلومات إضافية حول المعلمة θ. وبعبارة أخرى، إذا علمنا قيمة الإحصاء الكافي فإننا لسنا بحاجة لقيم العينة العشوائية، لأنها لا تقدم لنا معلومات أكثر عن θ.

مثـــــال 1.2.4

إذا كـانت $\boldsymbol{X}=\left(X_1,X_2,\ldots,X_n\right)$ عينــة عشـــوائية مـــن مجتمـــع يتبـــع توزيـــع بواسـون بمعلمـة θ، فهـل $T(\boldsymbol{X})=n\overline{X}$ يعتـبر إحصـاءً كافيــاً للمعلمـة θ؟

تتطلـب الإجابة على هـذا السـؤال إثبـات أن دالـة الاحتمـال الشـرطية $f\left(x|t;\theta\right)$ مسـتقلة عـن θ.

$$f\left(x|t;\theta\right)=P_\theta\left(\boldsymbol{X}=x|t=n\bar{x};\theta\right)$$
$$=\frac{P_\theta\left(\boldsymbol{X}=x,T=n\bar{x}\right)}{P_\theta\left(T=n\bar{x}\right)}=\frac{f\left(x_1,\ldots,x_n;\theta\right)}{P_\theta\left(T=n\bar{x}\right)}$$

لأن $\{\boldsymbol{X}=x\}\subseteq\{T(\boldsymbol{X})=t\}$.

بما أن المتغير العشـوائي الملاحـظ ع يخضـع لتوزيـع بواسـون بالمعلمـة θ، فـإن:

$$f(x;\theta)=\frac{\theta^x}{x!}e^{-\theta}\qquad;\qquad x=0,1,2,\ldots$$

وعليـه:

$$f\left(x_1,x_2,\ldots,x_n;\theta\right)=\prod_{i=1}^{n}\frac{\theta^{x_i}}{x_i\,!}e^{-\theta}=\frac{\theta^{\sum x_i}}{\prod_1^n(x_i\,!)}e^{-n\theta}=\frac{\theta^{n\bar{x}}}{\prod_1^n(x_i\,!)}e^{-n\theta}$$

وحيـث إن $T(\boldsymbol{X})=n\overline{X}=\sum_1^n X_i$ مجمـوع n متغـير عشـــوائي مســتقل لكـــل منهـا نفـس توزيـع المجتمـع الـذي سـحبت منـه العينـة العشـوائية، فـإن T يتبـع توزيـع بواسـون بمعلمـة $n\theta$ [راجـع الفقـرة (3.8.3) علاقـة رقـم (11.8.3)]:

$$P_\theta\left(T=n\bar{x}\right)=\frac{(n\theta)^{n\bar{x}}}{(n\bar{x})!}e^{-n\theta}\qquad;\qquad n\bar{x}=0,1,2,\ldots$$

وبالتـالي:

$$f\left(x\middle|t;\theta\right)=\frac{\theta^{n\bar{x}}e^{-n\theta}}{\prod_{1}^{n}(x_i!)}\Bigg/\left(\frac{(n\theta)^{n\bar{x}}}{(n\bar{x})!}e^{-n\theta}\right)$$

$$=\frac{(n\bar{x})!}{n^{n\bar{x}}\prod_{i=1}^{n}(x_i!)}=\frac{t!}{n^t\prod_{1}^{n}(x_i!)}\qquad;\qquad t=n\bar{x}$$

لا يعتمد على θ، وهذا يعني أن $T=n\bar{X}$ إحصاء كافي لـ θ. ونلاحظ أن التوزيع الشرطي $f\left(x\middle|t;\theta\right)$ عبارة عن توزيع كثير الحدود الموافق لـ n ناتج متساوي الإمكانية، أي أن:

$$\theta=\left(\theta_1,\ldots,\theta_n\right)\quad,\quad\theta_i=\frac{1}{n}\quad,\quad i=1,2,\ldots,n$$

مثال 2.2.4

إذا كانت $X=\left(X_1,X_2,\ldots,X_n\right)$ عينة عشوائية من توزيع $N\left(\theta,\sigma^2\right)$، أي عينة عشوائية من مجتمع يتبع توزيعاً طبيعياً بمعلمة θ، فهل يمكن اعتبار $T(X)=\bar{X}$ إحصاءً كافياً للمعلمة θ؟

بما أن المتغير العشوائي ξ قيد الدراسة يتبع توزيعاً طبيعياً بمعلمة مجهولة θ، فإن:

$$f(x,\theta)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2}\quad,\quad-\infty<x<+\infty$$

وعليه:

$$f(\mathrm{x},\theta)=\prod_{1}^{n}f\left(x_i;\theta\right)=\prod_{1}^{n}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_i-\theta)^2}{2\sigma^2}}$$

$$=\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left[\frac{n(\bar{x}-\theta)^2+\sum_{1}^{n}\left(x_i-\bar{x}\right)^2}{2\sigma^2}\right]$$

لأن:

$$\sum_1^n (x_i-\theta)^2 = n(\bar{x}-\theta)^2 + \sum_1^n (x_i-\bar{x})^2$$

وكما نعلم، أن متوسط عينة عشوائية مأخوذة من مجتمع طبيعي $N(\theta,\sigma^2)$ يخضع بدوره لتوزيع طبيعي $N(\theta,\sigma^2/n)$، أي أن:

$$g(t;\theta) = \left(\frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\right)\exp\left[-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}\right] \quad ; \quad t=\bar{x}$$

لنحسب الآن التوزيع الشرطي $f(x|t;\theta)$:

$$f(x|t;\theta) = \frac{f(x|t;\theta)}{g(t;\theta)} = \frac{1}{\sqrt{n}}\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n-1}{2}} \exp\left[-\frac{\sum_1^n (x_i-\bar{x})^2}{2\sigma^2}\right]$$

وهذا خالي من المعلمة θ، وعليه فإن $T(X)=\bar{X}$ إحصاء كافي للمعلمة θ في النموذج $N(\theta,\sigma^2)$.

مثال 3.2.4

إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من مجتمع دالة كثافته الاحتمالية:

$$f(x;\theta)=\theta x^{\theta-1} \quad ; \quad 0<x<1$$

فأثبت أن $T(X)=-\sum_{i=1}^n \ln X_i$ إحصاء كافي للمعلمة θ.

بما أن:

$$f(x;\theta)=\prod_{i=1}^n f(x_i;\theta)=\theta^n\prod_{i=1}^n x_i^{\theta-1}=\theta^n\frac{\prod_{i=1}^n x_i^{\theta}}{\prod_{i=1}^n x_i}$$

وإذا كان المتغير العشوائي وحيد البعد X يتبع التوزيع $f(x;\theta)$، فإن المتغير العشوائي $Y = -\ln X = \varphi(X)$ يتبع التوزيع:

$$f_Y(y,\theta) = f_X\left(\varphi^{-1}(y);\theta\right)\left|\frac{dx}{dy}\right| = \theta e^{-\theta y} \quad ; \quad y > 0$$

وهذا ما هو إلا التوزيع الأسي $\Gamma\left(1,\frac{1}{\theta}\right)$، وبالتالي [مبرهنة (4.3.2)] فإن توزيع المتغير العشوائي $T = -\sum_{i=1}^{n}\ln X_i$ هو عبارة عن توزيع جاما $\Gamma\left(n,\frac{1}{\theta}\right)$، أي أن:

$$\begin{aligned} g(t;\theta) &= \frac{\theta^n}{\Gamma(n)} t^{n-1}\, e^{-\theta t} \\ &= \frac{\theta^n}{\Gamma(n)} t^{n-1}\, e^{\theta \ln \prod_1^n x_i} \\ &= \frac{\theta^n}{\Gamma(n)} t^{n-1} \left(\prod_1^n x_i\right)^{\theta} \\ &= \frac{\theta^n}{\Gamma(n)} t^{n-1} \prod_1^n x_i^{\theta} \end{aligned}$$

وعلى ذلك فإن:

$$f\left(x\middle|t;\theta\right) = \frac{f(x;\theta)}{g(t;\theta)} = \frac{\theta^n \prod_1^n x_i^{\theta}\, \Gamma(n)}{\theta^n t^{n-1} \prod_1^n x_i^{\theta}\left(\prod_1^n x_i\right)} = \frac{\Gamma(n)}{\left(\prod_1^n x_i\right) t^{n-1}} \quad ; \quad t = -\sum_1^n \ln x_i$$

أي أن $f\left(x\middle|t;\theta\right)$ لا يعتمد على θ، وبالتالي فإن $T = -\sum_1^n \ln X_i$ إحصاءً كافياً للمعلمة θ.

مثال 4.2.4

إذا كانت $X = \left(X_1,\ldots,X_n\right)$ عينة عشوائية من مجتمع $\mathcal{L}(\xi) \in R(o,\theta)$

فـأثبت أن $T = X_{(n)}$ إحصـاءً كافيـاً للمعلمـة θ.

بمـا أن:

$$f(x;\theta) = \frac{1}{\theta} \qquad ; \qquad 0 < x \le \theta$$

والتوزيـع الاحتمـالي للإحصـاء المرتـب $X_{(n)}$ [راجـع العلاقـة (3.10.3)] هــو:

$$h(t;\theta) = n\frac{t^{n-1}}{\theta^n} \qquad ; \qquad 0 < t \le \theta \quad , \quad t = x_{(n)}$$

وعلى ذلـك فـإن:

$$L(x|t;\theta) = \frac{f(x;\theta)}{h(t;\theta)} = \frac{1/\theta^n}{nt^{n-1}/\theta^n} = \frac{1}{nt^{n-1}}$$

وهـذا خـالي مـن θ، وبالتـالي فـإن $T = X_{(n)}$ إحصـاءً كافيـاً للمعلمـة θ.

إن الطريقـة السـابقة الـتي تعتمـد علـى تعريــف الإحصـاء الكـافي (العلاقـة 1.2.4)، تمكننـا فقـط مـن الإجابـة علـى التسـاؤل التـالي: هـل إحصـاء مـا معطـى $T(X)$ كـافي أم لا؟، وهـذه الطريقـة شـاقة وصعبـة في حـالات عـدة، بالإضافـة إلى أنهـا لا تقـدم لنـا أي إمكانية للحصـول علـى إحصـاء كـافي.

إذا أردنـا التـأكد مـن وجـود إحصـاء كـافي (أو التـأكد مـن إحصـاء مـا أنـه كـافي) والحصول عليه، فإن المبرهنـة الآتيـة تقـدم لنـا طريقـة سـهلة لتحقيـق ذلـك.

مـــبرهنة 1.2.4

لتكـن $X = (X_1, X_2, \dots, X_n)$ عينـة عشـوائية مــن مجتمــع توزيعــه $f(x;\theta)$.

إن الشــرط الــلازم والكــافي ليكــون الإحصــاء $T = T(X)$ كافيــاً مــن أجــل المعلمـة θ هـو إمكانيـة تحليـل دالـة كثافـة (أو دالـة احتمــال) العينــة العشــوائية $f(x;\theta)$ إلى مضروبـين علـى النحـو الآتـي:

$$f(x;\theta) = g(t;\theta)h(x) \qquad (2.2.4)$$

حيث إن:

$g(t;\theta)$: دالة كثافة (أو دالة احتمال) توزيع الإحصاء $T(X)$، وهي تعتمد على θ وعلى العينة X من خلال الدالة $t = T(x)$ فقط.

$h(x)$: دالة غير سالبة لا تعتمد على θ.

تدعى هذه المبرهنة بمبرهنة نيمان وفيشر (Fisher-Neyman theorem) كما تدعى بمبرهنة التحليل إلى مضاريب (factorization theorem).

الإثبات

سنقتصر على إثبات المبرهنة (1.2.4) في أهم حالتين خاصتين: المنقطع والأملس.

نورد فيما يلي الإثبات من أجل نموذج منقطع.

لتكن $X = (X_1, X_2, \ldots, X_n)$ عينة عشوائية مأخوذة من توزيع منقطع $f(x;\theta)$.

لنفترض أن الإحصاء $T(X)$ كافي، ولنبرهن أن العلاقة (2.2.4) صحيحة.

بما أن $T(X)$ إحصاء كافي، فإن الدالة $f(x|t;\theta)$ مستقلة عن θ من أجل أي قيمة t من مجموعة قيم $T(X)$ (حسب تعريف الإحصاء الكافي)، أي يمكننا كتابة الدالة $f(x|t;\theta)$ كدالة في x ولتكن $h(x)$، وعندها:

$$
\begin{aligned}
f(x;\theta) &= P_\theta\big(X = x, T(X) = t\big) = \\
&= P_\theta\big(T(X) = t\big)P_\theta\big(X = x | T(X) = t\big) = g(t;\theta)f(x|t;\theta) = \\
&= g(t;\theta)h(x)
\end{aligned}
$$

لأن الحادث $\{X = x\} \subseteq \{T(X) = t\}$، أي أن العلاقة (2.2.4) محققة.

لنفترض الآن أن العلاقة (2.2.4) محققة، ولنبرهن أن $T(X)$ إحصاء كافي لـ θ.

مهما تكن x، بحيث $T(x) = t$، فإن:

$$f(x|t;\theta) = P_\theta(X = x|T(X) = t) = \frac{P_\theta(X = x, T(X) = t)}{P_\theta(T(X) = t)} =$$

$$= \frac{f(x;\theta)}{\sum\limits_{x':T(x')=t} f(x';\theta)} = \frac{g(t;\theta)h(x)}{\sum\limits_{x':T(x')=t} g(t;\theta)h(x')} = \frac{h(x)}{\sum\limits_{x':T(x')=t} h(x')}$$

أي أن $f(x|t;\theta)$ مستقلة عن θ وذلك مهما تكن x، بحيث $T(x) = t$. وإذا كانت $T(x) \neq t$ فمن الواضح أن $f(x|t;\theta) = 0$. هكذا، في كل الحالات (مهما تكن x) فإن الاحتمال الشرطي $f(x|t;\theta)$ مستقل عن θ، وبالتالي الإحصاء $T(X)$ كافي.

إن إثبات المبرهنة (1.2.4) أكثر صعوبة في حالة نموذج مستمر، لذا سنقتصر على إثباتها في حالة خاصة، وذلك عندما تكون الدالة $T(X)$ ملساء في X، وأن المتغيرات البديلة:

$$Y_1 = T(X) \quad , \quad Y_2 = Y_2(X), \ldots\ldots, Y_n = Y_n(X)$$

موجودة وقابلة للحل بالنسبة للمتغيرات $X_i \; ; \; i = 1,2,\ldots,n$ أي أن المحدد الجاكوبي $0 \neq J = \left|\frac{\partial X_i}{\partial Y_i}\right|$.

بناءً على هذه الشروط، فإن دالة الكثافة الاحتمالية للمتغير العشوائي $Y = (Y_1, Y_2, \ldots, Y_n)$ يمكن إيجادها على النحو الآتي:

$$f(y;\theta) = f(x;\theta)|J| \quad ; \quad x_i = \varphi(y_1, y_2, \ldots, y_n)$$

وعليه، دالة كثافة توزيع المتغير العشوائي $Y_1 = T(X)$:

$$f_1(y_1;\theta) = \int\limits_{R^{n-1}} f(y;\theta)dy_2dy_3\ldots dy_n = \int\limits_{R^{n-1}} f(x;\theta)|J|dy_2dy_3\ldots dy_n$$

وكثافة الاحتمال الشرطي لـ Y بافتراض $T(X) = t$، تعطى بالعلاقة:

$$f(y|t;\theta) = \frac{f(y;\theta)}{f_1(t;\theta)} = \frac{f(x;\theta)|J|}{f_1(t;\theta)} \quad ; \quad y_1 = t$$

بعد هذه الملاحظات التمهيدية، فإن إثبات المبرهنة (1.2.4) يتم تماماً كما في حالة نموذج منقطع.

لنفترض أن العلاقة (2.2.4) محققة، ولنثبت أن $f\left(x|t;\theta\right)$ مستقل عن θ مهما تكن x، بحيث $T(x) = t$.

$$f\left(y|t;\theta\right) = \frac{f(x;\theta)|J|}{f_1(t;\theta)} = \frac{g(t;\theta)h(x)|J|}{\int\limits_{R^{n-1}} f(x;\theta)|J|dy_2 d_3 \ldots dy_n} =$$

$$= \frac{g(t;\theta)h(x)|J|}{g(t;\theta)\int\limits_{R^{n-1}} h(x)|J|dy_2 d_3 \ldots dy_n} = \frac{h(x)|J|}{\int\limits_{R^{n-1}} h(x)|J|dy_2 d_3 \ldots dy_n} = h^*(x)|J|$$

وهذا الأخير لا يعتمد على θ، أي أن دالة الكثافة الشرطية لـ Y بفرض $Y_1 = T(X)$ مستقلة عن θ، أي أن $f\left(x|t;\theta\right)$ مستقلة عن θ، وهذا يعني أن $T(X)$ إحصاء كافي للمعلمة θ.

لنفترض الآن أن الإحصاء $T(X)$ كاف لـ θ، ولنثبت أن العلاقة (2.2.4) محققة.

بما أن:

$$f(x;\theta) = \frac{f\left(y|t;\theta\right)f_1(t;\theta)}{|J|} \quad ; \quad T(X) = t$$

وعليه:

$$f(x;\theta) = g(t;\theta)h(x)$$

حيث $g(t;\theta) = f_1(t;\theta)$ و $h(x) = \dfrac{f\left(y|t;\theta\right)}{|J|}$ (لأن $T(X)$ إحصاء كاف لـ θ حسب الفرض).

لنوضح تطبيق المبرهنة (1.2.4)، أي العلاقة (2.2.4)، من خلال بعض الأمثلة.

مثـــــال 5.2.4

إذا كـانت $X=(X_1,X_2,\dots,X_n)$ عينـــة عشـــوائية مـــن مجتمـــع يتبـــع توزيـــع بواسـون بالمعلمـة θ، فـأوجد إحصـاءً كافيـاً للمعلمـــة θ.

بما أن:

$$f(x;\theta)=\frac{\theta^{x}}{x!}e^{-\theta} \quad ; \quad x=0,1,\dots$$

فإن دالـة الاحتمـال لتوزيـع العينـة X:

$$f(x;\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\prod_{i=1}^{n}\frac{\theta^{x_i}}{x_i!}e^{-\theta}=\frac{\theta^{\sum_1^n x_i}}{\prod_1^n (x_i!)}\,e^{-n\theta}$$

وبوضـــــع $t=\sum_{i=1}^{n}x_i$ و $g(t;\theta)=e^{-n\theta}\theta^{\sum x_i}$ و $h(x)=1\Big/\prod_1^n(x_i!)$ يمكننـــــا كتابـة الدالـة $f(x;\theta)$ علـى النحـو الآتـي:

$$f(x;\theta)=g(t;\theta)h(x)$$

أي ان العلاقـة (2.2.4) محققـة، ومـن ثـم $T(X)=\sum_{i=1}^{n}X_i$ يعتـبر إحصـاءً كافيـاً للمعلمـة θ.

مثـــــال 6.2.4

إذا كـــانت $X=(X_1,X_2,\dots,X_n)$ عينـــة عشـــوائية مـــن توزيـــع طبيعـــي $N(\theta,\sigma^2)$ فـأوجد إحصـاءً كافيـاً للمعلمـة θ.

حسـب الفـرض:

$$f(x;\theta)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2} \quad , \quad -\infty<x<\infty$$

وعليه:

$$f(\boldsymbol{x};\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_i-\theta)^2}{2\sigma^2}}$$

$$=\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2\right]$$

$$=\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\overline{x})^2\right]\exp\left[-\frac{n}{2\sigma^2}(\overline{x}-\theta)^2\right]$$

وبوضع:

$$g(t;\theta)=\exp\left[-\frac{n}{2\sigma^2}(t-\theta)^2\right]\quad;\quad t=\overline{x}$$

$$h(\boldsymbol{x})=\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\overline{x})^2\right]$$

نجد:

$$f(\boldsymbol{x};\theta)=g(t;\theta)h(\boldsymbol{x})$$

أي أن العلاقة (2.2.4) محققة. وهذا يعني أن $T(X)=\overline{X}$ إحصاء كافٍ لـ θ.

مثـــال 7.2.4

لتكن $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ عينة عشوائية من توزيع كثافته:

$$f(x;\theta)=\theta x^{\theta-1}\quad,\quad 0\le x\le 1$$

والمطلوب إيجاد إحصاء كافٍ للمعلمة θ.

كثافة توزيع العينة $\boldsymbol{X}$:

$$f(\boldsymbol{x};\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\prod_{i=1}^{n}\theta x_i^{\theta-1}=\theta^n\left(\prod_{i=1}^{n}x_i\right)^{\theta}\Big/\prod_{i=1}^{n}x_i$$

وبوضع:

$$g(t;\theta) = \theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta}$$

$$h(x) = 1\Big/\prod_{i=1}^{n} x_i \qquad , \qquad t = \prod_{i=1}^{n} x_i$$

نجد:

$$f(x;\theta) = g(t;\theta)h(x)$$

أي أن العلاقة (2.2.4) محققة، وبالتالي فإن $T(X) = \prod_{i=1}^{n} X_i$ يكون إحصاءً كافياً للمعلمة θ. ولقد رأينا في المثال (3.2.4) أن $T(X) = -\ln \prod_{i=1}^{n} X_i$ إحصاءً كافياً أيضاً لـ θ، فهل هذا يعني أن هنالك تناقض؟.

نلاحظ بالعودة إلى تعريف الإحصاء الكافي والمبرهنة (1.2.4) أننا لم نذكر ما يدل على أن الإحصاء الكافي للمعلمة θ وحيد (unique) أم لا. ويمكننا القول في هذا المجال أن لكل معلمة θ أكثر من إحصاء كافي، وهذا ما تؤكده المبرهنة التالية.

مبرهنة 2.2.4

إذا كان $T(X)$ إحصاءً كافياً لمعلمة θ في نموذج $F(x;\theta)$، وكانت $U = \psi(T)$ دالة تقابلية (تناظر واحد لواحد) في T، فإن U أيضاً إحصاء كافي للمعلمة θ والدالة $\psi(\theta)$.

يمكننا إثبات صحة هذه المبرهنة بسهولة وذلك باستبدال المتغير T بالمتغير U في العلاقة (2.2.4)، فنجد:

$$\begin{aligned} f(x;\theta) &= g(t;\theta)h(x) \\ &= g\left(\psi^{-1}(u);\theta\right)h(x) \\ &= g^{*}(u;\theta)h(x) \end{aligned}$$

وبالتالي يكون الإحصاء $U(X)$ إحصاءً كافياً للمعلمة θ، وهذا يعني غالباً أنه يوجد أكثر

من إحصاء كافي للمعلمة θ. فمثلاً، إذا كان $T=\sum_{i=1}^{n}X_i$ إحصاءً كافياً للمعلمة θ، فإن $T(X)=\ln\sum_{1}^{n}X_i$ يكون بدوره إحصاءً كافياً لـ θ و $\tau(\theta)=\ln\theta$. وإذا كان $T(X)=\overline{X}$ إحصاءً كافياً لـ θ، فإن $T_1(X)=a\overline{X}+b$ يكون إحصاءً كافياً لـ θ وللدالة المعلمية $\tau(\theta)=a\theta+b$. نترك إثبات ذلك للقارئ على سبيل المثال.

نلاحظ من خلال الأمثلة السابقة أن البحث اقتصر على إيجاد إحصاء كافي $T(X)$ وحيد البعد لمعلمة θ وحيدة البعد أيضاً، ومن ثم ربما يخلص القارئ إلى الآتي:

1. الإحصاء الكافي لمعلمة وحيدة البعد هو بدوره وحيد البعد.
2. الإحصاء الكافي يقتصر على المعالم وحيدة البعد.

لذا، لا بد من إيضاح عدم صحة ذلك التصور. فمثلاً، إذا كانت المعلمة θ، التي نرغب في إيجاد إحصاء كافي لها وحيدة البعد، فقد لا نستطيع تلخيص أو تكثيف المعلومات المتوفرة في العينة العشوائية $X=(X_1,X_2,...,X_n)$ المأخوذة من مجتمع دالة توزيعه $F(x;\theta)$ بإحصاء كافي وحيد البعد (إحصاء واحد) دون فقدان أو خسارة لبعض المعلومات. وبعبارة أخرى قد لا يوجد إحصاء كافي $T(X)$ وحيد البعد للمعلمة θ، بينما يمكن وجود إحصاء كافي متعدد الأبعاد (مؤلف من مجموعة من الإحصاءات). هذا بالإضافة إلى أننا في تعريف الإحصاء الكافي والمبرهنة (1.2.4) لم نذكر أي شيء عن أبعاد المعلمة θ (وحيدة البعد أم متعددة الأبعاد)، وعليه يمكن تعميم ما سبق على حالة معلمة $\theta=(\theta_1,\theta_2,...,\theta_k)$ ذات k بعد.

إذا رمزنا بـ $T(X)=(T_1,T_2,...,T_r)$ للإحصاء الكافي للمعلمة $\theta=(\theta_1,\theta_2,...,\theta_k)$ فإن $r\le n$ يمكن أن تكون أصغر أو أكبر أو مساوية لـ k. ويمكن إيضاح ذلك من خلال بعض الأمثلة.

مثـــال 8.2.4

تعتبر دائماً العينة العشوائية $\boldsymbol{X}=\left(X_1,X_2,\ldots,X_n\right)$، المأخوذة من مجتمع دالة توزيعه $F(x;\theta)$، إحصاءً كافياً للمعلمة θ (وحيدة البعد أو متعددة الأبعاد)، وهذا الإحصاء الكافي يدعى بالتافه لأنه لا يقدم أي تكثيف للمعلومات المتوفرة في العينة العشوائية $\boldsymbol{X}$ حول المعلمة θ.

مثـــال 9.2.4

إذا كانت $\boldsymbol{X}=\left(X_1,X_2,\ldots,X_n\right)$ عينة عشوائية من مجتمع دالة توزيعه $F(x;\theta)$، والإحصاءات المرتبة الموافقة لها:

$$X_{(1)} \le X_{(2)} \le \ldots. \le X_{(n)}$$

فيكون الإحصاء $\boldsymbol{T}(\boldsymbol{X})=\left(X_{(1)},X_{(2)},\ldots,X_{(n)}\right)$ دائماً إحصاءً كافياً للمعلمة θ (وحيدة البعد أو متعددة الأبعاد)، وذلك لأن:

$$f(x;\theta)=\prod_{i=1}^{n} f\left(x_i,\theta\right)$$

$$h(t;\theta)=n!\prod_{i=1}^{n} f\left(t_i,\theta\right) \quad , \quad t_1<t_2<\ldots<t_y$$

ويكون التوزيع الشرطي للعينة $\boldsymbol{X}$ بمعلومية $\boldsymbol{T}(\boldsymbol{X})=t$ هو:

$$f\left(x|t;\theta\right)=\frac{1}{n!} \tag{3.2.4}$$

لا يعتمد على θ، وهذا بشرط أن x هي تبديلة للإحصاءات $\left(t_1,t_2,\ldots,t_n\right)$ وتعني العلاقة (3.2.4) أنه بشرط معرفة القيم المرتبة، فإن القيمة الوحيدة التي يمكن أن تأخذها العينة $\boldsymbol{X}$ هي تباديل $x_{(1)},x_{(2)},\ldots,x_{(n)}$ وحيث إن عدد هذه التباديل هو $n!$ فيكون لكل منها نفس الفرصة وهي $1/n!$.

مثـــال 10.2.4

بفـــرض $X=(X_1,X_2,\dots,X_n)$ عينـــة عشـــوائية مـــن توزيـــع طبيعـــي $N=(\theta_1,\theta_2^2)$ أوجد إحصـــاءً كافيـاً لــ $\theta=(\theta_1,\theta_2)$.

كما نعلم، دالـة كثافة توزيـع العينـة العشـوائية X هـي:

$$f(x;\theta)=\prod_{i=1}^{n}\frac{1}{\theta_2\sqrt{2\pi}}e^{-\frac{(x_i-\theta_1)^2}{2\theta_2^2}}=\theta_2^{-n}\exp\left[-\frac{\sum_1^n x_i^2-2\theta_1 n\bar{x}+n\theta_1^2}{2\theta_2^2}\right](2\pi)^{-n/2}$$

وبوضـع:

$$t=(t_1,t_2)\quad ;\quad t_1=T_1(x)=n\bar{x}\quad ,\quad t_2=T_2(x)=\sum_{i=1}^{n}x_i^2$$

يمكـن كتابـة الدالـة $f(x;\theta)$ علـى الشـكل:

$$f(x;\theta)=g(t;\theta)h(x)$$

حيـث:

$$f(t;\theta)=\theta_2^{-n}\exp\left[-\frac{t_2-2\theta_1 t_1+\theta_1^2}{2\theta_2^2}\right]$$

$$h(x)=(2\pi)^{-n/2}$$

وهـذا يعـني حسـب المبرهنـة (1.2.4) أن الإحصـاء:

$$T=(T_1,T_2)\quad ;\quad T_1(X)=n\bar{X}\quad ,\quad T_2(X)=\sum_{i=1}^{n}X_i^2$$

هو إحصاء كافي للمعلمة $\theta=(\theta_1,\theta_2)$. هنا بالإمكان ضم المضـروب $(2\pi)^{-n/2}$ إلى الدالـة g ووضع $h(x)=1$. وبشكل مشابه، يمكن إثبات أن الإحصاء:

$$T=(T_1,T_2) \quad ; \quad T_1(X)=\overline{X} \quad , \quad T_2(X)=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$$

هو أيضاً إحصاء كــافٍ للمعلمــة $\theta=(\theta_1,\theta_2)$:

$$f(\boldsymbol{x};\theta)=\theta_2^{-n}\exp\left[-\frac{\sum_{1}^{n}(x_i-\overline{x})^2}{2\theta_2^2}-\frac{n(\overline{x}-\theta_1)^2}{2\theta_2^2}\right](2\pi)^{-n/2}$$

$$=\theta_2^{-n}\exp\left[-\frac{t_2}{2\theta_2^2}-\frac{n(t_1-\theta_1)^2}{2\theta_2^2}\right](2\pi)^{-n/2}$$

$$=g(\boldsymbol{t};\theta)h(\boldsymbol{x})$$

حيــث:

$$g(\boldsymbol{t};\theta)=\theta_2^{-n}\exp\left[-\frac{t_2}{2\theta_2^2}-\frac{n(t_1-\theta_1)^2}{2\theta_2^2}\right]$$

$$h(\boldsymbol{x})=(2\pi)^{-n/2}$$

هكذا، من كل المعلومات المتوفرة في العينة العشوائية $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ حــول المعلمة θ يكفينا معرفة $\overline{X}$ و $\sum_{i=1}^{n}X_i^2$.

مثـــال 11.2.4

إذا كانت $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ عينة عشوائية من توزيع منتظم $R(0,\theta)$، فالمطلوب إيجاد إحصاء كاف للمعلمة θ.

كما نعلم كثافة التوزيع المنتظم بالمعلمة θ.

$$f(x;\theta)=\begin{cases}1/\theta \;;\; 0<x\le\theta\\ 0 \;;\; x\le 0 \text{ or } x>\theta\end{cases}$$

وعليه كثافة توزيع العينة العشوائية $\boldsymbol{X}$:

$$f(\boldsymbol{x};\theta)=\begin{cases}1/\theta^n & ;\quad 0<x_i\le\theta\ ;\ i=1,2,\dots,n\\ 0 & ;\quad x_i\le 0 \text{ or } x_i>\theta\end{cases}$$

إذا رمزنا بـ $x_{(1)}=\min x_i$ و $x_{(n)}=\max x_i$ فيمكننا كتابة دالة الكثافة $f(\boldsymbol{x};\theta)$ كدالة في θ على النحو الآتي:

$$f(\boldsymbol{x};\theta)=\begin{cases}1/\theta^n & ,\quad \theta\ge x_{(n)}\\ 0 & ,\quad \theta<x_{(n)}\end{cases}$$

أو على الشكل:

$$f(\boldsymbol{x};\theta)=\frac{e\left(\theta-x_{(n)}\right)}{\theta^n}$$

حيث $e(x)$ دالة خوفسيد.

وبوضع:

$$h(\boldsymbol{x})=1\quad,\quad g(t;\theta)=\frac{e\left(\theta-x_{(n)}\right)}{\theta^n}\quad;\quad T(\boldsymbol{x})=t=x_{(n)}$$

يمكننا كتابة $f(\boldsymbol{x};\theta)$ على الصورة:

$$f(\boldsymbol{x};\theta)=g(t;\theta)h(\boldsymbol{x})$$

أي أن العلاقة (2.2.4) محققة، وهذا يعني أن الإحصاء المرتب $X_{(n)}$ في العينة $\boldsymbol{X}$ هو إحصاء كافي للمعلمة θ.

وبشكل مشابه، من أجل النموذج المنتظم العام $R(\theta_1,\theta_2)$ يمكن كتابة كثافة توزيع العينة $\boldsymbol{X}=(X_1,X_2,\dots,X_n)$ على الصورة:

$$f(\boldsymbol{x};\theta)=\frac{e\left(\theta_2-x_{(n)}\right)e\left(x_{(1)}-\theta_1\right)}{\left(\theta_2-\theta_1\right)^n}$$
$$=g(\boldsymbol{t};\theta)h(\boldsymbol{x})$$

حيث: $g(\boldsymbol{t};\theta)h(\boldsymbol{x})=\dfrac{e\left(\theta_2-x_{(n)}\right)e\left(x_{(1)}-\theta_1\right)}{\left(\theta_2-\theta_1\right)^n}$, $h(\boldsymbol{x})=1$، وهذا يعني أن الإحصاء $T_1=X_{(1)}$, $T_2=X_{(n)}$; $\boldsymbol{T}=\left(T_1,T_2\right)$ هو إحصاء كافي للمعلمة $\theta=\left(\theta_1,\theta_2\right)$.

مثال 12.2.4

لتكن $\boldsymbol{X}=\left(X_1,X_2,\ldots,X_n\right)$ عينة عشوائية من نموذج $R\left(a(\theta),b(\theta)\right)$، حيث $a(\theta)<b(\theta)$ دالتين مستمرتين في θ.

لإيجاد إحصاء كافٍ لـ θ، بشكل مشابه لما ورد في المثال السابق، نجد:

$$f(\boldsymbol{x};\theta)=\frac{e\left(x_{(1)}-a(\theta)\right)e\left(b(\theta)-x_{(n)}\right)}{\left(b(\theta)-a(\theta)\right)^n} \tag{4.2.4}$$
$$=g(\boldsymbol{t};\theta)h(\boldsymbol{x})$$

حيث إن:

$$h(\boldsymbol{x})=1 \quad , \quad g(\boldsymbol{t};\theta)=\frac{e\left(x_{(1)}-a(\theta)\right)e\left(b(\theta)-x_{(n)}\right)}{\left(b(\theta)-a(\theta)\right)^n} \quad , \quad \boldsymbol{t}=\left(x_{(1)},x_{(n)}\right)$$

وهذا يعني حسب المبرهنة (1.2.4) أن الإحصاء $\boldsymbol{T}=\left(X_{(1)},X_{(n)}\right)$ إحصاء كافٍ لـ θ.

والسؤال الذي يطرح نفسه الآن: هل يمكن إيجاد إحصاء كافٍ وحيد البعد للمعلمة θ في المسألة المطروحة، أي إحصاء كافٍ أكثر تكثيفاً؟

نلاحظ أن البسط (الصورة) في العلاقة (4.2.4) يساوي الواحد فقط عندما:

$$\left\{x_{(1)}\geq a(\theta)\ ,\ b(\theta)\geq x_{(n)}\right\} \tag{5.2.4}$$

وبفرض أن $a(\theta)$ متزايدة و $b(\theta)$ متناقصة بازدياد θ، فإن الشرط (5.2.4) يكون مكافئاً للشرط:

$$\left\{\theta \le a^{-1}\left(x_{(1)}\right)\ ,\ \theta \le b^{-1}\left(x_{(n)}\right)\right\} \Leftrightarrow \left\{\theta \le T_1(x) = \min\left(a^{-1}\left(x_{(n)}\right)\ ,\ b^{-1}\left(x_{(n)}\right)\right)\right\}$$

وعليه يمكن كتابة العلاقة (4.2.4) على الشكل:

$$f(x;\theta) = \frac{e\left(T_1(x) - \theta\right)}{\left(b(\theta) - a(\theta)\right)^n}$$

ومن ثم فإن:

$$T_1 = T_1(X) = \min\left(a^{-1}\left(X_{(n)}\right)\ ,\ b^{-1}\left(X_{(n)}\right)\right)$$

إحصاء كاف وحيد البعد للمعلمة θ، وهذا يعني أن $T_1(x)$ أكثر تكثيفاً من الإحصاء الكافي ذو البعدين $T = \left(X_{(1)}, X_{(2)}\right)$.

وبشكل مشابه، بفرض $a(\theta)$ متناقصة و $b(\theta)$ متزايدة بازدياد θ، فإن الإحصاء الكافي وحيد البعد يكون $T_2 = T_2(X) = \max\left(a^{-1}\left(X_{(n)}\right), b^{-1}\left(X_{(n)}\right)\right)$، وهذا أكثر تكثيفاً من $T = \left(X_{(1)}, X_{(2)}\right)$، وتختفي هاتان الحالتان عندما يوجد للنموذج $R\left(a(\theta), b(\theta)\right)$ إحصاء كافي وحيد البعد.

تجدر الإشارة هنا، إلى أن T_1 و T_2 دالتين في الإحصاء الكافي T، وهذه الحالة لها طبيعة عامة، أي أن الإحصاء الكافي الأصغر (الأكثر تكثيفا) هو دالة في أي إحصاء كافي آخر للمعلمة θ.

فمن أجل النموذج $R(-\theta,\theta)$ يأخذ الإحصاء الكافي وحيد البعد T_2 الشكل:

$$T_2 = \max\left(-X_{(1)}, X_{(n)}\right) = \max\left(\left|X_{(1)}\right|, \left|X_{(n)}\right|\right)$$

ومن أجل النموذج $R(\theta, \theta+1)$ فإن الإحصاء الكافي وحيد البعد غير موجود.

مثـــــال 13.2.4

إذا كـانت $\boldsymbol{X}=\left(X_1,X_2,\ldots,X_n\right)$ عينـــة عشـــوائية مـــن توزيـــع كوشـــي بالمعلمـــة θ، فدالـة الكثافـة:

$$f(x;\theta)=\frac{1}{\pi^n}\prod_{i=1}^{n}\frac{1}{1+\left(x_i-\theta\right)^2}$$

وهـذا لا يمكـن كتابتـه بصـورة العلاقـة (2.2.4)، ومـن ثـم لا يمكـــن إيجـاد إحصـــاء كـافٍ للمعلمـة θ، بحجـم أقـــل مـــن n، أي لا يوجـــد ســـوى الإحصـــاء الكـــافي التافـه $\boldsymbol{T}=\boldsymbol{X}$.

نلاحـظ مما سبق إمكانيـة وجـود أكـثر مـن إحصـاء كـافٍ لمعلمـة θ (وحيـدة البعـد أو متعـددة الأبعـاد) في نمـوذج مفـروض، بـل يمكـــن أن تكـــون كثـــيرة جـــداً، فمثـــلاً بـــالعودة إلى المثـــال (9.2.4) نجـــد أن المعلمـــة $\theta=\left(\theta_1,\theta_2\right)$ لهـــا أكـــثر مـــن إحصـــاء كافٍ، منهـــا:

$$\boldsymbol{T}=\left(T_1,T_2\right)\quad;\quad T_1(\boldsymbol{X})=n\overline{X}\quad,\quad T_2(\boldsymbol{X})=\sum_{i=1}^{n}X_i^2$$

$$\boldsymbol{T}=\left(T_1,T_2\right)\quad;\quad T_1(\boldsymbol{X})=\overline{X}\quad,\quad T_2(\boldsymbol{X})=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$$

بالإضافـة إلى الإحصـــاء التافـــة $\boldsymbol{T}=\boldsymbol{X}$.

مـن الواضـح أننـا نرغـــب في إيجـاد الإحصـــاء الكـــافي الأكـــثر تكثيفـــاً للمعلومـــات المتوفـــرة في العينـــة العشـــوائية $\boldsymbol{X}=\left(X_1,X_2,\ldots,X_n\right)$ حـــول المعلمـــة θ في النمـــوذج المفـــروض $F(x;\theta)$، غـــير أن ذلـك غـير ممكـن دائمـــاً. حيـــث نجـــد في حـــالات عـــدة أن الإحصـــاء الكـــافي الوحيـــد في نمـــوذج مفـــروض $F(x;\theta)$ هـــو الإحصـــاء التافـــه. فمثـــلاً نمـــوذج كوشـــي والنمـــوذج المنتظـــم $R(\theta,\theta+1)$ [راجـــع المثـــالين (12.2.4) ، (13.2.4)]. يدعـــى الإحصـــاء الكـــافي الأكـــثر تكثيفـــاً لمعلمـــة θ في نمـــوذج $F(x;\theta)$ بالإحصـــاء الكـــافي الأصغـــر (minimum sufficient statistic).

تعـــريف 3.2.4: الإحصــاء الكــافي الأصغــر

Minimum Sufficient Statistic

إذا كـــان T_1 و T_2 إحصــــائين كــــافيين للمعلمــــة θ في النمــــوذج $F(x;\theta)$، فيقــال إن الإحصــاء الكــافي T_1 يتبــع الإحصــاء الكــافي T_2 عندمــا تكــون T_1 دالــة مقيســة في T_2، أي أن $T_1 = \varphi(T_2)$. وهـــذا يعـــني أن الإحصــــاء الكــــافي T_1 أكــثر تكثيفــاً للمعلومـات المتوفـــرة في العينـــة X حـــول المعلمـــة θ مـــن الإحصــــاء الكــــافي T_2. وإذا كـــان الإحصــاء الكــافي T_0 يتبــع أي إحصــاء كــافٍ آخــر للمعلمــة θ، فيدعــى بالإحصــاء الكــافي الأصغــر أو الأصغــري.

هكـذا، فالإحصـاء الكـافي الأصغـر هـو عبـارة عـن الإحصـاء الكـافي الأكـثر تكثيفـاً علــى الإطـــلاق للمعلومـــات المتوفــرة في العينـــة العشــوائية X حـــول المعلمـــة θ في النمـوذج $F(x;\theta)$. وهـذا يعــني أنـه إذا وجـد إحصـاء كـاف أصغـر T_0 لمعلمـة θ، فـإن أي تكثيـف أكـبر للمعطيـــات المتوفـرة في العينـــة العشــوائية X حـــول θ، بالمقارنــة مـع T_0، غـير ممكـن. وعندهـــا تعتـــبر المعلومـــات الأخـــرى المتوفـــرة في العينـــة العشـــوائية X، والـتي لم يمتصهـا الإحصـاء الكـافي T_0، ناتجـــة عـــن آليـــة عشــوائية مســتقلة عـــن θ (لا تتضمـــن أيـــة معلومـــات حـــول θ).

3.4 توزيــع إحصــاء بشــرط معرفــة إحصــاء كــاف

DISTRIBUTION OF STATISTIC CONDITIONAL ON A SUFFICIENT STATISTIC

رأينـا عنـد إثبـــات المبرهنـــة (1.2.4)، وباعتبـــار $Y_1 = T$ إحصــــاء كــــاف للمعلمـــة θ في نمـــوذج $F(x;\theta)$، أن:

$$f(y|t,\theta) = \frac{f(y_1, y_2, \dots, y_n; \theta)}{f_1(t;\theta)} = h^*(x)|J| \qquad ; \qquad y = (y_1, \dots, y_n)$$

أي أن التوزيـــع المشـــروط للإحصـــاءات $Y_2, \dots, Y_n$ بافـــتراض معلوميـــة $T = t$ لا

يعتمد على θ (مستقل عن θ). وبشكل مشابه، يمكن إثبات أن التوزيع الشرطي لأي مجموعة جزئية من الإحصاءات $Y_2,\dots,Y_n$ بمعلومية T لايعتمد على θ. وأهمية هذه الخاصة تكمن في أنها تؤخذ أحياناً على أنها تعريف للإحصاء الكافي.

تعريف 1.3.4

إذا كانت $X=(X_1,X_2,\dots,X_n)$ عينة عشوائية من توزيع $F(x;\theta)$، فيقال إن الإحصاء $T=T(X)$ إحصاء كاف للمعلمة θ إذا فقط إذا كان التوزيع الشرطي لأي إحصاء آخر $U=\psi(X)$ بافتراض معرفة $T=t$ مستقل عن θ.

إن هذا التعريف مفيد في معرفة ما إذا كان إحصاءٌ ما غير كافٍ. فمثلا، لإثبات أن الإحصاء $U=\psi(X)$ ليس كافياً فإننا نحتاج فقط إلى إيجاد إحصاء آخر $T=T(X)$ يكون توزيعه الشرطي بافتراض معرفة $U=u$ يعتمد على θ (غير مستقل عن θ).

مثال 1.3.4

إذا كانت $X=(X_1,X_2)$ عينة عشوائية من توزيع $N(\theta,1)$، فأثبت أن $U=X_1-X_2$ إحصاء غير كافي للمعلمة θ.

بما أن:

$$f(x;\theta)=\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-\theta)^2}{2}} \quad , \quad -\infty<x<+\infty$$

فإن:

$$f(x;\theta)=f(x_1,x_2;\theta)=f(x_1;\theta)f(x_2;\theta)$$

$$=\frac{1}{2\pi}\exp\left[-\frac{(x_1-\theta)^2+(x_2-\theta)^2}{2}\right]$$

ويمكن التأكد بسهولة من أن $T=X_1+X_2$ إحصاء كاف للمعلمة θ وتوزيع T هو $N(2\theta,2)$.

لنعين التوزيع المشترك للمتغيرات:

$$T = X_1 + X_2 \quad , \quad U = X_2 - X_1$$

وذلك باستخدام طريقة التحويلات

$$H(t,u) = f(x_1,x_2;\theta)|J| \quad ; \quad x_1 = \frac{t-u}{2} \ , \ x_2 = \frac{t+u}{2} \ , \ |J| = \frac{1}{2}$$

$$= \frac{1}{4\pi}\exp\left\{-\frac{1}{2}\left[\left(\frac{t-u}{2}-\theta\right)^2+\left(\frac{t+u}{2}-\theta\right)^2\right]\right\}$$

ولنحسب الآن التوزيع الشرطي للمتغير العشوائي U بمعلومية $T = t$ ولنرمز لذلك بـ $K(u|t)$ ، فنجد:

$$K(t|u) = \frac{H(t,u)}{h(u)} = \frac{\exp\left\{-\frac{1}{2}\left[\left(\frac{t-u}{2}-\theta\right)^2+\left(\frac{t+u}{2}-\theta\right)^2\right]\right\}}{4\pi\exp\left[-\frac{1}{4}u^2\right]}\left(\sqrt{4\pi}\right)$$

$$= \frac{1}{\sqrt{4\pi}}e^{-\frac{1}{2}(t-2\theta)^2}$$

وهذا الأخير يعتمد على θ ، وعليه حسب التعريف (1.2.4) فإن الإحصاء U إحصاء غير كافٍ للمعلمة θ .

4.4 عائلة التوزيعات الأسية والإحصاءات الكافية

EXPONENTIAL DISTRIBUTION FAMILY AND SUFFICIENT STATISTICS

سنتطرق فيما يلي لعائلة من النماذج المعلمية التي تدعى بالأسية، حيث يوجد دائماً لمثل هذه النماذج إحصاء كاف (غير الإحصاء الكافي التافه)، وهذا الإحصاء الكافي أصغري (minimal).

تعــــريف 1.4.4

يقـال عـن نمـوذج $F(x;\theta)$ إنـه أسـي، إذا أمكـن كتابـة الدالـة الموافقـة $f(x;\theta)$ علـى الشـكل الآتـي؛ بافـتراض θ معلمـة سـلمية:

$$f(x;\theta)=\exp\{A(\theta)B(x)+C(\theta)+D(x)\} \qquad (1.4.4)$$

وكثير مـن النمـاذج الـتي تصادفنـا في التطبيـق تحقـق هـذه الخاصـة، ونذكـر مـن هـذه النمـاذج:

$$N(\theta,\sigma^2)\ ,\ N(\mu,\theta^2)\ ,\ B(n,\theta)\ ,\ \Gamma(\lambda,\theta)\ ,\ \Pi(\theta)\ ,\ \overline{B}(r,\theta)$$

ومـن الواضـح أن توزيـع $K(\theta)$ غـير أسـي.

إذا كـانت $X=(X_1,X_2,...,X_n)$ عينـة عشـوائية مـأخوذة مـن مجتمـع توزيعـه الاحتـالي عضـو في عائلـة النمـاذج الأسـية، فـإن:

$$f(x;\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\prod_{i=1}^{n}e^{A(\theta)B(x_i)+C(\theta)+D(x_i)}=e^{A(\theta)\sum_1^n B(x_i)+nC(\theta)+\sum_1^n D(x_i)}$$

$$=e^{A(\theta)\sum_1^n B(x_i)+nC(\theta)}\ e^{\sum_1^n D(x_i)}=g(t;\theta)h(x) \quad ; \quad t=\sum_{i=1}^{n}B(x_i)$$

حيـث إن:

$$g(t;\theta)=e^{A(\theta)\sum_1^n B(x_i)+nC(\theta)}$$

$$h(x)=e^{\sum_1^n D(x_i)}$$

وبالتـالي، حسـب المبرهنـة (1.2.4) فـإن الإحصـاء $T(X)=\sum_{i=1}^{n}B(X_i)$ إحصـاء كـاف للمعلمـة θ، ويمكـن إثبـات أنـه الإحصـاء الكـافي الأصغـر (minimal).

مثـــال 1.4.4

إذا كـانت كثافـة توزيـع المتغـير العشـوائي الملاحـظ ع :

$$f(x;\theta)=\theta e^{-\theta x} \quad ; \quad x>0$$

و $\boldsymbol{X}=(X_1,X_2,\dots,X_n)$ عينة عشوائية من المجتمع $f(x;\theta)$، فأوجد الإحصاء الكافي الأصغر للمعلمة θ.

يمكن كتابة دالة الكثافة $f(x;\theta)$ على النحو الآتي:

$$f(x;\theta)=e^{-\theta x+\ln\theta}$$

وبذلك فهو نموذج أسي فيه:

$$A(\theta)=-\theta \ , \ B(x)=x \ , \ C(\theta)=\ln\theta \ , \ D(x)=0$$

وعليه فإن:

$$T(\boldsymbol{X})=\sum_{i=1}^{n}B(X_i)=\sum_{i=1}^{n}X_i$$

إحصاء كاف أصغر للمعلمة θ.

يمكن تعميم مفهوم عائلة النماذج الأسية ذات المعلمة السلمية على العائلة الأسية ذات المعالم متعددة الأبعاد **(multi-parameter exponential family)**

يقال عن نموذج $f(x;\theta)$، حيث $\theta=(\theta_1,\dots,\theta_k)$ إنه أسي إذا أمكن كتابته على الصورة:

$$f(x;\theta)=\exp\left[\sum_{j=1}^{k}A_j(\theta)B_j(x)+C(\theta)+D(x)\right]$$

وعليه:

$$f(\boldsymbol{x};\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\exp\left[\sum_{j=1}^{k}A_j(\theta)\sum_{i=1}^{n}B_j(x_i)+nC(\theta)+\sum_{i=1}^{n}D(x_i)\right]$$

وباستخدام المبرهنة (1.2.4) نجد أن الإحصاء:

$$T(X)=\left(T_1,T_2,\ldots,T_k\right) \quad ; \quad \begin{cases} T_1=\sum_{i=1}^{n} B_1(x_i) \\ T_2=\sum_{i=1}^{n} B_2(x_i) \\ \\ T_k=\sum_{i=1}^{n} B_k(x_i) \end{cases}$$

يعتبر إحصاءً كافياً للمعلمة $\theta=(\theta_1,\ldots,\theta_k)$ وأصغري.

مثال 2.4.4

إذا كانت $X=\left(X_1,X_2,\ldots,X_n\right)$ عينة عشوائية من توزيع بيتا بالمعلمة $\theta=\left(\theta_1,\theta_2\right)$، فأوجد إحصاءً كافياً للمعلمة θ.

كما نعلم كثافة توزيع بيتا تعطى كما يلي:

$$f(x;\theta)=\frac{1}{\beta\left(\theta_1,\theta_2\right)}x^{\theta_1-1}(1-x)^{\theta_2-1}$$

ويمكن كتابتها على النحو الآتي:

$$f(x;\theta)=\exp\left[-\ln\beta\left(\theta_1,\theta_2\right)-\ln x-\ln(1-x)+\theta_1\ln x+\theta_2\ln(x-1)\right]$$

أي نموذج أسي، وبالمقارنة بالصيغة (2.4.4)، نجد:

$$A_1(\theta)=\theta_1 \ , \ A_2(\theta)=\theta_2 \quad , \quad B_1(x)=\ln x \ , \ B_2(x)=\ln(1-x)$$
$$C(\theta)=-\ln\beta\left(\theta_1,\theta_2\right) \quad , \quad D(x)=-\ln x(1-x)$$

وعليه فإن:

$$T=\left(T_1,T_2\right) \quad ; \quad T_1=\sum_{i=1}^{n}\ln x_i \ , \ T_2=\sum_{i=1}^{n}\ln\left(1-x_i\right)$$

إحصاء كاف للمعلمة $\theta=\left(\theta_1,\theta_2\right)$ وأصغري.

مثــــال 3.4.4

إذا كـانت $X = (X_1, X_2, \dots, X_n)$ عينــة عشــوائية مــن مجتمــع يتبــع توزيــع $N(\theta_1, \theta_2^2)$، فأوجد الإحصــاء الكــافي الأصغــر للمعلمــة $\theta = (\theta_1, \theta_2^2)$.

بما أن النمـوذج الطبيعي $N(\theta_1, \theta_2^2)$ أسـي، فيمكننا كتابته على النحـو الآتــي:

$$f(x;\theta) = \frac{1}{\theta_2 \sqrt{2\pi}}\, e^{-\frac{(x-\theta_1)^2}{2\theta_2^2}}$$

$$= \exp\left[-\frac{1}{2}\ln 2\pi - \frac{1}{2}\left(\ln\theta_2^2 + \frac{\theta_1^2}{\theta_2^2}\right) + \frac{\theta_1}{\theta_2^2}x - \frac{x^2}{2\theta_2^2}\right]$$

وبالمقارنـة مـع الصيغـة الأسـية (2.4.4)، نجـد:

$$A_1(\theta) = \frac{\theta_1}{\theta_2^2} \quad , \quad A_2(\theta) = -\frac{1}{2\theta_2^2} \quad , \quad B_1(x) = x \quad , \quad B_2(x) = x^2$$

$$C(\theta) = -\frac{1}{2}\left(\ln\theta_2^2 + \frac{\theta_1^2}{\theta_2^2}\right) \quad , \quad D(x) = -\frac{1}{2}\ln 2\pi$$

وعليه فـإن:

$$T = (T_1, T_2) \quad ; \quad T_1 = \sum_{i=1}^{n} B_1(X_i) = \sum_{i=1}^{n} X_i \quad , \quad T_2 = \sum_{i=1}^{n} B_2(X_i) = \sum_{i=1}^{n} X_i^2$$

إحصـاء كـاف وأصغـر للمعلمــة $\theta = (\theta_1, \theta_2^2)$.

إن أهميـة النمـوذج الأسـي ليس فقـط في وجـود إحصـاء كـاف أصغـر لـه، بـل في وجـود الإحصـاء التـام (الكـامل) لـه وسـهولة الحصـول عليـه، ويلعـب الإحصــاء التــام دوراً هامـاً في الوصـول إلى المقـدّرات الأكفـأ، كمـا سنرى ذلـك في فصـل لاحـق.

5.4 الكفاية والتمام SUFFICIENCY AND COMPLETENESS

يستخدم مفهوم الكفاية (sufficiency) في أحيان كثيرة مع مفهوم آخر يدعى بالكمال أو التمام (completeness)، لذا لابد من تعريف الإحصاء الكافي والتام أو الإحصاء الكافي التام (complete and sufficient statistic).

تعريف 1.5.4

إذا كان $T = T(X)$ إحصاءً كافياً للنموذج $f(x;\theta)$، وكان الإحصاء $\varphi(T)$ يحقق الشرط:

$$E_\theta \varphi(T) = 0 \quad ; \quad \forall\, \theta \in \Theta$$

فإن $\varphi(t) = 0$ على كامل مجموعة قيم الإحصاء T، وهذا يعني أن:

$$P_\theta\big(X \in \{x; \varphi(T) \neq 0\}\big) = 0 \quad ; \quad \forall\, \theta \in \Theta \quad , \quad t = T(x)$$

يدعى عندئذٍ الإحصاء الكافي T بالتام.

نلاحظ من هذا التعريف أنه إذا كان الإحصاء $T(X)$ تاماً للمعلمة θ فهو أيضاً إحصاء كاف، لذا بدلاً من استخدام مصطلح "إحصاء كاف وتام" يمكننا اختصاراً القول "إحصاء تام".

مثال 1.5.4

إذا كانت $X = (X_1, X_2, \ldots, X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta) = \frac{1}{\theta} \quad ; \quad 0 < x < \theta \; , \; \theta > 0$$

وعلمت أن $T = \max(X)$ إحصاء كاف للمعلمة θ، فأثبت أن T إحصاء تام.

التوزيع الاحتمالي لـ T:

$$g(t;\theta) = \frac{nt^{n-1}}{\theta^n} \quad ; \quad 0 < t < \theta \; , \; \theta > 0$$

وبفرض $\varphi(T)$ إحصاء ما في T، فإن توقعه الرياضي:

$$E_\theta\left[\varphi(T)\right] = \int_R \varphi(t)g(t;\theta)dt =$$

$$= \int_0^\theta \varphi(t)\frac{nt^{n-1}}{\theta^n}dt$$

إذا كان الشرط $E_\theta\left[\varphi(T)\right] = 0 \; ; \; \forall\, \theta \in \Theta$ محقق، فإن:

$$\int_0^\theta \varphi(t)\frac{nt^{n-1}}{\theta^n}dt = 0 \quad ; \quad \forall\, \theta \in \Theta$$

وباشتقاق الطرفين بالنسبة لـ θ، نجد أن:

$$\varphi(\theta)\frac{n\theta^{n-1}}{\theta^n} = 0 \Rightarrow \varphi(\theta) = 0 \quad ; \quad \forall\, \theta \in \Theta$$

أي أن $\varphi(t) = 0$ مهما تكن t من مجموعة قيم T، وهذا يعني حسب التعريف (1.5.4) أن T إحصاء تام.

مثال 2.5.4

إذا كانت $X = \left(X_1, \ldots, X_n\right)$ عينة عشوائية من توزيع بيرنولي $B(1,\theta)$، فاثبت أن $T(X) = \sum_{i=1}^{n} X_i$ إحصاء تام.

كما نعلم أن توزيع الإحصاء T هو $B(n,\theta)$، أي أن:

$$g(t;\theta) = C_n^t \theta^t (1-\theta)^{n-t} \quad ; \quad t = 0,1,\ldots,n \;,\; 0 < \theta < 1$$

لنفترض أن $\varphi(T)$ دالة ما في T، ومن ثم فإن:

$$E_\theta\varphi(T) = \sum_{t=0}^{n} \varphi(t) C_n^t \theta^t (1-\theta)^{n-t}$$

$$= (1-\theta)^n \sum_{t=0}^{n} \varphi(t) C_n^t \left(\frac{\theta}{1-\theta}\right)^t$$

إن كون $E_\theta\varphi(T) = 0$ على الفترة $0 < \theta < 1$ يعني أن:

$$\sum_{t=0}^{n}\varphi(t)C_n^t\left(\frac{\theta}{1-\theta}\right)^t = \sum_{t=0}^{n}\varphi(t)C_n^t d^t = 0$$

لجميع قيم $d = \frac{\theta}{1-\theta}$، وهذه كثيرة حدود في d ولكي تتطابق مع الصفر لابد أن تكون معاملات (أمثال) d^t حيث $t = 0,1,\ldots,n$ مساوية للصفر، أي أن:

$$\varphi(t)C_n^t = 0 \quad ; \quad t = 0,1,\ldots,n$$

وحيث إن $C_n^t \neq 0$ فإن $\varphi(t) = 0 \; ; \; t = 0,1,\ldots,n$ وهذا يعني أن $T = \sum_1^n X_i$ إحصاء تام.

مثال 3.5.4

إذا كانت X عينة عشوائية من ملاحظة واحدة من توزيع $N(0,\theta^2)$، فأثبت أن $T(X) = X$ إحصاء غير تام.

إن توزيع $T = X$ هو:

$$g(t;\theta) = \frac{1}{\theta\sqrt{2\pi}} e^{-\frac{1}{2\theta^2}t^2} \quad ; \quad -\infty < t < +\infty$$

وبافتراض $\varphi(T)$ دالة ما في T، فإن:

$$E_\theta\varphi(T) = \frac{1}{\theta\sqrt{2\pi}}\int_{-\infty}^{+\infty}\varphi(t)e^{-\frac{1}{2\theta^2}t^2} = 0 \quad ; \quad -\infty < t < +\infty$$

وهذا لا يعني أن $\varphi(t)$ مطابق للصفر، وبالتالي فالإحصاء $T = X$ غير تام.

إن التحقق بصورة عامة من تمتع إحصاء كاف ما بخاصة التام يتطلب رياضيات متقدمة باستثناء بعض الحالات الخاصة جداً، والمبرهنة الآتية التي سنقبلها بدون برهان تساعد في الحصول على الإحصاءات التامة في معظم المسائل التطبيقية.

مبرهنة 4.5.4

إذا كانت $X = (X_1, \ldots, X_n)$ عينة عشوائية من توزيع ينتمي لعائلة النماذج الأسية:

$$f(x;\theta) = \exp\left[A(\theta)B(x) + C(\theta) + D(x)\right]$$

فإن $T = \sum_{i=1}^{n} B(X_i)$ إحصاء كاف أصغر وتام.

مثال 5.5.4

إذا كانت $X = (X_1, \ldots, X_n)$ عينة عشوائية من توزيع $N(\theta_1, \theta_2^2)$، فإن:

$$T = (T_1, T_2) \quad ; \quad T_1 = \sum_{i=1}^{n} X_i \ , \ T_2 = \sum_{i=1}^{n} X_i^2$$

إحصاء كاف تام.

وكذلك بناءً على نتائج المثال (1.4.4) فإن الإحصاء $T(X) = \sum_{1}^{n} X_i$ يعتبر إحصاءً كافياً وتاماً للمعلمة θ في النموذج $\Gamma\left(1, \frac{1}{\theta}\right)$.

6.4 دالة المعقولية *LIKELIHOOD FUNCTION*

لنفترض كالعادة $f(x;\theta)$ دالة الكثافة (أو دالة الاحتمال في حالة المنقطع) لتوزيع المتغير العشوائي الملاحظ ξ، $X = (X_1, X_2, \ldots, X_n)$ عينة عشوائية من التوزيع $\mathcal{L}(\xi) \in \boldsymbol{F}$ و $x = (x_1, x_2, \ldots, x_n)$ العينة المشاهدة (نقطة ملاحظة لـ X).

كما نعلم أن توزيع المتغير العشوائي $X = (X_1, X_2, \ldots, X_n)$ يرمز له بـ $f(x;\theta) = f(x_1, x_2, \ldots, x_n; \theta)$، وإذا اعتبرنا هذه الدالة، عند مشاهدة x، كدالة

في المعلمة θ، فإنها تدعى بدالة المعقولية (وفي بعض المراجع دالة الأرجحية، أو دالة الإمكان)، ويرمز لها بـ $L = L(x;\theta)$، أي أن:

$$L(x;\theta) = f(x_1, x_2, \ldots, x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

وسنعتمد فيما يلي مصطلح "دالة المعقولية".

نلاحظ أن الدالة $L(x;\theta)$ هي عبارة عن دالة كثافة (أو دالة احتمال) المتغير العشوائي X وأن $f(x;\theta)$، عند قيمة معينة x لـ X، ينظر إليها كدالة في $\theta \in \Theta$، وتدعى عندئذٍ بدالة المعقولية.

نفترض فيما يلي، أن الدالة $L(x;\theta) > 0$ من أجل كل $x \in G$ ومن أجل كل $\theta \in \Theta$، وقابلة للاشتقاق بالنسبة لـ θ. إذا كانت المعلمة θ سلمية (وحيدة البعد)، فإن المتغير العشوائي:

$$U(X;\theta) = \frac{\partial \ln L(X;\theta)}{\partial \theta} = \sum_{i=1}^{n} \frac{\partial \ln f(X_i;\theta)}{\partial \theta} \qquad (1.6.4)$$

يدعى بدالة مساهمة (أو مساهمة) العينة X، كما يدعى المقدار:

$$\frac{\partial \ln f(X_i;\theta)}{\partial \theta}$$

بمساهمة الملاحظة رقم i على المتغير العشوائي ξ الذي يمثل الصفة المدروسة في المجتمع الإحصائي.

ونفترض أيضاً فيما يلي أن:

$$0 < E_\theta U^2(X;\theta) < \infty \quad ; \quad \forall \theta \in \Theta$$

ونفهم دالة المساهمة (أو المساهمة)، في حالة معلمة متعددة الأبعاد $\theta = (\theta_1, \theta_2, \ldots, \theta_r)$ على أنها متجه عشوائي $U(X;\theta) = (U_1(X;\theta), U_2(X;\theta), \ldots, U_r(X;\theta))$، حيث:

$$U_i(X;\theta) = \frac{\partial \ln L(X;\theta)}{\partial \theta_i} \quad ; \quad i = 1,2,\ldots,r$$

سنحتاج لاحقاً، وأكثر من مرة، للاشتقاق بالنسبة لـ θ لتكاملات دوال معرفة على فضاء العينة G، وكذلك أيضاً إلى تغيير ترتيب عمليتي التكامل والاشتقاق. تدعى النماذج التي تحقق تلك الشروط بالإضافة إلى شروط أخرى بالنماذج النظامية (regularity mogeles)، والشروط الدقيقة، التي تتضمن نظامية النموذج معروفة من التحليل الرياضي وشكلها يتحدد في كل حالة معينة.

لنرى الآن بعض خواص دالة المساهمة $U(X;\theta)$ في حالة نماذج نظامية.

بما أن الدالة $L(x;\theta)$ عبارة عن كثافة (أو دالة احتمال) المتغير العشوائي $X = (X_1, X_2, \ldots, X_n)$، فإن المطابقة الآتية محققة دوماً بالنسبة لـ θ.

$$\int L(x;\theta)dx \equiv 1 \quad , \quad dx = \prod_{i=1}^{n} dx_i \tag{2.6.4}$$

(هنا ولاحقاً من أجل النماذج النظامية المنقطعة يستبدل التكامل بالمجموع).

إذا كان النموذج نظامياً، فباشتقاق المطابقة (2.6.4) بالنسبة لـ θ، نجد:

$$\frac{\partial}{\partial\theta}\int L(x;\theta)dx = \int\frac{\partial L(x;\theta)}{\partial\theta}dx = \int\frac{\partial \ln L(x;\theta)}{\partial\theta}L(x;\theta)dx = E_\theta U(X;\theta) = 0$$

لأن:

$$\frac{\partial \ln L(x;\theta)}{\partial\theta} = \frac{\partial L(x;\theta)}{\partial\theta}\Big/ L(x;\theta)$$

وهذا يعني، من أجل نموذج نظامي، أن التوقع الرياضي (القيمة المتوقعة) لدالة المساهمة يساوي الصفر مهما تكن $\theta \in \Theta$، أي أن:

$$E_\theta U(X;\theta) = 0 \quad ; \quad \forall\, \theta \in \Theta \tag{3.6.4}$$

وباشتقاق طرفي المطابقة (3.6.4) بالنسبة لـ θ، نجد:

$$\int\left[\frac{\partial^2 \ln L(x;\theta)}{\partial\theta^2}L(x;\theta) + \left(\frac{\partial \ln L(x;\theta)}{\partial\theta}\right)^2 L(x;\theta)\right]dx \equiv 0$$

وعلى ذلك، فإن:

$$\int\left(\frac{\partial \ln L(x;\theta)}{\partial\theta}\right)^2 L(x;\theta)dx = -\int\frac{\partial^2 \ln L(x;\theta)}{\partial\theta^2}dx \quad ; \quad \forall\, \theta \in \Theta$$

وهذا يعـني:

$$E_\theta\left(\frac{\partial \ln L(x;\theta)}{\partial\theta}\right)^2 = -E_\theta\left(\frac{\partial^2 \ln L(x;\theta)}{\partial\theta^2}\right)$$

أي:

$$E_\theta[U(\boldsymbol{X};\theta)]^2 = -E_\theta\left(\frac{\partial U(\boldsymbol{X};\theta)}{\partial\theta}\right) \quad ; \quad \forall\, \theta \in \Theta \qquad (4.6.4)$$

وبناءً على تعريف دالة المساهمة (1.6.4)، يمكن كتابة العلاقة (4.6.4) على النحو الآتـي:

$$E_\theta\left(\sum_{i=1}^{n}\frac{\partial \ln f(X_i;\theta)}{\partial\theta}\right)^2 = -nE_\theta\left(\frac{\partial^2 \ln f(X_i;\theta)}{\partial\theta^2}\right) \quad ; \quad \forall\, \theta \in \Theta \qquad (5.6.4)$$

مثـال 1.6.4

إذا كـانت $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ عينـة عشـوائية مـن توزيـع بواسـون $\Pi(\theta)$، فـأوجد دالـة المسـاهمة $U(\boldsymbol{X};\theta)$، ثـم تحقق مـن صحـة العلاقـة (5.6.4).

كما نعلـم:

$$f(x;\theta) = \frac{\theta^x}{x!}e^{-\theta} \quad , \quad x = 0,1,\ldots \;,\; \theta > 0$$

وعلى ذلـك، فـإن:

$$L(\boldsymbol{X};\theta) = \prod_{i=1}^{n} f(X_i;\theta) = \prod_{i=1}^{n}\frac{\theta^{X_i}}{X_i\,!}e^{-\theta} = \frac{\theta^{\sum X_i}}{\prod X_i\,!}e^{-n\theta}$$

$$\ln L(\boldsymbol{X};\theta) = \sum_{i=1}^{n} X_i \ln\theta - \sum_{i=1}^{n}\ln(X_i\,!) - n\theta$$

ومن ثـم:

$$U(\boldsymbol{X};\theta)=\frac{\partial \ln L(\boldsymbol{X};\theta)}{\partial\theta}=\frac{1}{\theta}\sum_{i=1}^{n}X_i-n=n\left(\frac{\overline{X}}{\theta}-1\right) \quad ; \quad \theta>0$$

$$\frac{\partial U(\boldsymbol{X};\theta)}{\partial\theta}=-\frac{n}{\theta^2}\overline{X} \quad ; \quad \theta>0$$

وبالتالي:

$$E_\theta[U(\boldsymbol{X};\theta)]=\frac{n}{\theta}E_\theta\overline{X}-n=n-n=0$$

$$E_\theta[U(\boldsymbol{X};\theta)]^2=E_\theta\left[n\left(\frac{\overline{X}}{\theta}-1\right)\right]^2=n^2E\left(\frac{\overline{X}^2}{\theta^2}-\frac{2}{\theta}\overline{X}+1\right)=$$

$$=\frac{n^2}{\theta^2}E\left(\frac{\sum X_i}{n}\right)^2-\frac{2n^2}{\theta}E\overline{X}+n^2$$

$$=\frac{1}{\theta^2}E\left[\sum_{i=1}^{n}X_i^2+2\sum_{i<j}^{n}X_iX_j\right]-2n^2+n^2$$

$$=\frac{1}{\theta^2}\left[n\left(\theta+\theta^2\right)+n(n-1)\theta^2\right]-2n^2+n^2=\frac{n}{\theta}$$

$$E_\theta\left(\frac{\partial U(\boldsymbol{X};\theta)}{\partial\theta}\right)=E_\theta\left(-\frac{n}{\theta^2}\overline{X}\right)=-\frac{n}{\theta^2}E\overline{X}=-\frac{n}{\theta}$$

أي أن:

$$E_\theta\left(U(\boldsymbol{X};\theta)\right)^2=-E_\theta\left(\frac{\partial U(\boldsymbol{X};\theta)}{\partial\theta}\right)=\frac{n}{\theta} \quad , \quad \theta>0$$

ويمكننا الوصول إلى نفس النتيجة بطريقة أخرى على النحو الآتي:

$$\ln f(X_i;\theta)=-\theta+X_i\ln\theta-\ln(X_i!)$$

$$\frac{\partial \ln f(X_i;\theta)}{\partial\theta}=-1+\frac{X_i}{\theta} \quad , \quad \frac{\partial^2 \ln f(X_i;\theta)}{\partial\theta^2}=-\frac{X_i}{\theta^2}$$

$$E_\theta\left[\sum_{i=1}^{n}\frac{\partial \ln f(X_i;\theta)}{\partial\theta}\right]^2=E_\theta\left[\sum_{i=1}^{n}\left(-1+\frac{X_i}{\theta}\right)\right]^2=\frac{n}{\theta}$$

$$-nE_\theta\left[\frac{\partial^2 \ln f(X_i;\theta)}{\partial\theta^2}\right]=-nE_\theta\left(-\frac{X_i}{\theta^2}\right)=\frac{n}{\theta}$$

أي أن:

$$E_\theta\left[\sum_{i=1}^{n}\frac{\partial \ln f(X_i;\theta)}{\partial\theta}\right]^2 = -nE_\theta\left[\frac{\partial^2 \ln f(X_i;\theta)}{\partial\theta^2}\right] = \frac{n}{\theta}$$

7.4 معلومات فيشر *FISHER INFORMATIONS*

سنناقش في هذا البند نوعاً من دوال المساهمة التي تلعب دوراً هاماً في الإحصاء وخاصة في مسائل التقدير وتسمى دالة معلومات فيشر (Fisher informations function) أو فقط معلومات فيشر حول المعلمة θ المتوفرة في العينة X ونرمز لها بـ $i_n(\theta)$.

إذا كانت $X=(X_1,X_2,\dots,X_n)$ عينة عشوائية من توزيع نظامي $f(x;\theta)$ ويعتمد على معلمة وحيدة البعد θ، فإن $i_n(\theta)$ تعرف بالعلاقة:

$$i_n(\theta) = \underset{\theta}{\operatorname{var}} U(X;\theta) = E_\theta[U(X;\theta)]^2 \tag{1.7.4}$$

وهي تقيس كمية المعلومات المحتواة في العينة X حول المعلمة θ. ويدعى المقدار:

$$i(\theta) = i_1(\theta) = E_\theta\left[\frac{\partial \ln f(X_i,\theta)}{\partial\theta}\right]^2 \tag{2.7.4}$$

بكمية المعلومات (معلومات فيشر) حول المعلمة θ المتوفرة في ملاحظة واحدة، ونستخلص من العلاقات (1.6.4)، (3.6.4)، (1.7.4) و (2.7.4) أن:

$$i_n(\theta) = ni(\theta) \tag{3.7.4}$$

وهذا يعني أن كمية المعلومات المحتواة في عينة عشوائية حول المعلمة θ متناسبة طرداً مع حجم العينة n. وبعبارة أخرى أن كمية المعلومات $i_n(\theta)$ التي تقدمها العينة $X=(X_1,X_2,\dots,X_n)$ حول المعلمة θ تساوي n ضعف كمية المعلومات التي تعطيها ملاحظة واحدة.

وبما أن النموذج $f(x;\theta)$ نظامي، فهو قابل للاشتقاق مرتين بالنسبة لـ θ، وعلى ذلك فباشتقاق العلاقة (3.6.4) مرتين بالنسبة لـ θ عندما $n=1$، فنحصل على صيغة مكافئة من أجل $i(\theta)$:

$$0=\int\frac{\partial^2\ln f(x;\theta)}{\partial\theta^2}f(x;\theta)dx+\int\left[\frac{\partial\ln f(x;\theta)}{\partial\theta}\right]f(x;\theta)dx$$

أي أن:

$$i(\theta)=-E_\theta\left[\frac{\partial^2\ln f(X_1;\theta)}{\partial\theta^2}\right] \tag{4.7.4}$$

وهذه العلاقة غالباً أسهل في إيجاد $i(\theta)$.

مثال 1.7.4

إذا كانت $X=(X_1,X_2,\dots,X_n)$ عينة عشوائية من نموذج بيرنولي $B(1,\theta)$، فأوجد معلومات فيشر المحتواة في ملاحظة واحدة ومن ثم في العينة حول المعلمة θ.

كما نعلم:

$$f(x;\theta)=\theta^x(1-\theta)^{1-x}\quad;\quad x=0,1\quad,\quad 0<\theta<1$$

وعلى ذلك، فإن:

$$\ln f(x;\theta)=x\ln\theta+(1-x)\ln(1-\theta)$$
$$\frac{\partial\ln f(x;\theta)}{\partial\theta}=\frac{x}{\theta}-\frac{1-x}{1-\theta}$$
$$\frac{\partial^2\ln f(x;\theta)}{\partial\theta^2}=-\frac{x}{\theta^2}-\frac{1-x}{(1-\theta)^2}$$
$$E_\theta\left[\frac{\partial^2\ln f(x;\theta)}{\partial\theta^2}\right]=-\frac{\theta}{\theta^2}-\frac{1-\theta}{(1-\theta)^2}=-\frac{1}{\theta}-\frac{1}{1-\theta}$$

وبالتالي فكمية المعلومات المتوفرة في ملاحظة واحدة هي:

$$i_1(\theta) = -E_\theta\left[\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}\right] = \frac{1}{\theta} + \frac{1}{1-\theta} = \frac{1}{\theta(1-\theta)}$$

وكمية المعلومات المتوفرة في العينة X حول المعلمة θ هي:

$$i_n(\theta) = ni(\theta) = \frac{n}{\theta(1-\theta)}$$

وبما أن تباين المتغير العشوائي ξ الذي يتبع توزيع بيرنولي هو $\theta(1-\theta)$، إذن:

$$i(\theta) = \frac{1}{\underset{\theta}{\operatorname{var}}\,\xi}$$

مثال 2.7.4

إذا كانت $X = (X_1, X_2, \ldots, X_n)$ عينة عشوائية من $N(\theta, \sigma^2)$، فأوجد معلومات فيشر في ملاحظة واحدة X_i ومن ثم في العينة العشوائية X.

هنا:

$$f(x;\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\theta)^2} \quad ; \quad -\infty < x < +\infty\ ,\ -\infty < \theta < +\infty$$

ومن ثم:

$$\ln f(x;\theta) = \ln\frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2\sigma^2}(x-\theta)^2$$

$$U(X_1;\theta) = \frac{\partial \ln f(X_1,\theta)}{\partial\theta} = \frac{X_1-\theta}{\sigma^2}$$

$$\frac{\partial^2 \ln f(X_1,\theta)}{\partial\theta^2} = -\frac{1}{\sigma^2}$$

وعلى ذلك، فإن:

$$i_1(\theta) = i(\theta) = -E_\theta\left[\frac{\partial^2 \ln f(X_1,\theta)}{\partial\theta^2}\right] = \frac{1}{\sigma^2}$$

$$i_n(\theta) = ni(\theta) = \frac{n}{\sigma^2}$$

وشكل الدالة $i(\theta)$ من أجل بعض النمـــاذج وارد في الجـــدول (1.7.4)، ونـــترك إثبات ذلك للقارئ على سبيل المثـــال.

جـــدول (1.7.4)

النموذج	$N(\mu,\theta^2)$	$\Gamma(\lambda,\theta)$	$K(\theta)$	$B(n,\theta)$	$\Pi(\theta)$	$\overline{B}(r,\theta)$
$i(\theta)$	$2/\theta^2$	λ/θ^2	$1/2$	$n/[\theta(1-\theta)]$	$1/\theta$	$r/[\theta(1-\theta)^2]$

وكمثال على النموذج غير النظامي نذكر التوزيع المنتظم $R(0,\theta)$ ، هنا من المطابقـــة $\int_0^\theta \frac{1}{\theta}dx \equiv 1$ لا ينتج أن $\int_0^\theta \frac{\partial}{\partial\theta}\left(\frac{1}{\theta}\right)dx = 0$ لأنه عند اشتقاق التكامل بالنسبة لحده الأعلـــى يظهر أيضاً حد آخر. وفي هذه الحالة تكمن عدم النظامية في أن فضاء العينة G يعتمد علـــى المعلمة المجهولة θ.

والتساؤل الذي يطرح نفسه الآن: ما هي العلاقـــة بـــين كميـــة المعلومـــات الـــتي يقدمها إحصاء ما حول المعلمة θ وكمية المعلومـــات الـــتي تعطيـــها العينـــة العشـــوائية حول تلك المعلمة؟ وهذا ما تجيب عنه المبرهنـــة الآتيـــة.

مـــبرهنة 1.7.4

إذا كانت $X = (X_1, X_2, \ldots, X_n)$ عينة عشـــوائية مـــن نمـــوذج $f(x;\theta)$، وكـــان $T = T(X)$ إحصاء ما، فـــإن:

$$i_T(\theta) \le i_n(\theta)$$

حيث $i_T(\theta)$ كمية المعلـــومات التي يعطيـــها الإحصـــاء $T(X)$ حـــول المعلمـــة θ، وتتحقق المساواة عندما يكون T إحصاء كـــاف للمعلمـــة θ.

الإثبـــات

كما نعلم من نظرية الاحتمالات أن:

$$L(x;\theta) = h(t;\theta)g(x|t;\theta) \tag{5.7.4}$$

حيث إن:

$h(t;\theta)$ دالة الكثافة (دالة الاحتمال) لتوزيع الإحصاء T.

$g(x|t;\theta)$ دالة الكثافة (دالة الاحتمال) الشرطية للعينة X بشرط أن $T = t$.

بأخذ لوغاريتم الطرفين في (5.7.4)، نجد:

$$\ln L(x;\theta) = \ln h(t;\theta) \ln g(x|t;\theta)$$

وبإجراء الاشتقاق مرتين بالنسبة لـــ θ وأخذ توقع الطرفين نحصل على:

$$E_\theta\left[\frac{\partial^2 \ln L(X;\theta)}{\partial\theta^2}\right] = E_\theta\left[\frac{\partial^2 \ln h(T;\theta)}{\partial\theta^2}\right] + E_\theta\left[\frac{\partial^2 \ln g(X|t;\theta)}{\partial\theta^2}\right]$$

وعلى ذلك فإن:

$$i_n(\theta) = i_T(\theta) + E_\theta\left[-\frac{\partial^2 \ln g(X|t;\theta)}{\partial\theta^2}\right]$$

أو:

$$i_n(\theta) = i_T(\theta) + E_\theta\left[\frac{\partial \ln g(X|t;\theta)}{\partial\theta}\right]^2$$

وبملاحظة أن المقدار الثاني في الطرف الأيمن لا يمكن أن يكون سالباً (صفر أو موجـــب) فإن:

$$i_T(\theta) \le i_n(\theta)$$

وتتحقق المساواة عندما

$$E_{\theta}\left[\frac{\partial \ln g(X|t;\theta)}{\partial \theta}\right]^{2}=0$$

وهذا لا يمكن إلا إذا كانت الدالة $g(X|t;\theta)$ مستقلة عن θ، أي أن $T=T(X)$ إحصاء كاف.

هكذا، تبين المبرهنة (1.7.4) أن كمية المعلومات التي تعطيها العينة العشوائية X حول المعلمة المجهولة θ أكبر من كمية المعلومات التي يعطيها أي إحصاء $T=T(X)$، باستثناء الإحصاء الكافي فيعطي نفس كمية المعلومات التي تعطيها العينة.

مثال 2.7.5

إذا كانت $X=(X_1,X_2,...,X_n)$ عينة عشوائية من مجتمع توزيعه:

$$f(x;\theta)=\theta e^{-\theta x} \quad , \quad x\ge 0$$

وعلمت أن $T(X)=\sum_{i=1}^{n}X_i$ إحصاء كافياً لـ θ، فأوجد معلومات فيشر في العينة وفي الإحصاء T.

بما أن:

$$L(x;\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\theta^{n}e^{-\theta\sum x_i}$$

فإن:

$$\ln L(x;\theta)=n\ln\theta-\theta\sum x_i$$

$$\frac{\partial \ln L(x;\theta)}{\partial\theta}=\frac{n}{\theta}-\sum x_i \quad , \quad \frac{\partial^2 \ln L(x;\theta)}{\partial\theta^2}=-\frac{n}{\theta^2}$$

إذن:

$$i_n(\theta) = -E_\theta\left[\frac{\partial^2 \ln L(X;\theta)}{\partial\theta^2}\right] = -E_\theta\left(-\frac{n}{\theta^2}\right) = \frac{n}{\theta^2}$$

وبما أن $T(X)$ متغير عشوائي مؤلف من مجموع n من المتغيرات العشوائية المستقلة ولكل منها التوزيع $f(x;\theta)$، فإن توزيعه:

$$h(t;\theta) = \frac{\theta^n}{\Gamma(n)} t^{n-1} e^{-\theta t} \quad , \quad t > 0$$

وعلى ذلك:

$$\ln h(t;\theta) = n\ln\theta - \ln\Gamma(n) + (n-1)\ln t - \theta t$$

$$\frac{\partial \ln h(t;\theta)}{\partial\theta} = \frac{n}{\theta} - t \quad , \quad \frac{\partial^2 \ln h(t;\theta)}{\partial\theta^2} = -\frac{n}{\theta^2}$$

إذن:

$$i_T(\theta) = -E_\theta\left[\frac{\partial^2 \ln h(T;\theta)}{\partial\theta^2}\right] = -E_\theta\left(-\frac{n}{\theta^2}\right) = \frac{n}{\theta^2}$$

وبالتالي:

$$i_n(\theta) = i_T(\theta) = \frac{n}{\theta^2}$$

بينما لو أخذنا الإحصاء $T(X) = X_1 + X_2$ (غير كاف طبعاً)، فنجد:

$$i_T(\theta) = \frac{2}{\theta} < i_n(\theta) = \frac{n}{\theta^2} \quad ; \quad n > 2$$

مثال 3.7.4

لتكن $X = (X_1, X_2, \ldots, X_n)$ عينة عشوائية من $\Pi(\theta)$ و $T = \sum_1^n X_i$.
أوجد معلومات فيشر في العينة وفي الإحصاء T.

نلاحظ من الجدول (1.7.4) أن $i_1(\theta) = i(\theta) = 1/\theta$ وعليه $i_n(\theta) = ni(\theta) = n/\theta$
وبما أن $T = \sum_{i=1}^n X_i$ إحصاء كاف وتوزيعه هو:

$$h(t;\theta)=\frac{(n\theta)^t}{t!}e^{-n\theta} \qquad ; \qquad t=0,1,2,\ldots$$

فإن:

$$\ln h(t;\theta)=t\ln(n\theta)-\ln(t!)-n\theta$$

$$\frac{\partial \ln h(t;\theta)}{\partial\theta}=\frac{t}{\theta}-n \quad , \quad \frac{\partial^2 \ln h(t;\theta)}{\partial\theta^2}=-\frac{t}{\theta^2}$$

$$i_T(\theta)=-E_\theta\left[\frac{\partial^2 \ln h(T;\theta)}{\partial\theta^2}\right]=-E_\theta\left(-\frac{T}{\theta^2}\right)=\frac{n\theta}{\theta^2}=\frac{n}{\theta}$$

إذن:

$$i_n(\theta)=i_T(\theta)=\frac{n}{\theta}$$

يمكن تعميم مفهوم معلومات فيشر إلى الحالة التي يكون فيها النموذج الإحصائي، الذي يخضع له توزيع المتغير العشوائي الملاحظ ع، معتمداً على معلمة ذات r بعد $\theta=(\theta_1,\theta_2,\ldots,\theta_r)$، أي r معلمة. فنحصل عندئذٍ على ما يسمى بمصفوفة معلومات فيشر أو فقط مصفوفة المعلومات (informations matrix). ونرمز لها بـ $I(\theta)=I$ في حالة ملاحظة واحدة وبـ $I_n=I_n(\theta)$ في حالة عينة بحجم n، وتعرف على النحو الآتي:

$$I=\|g_{ij}\|_1^r \qquad I_n=nI \tag{6.7.4}$$

حيث إن:

$$g_{ij}=g_{ij}(\theta)=-E_\theta\left[\frac{\partial^2 \ln f(X;\theta)}{\partial\theta_i\partial\theta_j}\right] \quad ; \quad i,j=1,2,\ldots,r$$

مثال 4.7.4

إذا كانت $X=(X_1,X_2,\ldots,X_n)$ عينة عشوائية من التوزيع $N(\theta_1,\theta_2^2)$، فأوجد مصفوفة المعلومات.

هنا $\theta=(\theta_1,\theta_2)$ ودالة الكثافة:

$$f(x;\theta)=\frac{1}{\theta_2\sqrt{2\pi}}e^{-\frac{1}{2\theta_2^2}(x-\theta_1)^2} \quad , \qquad -\infty<x<\infty$$

وعلى ذلك فإن:

$$\ln f(x;\theta)=-\frac{1}{2}\ln 2\pi-\frac{1}{2}\ln\theta_2-\frac{1}{2\theta_2^2}(x-\theta_1)^2$$

$$\frac{\partial \ln f(x;\theta)}{\partial\theta_1}=\frac{(x-\theta_1)}{\theta_2^2} \quad , \qquad \frac{\partial^2 \ln f(x;\theta)}{\partial\theta_1^2}=-\frac{1}{\theta_2^2}$$

$$\frac{\partial \ln f(x;\theta)}{\partial\theta_2}=+\frac{1}{2\theta_2^2}-\frac{(x-\theta_1)^2}{2\theta_2^3}$$

$$\frac{\partial^2 \ln f(x;\theta)}{\partial\theta_2^2}=\frac{1}{\theta_2^2}-\frac{(x-\theta_1)^2}{\theta_2^4}$$

$$\frac{\partial^2 \ln f(x;\theta)}{\partial\theta_1\partial\theta_2}=\frac{-2(x-\theta_1)}{\theta_2^3}$$

$$g_{11}=-E_\theta\left(-\frac{1}{\theta_2^2}\right)=\frac{1}{\theta_2^2} \quad , \qquad g_{12}=g_{21}=-E_\theta\left[-\frac{2(X_1-\theta_1)}{\theta_2^3}\right]=0$$

$$g_{22}=-E_\theta\left[\frac{1}{\theta_2^2}-\frac{3(X_1-\theta_1)^2}{\theta_2^4}\right]=-\frac{1}{\theta_2^2}+\frac{3}{\theta_2^4}E(X_1-\theta_1)^2$$

$$=-\frac{1}{\theta_2^2}+\frac{3}{\theta_2^4}\theta_2^2=\frac{2}{\theta_2^2}$$

إذن مصفوفة المعلومات الموافقة لملاحظة واحدة:

$$I(\theta)=\begin{Vmatrix}1/\theta_2^2 & 0\\ 0 & 2/\theta_2^2\end{Vmatrix}$$

والموافقة للعينة $\boldsymbol{X}$:

$$I(\theta)=nI(\theta)=\begin{Vmatrix}n/\theta_2^2 & 0\\ 0 & 2n/\theta_2^2\end{Vmatrix}$$

تمـــارين

1. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع بواسون $\Pi(\theta)$، فهل $T(X)=\sum_{i=1}^{n-1} X_i$ إحصاء كافي لـ θ؟

2. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $B(1,\theta)$، فأثبت أن $T=\sum_{1}^{n} X_i$ إحصاء كافي لـ θ.

3. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع هندسي بمعلمة θ، فأثبت أن $T=\sum_{i=1}^{n} X_i$ إحصاء كافي لـ θ.

4. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $N(2,\theta^2)$، فأثبت أن $T(X)=\sum_{i=1}^{n}(X_i-2)^2$ إحصاء كاف للمعلمة θ^2.

5. إذا كانت $X=(X_1,X_2)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\theta x^{\theta-1} \quad , \quad 0<x<1$$

فبين الإحصاء الكافي من الإحصاءات الآتية:

$$X_1+X_2 \quad , \quad X_2-X_1 \quad , \quad X_1 \quad , \quad X_1X_2$$

6. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع

$$f(x;\theta)=e^{2\theta-x} \quad ; \quad x\ge 2\theta$$

فاثبت أن $T=X_{(1)}$ إحصاء كاف للمعلمة θ.

7. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع

$$f(x;\theta)=\frac{\theta}{x^2} \quad , \quad x>\theta$$

فهل $T=X_{(1)}$ إحصاء كاف للمعلمة θ؟

8. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع

$$f(x;\theta)=\frac{2}{\theta^2-1}\theta^x \quad ; \quad 0<x<1 \ , \ \theta>1$$

فأوجد إحصاءً كافياً لـ θ.

9. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $\mathcal{L}(\xi)\in\Gamma(\theta,\lambda)$، فأثبت أن $T=\sum_{i=1}^{n}X_i$ إحصاء تام للمعلمة θ.

10. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $N(\theta,4\theta^2)$، فأثبت أن $T=(\overline{X},S^2)$ إحصاء كافي للمعلمة θ لكن غير تام.

11. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع منتظم $R(\theta,2\theta)$، فأثبت أنه في هذه الحالة لا يوجد الإحصاء الكافي وحيد البعد للمعلمة θ، ثم بين أن الإحصاء $T=(X_{(1)},X_{(n)})$ كاف لكن غير تام للمعلمة θ.

12. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\theta^x(1-\theta)^{1-x} \quad , \quad x=0,1 \ , \ 0<\theta<1$$

فأوجد إحصاءً كافياً للمعلمة θ أولاً ثم أوجد دالة مساهمة العينة X.

13. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع

$$f(x;\theta_1,\theta_2)=\frac{1}{\theta_1}e^{\frac{\theta_2-x}{\theta_1}} \quad ; \quad x>\theta_2 \ , \ \theta_1>0$$

أثبت أن $T=(X_{(1)},n\overline{X})$ إحصاء كافي للمعلمة $\theta=(\theta_1,\theta_2)$.

14. ليكن المتغير العشوائي الملاحظ ع يفترض قيمه في الفئة $[a(\theta),b]$، حيث إن $a(\theta)$ دالة مضطردة في θ، أثبت أن القيمة الصغرى $X_{(1)}$ في العينة $X=(X_1,\ldots,X_n)$ تعتبر إحصاءً كافياً للمعلمة θ عندما وفقط عندما تكون لدالة الكثافة $f(x;\theta)$ الشكل:

$$f(x;\theta)=\frac{g(x)}{h(\theta)} \quad ; \quad a(\theta)\le x\le b$$

ثم أثبت أن هذه النتيجة تعتبر صحيحة من أجل الإحصاء $X_{(n)}$ وفئة قيم ع هي $[a,b(\theta)]$، حيث إن $b(\theta)$ دالة مضطردة في θ.

15. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $N(\theta,\sigma^2)$، فأثبت أن $T=\overline{X}$ إحصاء كاف تام للمعلمة θ، ثم أوجد دالة مساهمة العينة X.

16. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من التوزيع $R(\theta_1,\theta_2)$ فأثبت أن $T=(X_{(1)},X_{(2)})$ إحصاء تام للمعلمة $\theta=(\theta_1,\theta_2)$.

17. لتكن $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع

$$f(x;\theta)=\frac{e^{-\frac{(x-\theta_1)}{\theta_2}}}{\theta_2} \quad ; \quad x\ge\theta_1 \ , \ \theta=(\theta_1,\theta_2)$$

فأثبت أن $T=(X_{(1)},\overline{X})$ إحصاء كاف للمعلمة θ.

18. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع

$$f(x;\theta)=e^{-x+\theta}\left(1-e^{-x+\theta}\right)^{-2} \quad ; \quad -\infty<x<\infty$$

فاثبت أن دالة المعلومات $i(\theta)=\frac{1}{3}$، ثم أوجد $i_n(\theta)$.

19. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع

$$f(x;\theta)=\frac{\ln\theta}{\theta-1}\theta^x \quad ; \quad 0<x<1 \ , \ \theta>1$$

فـالمطلوب:

1. أوجد إحصاءً كافياً لـ θ.
2. أوجد دالة مساهمة العينـة $\boldsymbol{X}$.
3. احسب معلومات فيشر المتوفرة في العينة $\boldsymbol{X}$ حول θ.

20 إذا كانت $\boldsymbol{X}=\left(X_1,\ldots,X_n\right)$ عينة عشوائية مـن توزيـع

$$f(x;\theta)=2\theta e^{-2\theta x} \quad ; \quad x\geq 0$$

1. أوجد إحصاءً كافياً لـ θ.
2. أوجد دالة مساهمة العينـة $\boldsymbol{X}$.
3. احسب معلومات فيشر المتوفرة في العينة $\boldsymbol{X}$ وفي الإحصـاء الكـافي الـذي تحصلت عليـه في (1).

21 إذا كانت $\boldsymbol{X}=\left(X_1,\ldots,X_n\right)$ عينة عشوائية من توزيع $N\left(2\theta,\sigma^2\right)$ فالمطلوب:

1. أوجد إحصاءً كافياً للمعلمة θ.
2. أوجد دالة مساهمة العينـة $\boldsymbol{X}$.
3. احسب معلومات فيشر المتوفرة في العينـة $\boldsymbol{X}$ حـول θ.

22 إذا كـانت $\boldsymbol{X}=\left(X_1,\ldots,X_n\right)$ عينة عشـوائية من تـوزيع $\Gamma(\theta,2)$، وكـان $T=\sum_{i=1}^{n}X_i$ إحصاء تام لـ θ، فأوجد الإحصاء الكافي لـ $\tau(\theta)=2\theta+3$.

الفصل الخامس:

خواص مقدرات النقطة

1.5 مقدمة

يهتم علم الإحصاء بدراسة المجتمعات الإحصائية، وفي حالات عدة يقتصر الاهتمام على دراسة صفة مميزة لمفردات المجتمع الإحصائي، التي نعبر عنها بمتغير عشوائي حقيقي ξ، وبمعرفة التوزيع الاحتمالي لـ ξ يمكن الحصول على كافة المميزات العددية المختلفة (المعالم) لمجتمع الصفة المدروسة، وذلك باستخدام طرق نظرية الاحتمالات. وبعبارة أخرى، بمعرفة التوزيع $\mathcal{L}(\xi)$ نعلم كل شيء عن المجتمع ξ. لكن عادة في التطبيقات الإحصائية التوزيع $\mathcal{L}(\xi)$ غير معلوم، والهدف الرئيسي للمعاينة هو الحصول على المعلومات واستخلاص النتائج التي تعتبر الأساس لإعطاء استدلالات حول مجتمع المتغير العشوائي ξ، أي حول نوع توزيعه ومعالمه المختلفة.

لنفترض بناءً على معطيات العينة العشوائية $X = (X_1,...,X_n)$ المأخوذة من مجتمع الصفة المدروسة ξ، أو بناءً على التصورات النظرية لآلية تشكل الصفة ξ واستخدام مبرهنات النهاية المركزية، أمكن معرفة النموذج الإحصائي الذي يخضع له ξ (معرفة عائلة التوزيعات الاحتمالية التي ينتمي إليها توزيع ξ). وبعبارة أخرى، نفترض أن شكل دالة توزيع ξ معلومة وهي $F(x;\theta)$، لكنها تعتمد على معلمة غير معلومة θ. يصبح المطلوب هو تقدير القيمة الحقيقية لـ θ الموافقة للتوزيع الاحتمالي لـ ξ.

يمكن التمييز بين نوعين من التقدير: التقدير بنقطة (point estimation) والتقدير بفترة (interval estimation). في حالة التقدير بنقطة تستخدم المعلومات المتوفرة في العينة العشوائية الملاحظة $x = (x_1,...,x_n)$ للوصول إلى عدد واحد أو نقطة واحدة تكون تقديراً للمعلمة θ. أما في حالة التقدير بفترة، تستخدم المعلومات المتوفرة في العينة x للوصول إلى عددين أو نقطتين يفترض أن القيمة الحقيقية للمعلمة θ تقع بينهما باحتمال معين.

تعطى التقديرات النقطية عادة بدلالة دالة في مركبات العينة العشوائية X، التي تدعى بالمقدِّر وقيمتها عند عينة مشاهدة x بالتقدير، ولكل معلمة θ مقدِّرات مختلفة، وبالتالي لابد من معايير للمفاضلة بينها واختيار أفضلها من أجل تقدير المعلمة θ.

سنتعرض في هذا الفصل في إطار نموذج إحصائي للخواص الأساسية التي نرغب تحققها في المقدِّر الجيد للمعلمة θ، أو لدالة لها. وبشكل خاص سنبحث في المقدِّرات غير المتحيزة (unbiased estimators)، ونعطي الاهتمام الأكبر والأساسي للمقدِّرات الأكفأ (efficient estimators)، وذلك عندما تكون θ وحيدة البعد أو متعددة الأبعاد.

2.5 التقدير الإحصائي بنقطة

STATISTICAL POINT ESTIMATION

لقد سبق أن تعرضنا لهذا المفهوم في الفصل الثالث. فمثلاً، ذكرنا أن قيمة دالة التوزيع التجريبي في نقطة يمكن اعتبارها تقديراً لقيمة دالة التوزيع النظري في تلك النقطة، وأيضاً تعتبر مميزات عينة عشوائية (الوسط الحسابي، الوسيط، التباين، ...الخ) تقديرات للميزات النظرية الموافقة لها (ثوابت أو معالم المجتمع الذي سحبت منه العينة). ومصطلح "تقدير" (estimation) بني أساساً على أنه من أجل عينة كبيرة الحجم، احتمال أن يكون الفرق كبيراً بين قيم المميزات المختلفة للعينة وقيم

المميزات النظرية الموافقة لها صغير. ولهذا منطقياً (في المتوسط من أجل عينات كبيرة الحجم) اعتبار مميزات العينة المشاهدة كقيم تقريبية جيدة للمميزات النظرية الموافقة، عندما تكون هذه الأخيرة غير معلومة. هكذا، في هذا المصطلح (مصطلح تقدير) يتموضع المدلول المحدد للتقارب. وعند استخدام النظرية الإحصائية في التطبيق غالباً نحتاج لإيجاد القيم التقريبية للميزات النظرية المختلفة للتوزيع المقترح (معالم الصفة المدروسة ξ) عند أي حجم للعينة، ومنها المحدود.

فيما يلي نفترض لدينا نموذج إحصائي ما (عائلة التوزيعات من نوع معين):

$$\boldsymbol{F} = \{F(x;\theta)\}$$

الموافق لمنظومة تكرارات مستقلة من الملاحظات على متغير عشوائي ما ξ، كما نفترض شروط الانتظام على النموذج $\boldsymbol{F}(x;\theta)$ محققة، التي تضمن صحة النتائج والإثباتات الموافقة.

لتكن $\boldsymbol{X} = (X_1,\ldots,X_n)$ عينة من التوزيع $\mathcal{L}(\xi) \in \boldsymbol{F}$. هكذا المعلومة المسبقة عن المتغير العشوائي الملاحظ ξ تتمثل في أن دالة توزيعه $F(x;\theta)$ عضوة في عائلة دوال التوزيع المعلمية المعطاة $\boldsymbol{F}$، أي أن شكل (صيغة) دالة توزيع ξ معلوم (إلا أنه يعتمد على المعلمة θ). وبمعرفة قيمة المعلمة θ يتحدد تماماً توزيع ξ. وتصاغ بشكل عام مسألة تقدير المعلمة θ على النحو الآتي: باستخدام المعطيات الإحصائية التي توفرها العينة العشوائية الملاحظة x الوصول إلى تقدير للقيمة الحقيقية ولنرمز لها بـ θ_0 للمعلمة θ غير المعلومة.

ويمكن تلخيص الخطوات المتبعة لإيجاد التقدير النقطي (التقدير بنقطة) للمعلمة المجهولة θ بما يلي:

1. تعيين فئة القيم الممكنة للمعلمة غير المعلومة θ، وتسمى هذه الفئة عادة بفضاء العينة (parametric space) ونرمز له بـ Θ.
2. اختيار الإحصاء المناسب $T = T(\boldsymbol{X})$، الذي قيمته عند ملاحظة معينة

$x=(x_1,\ldots,x_2)$ للعينة العشوائية X تتخذ كقيمة تقريبية جيدة لـ θ_0 (القيمة الحقيقية لـ θ الموافقة لـ $\mathcal{L}(\xi)$)، وفي هذه الحالة يقال: أن الإحصاء $T=T(X)$ يقدّر θ أو الإحصاء T مقدّر لـ θ، وتدعى الدالة $T(X)$ بمقدّر (estimator) المعلمة θ، وقيمته المحسوبة من بيانات عينة مشاهدة $x=(x_1,\ldots,x_2)$ هي $t=T(x)$ تسمى تقدير (estimate) للمعلمة θ.

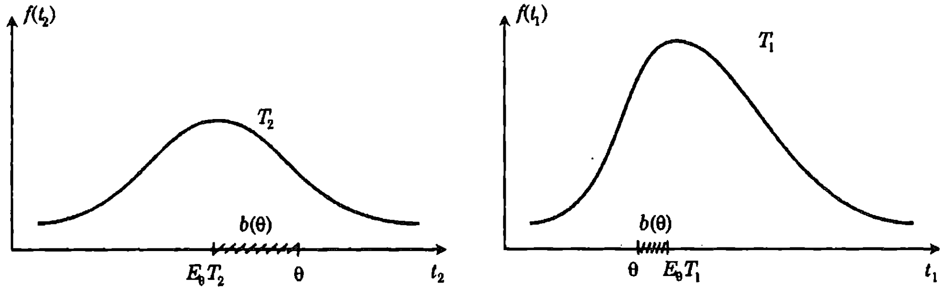
يوجد، بشكل عام، أكثر من إحصاء يصلح كمقدّر لمعلمة التوزيع θ، أي أن المقدّر T لمعلمة θ يمكن أن يأخذ أشكالاً مختلفة. فمثلاً، $T=\prod_{i=1}^{n}X_i$ ، $T=\sum_{i=1}^{n}X_i$ ، $T=\dfrac{X_{(n)}-X_{(1)}}{2}$ ، $T=X_{(n)}$ و... الخ. إن كل مقدّر يفترض فئة من القيم، وهذه الفئة عادة (ليس دائماً) تساوي فئة القيم الممكنة للمعلمة θ، أي فضاء المعلمة θ، وإذا رمزنا لفئة القيم الممكنة لـ T بـ $\Im$ فإن $\Im\subseteq\theta$.

والتساؤل الذي يطرح نفسه الآن: كيف نقارن المقدّرات المختلفة لمعلمة θ لاختيار أفضلها؟ فمثلاً، إذا كان $T_1=T_1(X)$ و $T_2=T_2(X)$ مقدّرين لمعلمة θ، فأيهما نفضل؟ ولماذا؟.

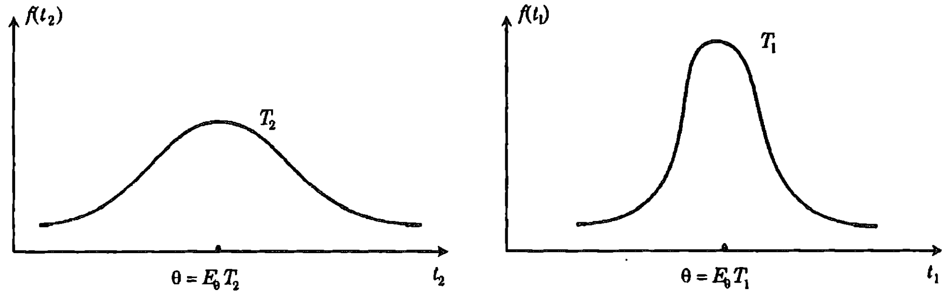
للإجابة على هذا التساؤل يلزمنا معيار للجودة، وذلك لمقارنة المقدّرات المختلفة للمعلمة θ، وهذا المعيار بدوره ليس بالضرورة وحيداً، بل يمكن استخدام معايير مختلفة مرتبطة بالأهداف التي يبنى المقدّر من أجلها، إلا أن أي معيار لجودة مقدّر يتحدد بالمقياس المتخذ لقياس قرب قيم المقدّر من القيمة الحقيقية θ_0 للمعلمة المقدَّرة. بالإضافة إلى ذلك، نقيد عائلة المقدّرات ببعض الشروط. وعلى ذلك، إذا تم تعيين عائلة المقدّرات لمعلمة θ في نموذج إحصائي معطى وتم أيضاً اختيار مقياس القرب، فإن المقدّر T الذي يوافق أصغر قيمة لذلك المقياس يعتبر الأفضل ضمن هذه العائلة من المقدّرات.

كما نعلم أن كل مقدّر $T=T(X)$ هو عبارة عن إحصاء، أي هو متغير عشوائي، وبالتالي له توزيع احتمالي نحصل عليه من عينات عشوائية متكررة مسحوبة من توزيع المتغير العشوائي الملاحظ ξ، ولحسن الحظ تتوفر طرق رياضية

لاشتقاق مثل هذا التوزيع (توزيع T). لذا، الشرط العام لبناء المقدِّرات النقطية $T(X)$ يتمثل في شرط التمركز (بهذا المفهوم أو ذاك) لتوزيع T حول القيمة الحقيقية للمعلمة المقدَّرة. وكلما كانت درجة التمركز هذه أكبر كلما كان المقدِّر أفضل. فمثلا، ليكن المقدِّرين $T_1(X)$ و $T_2(X)$ للمعلمة θ، وتوزيع المعاينة لكل منهما كما هو مبين على الشكلين الآتيين:



شكل (1.2.5)



شكل (2.2.5)

نلاحظ من الشكل (1.2.5) أن:

$$ET_1 \approx \theta \quad , \quad ET_2 \neq \theta$$

وأن معظم قيم T_1 أقرب إلى θ إذا ما قورنت بقيم T_2. ونلاحظ من الشكل (2.2.5) أن معظم قيم T_1 متمركزة حول θ على عكس قيم T_2 وعلى ذلك فإن المقدِّر T_1 أفضل من المقدِّر T_2 لتقدير المعلمة θ.

3.5 متوسط مربع الخطأ MEAN SQUARED ERROR

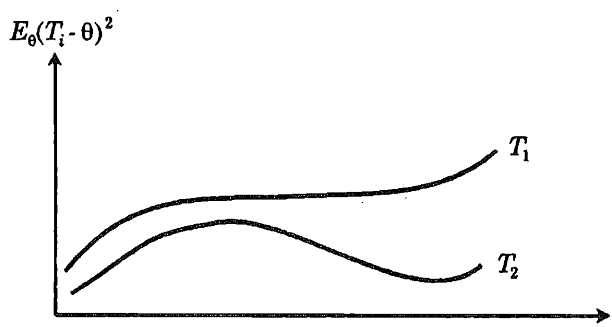
إذا كانت $X=(X_1,\ldots,X_n)$ عينـــة عشـــوائية مســـحوبة مـــن مجتمـــع $\mathcal{L}(\xi)$، صيغة دالة توزيعه $F(x;\theta)$ معلومة، لكنها تعتمد على معلمـــة وحيـــدة البعـــد θ غـــير معلومة، وكان $T=T(X)$ مقدِّراً لـــ θ، فـــإن المقـــدار:

$$E_\theta(T-\theta)^2$$

يدعى بمتوسط مربع الخطأ (mean squared error) للمقدِّر $T=T(X)$ للمعلمة θ أو بمتوسط الأخطاء المربعة للمقدِّر T ويرمز له بـــ $\mathrm{MSE}(T)$. وعلى ذلك فإننا نفضل المقدِّر الذي له أقل متوسط مربع خطأ. فمثلاً، إذا كان T_1,T_2 مقدِّرين للمعلمة θ، فإننا نفضـــل T_1 على T_2 إذا كان:

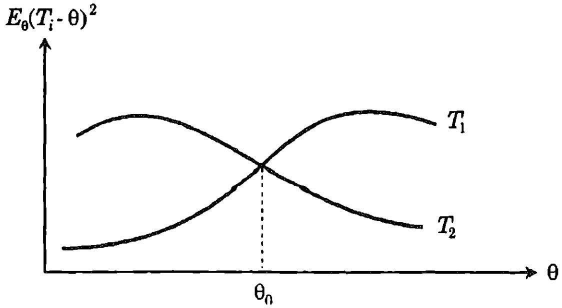
$$E_\theta(T_1-\theta)^2 < E_\theta(T_2-\theta)^2 \quad ; \quad \forall\, \theta \in \Theta$$

لكن نادراً ما نجد مقدِّراً له أقل متوسط مربع خطأ لجميع قيم $\theta\in\Theta$. فمثـــلاً، نادراً ما نجد مثل الحالة المبينة علـــى الشـــكل (1.3.5).



شكل (1.3.5)

بينما على الغالب تصادفنا حالات من تلك المبينـــة علـــى الشـــكل (2.3.5).



شكل (2.3.5)

أي أن T_1 أفضل من T_2 لبعض قيم θ $(\theta < \theta_0)$ وأن T_2 أفضل من T_1 لبعض القيم الأخرى لـ θ $(\theta > \theta_0)$.

بناءً على ما سبق يجب البحث عن معايير أخرى للمفاضلة بين المقدِّرات، وهنالك بعض الصفات التي نرغب توفرها في المقدِّر وهي:

1. الكفاية (sufficiency).
2. عدم التحيز (unbiasedness).
3. الاتساق (consistency).
4. الكفاءة (efficiency).

درسنا في الفصل الرابع الصفة الأولى، وسنتطرق في البنود التالية بشيء من التفصيل لبقية الصفات التي يفضل بشكل عام توفرها في المقدِّر الجيد.

4.5 المقدِّرات غير المتحيزة *UNBAISED ESTIMATORS*

إذا كان $T = T(X)$ مقدِّراً لمعلمة θ في نموذج مفروض، فإن متوسط مربع الخطأ:

$$\text{MES}(T) = E_\theta(T - \theta)^2$$

يمكن كتابته على النحو الآتي:

$$E_\theta(T-\theta)^2 = E_\theta\left[(T-E_\theta T)+(E_\theta T-\theta)\right]^2$$
$$= E_\theta(T-E_\theta T)^2 + (E_\theta T-\theta)^2 + 2(E_\theta T-\theta)E_\theta(T-E_\theta T)$$

أي أن:

$$\text{MES}(T) = V_\theta T + b^2(\theta) \qquad (1.4.5)$$

لأن $E_\theta(T-E_\theta T) = E_\theta T - E_\theta T = 0$ و $b(\theta) = (E_\theta T - \theta)$.

يسمى $b(\theta)$ بمقدار التحيز (biased) في الإحصاء T كمقدِّر للمعلمة θ، وهو يبين الفرق بين متوسط المقدِّر T $(E_\theta T)$ والمعلمة المقدَّرة θ [أنظر شكل (1.2.5)].

هكذا، نلاحظ أن متوسط مربع الخطأ لمقدِّر T للمعلمة θ عبارة عن تباين المقدِّر $V_\theta T$ مضافاً إليه مربع التحيز $b^2(\theta)$. وعلى ذلك، فإن متوسط مربع خطأ مقدِّر T للمعلمة θ يكون أصغر ما يمكن إذا كان كل من $V_\theta T$ و $b^2(\theta)$ أصغر ما يمكن، أي أن قيم المقدِّر أقرب ما يمكن من المعلمة θ.

بما أن $b^2(\theta)$ مقدار غير سالب وأصغر قيمة ممكنة له هي الصفر، ويحدث ذلك عندما $E_\theta T = \theta$ ، أي أن متوسط T ينطبق على المعلمة θ. وهذا يعني أنه لا يوجد تحيز في المقدِّر T (لا توجد أخطاء نظامية)، وعندئذٍ يقال أن T مقدِّر غير متحيز (unbiased estimator) للمعلمة θ. وفي هذه الحالة متوسط مربع خطأ المقدِّر T مساوٍ لتباينه، أي أن $E_\theta(T-\theta)^2 = V_\theta T$ [أنظر الشكل (2.2.5)].

تعريف 1.4.5: المقدِّر غير المتحيز Unbiased Estimator

يقال عن الإحصاء $T = T(X)$ أنه مقدِّر غير متحيز للمعلمة θ (وحيدة البعد)، إذا حقق الشرط:

$$E_\theta T = \theta \quad ; \quad \forall\, \theta \in \Theta \qquad (2.4.5)$$

وإذا لم يحقق الإحصاء $T = T(X)$ الشرط (2.4.5) فيدعى بالمقدِّر المتحيز (biased estimator) للمعلمة θ، أي أن التقدير يشتمل على أخطاء نظامية.

أحياناً لا يطلب تقدير المعلمة θ فقط، وإنما بشكل عام دالـة مـا في θ نرمـز لهـا بـ $\tau(\theta)$ وتدعى بدالة المعلمة. وفي هذه الحالة متوسط مربـع خطـأ المقـدّر T لدالـة المعلمـة $\tau(\theta)$:

$$E_\theta\left[T - \tau(\theta)\right]^2 = V_\theta T + b^2(\theta)$$

حيث $b(\theta) = \left[E_\theta T - \tau(\theta)\right]$. ويقـال أن الإحصـاء $T = T(X)$ مقـدّر غـير متحيز لـ $\tau(\theta)$، إذا تحققت العلاقـة:

$$E_\theta T = \tau(\theta) \quad ; \quad \forall\, \theta \in \Theta \qquad (3.4.5)$$

نلاحظ مما سبق أن المقدّر غير المتحيز يعطي إلى حـد مـا في المتوسـط تقديـرات مقبولة، لكن هذا لا يعني أن قيم المقدّر غير المتحـيز، بشـكل عـام، تسـاوي $\tau(\theta)$ أو قريبة منها، وهذا يتوقف على درجة تشتت T حـول $\tau(\theta)$، أي علـى تبـاين المقـدّر T [أنظر شكل (2.2.5)]. هذا بالإضافة، كمـا سـنبين لاحقـاً، مـن أجـل عائلـة المقدّرات غير المتحيزة يمكن بناء مبرهنة مفيدة تطبيقياً وسـهلة بكفايـة، فيـها مقيـاس تمركز (دقة) المقدّر عبارة عن تباينه. وهذا يعني أن نختار بـين المقـدّرات غـير المتحـيزة على أساس تبايناتها ونفضل المقدّر غير المتحيز بـأقل تبـاين.

مثـال 1.4.5

لتكن $X = (X_1, \dots, X_n)$ عينة عشوائية من مجتمع توزيعه $\mathcal{L}(\xi) \in N(\theta, \sigma^2)$. والمطلوب:

1. أثبت أن المقدّرات الآتية للمعلمة θ غـير متحـيزة:

$$T_1 = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i \quad , \quad T_2 = \frac{1}{2}\left(\overline{X} + X_1\right) \quad , \quad T_3 = \frac{1}{3}\left(\overline{X} + X_1 + X_2\right)$$

2. أي المقدّرات الثلاث $T_i \, , \, i = 1,2,3$ أفضل كمقدّر لـ θ.

لإثبات أن مقدّر ما للمعلمة θ غير متحيز، يجب إثبـات تحقـق العلاقـة (2.4.5).

بما أن $E_\theta \xi = \theta$، نجد:

$$E_\theta T_1 = E_\theta\left(\frac{1}{n-1}\sum_{i=1}^{n-1} X_i\right) = \frac{1}{n-1}\sum_{i=1}^{n-1} E_\theta X_i = \frac{n-1}{n-1}\theta = \theta$$

$$E_\theta T_2 = E_\theta\left[\frac{1}{2}\left(\overline{X}+X_1\right)\right] = \frac{1}{2}\left(E_\theta\overline{X}+E_\theta X_1\right) = \frac{1}{2}(\theta+\theta) = \theta$$

$$E_\theta T_3 = E_\theta\left[\frac{1}{3}\left(\overline{X}+X_1+X_2\right)\right] = \frac{1}{3}\left(E_\theta\overline{X}+E_\theta X_1+E_\theta X_2\right) = \frac{3\theta}{3} = \theta$$

إذن المقدّرات الثلاث غير متحيزة لمتوسط المجتمع θ.

بما أن المقدّرات الثلاث غير متحيزة، فإن المقدّر الأفضل هو الذي له أقل تباين، لذا نحسب التباينات $V_\theta T_i\ ;\ i = 1,2,3$:

$$V_\theta T_1 = V_\theta\left(\frac{1}{n-1}\sum_{i=1}^{n-1}X_i\right) = \frac{1}{(n-1)^2}\sum_{i=1}^{n-1}V_\theta X_i = \frac{n-1}{(n-1)^2}\sigma^2 = \frac{1}{n-1}\sigma^2$$

$$V_\theta T_2 = V_\theta\left[\frac{1}{2}\left(\overline{X}+X_1\right)\right] = \frac{1}{4}V_\theta\left(\frac{1}{n}\sum_{i=2}^{n}X_i + \frac{n+1}{n}X_1\right) =$$

$$= \frac{1}{4}\left(\frac{1}{n^2}\sum_{i=2}^{n}V_\theta X_i + \frac{(n+1)^2}{n^2}V_\theta X_1\right) = \frac{1}{4}\left(\frac{n-1}{n^2}\sigma^2 + \frac{(n+1)^2}{n^2}\sigma^2\right)$$

$$= \frac{n+3}{4n}\sigma^2$$

$$V_\theta T_3 = V_\theta\left[\frac{1}{3}\left(\overline{X}+X_1+X_2\right)\right] = \frac{1}{9}V_\theta\left(\overline{X}+X_1+X_2\right) =$$

$$= \frac{1}{9}V_\theta\left(\frac{1}{n}\sum_{i=3}^{n}X_i + \frac{n+1}{n}X_1 + \frac{n+1}{n}X_2\right) =$$

$$= \frac{1}{9}\left(\frac{1}{n^2}\sum_{i=3}^{n}V_\theta X_i + \frac{(n+1)^2}{n^2}V_\theta X_1 + \frac{(n+1)^2}{n^2}V_\theta X_2\right) =$$

$$= \frac{1}{9}\left(\frac{(n-2)}{n^2}\sigma^2 + \frac{(n+1)^2}{n^2}\sigma^2 + \frac{(n+1)^2}{n^2}\sigma^2\right) = \frac{(2n+5)}{9n}\sigma^2$$

بمقارنة التباينات الثلاث نجد أن تباين المقدّر T_1 هو الأصغر، وبالتالي فهو المقدّر الأفضل ضمن تلك المقدّرات.

مثـــال 2.4.5

إذا كـانت $X=(X_1,\ldots,X_n)$ عينـة عشـوائية مـن مجتمـع يتبـع التوزيـع الأسـي $\Gamma(1,\theta)$ فـالمطلوب:

1. أثبت أن $T_1=\frac{1}{2}\left(\overline{X}+X_1\right)$ مقـدّر غـير متحيز للمعلمـة θ واحسـب تباينـه.
2. أثبـت أن $T_2=\frac{n}{n+1}\overline{X}^2$ مقـدّر غـير متحـيز للدالـة $\tau(\theta)=\theta^2$.
3. أوجـد قيمـة c الـتي تجعـل $T_3=\frac{c}{\overline{X}}$ مقـدّراً غـير متحيز لــ $\tau(\theta)=\frac{1}{\theta}$.

كمـا نعلـم إذا كـان ξ متغـيراً عشـوائياً خاضعـاً للتوزيـع $\Gamma(1,\theta)$ فـإن كثافـة توزيعـه الاحتماليـة [العلاقـة (24.3.2)]:

$$f(x;\theta)=\frac{1}{\theta}e^{-\frac{x}{\theta}} \quad ; \quad x>0\ ,\ \theta>0$$

وأن:

$$E_\theta\xi=\theta \quad , \quad E_\theta\xi^2=2\theta^2 \quad , \quad V_\theta\xi=\theta^2$$

وعلى ذلـك:

$$E_\theta T_1=\frac{1}{2}E_\theta\left(\overline{X}+X_1\right)=\frac{1}{2}\left(E_\theta\overline{X}+E_\theta X_1\right)=\frac{1}{2}(\theta+\theta)=\theta$$

أي أن T_1 مقـدّر غير متحيز للمعلمـة θ.

$$V_\theta T_1=\frac{1}{4}V_\theta\left(\overline{X}+X_1\right)=\frac{1}{4}V_\theta\left[\frac{1}{n}\sum_{i=2}^{n}X_i+\frac{n+1}{n}X_1\right]$$

$$=\frac{1}{4}\left[\frac{1}{n^2}\sum_{i=2}^{n}V_\theta X_i+\frac{(n+1)^2}{n^2}V_\theta X_1\right]$$

$$=\frac{1}{4}\left[\frac{(n-1)}{n^2}\theta^2+\frac{(n+1)^2}{n^2}\theta^2\right]=\frac{n+3}{4n}\theta$$

$$E_\theta T_2=\frac{n}{n+1}E_\theta\overline{X}^2=\frac{n}{n+1}\left[V_\theta\overline{X}+\left(E\overline{X}\right)^2\right]=\frac{n}{n+1}\left(\frac{\theta^2}{n}+\theta^2\right)$$

$$=\frac{n}{n+1}\ \frac{n+1}{n}\theta^2=\theta^2$$

وهـذا يعـني أن T_2 مقـدِّر غـير متحـيز لــ $\tau(\theta)=\theta^2$.

نلاحــظ أن:

$$T_3=\frac{c}{\overline{\overline{X}}}=\frac{nc}{\sum X_i}=\frac{nc}{Y} \qquad ; \qquad Y=\sum_{i=1}^{n} X_i$$

وكمـا أشـرنا سـابقاً [مبرهنـة (4.3.2)]، فـإن المتغـير العشـوائي Y يخضـع للتوزيـع $\Gamma(n;\theta)$، أي أن:

$$h(y,\theta)=\frac{1}{\theta^n\Gamma(n)}y^{n-1}e^{-y/\theta} \qquad , \qquad y>0$$

ومنها نجـد:

$$E_\theta\left(\frac{1}{Y}\right)=\frac{1}{\theta^n\Gamma(n)}\int_0^\infty \frac{1}{y}y^{n-1}e^{-y/\theta}d\theta=\frac{1}{\theta^n\Gamma(n)}\int_0^\infty y^{n-2}e^{-y/\theta}d\theta$$

وبوضـع $z=\dfrac{y}{\theta}$ نحصـل علـى:

$$E_\theta\left(\frac{1}{Y}\right)=\frac{1}{\theta^n\Gamma(n)}\int_0^\infty (\theta z)^{n-2}e^{-z}d\theta$$

$$=\frac{1}{\theta^n\Gamma(n)}\theta^{n-1}\int_0^\infty z^{n-2}e^{-z}dz=\frac{1}{\theta^n\Gamma(n)}\theta^{n-1}\Gamma(n-1)=\frac{1}{n-1}\ \frac{1}{\theta}$$

وعلى ذلـك فـإن:

$$E_\theta\left(\frac{c}{\overline{\overline{X}}}\right)=E_\theta\left(\frac{nc}{Y}\right)=ncE_\theta\left(\frac{1}{Y}\right)=\frac{nc}{n-1}\ \frac{1}{\theta}$$

ولكــي يكــون $T_3=\dfrac{c}{\overline{\overline{X}}}$ مقــدِّراً غــير متحــيز لــ $\tau(\theta)=\dfrac{1}{\theta}$ ، يجــب أن يتحقــق شـرط عـدم التحــيز (3.4.5)، أي أن:

$$\frac{cn}{(n-1)\theta}=\frac{1}{\theta}$$

ومنها نجـد:

$$c=\frac{n-1}{n}$$

أي أن:

$$T_3 = \frac{n-1}{n\overline{X}} = \frac{n-1}{\sum_{i=1}^{n} X_i}$$

مثـــال 3.4.5

إذا كانت $X = (X_1,...,X_n)$ عينة عشوائية مـن مجتمـع يخضـع لتوزيـع بـيرنولي $B(1,\theta)$، فـالمطلوب:

1. بـين أي مـن المقـدِّرات الآتيـة لـ θ غـير متحـيز، ثـم أوجـد مقـدار التحـيز للمقـدِّرات المتحـيزة:

$$T_1 = 2\overline{X} - X_1 \quad , \quad T_2 = \overline{X} + X_1 \quad , \quad T_3 = \frac{X_1 + X_2}{2} \quad , \quad T_4 = \frac{\overline{X}}{2}$$

2. ما هو المقدِّر الأفضـل مـن بين المقـدِّرات غـير المتحيزة؟

كمـا نعلـم، إذا كـان ξ متغـير عشـوائي خـاضع لتوزيـع $B(1,\theta)$، فـإن:

$$f(x;\theta) = \theta^x (1-\theta)^{1-x} \quad ; \quad x = 1,2 \ , \ 0 < \theta < 1$$

$$E_\theta \xi = \theta \quad , \quad V_\theta \xi = \theta(1-\theta)$$

وعلى ذلـك:

$$E_\theta T_1 = E_\theta\left(2\overline{X} - X_1\right) = 2E_\theta \overline{X} - EX_1 = 2\theta - \theta = \theta$$

$$E_\theta T_2 = E_\theta\left(\overline{X} + X_1\right) = E_\theta \overline{X} + EX_1 = \theta + \theta = 2\theta \neq \theta \quad ; \quad b(\theta) = 2\theta - \theta = \theta$$

$$E_\theta T_3 = E_\theta\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{2} E_\theta X_1 + \frac{1}{2} E_\theta X_2 = \frac{1}{2}\theta + \frac{1}{2}\theta = \theta$$

$$E_\theta T_4 = \frac{1}{2} E\overline{X} = \frac{1}{2}\theta \neq \theta \quad ; \quad b(\theta) = E_\theta T_4 - \theta = \frac{1}{2}\theta - \theta = -\frac{1}{2}\theta$$

وهـذا يعـني أن كـل مـن المقدِّريـن T_1, T_3 غـير متحـيز، بينمـا كـل مـن T_2, T_4 متحـيز.

لنحسـب الآن التباينـات:

$$V_\theta T_1 = V_\theta\left(2\overline{X} - X_1\right) = V_\theta\left(\frac{2}{n}\sum_{i=2}^{n} X_i - \frac{n-2}{n} X_1\right)$$

$$= \frac{4}{n^2}\sum_{i=2}^{n} V_\theta X_i + \frac{(n-2)^2}{n^2} V_\theta X_1 = \frac{4(n-1)}{n^2}\left[\theta(1-\theta)\right] + \frac{(n-2)^2}{n^2}\left[\theta(1-\theta)\right]$$

$$= \frac{4(n-1)+(n-2)^2}{n^2}\left[\theta(1-\theta)\right] = \theta(1-\theta)$$

$$V_\theta T_2 = V_\theta\left(\overline{X} + X_1\right) = V_\theta\left(\frac{1}{n}\sum_{i=2}^{n} X_i + \frac{n+1}{n} X_1\right)$$

$$= \frac{n-1}{n^2}\left[\theta(1-\theta)\right] + \frac{(n+1)^2}{n^2}\left[\theta(1-\theta)\right] = \frac{n+3}{n}\left[\theta(1-\theta)\right]$$

$$V_\theta T_3 = V_\theta\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{4}\left(V_\theta X_1 + V_\theta X_2\right) = \frac{1}{2}\theta(1-\theta)$$

$$V_\theta T_4 = V_\theta\left(\frac{\overline{X}}{2}\right) = \frac{1}{4} V_\theta \overline{X} = \frac{1}{4} V_\theta\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

$$= \frac{1}{4n^2}\sum_{i=2}^{n} V_\theta X_i = \frac{n}{4n^2}\left[\theta(1-\theta)\right] = \frac{1}{4n}\left[\theta(1-\theta)\right]$$

بمقارنة تباين المقدِّرين غير المتحيزين T_1, T_3 نجد أن تباين T_3 أقل من تباين T_1 مهما تكن $\theta \in \Theta = (0,1)$، وبالتالي فالمقدِّر T_3 الأفضل لتقدير θ.

إذا كان T مقدِّراً غير متحيز لمعلمة θ، فإن قيمته عند عينة مشاهدة $x = \left(x_1, \ldots, x_n\right)$ تدعى بتقدير غير متحيز لـ θ.

كما أشرنا سابقاً، أن المقدِّر غير المتحيز T لـ $\tau(\theta)$ (كحالة خاصة $\tau(\theta) = \theta$) يعطي إلى حد ما في المتوسط تقديرات معقولة، لكن هذا لا يعني بالضرورة أن قيمة T عن عينة مشاهدة تساوي $\tau(\theta)$ أو قريبة منها، وهذا يتعلق بدرجة تشتت T حول $\tau(\theta)$، أي على تباينه. وبعبارة أخرى، على درجة تجانس قيم T. وبناءً على ذلك يجب عدم التمسك والمبالغة في قيمة صفة عدم التحيز (unbiasedness)، حيث يمكن أن يكون شرط عدم التحيز قاسياً في بعض

الحالات ويؤدي إلى نتائج غير مرغوبة. بالإضافة إلى ذلك يمكن أن تكون المقدِّرات غير المتحيزة لدالة معلمية $\tau(\theta)$ (ومنها $\tau(\theta)=\theta$) من أجل نموذج مفروض غير موجود إطلاقاً، وفي حالات أخرى، بالرغم من أنها موجودة، لكنها تطبيقياً غير مفيدة. لنوضح ذلك من خلال المثالين التاليين:

مثال 4.4.5

لتكن X عينة عشوائية حجمها $n=1$، مأخوذة من توزيع $\mathcal{L}(\xi)\in\Pi(\theta)$، ويطلب إيجاد المقدِّر غير المتحيز لـ $\tau(\theta)=\dfrac{1}{\theta}$.

إذا كان $T(X)$ مقدِّراً غير متحيز لـ $\tau(\theta)$، فإن الشرط (3.4.5) يكتب على النحو الآتي:

$$\sum_{x=0}^{\infty}T(x)\frac{\theta^{x}}{x!}e^{-\theta}=\frac{1}{\theta}\quad;\quad\forall\,\theta\in\Theta=(0,\infty)$$

أو

$$\sum_{x=0}^{\infty}T(x)\frac{\theta^{x+1}}{x!}=e^{\theta}=\sum_{r=0}^{\infty}\frac{\theta^{r}}{r!}\quad;\quad\forall\,\theta\in(0,\infty)$$

ومن الواضح أن الدالة $T(X)$ المحققة للشرط الأخير والمستقلة عن θ غير موجودة (شرط الاستقلال عن θ يكمن في تعريف المقدِّر لأنه إحصاء)، أي أن المقدِّر غير المتحيز لـ $\tau(\theta)$ في الحالة المفروضة غير موجود.

مثال 5.4.5

لتكن X عينة عشوائية بحجم $n=1$، مأخوذة من توزيع $\mathcal{L}(\xi)\in\overline{B}(1,\theta)$ ويطلب إيجاد المقدِّر غير المتحيز لـ $\tau(\theta)=\theta$.

بتطبيق العلاقة (2.4.5) نجد:

$$\sum_{x=0}^{\infty}T(x)\theta^{x}=\frac{\theta}{1-\theta}=\sum_{r=1}^{\infty}\theta^{r}\quad;\quad\forall\,\theta\in(0,1)$$

وبمساواة الأمثال (المعاملات) حسب درجة θ، نلاحظ أن المقدِّر غير المتحيز الوحيد لـ θ في هذه الحالة يتمثل في الإحصاء:

$$T(X)=\begin{cases}0 & ; \quad X=0\\ 1 & ; \quad X\geq 1\end{cases}$$

لكن قيمتي هذا الإحصاء لا ينتميان لفضاء العينة $\theta=(0,1)$ للنموذج المعطى. لذا مثل هذا المقدِّر عملياً غير مفيد.

يبين هذان المثالان، أنه ليس بالضرورة دائماً البحث فقط عن المقدِّرات غير المتحيزة. بل أحياناً وكما يبين المثال التالي، إن المقدِّر ذو التحيز الصغير $b(\theta)$ ومتوسط مربع الخطأ صغير (تحيز صغير وتباين صغير) أفضل من مقدِّر غير متحيز ذو تباين كبير لتقدير دالة $\tau(\theta)$.

مثال 6.4.5

لتكن $X=(X_1,...,X_n)$ عينة من توزيع $\mathcal{L}(\xi)\in N(\theta_1,\theta_2^2)$، حيث $n\geq 2$. ويطلب تقدير $\tau(\theta)=\tau(\theta_1,\theta_2^2)=\theta_2^2$; $\theta=(\theta_1,\theta_2^2)$.

نجد من العلاقة (1.7.3)، إذا كان $S^2=S^2(X)$ تباين العينة فإن $E_\theta(S^2)=\frac{n-1}{n}\theta_2^2\neq\theta_2^2$، أي أنه مقدِّر متحيز لـ $\tau(\theta)=\theta_2^2$. وعلى ذلك فإن المقدِّر غير المتحيز لـ $\tau(\theta)=\theta_2^2$ هو:

$$S^{*2}=\frac{n}{n-1}S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \tag{4.4.5}$$

وحسب المبرهنة (5.9.3)، نجد:

$$\mathcal{L}_\theta\left(\frac{n-1}{\sigma^2}S^{*2}\right)=\chi^2_{(n-1)}$$

وبالتالي حسب العلاقة (32.3.2):

$$E_\theta\left(\frac{n-1}{\theta_2^2}S^{*2}\right)=(n-1) \quad , \quad V_\theta\left(\frac{n-1}{\theta_2^2}S^{*2}\right)=2(n-1)$$

أو

$$E_\theta\left(S^{*2}\right)=\theta_2^2 \quad , \quad V_\theta\left(S^{*2}\right)=\frac{2}{n-1}\theta_2^4 \tag{5.4.5}$$

سنبين فيما بعد (فقرة 4.7.5)، حسب معيار أقل تباين، أن المقدّر S^{*2} يعتبر الأفضل ضمن كل المقدّرات غير المتحيزة للدالة $\tau(\theta)=\theta_2^2$ في معلمتي النموذج $N\left(\theta_1,\theta_2^2\right)$، أي أن تباينه المعطى بالعلاقة (5.4.5) أصغر من تباين أي مقدّر غير متحيز آخر لـ $\tau(\theta)=\theta_2^2$ وذلك مهما تكن $\theta\in\Theta$.

لنبحث الآن في عائلة المقدّرات من الشكل $T_\lambda=\lambda S^{*2}$، حيث λ عدد حقيقي. بما أن $E_\theta T_\lambda=\lambda E_\theta\left(S^{*2}\right)$، فيوجد في هذه العائلة مقدّر غير متحيز وحيد لـ $\tau(\theta)=\theta_2^2$ وهو S^{*2}، أي الموافق لـ $\lambda=1$.

لنحسب متوسط مربع الخطأ لمقدّر ما $T_\lambda\neq S^{*2}$، أي $\lambda\neq 1$:

$$\begin{aligned}E_\theta\left(T_\lambda-\theta_2^2\right)^2&=E_\theta\left[\lambda\left(S^{*2}-\theta_2^2\right)+(\lambda-1)\theta_2^2\right]^2\\&=\lambda^2V_\theta\left(S^{*2}\right)+(\lambda-1)^2\theta_2^4\end{aligned}$$

من ذلك، وحسب العلاقة (5.4.5) نجد:

$$E_\theta\left(T_\lambda-\theta_2^2\right)^2=\left[\frac{2\lambda^2}{n-1}+(\lambda-1)^2\right]\theta_2^4=\varphi(\lambda)\,\theta_2^4$$

تبلغ الدالة $\varphi(\lambda)$ قيمة صغرى عندما $\lambda^*=\dfrac{n-1}{n+1}$ وتساوي $\varphi(\lambda^*)=\dfrac{2}{n+1}$. وبأخذ العلاقة (5.4.5) بعين الاعتبار، نجد:

$$E_\theta\left(T_{\lambda^*}-\theta_2^2\right)^2=\frac{2}{n+1}\theta_2^4<\frac{2}{n-1}\theta_2^4=E_\theta\left(S^{*2}-\theta_2^2\right)^2$$

هكذا، وفق معيار متوسط مربع الخطأ، عندما $n \geq 2$، فــإن المقــدِّر المتحــيز

$$T_{\lambda^*} = \lambda^* S^{*2} = \frac{n-1}{n+1}\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{1}{n+1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

الذي انحرافه (تحيزه) $b(\theta) = \left(E_\theta T_{\lambda^*} - \theta_2^2\right) = -\dfrac{2}{n+1}\theta_2^2$ صغير عندما يكون حجم العينـة n كبيراً بشكل كــاف، يعتبر مقــدِّراً أفضل من المقدِّر غير المتحيز S^{*2} لـــــ $\tau(\theta) = \theta_2^2$ في المعلمة θ_2^2 للنموذج $N\left(\theta_1, \theta_2^2\right)$. وهذا يعني غالبية قيم T_{λ^*} أقرب إلى القيمة الحقيقيـــة لــ θ_2^2 من قيم S^{*2} المناظرة لها.

يبين هذا المثال، أنه لا يمكن اعتبار معيـــــار مقارنـــة المقــدِّرات وحيــداً، كمـــا لا يوجد مقدِّر وحيد لمعلمة مفروضة، يلائم كـــل الحـــالات.

5.5 المقــدِّرات المتســقة CONSISTENT ESTIMATORS

يبدو لنا بوضوح من تعريف العينة العشوائية أن المقدِّر المبنى على أساس عينـــة تحـــوي n قياساً (أو ملاحظة) مثلاً يجب أن يكون أفضل بشكل عام، من المقدِّر المبنى على أســـاس عينة عشوائية تحوي عدداً أقل من n. ولإيضاح الفكرة أكـــثر، لنفـــترض أن T_1 مقـــدِّر للمعلمة θ في مجتمع معطى مبنى على عينة حجمها $n = 1$، أي تحـــوي ملاحظة أو قياســاً واحداً للمتغير العشوائي الملاحظ ξ، وليكن T_2 مقدِّراً لنفس المعلمة مبنياً علــــى أســـاس عينة عشوائية حجمها $n = 2$، وبشكل عام لنفترض أن T_n هو مقدِّر لــ θ مبنى علـــى أساس عينة حجمها n، أي أن $T_1, T_2, \ldots, T_n, \ldots$ متتالية من المقدِّرات لــ θ نرمـــز لهـــا بـــ $\{T_n\}$. تعني الفكـــرة الســـابقة أن المتتالية $\{T_n\}$ تتقارب احتمالياً إلى θ (أو تنتـــهي احتمالياً إلى θ) عندما $n \to \infty$، ولهذا يهمنا معرفة سلوك المقـــدِّر T_n بازديـــاد حجـــم العينة.

تعريف 1.5.5: المقدِّر المتســـق

يقال عن مقدِّر T_n لدالة معلمية $\tau(\theta)$ أنــه متســق، إذا كــان مــن أجــل أي مقدار ε صغير موجــب:

$$\lim_{n\to\infty} P\left[\left|T_n - \tau(\theta)\right| < \varepsilon\right) = 1 \qquad (1.5.5)$$

وعندما $\tau(\theta) = \theta$ فالعلاقة (1.5.5) تكتب علــى النحــو الآتي:

$$\lim_{n\to\infty} P\left[\left|T_n - \theta\right| < \varepsilon\right) = 1$$

أو

$$\lim_{n\to\infty} P\left[\left|T_n - \theta\right| \geq \varepsilon\right) = 0$$

وهــذا يعني أن الإحصاء T_n يكون مقدِّراً متسقاً لــ $\tau(\theta)$ إذا اقترب مــن $\tau(\theta)$ بازدياد حجم العينة n، أي أن T_n يتقارب بالاحتمــال (coverges in probability) من $\tau(\theta)$.

مثـــال 1.5.5

إذا كــانت $X = (X_1,\dots,X_n)$ عينــة عشــوائية مــن مجتمــع يتبــع التوزيــع $\mathcal{L}(\xi) \in N(\theta,\sigma^2)$، فهل $T_n = \overline{X}$ مقدِّر متسق للمعلمة θ؟

كما نعلـم:

$$\mathcal{L}(X_i) \in N(\theta,\sigma^2) \quad ; \quad i = 1,\dots,n$$

$$\mathcal{L}(\overline{X}) \in N\left(\theta, \frac{\sigma^2}{n}\right)$$

ومـن ثم:

$$\mathcal{L}\left(\frac{\overline{X} - \theta}{\sigma/\sqrt{n}}\right) = N(0,1)$$

لنحســب الآن:

$$P_\theta\left(\left|\overline{X}-\theta\right|>\varepsilon\right)=1-P_\theta\left(\left|\overline{X}-\theta\right|\le\varepsilon\right)$$
$$=1-P_\theta\left(-\varepsilon\le\overline{X}-\theta\le\varepsilon\right)$$
$$=1-P_\theta\left(-\frac{\varepsilon\sqrt{n}}{\sigma}\le\frac{\overline{X}-\theta}{\sigma/\sqrt{n}}\le\frac{\varepsilon\sqrt{n}}{\sigma}\right)$$
$$=2-2\phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right)$$

حيث أن $\Phi(z)$ دالة التوزيع الطبيعي المعياري $N(0,1)$. وبأخذ نهاية الطرفين عندما $n\to\infty$ نجد:

$$\lim_{n\to\infty}P\left(\left|\overline{X}-\theta\right|>\varepsilon\right)=2-2\Phi(\infty)=2-2=0$$

وحسب التعريف (1.5.5) فإن $T_n=\overline{X}$ مقدّر متسق للمعلمة θ. وبما أن $E_\theta\overline{X}=\theta$، فإن $T_n=\overline{X}$ مقدّر غير متحيز لـ θ. وهذا يعني أن المقدّر المتسق يمكن أن يكون غير متحيز أيضاً.

إذا كان $T_n=T_n(X)$ مقدّراً متسقاً لـ $\tau(\theta)$، فأية قيمة ملاحظة له $t=T_n(x)$ تدعى بالتقدير المتسق (consistent estimate) لـ $\tau(\theta)$.

في حالات عدة تكون المبرهنة التالية مفيدة للتأكد من أن مقدّراً ما T_n متسق لـ $\tau(\theta)$، حيث أن طريقة التأكد من أن إحصاءً ما متسق بتطبيق العلاقة (1.5.5) تكون أحياناً عملية صعبة.

مبرهنة 1.5.5

إذا كان T_n مقدّراً لـ $\tau(\theta)$ وحقق الشرط:

$$\lim_{n\to\infty}E_\theta\left[T_n-\tau(\theta)\right]^2=0 \qquad (2.5.5)$$

فإن T_n مقدّر متسق لـ $\tau(\theta)$.

الإثبـــات

إن الشـرط المذكـور في المبرهنـــة كـــافٍ وليـــس ضروريـــاً، أي إذا كـــان المقـدّر T_n متسـقاً ليـس بـالضرورة محققـاً للشـرط (2.5.5)، لكـن العكـس صحيـح. والإثبـات شـبيه بإثبـات متباينـة تشيبيشـف. لنفـترض للتبسـيط أن $\tau(\theta) = \theta$ ، وعندئـذٍ:

$$E_\theta(T_n - \theta)^2 = \int_{-\infty}^{+\infty}(t-\theta)^2 f(t)dt =$$

$$= \int_{-\infty}^{\theta-\delta}(t-\theta)^2 f(t)dt + \int_{\theta-\delta}^{\theta+\delta}(t-\theta)^2 f(t)dt + \int_{\theta+\delta}^{+\infty}(t-\theta)^2 f(t)dt$$

ولدينا في التكاملين الأول والثالث $(t-\theta)^2 \geq \delta^2$ والتكامل الثاني غير سالب، إذن:

$$E_\theta(T_n - \theta)^2 \geq \int_{-\infty}^{\theta-\delta}(t-\theta)^2 f(t)dt + \int_{\theta+\delta}^{+\infty}(t-\theta)^2 f(t)dt \geq$$

$$\geq \delta^2\left[P_\theta(T_n \leq \theta - \delta) + P_\theta(T_n \geq \theta - \delta)\right] = \delta^2 P_\theta\left(|T_n - \theta| \geq \delta\right)$$

وعلى ذلـك نجـد:

$$P_\theta\left(|T_n - \theta| \geq \delta\right) \leq \frac{E(T_n - \theta)^2}{\delta^2}$$

بما أن الطـرف الأيمـن في هـذه المتباينـة ينتهـي، حســب الفـرض، إلى الصفـر عندمـا $n \to \infty$ ، بالإضافـة إلى أن الاحتمـال لا يمكـن أن يكـون ســالباً، فـإن:

$$\lim_{n\to\infty} P_\theta\left(|T_n - \theta| \geq \delta\right) = 0$$

أي أن T_n مقـدّر متسـق لـــ θ.

بما أن $E_\theta(T_n - \theta)^2 = V_\theta T_n + b^2(\theta)$ ، فـإن تحقـــق الشـــرط (2.5.5) يكـــافئ تحقق الشـرطين:

$$\lim_{n\to\infty} V_\theta T_n = 0$$

$$\lim_{n\to\infty} b^2(\theta) = 0 \;\Leftrightarrow\; \lim_{n\to\infty} E_\theta T_n = \theta \qquad (3.5.5)$$

وهذا يعني أن كل مقدّر متسق غير متحيز تقاربياً.

إذا كان T_n مقدّر غير متحيز لـ θ فإن الشرط الثاني في (3.5.5) محقق، وبالتالي يكفي التأكد من أن تباينه ينتهي إلى الصفر عندما $n \to \infty$ لكي يكون مقدّراً متسقاً. أي نقول عن مقدّر غير متحيز T_n لـ θ أنه متسق أيضاً إذا كان:

$$\lim_{n\to\infty} V_\theta T_n = 0$$

مثال 2.5.5

إذا كان X متغيراً عشوائياً يتبع توزيع ذي الحدين $B(n,\theta)$ ، فإن $T_n = \dfrac{X}{n}$ مقدّر متسق للمعلمة θ ، حيث X عدد النجاحات في n تكرار بيرنولي.

بما أن $\mathcal{L}(X) \in B(n,\theta)$ ، فإن:

$$f(x;\theta) = C_n^x \theta^x (1-\theta)^{n-x} \quad , \quad x = 0,1,\dots,n$$

وعلى ذلك نجد:

$$E_\theta T_n = \frac{1}{n} E_\theta X = \frac{1}{n} n\theta = \theta$$

وبالتالي T_n مقدّر غير متحيز للمعلمة θ ، وكذلك فإن:

$$V_\theta T_n = \frac{1}{n^2} V_\theta X = \frac{1}{n^2} n\theta(1-\theta) = \frac{\theta(1-\theta)}{n}$$

وعلى ذلك نجد:

$$\lim_{n\to\infty} V_\theta T_n = 0$$

وبما أن المقدّر T_n غير متحيز وتباينه ينتهي إلى الصفر عندما $n \to \infty$ ، فحسب المبرهنة (1.5.5) يكون T_n مقدّراً متسقاً للمعلمة θ .

مثال 3.5.5

إذا كانت $X = (X_1,\dots,X_n)$ عينة عشوائية مسحوبة من مجتمع يتبع توزيع

كوشي بالمعلمة θ، أي أن:

$$f(x;\theta)=\frac{1}{\pi\left[1+(x-\theta)^2\right]} \quad , \quad -\infty<x<+\infty$$

فإن متوسط العينة $T_n=\overline{X}$ يعتبر مقدّراً غير متسق للمعلمة θ.

كما نعلم أن توزيع متوسط العينة يخضع بدوره لتوزيع كوشي، أي أن:

$$f(\overline{x};\theta)=\frac{1}{\pi\left[1+(\overline{x}-\theta)^2\right]} \quad , \quad -\infty<\overline{x}<+\infty$$

ولكي يكون $T_n=\overline{X}$ مقدّراً متسقاً لـ θ يجب أن يكون:

$$\lim_{n\to\infty} P_\theta\left(\left|\overline{X}-\theta\right|\le\varepsilon\right)=1$$

وبوضع $Y=\overline{X}-\theta$ نجد:

$$P_\theta\left(\left|\overline{X}-\theta\right|\le\varepsilon\right)=P_\theta\left(|Y|\le\varepsilon\right)$$

$$=\frac{1}{\pi}\int_{-\varepsilon}^{+\varepsilon}\frac{dy}{1+y^2}=\frac{2}{\pi}\tan^{-1}y\Big|_0^{\varepsilon}$$

بما أن هذا المقدار مستقل عن حجم العينة، أي أن $P_\theta\left(\left|\overline{X}-\theta\right|\le\varepsilon\right)$ لا يتأثر بزيادة قيمة n، فإن $T_n=\overline{X}$ مقدّر غير متسق لـ θ.

بشكل عام، فإن متوسط العينة $T_n=\overline{X}$ المأخوذة من مجتمع متوسطه θ وتباينه منتهي يعتبر مقدّراً متسقاً لـ θ [راجع مبرهنة تشيبيشف (1.3.1)].

مثال 4.5.5

إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من مجتمع يتبع توزيع منتظم $R(0,\theta)$ و $T_n=X_{(n)}$، فالمطلوب:

1. هل T_n مقدّر متسق للمعلمة θ؟
2. هل T_n مقدّر غير متحيز للمعلمة θ؟

بتطبيق العلاقة (3.10.3) نجد أن توزيع المعاينة للإحصاء $Y = X_{(n)}$ هو:

$$h(y;\theta) = \frac{ny^{n-1}}{\theta^n} \quad ; \quad 0 < y < \theta$$

بما أن:

$$\begin{aligned} P_\theta\left(|Y-\theta| > \varepsilon\right) &= P_\theta\left[(Y-\theta) < -\varepsilon\right] + P_\theta\left[(Y-\theta) > \varepsilon\right] \\ &= P_\theta(Y < \theta - \varepsilon) + P_\theta(Y > \theta + \varepsilon) \end{aligned}$$

و $P_\theta(Y > \theta + \varepsilon) = 0$ لأن Y تأخذ قيمها في الفترة $(0,\theta)$، فإن:

$$P_\theta\left(|Y-\theta| > \varepsilon\right) = P_\theta(Y < \theta - \varepsilon) = \int_0^{\theta-\varepsilon} \frac{ny^{n-1}}{\theta^n} dy = \left(\frac{\theta-\varepsilon}{\theta}\right)^n$$

بأخذ نهاية الطرفين عندما $n \to \infty$ نجد:

$$\lim_{n\to\infty} P_\theta\left(|Y-\theta| > \varepsilon\right) = \lim_{n\to\infty}\left(\frac{\theta-\varepsilon}{\theta}\right)^n = 0$$

أي أن $Y = X_{(n)}$ مقدّر متسق للمعلمة θ. وحيث أن:

$$E_\theta Y = \frac{n\theta}{n+1} \neq \theta$$

فهذا يعني أن المقدّر $Y = X_{(n)}$ متسق لكن متحيز، وبالتالي فالمقدّر المتسق ليس بالضرورة غير متحيز.

مثال 5.5.5

إذا كانت $X = (X_1,\ldots,X_n)$ عينة عشوائية من مجتمع توزيعه $\mathcal{L}(\xi) \in \Gamma(1,\theta)$، فإن $T_n = X_{(1)}$ مقدّر غير متحيز للمعلمة θ، لأن:

$$EX_{(1)} = \frac{1}{\theta}\int_0^\infty e^{-x/\theta}dx = \theta\int_0^\infty e^{-x/\theta}dx = \theta$$

إلا أنه غير متسق لأنه لايعتمد على حجم العينة n، أي لا يتأثر بزيادة n. وهذا يعني أن المقدّر غير المتحيز ليس بالضرورة متسقاً.

6.5 المقدِّر الأكفأ THE EFFICIENT ESTIMATOR

لا يجوز بناءً على الملاحظات السابقة اعتبار صفة عدم التحيز (unbiasedness) شرطٌ لابد من تحققه في أي مقدِّر، ومع ذلك ففي كثير من الحالات التي نصادفها في التطبيق يكون هذا الشرط (صفة عدم التحيز) منطقياً ومبرراً، ولهذا سيتم البحث لاحقاً وبشكل خاص في المقدِّرات غير المتحيزة.

ليكن المطلوب تقدير دالة معلمية معطاة $\tau(\theta)$ للمعلمة θ في عائلة التوزيعات $\boldsymbol{F} = \{F(x;\theta) \ ; \ \theta \in \Theta\}$ بناءً على معطيات العينة العشوائية $\boldsymbol{X} = (X_1,\dots,X_n)$ المسحوبة من مجتمع توزيعه ينتمي لهذه العائلة، ولنفترض وجود مقدِّرات غير متحيزة لـ $\tau(\theta)$ في المسألة المطروحة، أي وجود إحصاءات $T = T(\boldsymbol{X})$ محققة للشرط (3.4.5). لنرمز لعائلة المقدِّرات غير المتحيزة لـ $\tau(\theta)$ بـ U_τ، أي أن $T \in U_\tau$ عندما المقدِّر T يحقق الشرط (3.4.5). بالإضافة لذلك نفترض أن تباينات المقدِّرات $T \in U_\tau$ محدودة، أي أن:

$$V_\theta T = E_\theta[T - \tau(\theta)]^2 < \infty \quad ; \quad \forall \ \theta \in \Theta$$

يمكن في هذه الحالة قياس جودة المقدِّرات $T \in U_\tau$ بمقدار تبايناتها، ونحصل بذلك على مقياس بسيط لمقارنة المقدِّرات غير المتحيزة.

إذا كان T، T^* مقدِّرين من العائلة U_τ، وكان:

$$V_\theta T^* \le V_\theta T \quad ; \quad \forall \ \theta \in \Theta \qquad (1.6.5)$$

فحسب مقياس أقل تباين بانتظام (وفق المعلمة θ) للمقدِّر T^*، يعتبر هذا الأخير مقدِّراً ليس أسوأ من المقدِّر T. وإذا تحققت المساواة الصارمة في العلاقة (1.6.5) من أجل قيمة واحدة على الأقل لـ θ، فيتطلب ذلك إعطاء المقدِّر T^* الأفضلية كمقدِّر أكثر دقة لتقدير $\tau(\theta)$ بالمقارنة مع المقدِّر T. ويقال عندئذ أن T^* أفضل من T (الأكفأ نسبيا relatively efficient) في تقدير $\tau(\theta)$.

تعريف 1.6.5 المقدِّر الأكفأ نسبيا Relatively Efficient Estimator

إذا كان كل من T و T^* مقدِّر غير متحيز لـ $\tau(\theta)$، وكانت المتباينة (1.6.5) محققة مهما تكن $\theta \in \Theta$، بالإضافة إلى تحقق المتباينة الصارمة من أجل قيمة واحدة على الأقل لـ θ، فيقال أن T^* الأكفأ نسبياً من T لتقدير $\tau(\theta)$.

وتقاس الكفاءة النسبية للمقدِّر T^* مقارنة بـ T بالنسبة بين تباينهما

$$e = \frac{V_\theta T^*}{V_\theta T} \le 1 \quad ; \quad \forall\, \theta \in \Theta \tag{2.6.5}$$

مثال 1.6.5

إذا كانت $X = (X_1, X_2, X_3, X_4)$ عينة عشوائية من توزيع $\mathcal{L}(\xi) \in \Gamma(1,\theta)$، وكان:

$$T_1 = X_1 \quad , \quad T_2 = \frac{(X_1 + X_2)}{2} \quad , \quad T_3 = \frac{(X_1 + X_2 + X_3)}{3} \quad , \quad T_4 = \overline{X}$$

فالمطلوب:

1. ما هو المقدِّر الأكفأ بين المقدِّرات $T_i\ ,\ i = 1,2,3,4$ لـ θ؟
2. احسب الكفاءة النسبية للمقدِّر الأكفأ مقارنة ببقية المقدِّرات الأخرى.

بما أن $\mathcal{L}(\xi) \in \Gamma(1,\theta)$، فإن:

$$E_\theta \xi = \theta \quad , \quad V_\theta \xi = \theta^2$$

وعلى ذلك:

$$E_\theta X_i = \theta \quad , \quad V_\theta X_i = \theta^2 \quad , \quad i = 1,2,3,4$$

ومن ثم:

$$E_\theta T_1 = E_\theta X_1 = \theta \quad , \quad E_\theta T_2 = \frac{1}{2}\left(E_\theta X_1 + E_\theta X_2\right) = \theta$$

$$E_\theta T_3 = \frac{1}{3}\left(E_\theta X_1 + E_\theta X_2 + E_\theta X_3\right) = \theta \quad , \quad E_\theta T_4 = E_\theta \overline{X} = \theta$$

1. ما هو المقدِّر الأكفأ بين المقدِّرات $T_i \;,\; i = 1,2,3,4$ لــ θ؟
2. احسب الكفاءة النسبية للمقدِّر الأكفأ مقارنة ببقية المقدِّرات الأخرى.

بما أن $\mathcal{L}(\xi) \in \Gamma(1,\theta)$ ، فإن:

$$E_\theta \xi = \theta \quad , \quad V_\theta \xi = \theta^2$$

وعلى ذلك:

$$E_\theta X_i = \theta \quad , \quad V_\theta X_i = \theta^2 \quad , \quad i = 1,2,3,4$$

ومن ثم:

$$E_\theta T_1 = E_\theta X_1 = \theta \quad , \quad E_\theta T_2 = \frac{1}{2}\left(E_\theta X_1 + E_\theta X_2\right) = \theta$$

$$E_\theta T_3 = \frac{1}{3}\left(E_\theta X_1 + E_\theta X_2 + E_\theta X_3\right) = \theta \quad , \quad E_\theta T_4 = E_\theta \overline{X} = \theta$$

وعلى ذلك فإن الإحصاءات $T_i \;,\; i = \overline{1,4}$ مقــدِّرات غير متحيزة لــ θ. وبالتـــالي فالمقدِّر الأكفأ لتقدير $\tau(\theta) = \theta$ ضمن تلك المقدِّرات هو المقدِّر الذي لـــه أصغـــر تبـــاين (smallest variance)، لذا يجب حساب التباينات:

$$V_\theta T_1 = \theta^2 \qquad V_\theta T_2 = \frac{\theta^2}{2} \quad , \quad V_\theta T_3 = \frac{\theta^2}{3} \quad , \quad V_\theta T_4 = \frac{\theta^2}{4}$$

نلاحــظ أن

$$V_\theta T_4 < V_\theta T_3 < V_\theta T_2 < V_\theta T_1$$

وهذا يعني أن T_4 هو المقدِّر الأكفأ ضمن هذه المجموعة من المقدِّرات.

تقاس الكفاءة النسبية للمقدِّر T_4 بالمقارنة مع كل من المقدِّرات T_1, T_2, T_3 كالآتي:

$$\frac{V_\theta T_4}{V_\theta T_1} = \frac{\theta^2/4}{\theta^2} = 0.25 \quad , \quad \frac{V_\theta T_4}{V_\theta T_2} = \frac{\theta^2/4}{\theta^2/2} = 0.5 \quad , \quad \frac{V_\theta T_4}{V_\theta T_3} = \frac{\theta^2/4}{\theta^2/3} = \frac{3}{4} = 0.75$$

وربما توجد مقدِّرات غير متحيزة أخرى لــ θ أفضل (أكفأ) من المقدِّر T_4.

هكذا، يعتبر المقدّر $T^* \in U_\tau$ الأكفأ لتقدير $\tau(\theta)$ إذا حقق الشرط:

$$V_\theta T^* = \inf_{T \in U_\tau} V_\theta T \quad ; \quad \forall\, \theta \in \Theta \tag{3.6.5}$$

نلاحظ أن المقدّر الأكفأ T^* ليس بالضرورة موجوداً دائماً ضمن عائلة المقدّرات غير المتحيزة U_τ لـ $\tau(\theta)$. حيث يمكن وجود مقدّرين $T_1, T_2 \in U_\tau$ بحيث أن T_1 يوافق أصغر تباين ضمن عائلة المقدّرات غير المتحيزة من أجل قيمة (أو قيم) معينة لـ $\theta \in \Theta$، بينما T_2 يوافق أصغر تباين ضمن عائلة المقدّرات U_τ من أجل قيمة (أو قيم) أخرى لـ $\theta \in \Theta$. في مثل هذه الحالات، مقارنة المقدّرات باستخدام مقياس واحد (التباين الأصغر) غير ممكنة. إلا أن الشرط (3.6.5) يبرز وبشكل وحيد المقدّر الأكفأ لـ $\tau(\theta)$ ضمن عائلة المقدّرات غير المتحيزة U_τ، إذا وجد مثل هذا المقدّر. وهذا ما تؤكده المبرهنة التالية.

مبرهنة 1.6.5

إذا وجد المقدّر الأكفأ ضمن عائلة المقدّرات غير المتحيزة U_τ لـ $\tau(\theta)$، فإنه يكون وحيداً (unique).

الإثبات

تعني المبرهنة أنه إذا كان كل من T_1 و T_2 مقدّر غير متحيز بأقل تباين بانتظام في θ لـ $\tau(\theta)$ (مقدّر أكفأ لـ $\tau(\theta)$)، فإن $T_1 = T_2$، أي أن:

$$P_\theta\big(X \in \{x :\ T_1(x) \neq T_2(x)\}\big) = 0 \quad ; \quad \forall\, \theta \in \Theta$$

لنفترض أن لدينا مقدّراً جديداً هو:

$$T_3 = \frac{T_1 + T_2}{2}$$

من الواضح أن:

$$E_\theta T_3 = \frac{1}{2}\big(E_\theta T_1 + E_\theta T_2\big) = \tau(\theta) \quad ; \quad \forall\, \theta \in \Theta$$

لأنـه حسـب الفـرض T_1, T_2 مقدّريـن غـير متحـيزين لــ $\tau(\theta)$ ، أي أن:

$$E_\theta T_1 = E_\theta T_2 = \tau(\theta) \quad ; \quad \forall\, \theta \in \Theta$$

وبالتالي فـإن T_3 مقـدّر غـير متحـيز لـ $\tau(\theta)$.

كمـا أن:

$$V_\theta T_3 = \frac{1}{4} V_\theta(T_1 + T_2) = \frac{1}{4}\left[V_\theta T_1 + V_\theta T_2 + 2\underset{\theta}{\mathrm{cov}}(T_1, T_2)\right]$$

$$= \frac{1}{4}\left[V_\theta T_1 + V_\theta T_2 + 2\rho\sqrt{V_\theta T_1 \times V_\theta T_2}\right] \quad ; \quad \forall\, \theta \in \Theta$$

بفرض T_1, T_2 غـير مستقلين. وحسـب الفــرض:

$$V_\theta T_1 = V_\theta T_2 \quad ; \quad \forall\, \theta \in \Theta$$

إذن:

$$V_\theta T_3 = \frac{1}{2}(1+\rho) V_\theta T_1 \quad ; \quad \forall\, \theta \in \Theta$$

وحيـث أننـا نفـترض أن T_1 مقـدّر أكفـأ لــ $\tau(\theta)$ ، فـإن:

$$V_\theta T_3 \geq V_\theta T_1 \quad ; \quad \forall\, \theta \in \Theta$$

إذن

$$\frac{1}{2}(1+\rho) V_\theta T_1 \geq V_\theta T_1 \quad ; \quad \forall\, \theta \in \Theta$$

أي أن

$$\frac{1+\rho}{2} \geq 1 \quad \Rightarrow \quad \rho \geq 1 \quad \Rightarrow \quad \rho = 1$$

لأنـه كمـا نعلـم معـامل الارتبـاط $|\rho| \leq 1$. وهـذا يعــني وجــود علاقــة خطيــة بــين المقدّريـن T_1, T_2 ، أي يمكــن أن نكتــب:

$$T_2 = a + bT_1 \quad ; \quad b > 0 \qquad .....(1)$$

وعلى ذلك فــإن:

$$E_\theta T_2 = a + bE_\theta T_1 \;\Rightarrow\; \tau(\theta) = a + b\tau(\theta) \;\Rightarrow\; a = \tau(\theta)(1-b) \quad .....(2)$$

و

$$V_\theta T_2 = b^2 V_\theta T_1 \;\Rightarrow\; b^2 = 1 \;\Rightarrow\; b = 1$$

وبوضع $b = 1$ في العلاقة (2) نجد $a = 0$، وبالتالي فــإن $T_1 = T_2$ وهــو المطلــوب.

مثـــال 2.6.5

إذا كــانت $X = (X_1,...,X_n)$ عينــة عشــوائية مــن مجتمــع يتبــع التوزيــع $\mathcal{L}(\xi) \in B(1,\theta)$، فأثبت أن $T^* = T^*(X) = \overline{X}$ المقدِّر الأكفأ للمعلمة θ.

بما أن:

$$f(x;\theta) = \theta^x (1-\theta)^{1-x} \quad , \quad x = 0,1$$

فإن $V_\theta \xi = \theta(1-\theta) \;,\; E_\theta \xi = \theta$، ومن ثم $E_\theta T^* = E_\theta \overline{X} = \frac{1}{n}\sum_{i=1}^{n} E_\theta X_i = \theta$، وهــذا يعني أن $T^* = \overline{X}$ مقدِّر غير متحيز للمعلمة θ.

وتبــلين T^*

$$V_\theta T^* = V_\theta \overline{X} = \frac{1}{n^2}\sum_{i=1}^{n} V_\theta X_i = \frac{\theta(1-\theta)}{n}$$

لنفترض الآن أن T مقدِّر ما آخر غير متحيز لــ θ، ولنثبت تحقق العلاقــة (1.6.5)، أي تحقق المتباينة:

$$V_\theta T \ge \frac{\theta(1-\theta)}{n} \quad ; \quad \forall \theta \in \Theta = (0,1) \qquad ....(1)$$

بما أن توزيع المتغير العشوائي الملاحظ ξ يخضــع لتوزيــع بــيرنولي، فــإن توزيــع العينة العشوائية $X = (X_1,...,X_n)$ يعطــى بالاحتمــالات:

$$L(x;\theta)=\prod_{i=1}^{n}f\left(x_i;\theta\right)=\theta^{n\bar{x}}(1-\theta)^{n(1-\bar{x})} \qquad ....(2)$$

وكما نعلـم:

$$1\equiv\sum_x L(x;\theta) \quad , \quad \theta\equiv E_\theta T=\sum_x tL(x;\theta)$$

فباشتقاق هاتين المتطابقتين بالنسبة لـ θ، نجد:

$$0\equiv\sum_x\frac{\partial L(x;\theta)}{\partial\theta}=\sum_x\frac{\partial\ln L(x;\theta)}{\partial\theta}L(x;\theta)=E_\theta\left(\frac{\partial\ln L(X;\theta)}{\partial\theta}\right)$$

و

$$1\equiv\sum_x t\frac{\partial L(x;\theta)}{\partial\theta}=\sum_x t\frac{\partial\ln L(x;\theta)}{\partial\theta}L(x;\theta)=E_\theta\left(T(X)\ \frac{\partial\ln L(X;\theta)}{\partial\theta}\right)$$

لأن

$$\frac{\partial\ln L(x;\theta)}{\partial\theta}=\frac{1}{L}\frac{\partial L(x;\theta)}{\partial\theta}$$

وعلى ذلك يمكن كتابة:

$$\begin{aligned}1&\equiv E_\theta\left[T(X)\ \frac{\partial\ln L(X;\theta)}{\partial\theta}\right]-E_\theta\left[\ \frac{\partial\ln L(X;\theta)}{\partial\theta}\right]\\&=E_\theta\left\{\left[T(X)-\theta\ \right]\frac{\partial\ln L(X;\theta)}{\partial\theta}\right\}=\operatorname{cov}_\theta\left[T(X),\frac{\partial\ln L(X;\theta)}{\partial\theta}\right]\\&=\rho\ \left[V_\theta T\times V_\theta\left(\frac{\partial\ln L(X;\theta)}{\partial\theta}\right)\right]^{1/2}\end{aligned}$$

ومن ثـم:

$$\rho^2\left[V_\theta T\times V_\theta\left(\frac{\partial\ln L(X;\theta)}{\partial\theta}\right)\right]=\rho^2\left[V_\theta T\times E_\theta\left(\frac{\partial\ln L(X;\theta)}{\partial\theta}\right)^2\right]=1$$

أي أن:

$$V_\theta T\times E_\theta\left(\frac{\partial\ln L(X;\theta)}{\partial\theta}\right)^2\ \geq1$$

لأن معامل الارتباط ρ بالقيمة المطلقة لا يتجاوز الواحد، وبالتالي:

$$V_\theta T \geq \frac{1}{E_\theta\left(\frac{\partial \ln L(\boldsymbol{X};\theta)}{\partial\theta}\right)^2} \quad ; \quad \forall\, \theta \in (0,1) \qquad .....(3)$$

بأخذ لوغاريتم الطرفين في (2) والاشتقاق بالنسبة لـ θ، نجد:

$$\ln L(\boldsymbol{x};\theta) = n\bar{x}\ln\theta + n(1-\bar{x})\ln(1-\theta)$$

$$\frac{\partial \ln L(\boldsymbol{x};\theta)}{\partial\theta} = \frac{n\bar{x}}{\theta} - \frac{n(1-\bar{x})}{1-\theta} = \frac{n(\bar{x}-\theta)}{\theta(1-\theta)}$$

وعلى ذلك:

$$V_\theta\left[\frac{\partial \ln(\boldsymbol{X};\theta)}{\partial\theta}\right] = E_\theta\left[\frac{\partial \ln(\boldsymbol{X};\theta)}{\partial\theta}\right]^2 = \frac{n^2}{\theta^2(1-\theta)^2}V_\theta(\overline{X}-\theta)$$

$$= \frac{n^2}{\theta^2(1-\theta)^2}\ \frac{\theta(1-\theta)}{n} = \frac{n}{\theta(1-\theta)} \qquad .....(4)$$

وبالتعويض في المتباينة (3) نجد:

$$V_\theta T \geq \frac{1}{n/\theta(1-\theta)} = \frac{\theta(1-\theta)}{n}$$

أي أن المتباينة (1) محققة [المتباينة (1.6.5) محققة]. وهذا يعني أن المقدِّر T^* هو الأكفأ لتقدير المعلمة $\tau(\theta) = \theta$.

لأهمية نموذج بيرنولي $B(1,\theta)$ في التطبيق نصيغ النتيجة المثبتة بشكل مبرهنة.

مبرهنة 2.6.5

يعتبر التكرار النسبي لظهور الحادث A في n تكراراً مستقلاً المقدِّر الأكفأ لاحتمال هذا الحادث.

نشير كنتيجة لهذه المبرهنة إلى أن قيمة دالة التوزيع التجريبي $F_n^*(x)$، عند كل نقطة x (عدد حقيقي x)، تعتبر التقدير الأكفأ لدالة التوزيع النظري $F(x)$ (دالة توزيع المتغير العشوائي الملاحظ ξ) عند هذه النقطة.

سنبرهن فيما يلي على خاصة هامة من خواص المقدّر الأكفأ.

مبرهنة 3.6.5

إذا كان T_1^* المقدّر الأكفأ لتقدير $\tau_1(\theta)$ و T_2^* المقدّر الأكفأ لتقدير $\tau_2(\theta) \neq \tau_1(\theta)$، فعندئذٍ الإحصاء $T^* = a_1T_1^* + a_2T_2^*$ يعتبر المقدّر الأكفأ للدالة $\tau(\theta) = a_1\tau_1(\theta) + a_2\tau_2(\theta)$ مهما يكن الثابتان a_1, a_2.

الإثبات

لنثبت أولاً الخاصة الهامة التالية للمقدّر الأكفأ:

إذا كان $\psi = \psi(\boldsymbol{X})$ إحصاء قيمته المتوقعة:

$$E_\theta\psi = 0 \quad ; \quad \forall\theta \in \Theta$$

فإن:

$$\operatorname{cov}\left(T_i^*, \psi\right) = 0 \quad ; \quad \forall\theta \in \Theta \ , \ i = 1,2 \tag{4.6.5}$$

لإثبات ذلك، لنأخذ $\tilde{T}_i = T_i^* + \lambda\psi$. نلاحظ بسهولة أن $\tilde{T}_i$ مقدّر غير متحيز لـ $\tau_i(\theta)$ مهما تكن λ، لأن:

$$E_\theta\tilde{T}_i = E_\theta T_i^* + \lambda E_\theta\psi = ET_i^* = \tau_i \quad ; \quad \theta \in \Theta$$

وبما أن $T_i^*\ ,\ i = 1,2$ المقدّر الأكفأ لـ $\tau_i(\theta)$، فإن:

$$V_\theta\tilde{T}_i = V_\theta T_i^* + \lambda^2 V_\theta\psi + 2\lambda\operatorname{cov}\left(T_i^*, \psi\right) \geq V_\theta T_i^* \quad ; \quad \theta \in \Theta \quad ...(1)$$

أي أن:

$$\lambda^2 V_\theta\psi + 2\lambda\operatorname{cov}\left(T_i^*, \psi\right) \geq 0$$

وبحل هذه المتباينة بالنسبة لـ λ نجد:

$$-\frac{\left|\operatorname{cov}\left(T_i^*,\psi\right)\right|}{V_\theta\psi}-\frac{\operatorname{cov}_\theta\left(T_i^*,\psi\right)}{V_\theta\psi}\le\lambda\le+\frac{\left|\operatorname{cov}_\theta\left(T_i^*,\psi\right)\right|}{V_\theta\psi}-\frac{\operatorname{cov}\left(T_i^*,\psi\right)}{V_\theta\psi}$$

وهذا يعني: إما:

$$\operatorname{cov}\left(T_i^*,\psi\right)=0 \quad , \quad \forall\,\theta\in\Theta$$

وهو الحل المقبول، أو:

$$-\frac{\left|\operatorname{cov}\left(T_i^*,\psi\right)\right|}{V_\theta\psi}-\frac{\operatorname{cov}_\theta\left(T_i^*,\psi\right)}{V_\theta\psi}<\lambda<+\frac{\left|\operatorname{cov}\left(T_i^*,\psi\right)\right|}{V_\theta\psi}-\frac{\operatorname{cov}\left(T_i^*,\psi\right)}{V_\theta\psi}$$

وعندئذٍ $V_\theta\tilde{T}_i<V_\theta T_i^*$ وهذا يتعارض مع المتباينة (1)، وبالتالي فهو حل مرفوض.

لنفترض الآن أن T مقدّر ما غير متحيز لـ $\tau(\theta)$ و $\psi=T^*-T$. وبما أن $E_\theta\psi=E_\theta T^*-E_\theta T=0$، فحسب الخاصة السابقة نجد:

$$\begin{aligned}0=\operatorname{cov}_\theta\left(T^*,\psi\right)&=\operatorname{cov}\left(T^*,T^*-T\right)=E_\theta\left(T^*-\tau\right)\left(T^*-T\right)\\&=E_\theta\left(T^*-\tau\right)\left[\left(T^*-\tau\right)-\left(T-\tau\right)\right]\\&=E_\theta\left(T^*-\tau\right)^2-E_\theta\left(T^*-\tau\right)\left(T-\tau\right)\\&=V_\theta T^*-\operatorname{cov}_\theta\left(T^*,T\right)\end{aligned}$$

أي أن:

$$\left[V_\theta T^*\right]^2=\left[\operatorname{cov}_\theta\left(T^*,T\right)\right]^2=\rho^2V_\theta T^*V_\theta T\le V_\theta T^*V_\theta T$$

وبالتالي:

$$V_\theta T^*\le V_\theta T \quad ; \quad \forall\,\theta\in\Theta$$

وهذا يعني أن T^* المقدّر الأكفأ لـ $\tau(\theta)$ وهو المطلوب إثباته.

7.5 معايير كفاءة المقدِّرات المبنية على متباينة كرامر وراو وتعميماتها

هناك العديد من المعالم أو الدوال المعلمية في الإحصاء التي يوجد لها مقدِّرات غير متحيزة بأقل تباين (المقدِّر الأكفأ). وسنبحث في هذا البند المعايير العامة لوجود المقدِّر الأكفأ لدالة في معلمة وحيدة البعد $\tau(\theta)$ وطرق إيجاده.

نفترض كالعادة $f(x;\theta)$ كثافة الاحتمال (أو دالة الاحتمال في حالة المنقطع) للمتغير العشوائي الملاحظ ξ، $X=(X_1,\dots,X_n)$ عينة عشوائية مأخوذة من توزيع $\mathcal{L}(\xi)\in \boldsymbol{F}=\{F(x;\theta);\theta\in\Theta\}$، $x=(x_1,\dots,x_n)$ قيمة مشاهدة لـ X. وتحقق النماذج $\boldsymbol{F}$ شروط الانتظام preqularity conditions.

1.7.5 متباينة كرامر وراو والمقدِّر الأكفأ

Cramer-Roa Inequality and Efficient Estimator

سنناقش مسألة تقدير دالة معلمة معطاة $\tau(\theta)$ في معلمة (وحيدة البعد) نموذج $F(x;\theta)$. نفترض أن النموذج $F(x;\theta)$ نظامي والدالة $\tau(\theta)$ قابلة للاشتقاق بالنسبة θ، و U_τ عائلة المقدِّرات غير المتحيزة لـ $\tau(\theta)$. عندئذٍ المبرهنة التالية صحيحة.

مبرهنة 1.7.5: متباينة كرامر وراو **Cramer-Roa Inequality**

من أجل أي مقدِّر $T=T(X)\in U_\tau$، فالمتباينة:

$$V_\theta T \ge \frac{[\tau'(\theta)]^2}{ni(\theta)} \qquad \forall\,\theta\in\Theta \tag{1.7.5}$$

محققة. وتتحقق المساواة (1.7.5) إذا كان المقدِّر T دالة خطية في مساهمة العينة من الشكل:

$$T(\boldsymbol{X})-\tau(\theta)=a(\theta)U(\boldsymbol{X};\theta) \qquad (2.7.5)$$

حيث أن $a(\theta)$ دالة ما في المعلمة θ.

الإثبات

حسب الفرض $T \in U_\tau$، أي أن:

$$E_\theta T = \int_{R^n} T(\boldsymbol{x})L(\boldsymbol{x};\theta)d\boldsymbol{x}=\tau(\theta) \quad ; \quad \forall\, \theta \in \Theta \qquad (3.7.5)$$

وباشتقاق الطرفين بالنسبة لـ θ (وهذا ممكن لأن $F(x;\theta)$ نموذج نظامي حسب الفرض) نجد:

$$\begin{aligned}\tau'(\theta) &= \int T(\boldsymbol{x})\frac{\partial L(\boldsymbol{x};\theta)}{\partial\theta}d\boldsymbol{x} = \int T(\boldsymbol{x})\frac{\partial \ln L(\boldsymbol{x};\theta)}{\partial\theta}L(\boldsymbol{x};\theta)d\boldsymbol{x} \\ &= E_\theta\left[T(\boldsymbol{X})U(\boldsymbol{X};\theta)\right]=\operatorname{cov}_\theta\left[T(\boldsymbol{X})U(\boldsymbol{X};\theta)\right]\end{aligned} \qquad (4.7.5)$$

لأن $E_\theta U(\boldsymbol{X};\theta)=0$ [العلاقة (3.6.4)].

ولكن:

$$\operatorname{cov}_\theta\left[T(\boldsymbol{X}),U(\boldsymbol{X};\theta)\right]=\rho\sqrt{V_\theta T(\boldsymbol{X})V_\theta U(\boldsymbol{X};\theta)} \le \sqrt{V_\theta T(\boldsymbol{X})V_\theta U(\boldsymbol{X};\theta)} \qquad (5.7.5)$$

لأن $|\rho| \le 1$.

بناءً على العلاقتين (4.7.5) و (5.7.5) نجد:

$$\left[\tau'(\theta)\right]^2 \le V_\theta T(\boldsymbol{X})V_\theta U(\boldsymbol{X};\theta)=i_n(\theta)V_\theta T(\boldsymbol{X})$$

لأن $V_\theta U(\boldsymbol{X},\theta)=E_\theta\left[U(\boldsymbol{X};\theta)\right]^2=i_n(\theta)$ [راجع العلاقة (1.7.4)]. وبالتالي:

$$V_\theta T \ge \frac{\left[\tau'(\theta)\right]^2}{i_n(\theta)}=\frac{\left[\tau'(\theta)\right]^2}{ni(\theta)} \quad ; \quad \forall\, \theta \in \Theta$$

تــؤول المتباينــة في (5.7.5) إلى مســاواة، إذا كــانت العلاقــة بــين $T(X)$ و $U(X;\theta)$ خطية، أي أن $|\rho| = 1$، وعندئذٍ:

$$V_\theta T = \frac{[\tau'(\theta)]^2}{ni(\theta)} \quad ; \quad \forall\, \theta \in \Theta$$

يمكـن كتابـة العلاقـة الخطيـة بين T و $U(X;\theta)$ على الصـورة:

$$T(X) = a(\theta)U(X;\theta) + c(\theta)$$

ويـأخذ التوقـع الرياضي (القيمـة المتوقعـة) للطرفـين نجـد أن:

$$\tau(\theta) = c(\theta)$$

وبالتالي يمكـن كتابـة:

$$T(X) - \tau(\theta) = a(\theta)U(X;\theta)$$

وهـو المطلـوب.

تدعــى المتباينــة (1.7.5) بمتباينــة كرامــر وراو (Cramer-Roa Inequality) وهــي تعــين الحــد الأدنــى لتبــاين المقــدِّرات غــير المتحــيزة (lower bound of the variance of unbiased estimators) للدالــة المعلميــة $\tau(\theta)$ المعطـاة في معلمـة نمـوذج نظـامي.

لقـد تم إثبـات المبرهنـة (1.7.5) باعتبـار أن التوزيـع $f(x;\theta)$ مســتمر، ويمكــن إثبـات صحتهـا في حالـة التوزيعـات المنقطعـة باسـتبدال التكـامل بـالمجموع، وهـذا يعـني أن المبرهنـة (1.7.5) يمكـن تطبيقهـا في حالـة التوزيعـات مسـتمرة كـانت أم منقطعـة.

مــلاحظة

إذا كــان $T = T(X)$ مقــدِّر مــا متحــيز لــ $\tau(\theta)$ ومقـدار تحـيزه $b(\theta)$ قــابل للاشـتقاق بالنسـبة θ، فباسـتخدام المطابقــة:

$$E_\theta T(X) = \int_{R^n} T(x)L(x;\theta)dx = \tau(\theta) + b(\theta) \quad ; \quad \forall\, \theta \in \Theta$$

بدلاً من المطابقة (3.7.5). وبإجراء مناقشة مشابهة لحالة مقدّر غير متحيز، نجد:

$$V_\theta T \geq \frac{\left[\tau'(\theta) + b'(\theta)\right]^2}{ni(\theta)} \quad ; \quad \forall\theta \in \Theta$$

وهي تعميم للمتباينة (1.7.5)، أي تعطي الحد الأدنى لتباين المقدّرات المتحيزة وغير المتحيزة لـ $\tau(\theta)$ (الحد الأدنى لتباين مقدّر lower bound of variance of an estimator).

مبرهنة 2.7.5

إذا وجد مقدّر $T^* \in U_\tau$ لتقدير $\tau(\theta)$، تباينه يساوي الحد الأدنى لمتباينة كرامر وراو (5.7.5)، أي أن:

$$V_\theta T^* = \frac{\left[\tau'(\theta)\right]^2}{ni(\theta)} \quad , \quad \forall\theta \in \Theta \tag{7.7.5}$$

فإن T^* المقدّر الأكفأ لتقدير $\tau(\theta)$، ويدعى للتمييز أحياناً بالمقدّر الأمثل.

نلاحظ من تعريفي المقدّر الأمثل والمقدّر الأكفأ لتقدير $\tau(\theta)$، بأن المقدّر الأمثل هو الأكفأ في نفس الوقت، لكن العكس ليس بالضرورة صحيحاً. وهذا ما سنوضحه في أمثلة لاحقة. ونخلص من المبرهنة (2.7.5) إلى أن معيار الأمثلية لمقدّر لتقدير $\tau(\theta)$ يتمثل في تحقيق العلاقة (7.7.5)، وإذا وجد مقدّر أمثل لدالة $\tau(\theta)$ فهو وحيد حسب المبرهنة (1.6.5)، لاعتبار المقدّر الأمثل هو مقدّر أكفأ أيضاً لـ $\tau(\theta)$. ويطلق على المعيار (7.7.5) لكفاءة مقدّر معيار كرامر وراو.

إذا كان T^* المقدّر الأمثل لـ $\tau(\theta)$، أي أن:

$$T^* - \tau(\theta) = a(\theta)U(X;\theta)$$

ومـن ثم:

$$U(X;\theta)=\frac{T^{*}-\tau(\theta)}{a(\theta)}$$

وبالتعويض في العلاقة (4.7.5)، نجـــد:

$$\tau'(\theta)=\operatorname{cov}\left[T^{*}(X),U(X;\theta)\right]=E_{\theta}\left[(T^{*}-\tau)U(X;\theta)\right]$$

$$=E_{\theta}\left[(T^{*}-\tau)\left(\frac{T^{*}-\tau}{a(\theta)}\right)\right]=\frac{1}{a(\theta)}E_{0}\left(T^{*}-\tau\right)^{2}=\frac{1}{a(\theta)}V_{\theta}T^{*}$$

وعلى ذلك فإن تباين المقدِّر الأمثل T^{*} لتقدير $\tau(\theta)$ يســاوي:

$$V_{\theta}T^{*}=a(\theta)\tau'(\theta) \tag{8.7.5}$$

تجـــدر الإشارة هنا إلى الخاصة الهامة الآتيـــة: بمـا أن مســاهمة العينـــة $U(X;\theta)$ تُعيَّن وبشكل وحيد بـــدلالة النمـــوذج $F(x;\theta)$، لـــذا الصيغـــة (2.7.5) إن وجـــدت فهي وحيدة. وهذا يعني إذا وجد مقدِّر أمثل T^{*} لدالـــة معلميـــة $\tau(\theta)$، فـــلا يوجـــد مقـــدِّر أمثل من أجل أي دالة معلمية أخـــرى في θ مختلفـــة عـــن $a\tau(\theta)+b$، حيـــث أن a و b ثوابـــت.

يمكن إثبات صحة هذه الخاصة بسهولة. لنفترض T^{*} المقـــدِّر الأمثـــل لـــ $\tau(\theta)$ و T_{1}^{*} المقدِّر الأمثل لـــــ $\tau_{2}(\theta)=f(\tau)\ ;\ \tau_{2}\neq\tau$، حيـــث أن θ معلمـــة (وحيـــدة البعد) التوزيـــع $F(x;\theta)$، أي أن:

$$V_{\theta}T^{*}=\frac{\left[\tau'(\theta)\right]^{2}}{i_{n}(\theta)}\quad,\quad V_{\theta}T_{1}^{*}=\frac{\left[f_{\tau}'\,\tau'(\theta)\right]^{2}}{i_{n}(\theta)}$$

وعلى ذلك فـــإن:

$$\frac{f_{\tau}'^{\,2}}{V_{\theta}T_{1}^{*}}=\frac{1}{V_{\theta}T^{*}}$$

ولا يمكـــن أن تتحقق هـــذه المســـاواة إلا إذا كـــانت T_1^* دالـــة خطية في T^*، أي أن $T_1^* = aT^* + b$ (حيث a و b ثوابت). وهذا يعني:

$$\tau_2(\theta) = a\tau(\theta) + b$$

مثـــــال 1:7.5

إذا كانت $X = (X_1,\dots,X_n)$ عينة عشوائية من مجتمع كثافـــــة توزيعـــه:

$$f(x;\theta) = \theta e^{-\theta x} \quad ; \quad x > 0$$

فـالمطلوب:

1. أوجد الحد الأدنى لتباين المقدّرات غير المتحيزة لــ $\tau(\theta) = \theta$.

2. أثبت أن المقدّر $T^* = \overline{X}$ هو المقدّر الأمثل لتقدير $\tau(\theta) = \dfrac{1}{\theta}$.

كما نعلم من المثال (*2.7.5*) أن $i_n(\theta) = \dfrac{n}{\theta^2}$، وعندما $\tau(\theta) = \theta$ فـإن $\tau'(\theta) = 1$، وبالتالي الحد الأدنى لتباين المقدّرات غير المتحيزة لــ θ يعطي بالحد الأدنى لمتباينة كرامـــر وراو:

$$V_\theta T \geq \frac{[\tau'(\theta)]^2}{i_n(\theta)} = \frac{1}{n/\theta^2} = \frac{\theta^2}{n} \quad ; \quad \theta > 0 \;,\; T \in U_\tau$$

أي أن الحد الأدنى هـــو $\dfrac{\theta^2}{n}$.

ولإثبات أن المقدّر $T^* = \overline{X}$ هو المقدّر الأمثل لــ $\tau(\theta)$، يلزمنا أولاً حساب $E_\theta T^*$ و $V_\theta T^*$.

كما نعلم أن التوزيع $f(x;\theta)$ عبـــارة عـــن حالـــة خاصـــة مـــن توزيـــع جامـــا $\Gamma(\alpha,\beta)$ الموافق لــ $\alpha = 1 \;,\; \beta = \dfrac{1}{\theta}$ [راجـــع الفقـــرة (5.3.2)]. وبالتـــالي حســـب العلاقة (25.3.2)، وبوضع $\dfrac{1}{\theta}$ بدلاً مـــن θ نجـــد:

$$E_\theta X_i = \frac{1}{\theta} \quad , \quad V_\theta X_i = \frac{1}{\theta^2} \quad ; \quad i = \overline{1,n}$$

وعلى ذلـك:

$$E_\theta \overline{X} = \frac{1}{\theta} \quad , \quad V_\theta \overline{X} = \frac{1}{n\theta^2}$$

أي أن $T^* = \overline{X}$ مقدّر غير متحيز لـ $\tau(\theta) = \frac{1}{\theta}$. وأن تباينه ينطبق على الحد الأدنى لمتباينة كرامر وراو

$$\frac{[\tau'(\theta)]}{i_n(\theta)} = \frac{1/\theta^4}{n/\theta^2} = \frac{1}{n\theta^2}$$

وهذا يعني أن $T^* = \overline{X}$ المقدّر الأمثل وأيضاً الأكفأ لتقدير $\tau(\theta) = \frac{1}{\theta}$.

مثال 2.7.5

لتكن $X = (X_1, \dots, X_n)$ عينة عشوائية من مجتمع يتبع التوزيع $\mathcal{L}(\xi) \in \Gamma(2,\theta)$ ونرغب في إيجاد الحد الأدنى لتباين المقدّرات غير المتحيزة لكل من:

$$\theta \quad , \quad \frac{1}{1+\theta} \quad , \quad \ln\theta$$

بما أن:

$$f(x;\theta) = \frac{1}{\theta^2} x e^{-x/\theta} \qquad ; \qquad x > 0$$

$$L(x;\theta) = \theta^{-2n} \prod_{i=1}^{n} x_i e^{-\frac{x_i}{\theta}}$$

$$\ln L(x;\theta) = -2n\ln\theta + \sum_{i=1}^{n} \ln x_i - \frac{1}{\theta}\sum_{i=1}^{n} x_i$$

$$\frac{\partial \ln L(x;\theta)}{\partial\theta} = -\frac{2n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i$$

$$\frac{\partial^2 \ln L(x;\theta)}{\partial^2\theta} = +\frac{2n}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n} x_i$$

فـإن معلومـات فيشـر المتوفـرة في العينـة X حـول المعلمـة θ:

$$i_n(\theta)=E_\theta\left[-\frac{\partial^2 \ln L(x;\theta)}{\partial^2\theta}\right]=E_\theta\left(-\frac{2n}{\theta^2}+\frac{2}{\theta^3}\sum_{i=1}^{n}X_i\right)$$
$$=-\frac{2n}{\theta^2}+\frac{2}{\theta^3}\sum_{i=1}^{n}E_\theta X_i=-\frac{2n}{\theta^2}+\frac{2}{\theta^3}(2n\theta)=\frac{2n}{\theta^2}$$

لأن $E_\theta X_i=2\theta$ [راجع العلاقة (23.3.2)، حيث $\beta=\theta$, $\alpha=2$]. وكمـا نعلـم، الحد الأدنى لتباين المقدّر غير المتحيزة لـ $\tau(\theta)$ يعطى بـالحد الأدنـى لمتباينـة كرامـر وراو، وبالتالي:

1. الحـد الأدنـى لتبـاين المقـدّرات غـير المتحـيزة لــ $\tau(\theta)=\theta$:

$$V_\theta T\geq\frac{[\tau'(\theta)]^2}{i_n(\theta)}=\frac{1}{2n/\theta^2}=\frac{\theta^2}{2n}$$

2. الحـد الأدنـى لتبـاين المقـدّرات غـير المتحـيزة لــ $\tau(\theta)=\frac{1}{1+\theta}$:

$$V_\theta T\geq\frac{[\tau'(\theta)]^2}{i_n(\theta)}=\frac{1/(1+\theta)^4}{2n/\theta^2}=\frac{\theta^2}{2n(1+\theta)^4}$$

3. الحـد الأدنـى لتبـاين المقـدّرات غـير المتحـيزة لــ $\tau(\theta)=\ln\theta$:

$$V_\theta T\geq\frac{1/\theta^2}{2n/\theta^2}=\frac{1}{2n}$$

مثـــال 3.7.5

إذا كـــانت $X=(X_1,\ldots,X_n)$ عينـــة عشـــوائية مـــن مجتمـــع يتبـــع التوزيـــع $\mathcal{L}(ξ)\in N(\theta,4)$، فـأوجد الحـد الأدنـى لتبـاين المقـدّرات غـير المتحـيزة لكـل مـــن:

$$\theta\quad,\quad\theta^2\quad,\quad e^\theta$$

بمـا أن:

$$f(x;\theta)=\frac{1}{2\sqrt{2\pi}}\exp\left[-\frac{1}{8}(x-\theta)^2\right]$$

ومن ثـم:

$$L(x;\theta)=\left(2\sqrt{2\pi}\right)^{-n/2}\exp\left[-\frac{1}{8}\sum_{i=1}^{n}\left(x_i-\theta\right)^2\right]$$

$$\ln L(x;\theta)=-\frac{n}{2}\ln\left(2\sqrt{2\pi}\right)-\frac{1}{8}\sum_{i=1}^{n}\left(x_i-\theta\right)^2$$

$$\frac{\partial\ln L(x;\theta)}{\partial\theta}=\frac{1}{4}\sum_{i=1}^{n}\left(x_i-\theta\right)=\frac{n}{4}(\bar{x}-\theta)$$

$$\frac{\partial^2\ln L(x;\theta)}{\partial\theta^2}=-\frac{n}{4}$$

وبالتـالي:

$$i_n(\theta)=-E_\theta\left[\frac{\partial^2 L(x;\theta)}{\partial\theta^2}\right]=-E_\theta\left(-\frac{n}{4}\right)=\frac{n}{4}$$

وعلى ذلـك، وبنـاءً على متباينـة كرامـر وراو، فـإن:

1. الحـد الأدنـى لتبـاين المقـدّرات غـير المتحـيزة لـ $\tau(\theta)=\theta$:

$$V_\theta T\geq\frac{\left[\tau'(\theta)\right]^2}{i_n(\theta)}=\frac{1}{n/4}=\frac{4}{n}$$

2. الحـد الأدنـى لتبـاين المقـدّرات غـير المتحـيزة لـ $\tau(\theta)=\theta^2$:

$$V_\theta T\geq\frac{\left[\tau'(\theta)\right]^2}{i_n(\theta)}=\frac{4\theta^2}{n/4}=\frac{16}{n}\theta^2$$

3. الحـد الأدنـى لتبـاين المقـدّرات غـير المتحـيزة لـ $\tau(\theta)=e^\theta$:

$$V_\theta T\geq\frac{\left[\tau'(\theta)\right]^2}{i_n(\theta)}=\frac{e^{2\theta}}{n/4}=\frac{4}{n}e^{2\theta}$$

مثـــــال 4.7.5

إذا كـانت $X=(X_1,\dots,X_n)$ عينـة عشـوائية مـن مجتمـع دالـة كثافـة توزيعـــه:

$$f(x;\theta)=\theta(1+x)^{-(1+\theta)} \quad , \quad x\geq 0$$

فهـل توجـد دالـة في θ $(\tau(\theta))$ مقدّرهـا الأمثـل موجــود؟

بما أن:

$$L(x;\theta)=\prod_{i=1}^{n} f(x_i,\theta)=\theta^n\prod_{i=1}^{n}(1+x_i)^{-(1+\theta)}$$

$$\ln L(x;\theta)=n\ln\theta-(1+\theta)\sum_{i=1}^{n}\ln(1+x_i)$$

فإن مســاهمة العينـة X

$$U(X;\theta)=\frac{\partial \ln L(X;\theta)}{\partial\theta}=+\frac{n}{\theta}-\sum_{i=1}^{n}\ln(1+X_i)$$

أي أن:

$$-\frac{1}{n}U(X;\theta)=\frac{1}{n}\sum_{i=1}^{n}\ln(1+X_i)-\frac{1}{\theta}$$

وإذا رمزنا بـــ:

$$T=\frac{1}{n}\sum_{i=1}^{n}\ln(1+X_i) \quad , \quad \tau(\theta)=\frac{1}{\theta} \quad , \quad a(\theta)=-\frac{1}{n}$$

نجد:

$$a(\theta)U(X;\theta)=T-\tau(\theta)$$

وحســب المبرهنـة (1.7.5) [العلاقـة (2.7.5)]، فـإن $T=\frac{1}{n}\sum_{i=1}^{n}\ln(1+X_i)$ المقـــدّر الأمثل للدالة $\tau(\theta)=\frac{1}{\theta}$ ، وأن تباينه:

$$V_\theta T=\frac{[\tau'(\theta)]^2}{i_n(\theta)}=\frac{1/\theta^4}{n/\theta^2}=\frac{1}{n\theta^2}$$

مثـــال 5.7.5

إذا كانت $X = (X_1, \ldots, X_n)$ عينة عشوائية من مجتمع كثافـــة توزيعـــه:

$$f(x;\theta) = \frac{1}{1+\theta} e^{-x/(1+\theta)} \quad , \quad x > 0$$

فهل توجد دالة ما $\tau(\theta)$ مقدِّرها الأمثـــل موجـــود.

بما أن:

$$L(\boldsymbol{x};\theta) = (1+\theta)^{-n} \exp\left[\frac{-1}{1+\theta}\sum_1^n x_i\right]$$

$$\ln L(\boldsymbol{x};\theta) = -n\ln(1+\theta) - \frac{1}{1+\theta}\sum_1^n x_i$$

فإن:

$$U(\boldsymbol{X};\theta) = \frac{\partial \ln L(\boldsymbol{X};\theta)}{\partial\theta} = -\frac{n}{1+\theta} + \frac{1}{(1+\theta)^2}\sum_{i=1}^n X_i = -\frac{n}{(1+\theta)^2}\left[\overline{X} - (1+\theta)\right]$$

أي أن:

$$\overline{X} - (1+\theta) = -\frac{(1+\theta)^2}{n} U(\boldsymbol{X};\theta)$$

وبوضـــع:

$$T = \overline{X} \quad , \quad \tau(\theta) = (1+\theta) \quad , \quad a(\theta) = -\frac{(1+\theta)^2}{n}$$

نحصل على العـــلاقة (2.7.5)، أي أن $T = \overline{X}$ المقدِّر الأمثل لـــ $\tau(\theta) = (1+\theta)$، وتباينه:

$$V_\theta T = V_\theta \overline{X} = \frac{1}{n^2}\sum_{i=1}^n V_\theta X_i = \frac{1}{n^2} n(1+\theta)^2$$

لأن $V_\theta X_i = (1+\theta)^2$ [راجع العلاقة (23.3.2)، حيث $\alpha = 1$ و $\beta = (1+\theta)$].

مثـــال 6.7.5

لتكن $X = (X_1,\dots,X_n)$ عينة عشوائية من مجتمع كثافـــة توزيعـــه:

$$f(x;\theta) = \theta e^{-\theta x} \quad , \quad x > 0$$

وعلمت أن $T_1(X) = \dfrac{(n-1)}{n\bar{X}}$ مقدِّر غير متحيز للمعلمة θ، فهل يعتبر المقـــدِّر الأمثـــل لـــ θ؟

بما أن $\mathcal{L}(X_i) \in \Gamma\left(1, \dfrac{1}{\theta}\right)$، فإن $\mathcal{L}\left(Y = \sum X_i\right) \in \Gamma\left(n, \dfrac{1}{\theta}\right)$، أي أن:

$$g(y;\theta) = \frac{\theta^n}{\Gamma(n)} y^{n-1} e^{-\theta y} \quad ; \quad y > 0$$

وعلى ذلـك:

$$E_\theta\left(\frac{1}{Y}\right) = \int_0^\infty \frac{\theta^n}{\Gamma(n)} \frac{1}{y} y^{n-1} e^{-\theta y} dy$$

$$= \frac{\theta}{n-1} \int_0^\infty \frac{\theta^{n-1}}{\Gamma(n-1)} y^{n-2} e^{-\theta y} dy = \frac{\theta}{n-1}$$

لأن $h(y;\theta) = \dfrac{\theta^{n-1}}{\Gamma(n-1)} y^{n-2} e^{-\theta y}$ ما هي إلا دالـــة الكثـــافة الاحتمـــالية للتـــوزيع $\Gamma\left(n-1, \dfrac{1}{\theta}\right)$، ومـــن ثم تكـــامله من 0 إلى ∞ يســـاوي 1 (خاصـــة دالـــة الكثافـة).

وبشكل مشابه نجـــد:

$$E_\theta\left(\frac{1}{Y}\right)^2 = \frac{\theta^2}{(n-1)(n-2)} \int_0^\infty \frac{\theta^{n-2}}{\Gamma(n-2)} y^{n-3} e^{-\theta y} dy = \frac{\theta^2}{(n-1)(n-2)}$$

وعلى ذلـك:

$$V_\theta\left(\frac{1}{Y}\right) = E_\theta\left(\frac{1}{Y}\right)^2 - E^2\left(\frac{1}{Y}\right) = \frac{\theta^2}{(n-1)(n-2)} - \frac{\theta^2}{(n-1)^2} = \frac{\theta^2}{(n-2)(n-1)^2}$$

وبالتـالي:

$$V_\theta T_1 = (n-1)^2 V_\theta\left(\frac{1}{\sum_1^n X_i}\right) = (n-1)^2 V_\theta\left(\frac{1}{Y}\right)$$

$$= (n-1)^2 \frac{\theta^2}{(n-2)(n-1)^2} = \frac{\theta^2}{(n-2)}$$

لنحسب الآن الحد الأدنى لمتباينة كرامر وراو:

$$V_\theta T \geq \frac{[\tau'(\theta)]^2}{i_n(\theta)}$$

بما أن:

$$L(x;\theta) = \theta^n e^{-\theta\sum x_i}$$

$$\ln L(x;\theta) = n\ln\theta - \theta\sum_1^n x_i$$

$$\frac{\partial \ln L(x;\theta)}{\partial\theta} = \frac{n}{\theta} - \sum_1^n x_i \quad , \quad \frac{\partial^2 \ln L(x;\theta)}{\partial\theta^2} = -\frac{n}{\theta^2}$$

فإن:

$$i_n(\theta) = E_\theta\left(-\frac{\partial^2 \ln L(x;\theta)}{\partial\theta^2}\right) = \frac{n}{\theta^2}$$

إذن:

$$V_\theta T \geq \frac{1}{n/\theta^2} = \frac{\theta^2}{n}$$

وبالمقارنة نجد:

$$V_\theta T_1 = \frac{\theta^2}{(n-2)} > \frac{\theta^2}{n}$$

أي أن $T_1 = \frac{n-1}{n\bar{X}}$ ليس المقدِّر الأمثل للمعلمة θ، ومع ذلك فهو المقدِّر الأكفأ ضمن عائلة المقدِّرات غير المتحيزة لـ θ، وهذا ما سنثبته لاحقاً.

إذا وجد المقدِّر الأمثل T^* لتقدير $\tau(\theta)$، وإذا كان T أي مقدِّر آخر غير متحيز لـ $\tau(\theta)$، فعندئذٍ كفاءة (efficiency) المقدِّر T لتقدير $\tau(\theta)$ تعطى بالنسبة:

$$eff\,\frac{V_\theta T^*}{V_\theta T} < 1 \qquad (9.7.5)$$

نلاحظ مما سبق:

1. المقدِّر الأكفأ لدالة معلمية $\tau(\theta)$ في معلمة نموذج $F(x;\theta)$ غير موجود دائماً.
2. المقدِّر الأمثل لدالة معلمية $\tau(\theta)$ في معلمة نموذج $F(x;\theta)$ غير موجود دائماً، حتى إذا كان المقدِّر الأكفأ لـ $\tau(\theta)$ موجوداً.
3. المقدِّر الأكفأ لـ $\tau(\theta)$ في معلمة نموذج $F(x;\theta)$ ليس بالضرورة الأمثل، وبعبارة أخرى فالحد الأدنى لمتباينة كرامر وراو في حالات كثيرة لا يمكن بلوغه.
4. متباينة كرامر وراو تعطي الحد الأدنى لتباين المقدِّرات $T \in U_\tau$. وبالتالي لها استخدامين:
   - يمكن التأكد من مقدِّر ما $T \in U_\tau$ أنه الأمثل أم لا لتقدير $\tau(\theta)$.
   - قياس جودة المقدِّر $T \in U_\tau$ بمقدار قرب $V_\theta T$ من الحد الأدنى لمتباينة كرامر وراو، أي قرب النسبة (9.7.5) من الواحد، بحيث إذا كانت $eff \approx 1$ يمكن اعتبار T كمقدِّر جيد لـ $\tau(\theta)$.

هكذا، متباينة كرامر وراو تمكننا من التأكد من أن مقدِّر ما $T \in U_\tau$ الأمثل أم لا لتقدير $\tau(\theta)$، وفي حالات عدة تساعدنا على إيجاد المقدِّر الأمثل لـ $\tau(\theta)$

(إن وجد)، وخاصة إذا كان النموذج المفروض $F(x;\theta)$ ينتمي لعائلة النماذج الأسية. وهذا ما سنبينه في الفقرة التالية.

2.7.5 عائلة النماذج الأسية والمقدّر الأكفأ
Family of Exponential Models and Efficient Estimator

سنتطرق في هذه الفقرة لعائلة النماذج المعلمية، التي تدعى بالأسية الواردة في البند (4.4)، حيث إنه توجد المقدّرات الأمثل دائماً في مثل هذه النماذج، ومن ثم الأكفأ لبعض الدوال المعلمية $\tau(\theta)$.

مبرهنة 3.7.5

إذا كانت $X = (X_1,\dots,X_n)$ عينة عشوائية من مجتمع توزيعه $f(x;\theta)$ ينتمي لعائلة النماذج الأسية، فيوجد المقدّر الأمثل $T^* = T^*(X)$ لدالة معلمية ما $\tau(\theta)$. كما أن العكس صحيح، أي إذا كان $T^* = T^*(X)$ المقدّر الأمثل لدالة $\tau(\theta)$، فإن النموذج $f(x;\theta)$ أسي.

الإثبات

لنفترض أن النموذج $f(x;\theta)$ ينتمي لعائلة النماذج الأسية، أي يمكن كتابته على الصورة:

$$f(x;\theta) = \exp[A(\theta)B(x) + C(\theta) + D(x)]$$

ومن ثم:

$$L(x;\theta) = \exp\left[A(\theta)\sum_{i=1}^{n} B(x_i) + nC(\theta) + \sum_{i=1}^{n} D(x_i)\right]$$

$$\ln L(x;\theta) = A(\theta)\sum_{i=1}^{n} B(x_i) + nC(\theta) + \sum_{i=1}^{n} D(x_i)$$

وعلى ذلك فمساهمة العينة X

$$U(\boldsymbol{X};\theta)=\frac{\partial \ln L(\boldsymbol{X};\theta)}{\partial\theta}=A'(\theta)\sum_{i=1}^{n}B(X_i)+nC'(\theta)$$
$$=nA'(\theta)\left[\frac{1}{n}\sum_{i=1}^{n}B(X_i)+\frac{C'(\theta)}{A'(\theta)}\right]$$

وبوضع:

$$T^*=\frac{1}{n}\sum_{i=1}^{n}B(X_i)\quad,\quad \tau(\theta)=-\frac{C'(\theta)}{A'(\theta)}\quad,\quad a(\theta)=\frac{1}{nA'(\theta)}$$

نحصل على العلاقة (2.7.5)، أي أن $T^*=\frac{1}{n}\sum_{i=1}^{n}B(X_i)$ المقدّر الأمثل لـ $\tau(\theta)=-\frac{C'(\theta)}{nA'(\theta)}$، وحسب العلاقة (8.7.5) فإن:

$$V_\theta T^*=\frac{\tau'(\theta)}{nA'(\theta)} \tag{9.7.5}$$

ولإثبات صحة العكس، لنفترض $T^*=T^*(\boldsymbol{X})$ المقدّر الأمثل لدالة ما $\tau(\theta)$ في معلمة توزيع مفروض $F(x;\theta)$.

بما أن T^* المقدّر الأمثل لـ $\tau(\theta)$، فإن العلاقة (2.7.5) محققة:

$$T^*(\boldsymbol{X})-\tau(\theta)=a(\theta)U(\boldsymbol{X};\theta)$$
$$=a(\theta)\frac{\partial \ln L(\boldsymbol{X};\theta)}{\partial\theta}$$

أي أن:

$$\frac{\partial \ln L(\boldsymbol{X};\theta)}{\partial\theta}=\frac{1}{a(\theta)}T^*(\boldsymbol{X})-\frac{\tau(\theta)}{a(\theta)}$$

وبمكاملة الطرفين بالنسبة لـ θ نجد:

$$\ln L(\boldsymbol{X};\theta)=T^*(\boldsymbol{X})\int\frac{1}{a(\theta)}d\theta-\int\frac{\tau(\theta)}{a(\theta)}d\theta=$$
$$=A_1(\theta)T^*(\boldsymbol{X})+C_1(\theta)+D_1(\boldsymbol{X}) \qquad .....(1)$$

وفي حالة النماذج الأسية:

$$\ln L(\boldsymbol{X};\theta) = A(\theta)\sum_{1}^{n} B(X_i) + nC(\theta) + \sum_{1}^{n} D(X_i) \qquad .....(2)$$

وبمقارنة (1) بـ (2) نجد أن:

$$T^*(X) = \sum_{1}^{n} B(X_i) \ , \ A_1(\theta) = A(\theta) \ , \ C_1(\theta) = nC(\theta) \ , \ D_1(X) = \sum_{1}^{n} D(X_i)$$

أي أن $F(x;\theta)$ ينتمي لعائلة النماذج الأسية، وهو المطلوب إثباته.

هكذا، حسب المبرهنة (3.7.5)، المقدّر الأمثل موجود فقط من أجل بعض الدوال $\tau(\theta)$ في حالة نماذج أسية.

ويمكن حساب دالة المعلومات $i_n(\theta)$ في حالة نماذج أسية (نظامية طبعاً) وفق الصيغة:

$$i_n(\theta) = n\tau'(\theta)A'(\theta)$$

وصحة هذه العلاقة تنتج مباشرة من مقارنة العلاقتين (7.7.5) و (9.7.5).

مثال 7.7.5

إذا كانت $\boldsymbol{X} = (X_1,...,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta) = \theta e^{-\theta x} \quad ; \quad x > 0$$

فأوجد دالة ما $\tau(\theta)$، المقدّر الأمثل لها موجود.

يمكن كتابة الدالة $f(x;\theta)$ على الصورة:

$$f(x;\theta) = \exp(\ln\theta - \theta x)$$

وعلى ذلك فهي تنتمي إلى عائلة النماذج الأسية، حيث أن:

$$A(\theta) = -\theta \quad , \quad B(x) = x \quad , \quad C(\theta) = \ln\theta \quad , \quad D(x) = 0$$

وبالتـالي حسـب المبرهنـة (3.7.5) فـإن الدالـة:

$$\tau(\theta) = -\frac{C'(\theta)}{A'(\theta)} = -\frac{1/\theta}{-1} = \frac{1}{\theta}$$

والمقدّر الأمثـل لهـا:

$$T^*(\boldsymbol{X}) = \frac{1}{n}\sum B(X_i) = \frac{1}{n}\sum X_i = \overline{X}$$

وبنـاءً علـى العلاقـة (7.7.5) نجـد:

$$V_\theta T^* = \frac{\tau'(\theta)}{nA'(\theta)} = \frac{-1/\theta^2}{n(-1)} = \frac{1}{n\theta^2}$$

مثـــال 8.7.5

إذا كـانت $\boldsymbol{X} = (X_1,\ldots,X_n)$ عينـة عشـوائية مـــن توزيـــع:

$$f(x;\theta) = \theta x^{\theta-1} \quad ; \quad 0 < x < 1$$

فأوجد دالة ما $\tau(\theta)$ لهـا مقـدّر أمثـل، ثـم احسـب تبـاين هـذا المقـدّر.

يمكـن كتابـة الدالـة $f(x;\theta)$ علـى الصـورة:

$$f(x;\theta) = \exp(\theta \ln x + \ln\theta - \ln x)$$

وهو عنصـر مـن عائلـة النمـاذج الأسية حيـث أن:

$$A(\theta) = \theta \quad , \quad B(x) = \ln x \quad , \quad C(\theta) = \ln\theta \quad , \quad D(x) = \ln x$$

وعلـى ذلـك وحسـب المبرهنـة (3.7.5)، فـإن:

$$\tau(\theta) = -\frac{C'(\theta)}{A'(\theta)} = -\frac{1/\theta}{1} = -\frac{1}{\theta}$$

والمقدّر الأمثـل لهـا:

$$T^*(X) = \frac{1}{n}\sum B(X_i) = \frac{1}{n}\sum \ln X_i$$

وبناءً على العلاقة (7.7.5) نجد:

$$V_\theta T^* = \frac{\tau'(\theta)}{nA'(\theta)} = \frac{+1/\theta^2}{n(1)} = \frac{1}{n\theta^2}$$

يتضمن الجدول (1.7.5) الدالة $\tau(\theta)$، المقدّر الأمثل لها والتباين الموافق له، وذلك لستة نماذج أسية، ونترك التأكد من ذلك للقارئ كتمارين.

جدول 1.7.5

النموذج	$\tau(\theta)$	T^*	$V_\theta T^*$
$N(\theta,\sigma^2)$	θ	$\overline{X}$	σ^2/n
$N(\mu,\theta^2)$	θ^2	$1/n\sum_1^n (X_i-\mu)^2$	$2\theta^2/n$
$\Gamma(\lambda,\theta)$	θ	$\overline{X}/\lambda$	$\theta^2/n\lambda$
$B(k,\theta)$	θ	$\overline{X}/k$	$\theta(1-\theta)/nk$
$\Pi(\theta)$	θ	$\overline{X}$	θ/n
$\overline{B}(r,\theta)$	$\theta/(1-\theta)$	$\overline{X}/r$	$\theta/\left[rn(1-\theta)^2\right]$

3.7.5 متباينة بختشاري والمقدّر الأكفأ

Bakhtshari Inequality and The Efficient Estimator

إن متباينة كرامر وراو تعطي الحد الأدنى (الأصغر على الإطلاق) لتباين المقدّرات غير المتحيزة لدالة معلمية $\tau(\theta)$ في معلمة نموذج نظامي $F(x;\theta)$ عند وجود المقدّر الأمثل. لكن إذا كان المقدّر الأمثل لـ $\tau(\theta)$ غير موجود (وهذا

ممكن)، أي لا يوجد مقدِّر $T \in U_\tau$ يحقق العلاقة (2.7.5))، فعندئذٍ يتم البحث عن المقدِّر الأكفأ (المقدِّر غير المتحيز $T \in U_\tau$ بأقل تباين)، وهذا يتطلب تعيين التباين الأصغر ضمن مجموعة تباينات المقدِّرات غير المتحيزة U_τ لـ $\tau(\theta)$ وذلك مهما تكن $\theta \in \Theta$.

كما رأينا، إن الشرط الأساسي لبلوغ الحد الأدنى لمتباينة كرامر وراو هو وجود مقدِّر $T^* \in U_\tau$ بحيث أن $T^* - \tau(\theta)$ دالة خطية في دالة المساهمة:

$$U(X;\theta) = \frac{\partial \ln L(X;\theta)}{\partial\theta} = \frac{1}{L}\frac{\partial L}{\partial\theta}$$

مهما تكن $\theta \in \Theta$ (هنا ولاحقاً سنكتب أحياناً، وعلى سبيل الاختصار، L بدلاً من $L(X;\theta)$). لنفترض أن مثل هذا المقدِّر غير موجود، لكن ومع ذلك يمكن وجود مقدِّر $T^* \in U_\tau$، بحيث أن الفرق $T^* - \tau(\theta)$ دالة خطية في المقادير:

$$\frac{1}{L}\frac{\partial L}{\partial\theta}, \frac{1}{L}\frac{\partial^2 L}{\partial\theta^2}, \ldots, \frac{1}{L}\frac{\partial^r L}{\partial\theta^r}$$

من أجل عدد ما $r \geq 2$. عندئذٍ يعتبر T^* المقدِّر الأكفأ لتقدير $\tau(\theta)$.

يدعى هذا المعيار للكفاءة، الذي يعتبر تعميماً لمعيار كرامر وراو (2.7.5))، بمعيار بختشاري ومبني أيضاً على مبدأ بلوغ الحد الأدنى في المتباينة الموافقة التي تدعى بمتباينة بختشاري.

لنرمز بـ $\tau^{(k)} = \frac{d^{(k)}\tau(\theta)}{d\theta^k}$ و $L^{(k)} = \frac{\partial^k L}{\partial\theta^k}$.

مبرهنة 4.7.5: متباينة بختشاري Bakhtshari Inequality

ليكن $T = T(X)$ مقدِّر غير متحيز لـ $\tau(\theta)$ $(T \in U_\tau)$. عندها:

$$V_\theta T \geq \sum_{i,j=1}^{r} c_{ij} a_i a_j \quad ; \quad \forall\, \theta \in \Theta \tag{11.7.5}$$

حيث أن:

$$c_{ij} = c_{ij}(\theta) = E_\theta\left(\frac{L^{(i)}}{L}\frac{L^{(j)}}{L}\right) \quad ; \quad i,j = 1,\dots,r$$

وتعين المعاملات $a_i = a_i(\theta)$ بحل منظومة المعادلات:

$$\sum_{j=1}^{r} c_{ij}a_j = \tau^{(i)} \quad ; \quad i = 1,\dots,r \tag{12.7.5}$$

إذا كانت المصفوفة $C = \|c_{ij}\|_1^r$ نظامية $\left(|C| \neq 0\right)$ و $C^{-1} = \|c^{ij}\|_1^r$ فإن المتباينة (11.7.5) مكافئة للمتباينة:

$$V_\theta T \geq \sum_{i,j=1}^{r} c^{ij}\tau^{(i)}\tau^{(j)} \quad ; \quad \forall\theta \in \Theta \tag{12.7.5}$$

وتتحقق المساواة في المتباينتين (11.7.5) و (12.7.5) عند تحقق الشرط التالي:

$$T - \tau = \sum_{i=1}^{n} a_i L^{(i)}/L \tag{13.7.5}$$

يستخدم معيار بختشاري عملياً على النحو الآتي: انطلاقاً من مشتقات دالة المعقولية بالنسبة لـ θ يتم اختيار تلك التركيبة الخطية لهم، بحيث نحصل على صيغة من الشكل (13.7.5)، وهذا يتم بافتراض وبشكل متتالي $r = 2,3,\dots$. وإذا توصلنا عند قيمة ما لـ r إلى مثل هذه الصيغة، فإن الزوج الذي نحصل عليه عندئذٍ $\tau(\theta)$ و T هو عبارة عن الدالة المعلمية (المقدَّرة) والمقدِّر الأكفأ لها على الترتيب. وكحالة خاصة، عندما $r = 1$ نحصل على معيار كرامر وراو (2.7.4)، هكذا للتأكد من مقدِّر ما $T \in U_\tau$ أنه الأكفأ، يجب البحث عن إمكانية كتابة $T - \tau$ كدالة خطية في المقادير $\frac{L^{(1)}}{L},\dots,\frac{L^{(r)}}{L}$، حيث $r \geq 1$.

مثال 9.7.5

لتكن $X = (X_1,\dots,X_n)$ عينة عشوائية من توزيع $N(\theta,\sigma^2)$، ونرغب في إيجاد دالة $\tau(\theta)$ لها مقدِّر أكفأ.

بما أن:

$$L(x;\theta)=\frac{1}{\left(\sigma\sqrt{2\pi}\right)^{n}}\exp\left\{-\frac{1}{2\sigma^{2}}\sum_{i=1}^{n}\left(x_{i}-\theta\right)^{2}\right\}$$

ومنها نجد:

$$\frac{1}{L}L^{(1)}=\frac{\partial\ln L}{\partial\theta}=\frac{n}{\sigma^{2}}\left(\overline{X}-\theta\right)$$

$$\frac{1}{L}L^{(2)}=\frac{1}{L}\frac{\partial}{\partial\theta}\left(L\frac{\partial\ln L}{\partial\theta}\right)=\left(\frac{\partial\ln L}{\partial\theta}\right)^{2}+\frac{\partial^{2}\ln L}{\partial\theta^{2}}=$$

$$=\frac{n^{2}}{\sigma^{4}}\left(\overline{X}-\theta\right)^{2}-\frac{n}{\sigma^{2}}=\frac{n^{2}}{\sigma^{4}}\left[\left(\overline{X}-\theta\right)^{2}-\frac{\sigma^{2}}{n}\right]$$

$$\frac{1}{L}\left(\frac{2\theta\sigma^{2}}{n}L^{(1)}+\frac{\sigma^{4}}{n^{2}}L^{(2)}\right)=2\theta\left(\overline{X}-\theta\right)+\left[\left(\overline{X}-\theta\right)^{2}-\frac{\sigma^{2}}{n}\right]=$$

$$=2\theta\overline{X}-2\theta^{2}+\overline{X}^{2}-2\theta\overline{X}+\theta^{2}-\frac{\sigma^{2}}{n}=$$

$$=\left(\overline{X}^{2}-\frac{\sigma^{2}}{n}\right)-\theta^{2}$$

ونجد حسب معيار بختشاري أن الإحصاء:

$$T^{*}=T^{*}(X)=\overline{X}^{2}-\frac{\sigma^{2}}{n}$$

يعتبر المقدّر الأكفأ (غير متحيز بأقل تباين) لـ $\tau(\theta)=\theta^{2}$.

4.7.5 معيار الكفاءة في حالة معلمة متعددة الأبعاد

إن معياري كفاءة المقدّرات الواردين سابقا [العلاقتين (2.7.5) و (13.7.5)] يخصان حالة معلمة وحيدة البعد.

لنعمم ما سبق على حالة معلمة متعددة الأبعاد. سنفترض أن $\theta=\left(\theta_{1},\dots,\theta_{r}\right)$ متجه ذات r بعد. وليكن المطلوب البحث عن مقدّر من أجل

دالة معلمية عددية معطاة $\tau(\theta)$، القابلة للاشتقاق بالنسبة لـ $\theta_i \; ; \; i = 1,2,\ldots,r$.
ونفترض أن النموذج المعطى $F(x;\theta)$ نظامي، عندها المبرهنة التالي صحيحة.

مبرهنة 5.7.5

إذا كان $T = T(X)$ مقدّر غير متحيز لدالة ما $\tau(\theta)$، فإن:

$$V_\theta T \geq \sum_{i,j=1}^{r} g_{ij} c_i c_j \quad ; \quad \forall \; \theta \in \Theta \tag{14.7.5}$$

حيث أن:

$$g_{ij} = g_{ij}(\theta) = E_\theta\left(\frac{\partial \ln L}{\partial \theta_i}\frac{\partial \ln L}{\partial \theta_j}\right) \quad ; \quad i, j = 1,2,\ldots,r$$

وتعين المعاملات $c_i = c_i(\theta)$ بحل منظومة المعادلات:

$$\sum_{j=1}^{r} g_{ij} c_j = \tau'_i = \frac{\partial \tau(\theta)}{\partial \theta_i} \quad ; \quad i = 1,2,\ldots,r$$

إذا كانت المصفوفة $I_n = I_n(\theta) = \left\| g_{ij} \right\|_1^r$ نظامية و $I_n^{-1} = \left\| g^{ij} \right\|_1^r$، فإن المتباينة (14.7.5) مكافئة للمتباينة:

$$V_\theta T \geq \sum_{i,j=1}^{r} g^{ij} \tau'_i \tau'_j \tag{15.7.5}$$

وتتحقق المساواة في المتباينتين (14.7.5) و (15.7.5) عندما:

$$T - \tau = \sum_{i=1}^{r} c_i \frac{\partial \ln L}{\partial \theta_i} \tag{16.7.5}$$

إذا كانت دالة المعقولية $L(x;\theta)$ قابلة للاشتقاق مرتين بالنسبة لـ θ، فباشتقاق المطابقة:

$$\int L(x;\theta) dx \equiv 1$$

أولا بالنسبة لـ θ_i ثم بالنسبة لـ θ_j، نجد:

$$g_{ij} = -E_\theta\left(\frac{\partial^2 \ln L(X;\theta)}{\partial\theta_i \partial\theta_j}\right) = -nE_\theta\left(\frac{\partial^2 \ln f(X_1;\theta)}{\partial\theta_i \partial\theta_j}\right) \qquad (17.7.5)$$

تدعى المصفوفة $I_n(\theta)$ بمصفوفة المعلومات للعينة X، بينما $I(\theta) = I_1(\theta)$ تدعى بمصفوفة المعلومات لملاحظة واحدة. وتبين العلاقة (17.7.5) أنه من أجل منظومة ملاحظات متكررة ومستقلة $I_n(\theta) = nI(\theta)$.

بشكل مشابه لحالة معلمة وحيدة البعد (حقيقية) نسمي المقدّر غير المتحيز $T^* = T^*(X)$ من أجل $\tau(\theta)$ بالأمثل، إذا كان تباينه منطبقاً على الحد الأدنى للمتباينة (14.7.5) أو (15.7.5) وذلك مهما تكن $\theta \in \Theta$.

ينتج من المبرهنة (5.7.5) أن معيار الأمثلية (ويعني أيضاً الكفاءة) للمقدّر T يتمثل في المساواة (16.7.5). نورد فيما يلي مثالاً لاستخدام هذا المعيار.

مثال 10.7.5

لتكن $X = (X_1,\dots,X_n)$ عينة عشوائية من التوزيع $\mathcal{L}(\xi) \in N(\theta_1,\theta_2^2)$. ويطلب تقدير الدالة $\tau(\theta) = \tau(\theta_1,\theta_2) = \theta_1$، أي تقدير متوسط توزيع ξ.

يمكن أخذ متوسط العينة $\overline{X}$ كمقدّر غير متحيز لـ θ_1. وهذا ما سنبينه كالآتي:

بما أن دالة المعقولية في الحالة المطروحة:

$$L(x;\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \frac{1}{\left(\theta_2\sqrt{2\pi}\right)^n}\exp\left\{-\frac{1}{2\theta_2^2}\sum_{i=1}^{n}(x_i-\theta_1)^2\right\}$$

فإن:

$$\frac{\partial \ln L}{\partial\theta_1} = \frac{n}{\theta_2^2}\left(\overline{X}-\theta_1\right) \quad , \quad \frac{\partial \ln L}{\partial\theta_2} = -\frac{n}{\theta_2} + \frac{1}{\theta_2^3}\sum_{i=1}^{n}(x_i-\theta_1)^2$$

ولإيجـاد المعـاملين $c_j ; j = 1,2$ نحتـاج أولاً لحسـاب $g_{ij}\ ;\ i,j = 1,2$ بنـاءً علـى العلاقـة (17.7.5):

$$g_{11} = -E_\theta\left[\frac{\partial^2 \ln L(X;\theta)}{\partial\theta_1^2}\right] = -E_\theta\left(-\frac{n}{\theta_2^2}\right) = \frac{n}{\theta_2^2}$$

$$g_{12} = g_{21} = -E_\theta\left[\frac{\partial^2 \ln L(X;\theta)}{\partial\theta_1\partial\theta_2}\right] = -E_\theta\left[-\frac{2n}{\theta_2^3}\left(\overline{X} - \theta_1\right)\right] = \frac{2n}{\theta_2^3}E_\theta\left(\overline{X} - \theta_1\right) = 0$$

$$g_{22} = -E_\theta\left[\frac{\partial^2 \ln L(X;\theta)}{\partial\theta_2^2}\right] = -E_\theta\left[+\frac{n}{\theta_2^2} - \frac{3}{\theta_2^4}\sum_{i=1}^{n}\left(X_i - \theta_1\right)^2\right] =$$

$$= -\frac{n}{\theta_2^2} + \frac{3n}{\theta_2^2} = \frac{2n}{\theta_2^2}$$

وبحل جملـة المعـادلتين:

$$\frac{n}{\theta_2^2}c_1 + 0c_2 = 1$$

$$0c_1 + \frac{2n}{\theta_2^2}c_2 = 0$$

نجـد: $c_1 = \frac{\theta_2^2}{n}$ و $c_2 = 0$.

إذن:

$$\sum_{i=1}^{2} c_i \frac{\partial \ln L}{\partial\theta_i} = \frac{\theta_2^2}{n}\frac{\partial \ln L}{\partial\theta_1} = \frac{\theta_2^2}{n}\frac{n}{\theta_2^2}\left(\overline{X} - \theta_1\right) = \overline{X} - \theta_1$$

أي أن العلاقـة (16.7.5) محققـة، وبالتـالي الإحصـاء $T(X) = \overline{X}$ المقـدِّر الأمثـل، ومن ثـم الأكفأ لـ $\tau(\theta) = \theta_1$.

وبما أن $|I_n| \neq 0$ فإن $g^{12} = g^{21} = 0$ ، $g^{22} = \frac{\theta_2^2}{2n}$ ، $g^{11} = \frac{\theta_2^2}{n}$، ومن ثم الحد الأدنى في المتباينة (15.7.5) يساوي:

$$\sum_{i,j=1}^{2} g^{ij}\tau_i'\tau_j' = \frac{\theta_2^2}{n}$$

وكمـــا نعلـــم $V_\theta \overline{X} = \frac{\theta_2^2}{n}$ ، $V_\theta \overline{X} = \sum g^{ij}\tau'_i\tau'_j = \frac{\theta_2^2}{n}$.

لكـــن $\overline{X}$ المقـدّر الأمثـل أيضـاً لـــ θ في النمــوذج $N(\theta,\sigma^2)$ [راجــع الجــدول (1.7.5)]، وهـذا يعـني أن $\overline{X}$ المقـدّر الأمثـل لمتوسـط النمـوذج الطبيعـي وغـير مرتبـط بمعرفـة التبـاين أم لا.

كمـا في حالـة معلمـة وحيـدة البعـد، المقـدّرات الأمثـل غـير موجـودة دائمـاً. فمثـلاً، في مسـألة تقديـر التبـاين غـير المعلـوم في نمـوذج طبيعـي، إذا كـان متوسـط النمـوذج معلـوم، أي $N(\mu,\theta_2^2)$، فكمـا بينـا في الفقـرة (2.7.5) [أنظـر الجـدول (1.7.5)]، إن تباين العينـة:

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

يعتبر المقـدّر الأمثـل وبنفـس الوقـت الأكفـأ لتبـاين المجتمـع، بينمـا في حالـة نموذج طبيعـي عـام $N(\theta_1,\theta_2^2)$ هـذا غـير موجـود.

أولاً: كمـا بينـا سـابقا في البنـد (4.5) [مثـال (6.4.5)]، إن تبـاين العينـة S^2 مقـدّر متحيز للدالـة $\tau(\theta) = \theta_2^2$ ، بينمـا المقـدّر غـير المتحيز عندمـا $n \geq 2$ هـو:

$$S^{*2} = \frac{n}{n-1}S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \quad ; \quad n \geq 2$$

ثانياً: إن الإحصـاء S^{*2} ليس المقـدّر الأمثـل مـن أجـل تقديـر θ_2^2، وهـذا يمكـن إثباتـه بمقارنـة تباينـه $V_\theta S^{*2} = \frac{2\theta_2^4}{n-1}$ [راجـع العلاقـة (5.4.5)]، بـالحد الأدنـى في المتباينـة (15.7.5). فعلاً، حسـب العلاقة (17.7.5) نجـد [راجـع المثـال (10.7.5)]:

$$g_{11} = \frac{n}{\theta_2^2} \quad , \quad g_{12} = g_{21} = 0 \quad , \quad g_{22} = \frac{2n}{\theta_2^2}$$

أي من أجـل النمـوذج $N(\theta_1,\theta_2^2)$ مصفوفـة المعلومـات:

$$I_n(\theta)=\begin{Vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{Vmatrix}=\begin{Vmatrix} n/\theta_2^2 & 0 \\ 0 & 2n/\theta_2^2 \end{Vmatrix}=n\begin{Vmatrix} 1/\theta_2^2 & 0 \\ 0 & 2/\theta_2^2 \end{Vmatrix}=nI(\theta)$$

وعلى ذلك فإن:

$$\det I_n(\theta)=g_{11}g_{22}=\frac{2n^2}{\theta_2^4}\neq 0 \quad , \quad g^{22}=\frac{\theta_2^2}{2n}$$

والجزء الأيمن من المتباينة (15.7.5) يساوي:

$$g^{22}\left(\tau_2'\right)^2=\frac{\theta_2^2}{2n}\left(2\theta_2\right)^2=\frac{\theta_2^4}{n}$$

وهو أصغر من:

$$V_\theta S^{*2}=\frac{2\theta_2^4}{n-1}$$

وهذا غير كافٍ للقول بأن هنالك مقدّر متحيز لـ $\tau(\theta)=\theta_2^2$ أكفأ نسبياً من المقدّر S^{*2}.

سنبين بشكل عام في الحالة المطروحة أنه لا يوجد المقدّر الأمثل [لا يتم بلوغ الحد الأدنى في المتباينة (14.7.5) أو (15.7.5) من أجل أي مقدّر غير متحيز لـ θ_2^2]، بينما المقدّر الأكفأ T^* موجود وهو S^{*2}.

باستخدام معيار مشابه لمعيار بختشاري من أجل معلمة متعددة الأبعاد، الذي يأخذ الشكل الآتي:

إذا كانت الصيغة:

$$T-\tau(\theta)=\frac{1}{L}\left[\sum_i c_i\frac{\partial L}{\partial\theta_i}+\sum_{i,j}c_{ij}\frac{\partial^2 L}{\partial\theta_i\partial\theta_j}+\ldots\ldots+\sum_{i_1,i_2,\ldots,i_s}c_{i_1\ldots i_s}\frac{\partial^2 L}{\partial\theta_{i_1}\ldots\partial\theta_{i_s}}\right] \tag{18.7.5}$$

محققة من أجل عدد صحيح ما $s \geq 2$ ومعاملات $c_{.} = c_{.}(\theta)$، فإن الإحصاء T يعتبر المقدِّر الأكفأ للدالة $\tau = \tau(\theta)$.

بتطبيق هذا المعيار على الحالة المطروحة نجد:

$$\frac{1}{L}\frac{\partial L}{\partial \theta_1} = \frac{n}{\theta_2^2}\left(\overline{X} - \theta_1\right)$$

$$\frac{1}{L}\frac{\partial^2 L}{\partial \theta_1^2} = \frac{\partial^2 \ln L}{\partial \theta_1^2} + \left(\frac{\partial \ln L}{\partial \theta_1}\right)^2 = -\frac{n}{\theta_2^2} + \frac{n^2}{\theta_2^4}\left(\overline{X} - \theta_1\right)^2$$

$$\frac{1}{L}\frac{\partial L}{\partial \theta_2} = -\frac{n}{\theta_2} + \frac{1}{\theta_2^3}\sum_{i=1}^{n}\left(X_i - \theta_1\right)^2 = \frac{1}{\theta_2^3}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 + \frac{n}{\theta_2^3}\left(\overline{X} - \theta_1\right)^2 - \frac{n}{\theta_2}$$

وعلى ذلك فإن:

$$\frac{1}{L}\left(\frac{\theta_2^3}{n-1}\frac{\partial L}{\partial \theta_2} - \frac{\theta_2^4}{n(n-1)}\frac{\partial L}{\partial \theta_1^2}\right) = \frac{\theta_2^3}{n-1}\left[\frac{1}{\theta_2^3}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\right.$$

$$\left. + \frac{n}{\theta_2^3}\left(\overline{X} - \theta_1\right)^2 - \frac{n}{\theta_2}\right] - \frac{\theta_2^4}{n(n-1)}\left[-\frac{n}{\theta_2^2} + \frac{n^2}{\theta_2^2}\left(\overline{X} - \theta_1\right)^2\right]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 + \frac{n}{n-1}\left(\overline{X} - \theta_1\right)^2 - \frac{n}{n-1}\theta_2^2 + \frac{1}{n-1}\theta_2^2 - \frac{n}{n-1}\left(\overline{X} - \theta_1\right)^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 - \theta_2^2 = S^{*2} - \theta_2^2$$

وهذا يعني، بناءً على المعيار (18.7.5)، أن الإحصاء $T^* = S^{*2}$ المقدِّر الأكفأ لـ θ_2^2 في النموذج $N\left(\theta_1, \theta_2^2\right)$.

هكذا، في مسألة تقدير الدالة المعلمية $\tau(\theta) = \theta_2^2$ من أجل النموذج $N\left(\theta_1, \theta_2^2\right)$ لدينا معيار أكثر دقة من (15.7.5) للتقدير من أجل تباينات المقدِّرات غير المتحيزة لـ $\tau(\theta)$، وهذا المعيار هو:

$$V_\theta T \geq \frac{2\theta_2^4}{n-1}$$

وتتحقق المساواة عندما $T^* = S^{*2}$.

8.5 الكفاية، التمام والمقدِّرات الأكفأ

SUFFICIENCY, COMPLETENESS AND EFFICIENT ESTIMATORS

إن معايير الكفاءة للمقدِّرات الواردة سابقاً، لها تطبيقات محدودة نسبياً، لأنها تتطلب شروط قاسية تتمثل في نظامية النموذج، هذا بالإضافة إلى أنها في أفضل حالة تمكننا من إيجاد المقدِّرات الأكفأ من أجل دوال معلمية خاصة $\tau(\theta)$. فمثلاً، بينا أن المقدِّر الأكفأ من أجل التباين θ_2^2 في النموذج $N(\theta_1,\theta_2^2)$ هو عبارة عن $T(X)=S^{*2}$، إلا أن إيجاد المقدِّر الأكفأ للانحراف المعياري θ_2 باستخدام الطرق الواردة سابقاً غير ممكن. ويعتبر استخدام الإحصاءات الكافية الأسلوب الأنجع لبناء المقدِّرات الأكفأ، وبحثها يشكل موضوع هذا البند.

مبرهنة 1.8.5

إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $f(x;\theta)$ و $T=T(X)$ المقدِّر الأمثل لـ $\tau(\theta)$، حيث θ وحيدة البعد، فإن T إحصاء كافي لـ $\tau(\theta)$.

الإثبات

بما أن T المقدِّر الأمثل لـ $\tau(\theta)$، فهو يحقق العلاقة (2.7.5):

$$T(x)-\tau(\theta)=a(\theta)U(x;\theta)=a(\theta)\frac{\partial \ln L(x;\theta)}{\partial\theta}$$

وبمكاملة الطرفين بالنسبة لـ θ نجد:

$$\ln L(x;\theta)=T(x)\int\frac{d\theta}{a(\theta)}-\int\frac{\tau(\theta)}{a(\theta)}d\theta=\varphi\big(T(x);\theta\big)-\psi(\theta)+c(x)$$

حيث أن:

$$\varphi\big(T(x);\theta\big) = T(x)\int\frac{d\theta}{a(\theta)} \quad , \quad \psi(\theta) = \int\frac{\tau(\theta)}{a(\theta)}d\theta$$

وعلى ذلك فـإن:

$$L(x;\theta) = e^{\varphi(t;\theta)-\psi(\theta)+c(x)}$$

وبوضـع:

$$g(t;\theta) = e^{\varphi(t;\theta)-\psi(\theta)} \quad , \quad h(x) = c(x)$$

نجد أن:

$$L(x;\theta) = g(t;\theta)h(x)$$

أي أن العلاقة (2.2.4) محققة، وبالتـالي فـالمقدِّر الأمثـل $T = T(X)$ هـو إحصـاء كافي لـ $\tau(\theta)$.

هكذا المقدِّر الأمثل لدالة معلمية ما $\tau(\theta)$ موجود فقط عندمـا يكـون الإحصـاء الكافي لـ $\tau(\theta)$ موجوداً. لكن العكس ليس بـالضرورة صحيحـاً، أي يمكـن وجـود إحصاء كافي لـ $\tau(\theta)$ حتى في حالة عدم وجـود المقـدِّر الأمثـل لـ $\tau(\theta)$، أي أن شرط الإحصاء الكافي أقل تحديداً من شرط المقدِّر الأمثل. وهـذا يعـني للبحـث عـن المقدِّر الأمثل ضمن العائلة U_τ يجب اقتصار البحـث ضمـن عائلـة المقـدِّرات غـير المتحيزة والتي تعتبر في نفس الوقت إحصاءات كافيـة لـ $\tau(\theta)$، أي البحـث ضمـن عائلة أضيف من U_τ ولنرمز لها بـ U_τ^* $\left(U_\tau^* \subset U_\tau\right)$.

1.8.5 الكفاية والمقدِّرات الأكفـأ

Sufficiency and Efficient Estimators

إن للإحصاءات الكافيـة دوراً هامـاً في إيجـاد المقـدِّرات الأكفـأ (الأمثـل إن وجدت). وسنرى في هـذه الفقـرة كيـف أن الكفايـة (sufficiency) تسـاعد في الحصول على مثل هذه المقـدّرات.

مبرهنة 2.8.5: مبرهنة راو وبلاكويل Rao-Blackwell Theorem

إذا كان المقدِّر الأكفأ لدالة ما $\tau(\theta)$ موجوداً، فهو دالة في إحصاء كافي لهذه الدالة.

الإثبات

ليكن $T = T(X)$ إحصاءً كافياً و $T_1 = T_1(X)$ مقدِّر غير متحيز لدالة معلمية معطاة $\tau(\theta)$. لنرى الدالة:

$$H(t) = E_\theta\left(T_1 \middle| T = t\right) = \int T_1(x) L\left(x \middle| t;\theta\right) dx$$

وبما أن $L\left(x \middle| t;\theta\right)$ لا تعتمد على θ (لأن T إحصاء كافي لـ $\tau(\theta)$)، فالدالة $H(t)$ لا تعتمد على θ، أي أن الدالة $H(T)$ إحصاء. ويمكن أيضاً إثبات أن هذه الدالة هي عبارة عن مقدِّر غير متحيز لتقدير $\tau(\theta)$ على النحو الآتي:

إذا رمزنا بـ $g(t;\theta)$ لكثافة توزيع الإحصاء $T(X)$، فإن:

$$\begin{aligned}
E_\theta\left[H(T)\right] &= \int H(t) g(t;\theta) dt = \\
&= \int T_1(x)\left[\int L\left(x \middle| t;\theta\right) g(t;\theta) dt\right] dx \\
&= \int T_1(x) L(x;\theta) dx = E_\theta T_1(X) = \tau(\theta)
\end{aligned}$$

لنبرهن الآن صحة المتباينة:

$$V_\theta\left[H(T)\right] \le V_\theta T_1(X) \quad ; \quad \forall\, \theta \in \Theta$$

وأن المساواة محققة فقط عندما $T_1(X) = H(T)$. وذلك على النحو الآتي:

$$\begin{aligned}
V_\theta T_1(X) = E_\theta\left[T_1(X) - \tau(\theta)\right]^2 &= E_\theta\left[\left(T_1 - H(T)\right) + \left(H(T) - \tau(\theta)\right)\right]^2 \\
&= E_\theta\left[T_1 - H(T)\right]^2 + E_\theta\left[H(T) - \tau(\theta)\right]^2 \\
&\quad + 2E_\theta\left[\left(T_1 - H(T)\right)\left(H(T) - \tau(\theta)\right)\right]
\end{aligned}$$

بما أن:

$$E_\theta\left[(T_1 - H(T))(H(T) - \tau(\theta))\right] = E_\theta\left[(T_1 - H(T))H(T)\right] =$$
$$= \int\left[E_\theta\left(T_1 | T = t\right) - H(t)\right]H(t)g(t;\theta)dt = 0$$

فإن:

$$V_\theta T_1(X) = E\left[T_1 - H(T)\right]^2 + V_\theta H(T) \geq V_\theta H(T)$$

وتتحقق المساواة عندما $T_1 = H(T)$.

تنص هذه المبرهنة على أنه إذا كان لدينا مقدّر T_1 غير متحيز لـ $\tau(\theta)$، وليس دالة في إحصاء كافي، فيمكن إيجاد مقدّر غير متحيز لـ $\tau(\theta)$، دالة في إحصاء كافي وتباينه أقل من تباين T_1. وهذا المقدّر الأخير هو عبارة عن القيمة المتوقعة المشروطة $H(T) = E\left(T_1 | T\right)$. وبالتالي المقدّر الأكفأ (إذا وجد) لدالة معلمية $\tau(\theta)$ فهو حكماً دالة في إحصاء كافي لـ $\tau(\theta)$. وهذا يعني تطبيقياً، عند البحث عن المقدّر الأكفأ لدالة معلمية ما يجب البحث ضمن عائلة المقدّرات $T \in U_\tau^*$، التي تعتبر دوال في إحصاءات كافية لـ $\tau(\theta)$.

مثال 1.8.5

إذا كانت $X = \left(X_1, \ldots, X_n\right)$ عينة عشوائية من توزيع بواسون:

$$f(x;\theta) = \frac{\theta^x}{x!}e^{-\theta} \quad , \quad x = 0,1,\ldots \quad ; \quad \theta > 0$$

وكان $T_1 = X_1$ مقدّر غير متحيز لـ θ و $T = \sum_{i=1}^{n} X_i$ إحصاء كافي لـ θ، فأوجد مقدّراً لـ θ أفضل من T_1.

حسب مبرهنة راو وبلاكويل (2.8.5) فالمقدّر:

$$H(T) = E_\theta\left(T_1 | T\right)$$

أفضل من المقدّر T_1 للمعلمة θ. ولتعيينه يلزمنا معرفة التوزيع المشروط للمتغير T_1 بشرط معرفة T.

$$P_\theta\left(T_1 = x_1 \middle| T(X) = k\right) = \frac{P_\theta\left(T_1 = x_1, T(X) = k\right)}{P_\theta\left(T(X) = k\right)}$$

$$= \frac{P_\theta\left(X_1 = x_1, \sum_2^n X_i = k - x_1\right)}{P_\theta\left(\sum_1^n X_i = k\right)}$$

$$= \frac{P_\theta\left(X_1 = x_1\right) P_\theta\left(\sum_2^n X_i = k - x_1\right)}{P_\theta\left(\sum_1^n X_i = k\right)}$$

$$= \frac{\left(e^{-\theta}\theta^{x_1}/x_1!\right)\left\{\left[(n-1)\theta\right]^{k-x_1} e^{-(n-1)\theta}/(k-x_1)!\right\}}{(n\theta)^k e^{-n\theta}/k!}$$

$$= \frac{k!}{x_1!(k-x_1)!}\left(\frac{1}{n}\right)^{x_1}\left(1-\frac{1}{n}\right)^{k-x_1} = C_k^{x_1}\left(\frac{1}{n}\right)^{x_1}\left(1-\frac{1}{n}\right)^{k-x_1}$$

وهذا ما هو إلا توزيع ذي الحدين $B(k, 1/n)$. وبالتالي فإن:

$$H(T) = E\left(T_1 \middle| T\right) = \frac{k}{n} = \frac{1}{n}\sum_1^n X_i = \overline{X}$$

ونلاحظ بسهولة أن:

$$E_\theta\left[H(T)\right] = E\overline{X} = \theta$$

$$V_\theta\left[H(T)\right] = V_\theta \overline{X} = \frac{\theta}{n} \le V_\theta X_1 = \theta$$

أي أن المقدّر غير المتحيز $\overline{X}$ أفضل من المقدّر غير المتحيز X_1 من أجل تقدير θ.

نخلص مما سبق إلى الملاحظات الآتية:

1. إذا كان المقدّر غير المتحيز T_1 لدالة ما $\tau(\theta)$ هو أصلاً دالة في إحصاء كافي، فإن المقدّر $H(T)=E_\theta\left(T_1|T\right)$ يكون مطابقاً للمقدّر T_1، أي لا نتوقع أن يكون هناك أي تحسن في قيمة التباين.

2. نستطيع استخدام أي إحصاء كافي للحصول على المقدّر $H(T)$، ولكن يفضل استخدام الإحصاء الكافي الأصغر.

3. مبرهنة راو وبلاكويل تبين كيفية الحصول على مقدّر غير متحيز أفضل من المقدّر غير المتحيز المفروض، وذلك بالاشتراط على الإحصاء الكافي الأصغر للدالة المقدَّرة $\tau(\theta)$.

إن للإحصاء الكافي التام (الإحصاء التام) دوراً هاماً عند البحث عن شكل واضح للمقدّرات الأكفأ، وهذا ما تبينه المبرهنة الآتية.

مبرهنة 3.8.5: مبرهنة ليمان وشيفي Lemann-Scheffe Theorem

إذا وجد إحصاء كافي تام لمعلمة نموذج $f(x;\theta)$، فإن أي دالة فيه تعتبر المقدّر الأكفأ لقيمتها المتوقعة.

الإثبات

إذا كانت $X=\left(X_1,...,X_n\right)$ عينة عشوائية من توزيع $f(x;\theta)$، $T=T(X)$ إحصاء كافي تام لـ θ وكان $H(T)$ دالة ما في T، ورمزنا بـ:

$$E_\theta\left[H(T)\right]=\tau(\theta) \qquad (1.8.5)$$

فتعني المبرهنة أن $H(T)$ المقدّر الأكفأ من أجل تقدير $\tau(\theta)$.

ينتج من تعريف الإحصاء التام أن $H(T)$ الدالة الوحيدة في T التي تحقق العلاقة (1.8.5)، لأنه إذا وجدت دالة ما أخرى $H_1(T)$ وتحقق العلاقة (1.8.5)، فإن:

$$E_\theta\left[H(T)-H_1(T)\right]=0 \quad ; \quad \forall\, \theta \in \Theta$$

وحسب التعريف (1.5.4) نجد:

$$H(t) \equiv H_1(t)$$

لإيجاد المقدِّر الأكفأ لـ $\tau(\theta)$، بناءً على المبرهنة (2.8.5)، يجب البحث في صف الدوال في T. لكن الإحصاء $H(T)$ الدالة الوحيدة في T كما بينا سابقاً، الذي يعتبر مقدِّر غير متحيز لـ $\tau(\theta)$، وبالتالي فإنه المقدِّر الأكفأ المطلوب.

هكذا، لنفترض وجود إحصاء كافي تام T للنموذج المدروس $f(x;\theta)$ ويطلب تقدير دالة معلمية معطاة $\tau(\theta)$. عندها:

1. إذا وجد مقدِّر ما غير متحيز لـ $\tau(\theta)$، فيوجد أيضاً مقدِّر آخر غير متحيز لـ $\tau(\theta)$، الذي يعتبر دالة في T. ويمكن القول أيضاً، أنه إذا لم يوجد مقدِّر غير متحيز لـ $\tau(\theta)$ من الشكل $H(T)$ [أي العلاقة (1.8.5) ليس لها حل]، فإن عائلة المقدِّرات غير المتحيزة U_τ لـ $\tau(\theta)$ خالية.
2. المقدِّر الأكفأ (إذا وجد) يعتبر دائماً دالة في T، ويعين بشكل وحيد بالعلاقة (1.8.5).
3. المقدِّر الأكفأ τ^* يمكن إيجاده وفق الصيغة:

$$\tau^* = H(T) = E_\theta\left(T_1 \middle| T\right)$$

انطلاقاً من أي مقدِّر غير متحيز T_1 للدالة $\tau(\theta)$.

مثال 2.8.5

إذا كانت $X = \left(X_1,..,X_n\right)$ عينة عشوائية مأخوذة من مجتمع يخضع للتوزيع $R(0,\theta)$، ونرغب في إيجاد المقدِّر الأكفأ للمعلمة θ.

كما نعلم أن $T = X_{(n)}$ إحصاء كافي تام لـ θ [راجع مثال (1.5.4)]. وحسب العلاقة (3.10.3) فإن توزيع $X_{(n)}$ هو:

$$h(y) = \frac{n}{\theta^n} x^{n-1} \quad , \quad 0 < x < \theta \quad , \quad \theta > 0$$

وعلى ذلك فإن:

$$E_\theta X_{(n)} = \int_0^\theta \frac{n}{\theta^n} y^n dy = \frac{n}{n+1}\theta$$

وإذا رمزنا بـ:

$$H(T) = \frac{n+1}{n} X_{(n)}$$

فإن:

$$E_\theta[H(T)] = E_\theta\left[\frac{n+1}{n} X_{(n)}\right] = \theta$$

وحسب المبرهنة (3.8.5) فإن $H(T)$ المقدّر الأكفأ لـ θ. ويمكن بسهولة إثبات أن:

$$V_\theta[H(T)] = \frac{\theta^2}{n(n+2)}$$

مثال 3.8.5

لتكن $X = (X_1, \dots, X_n)$ عينة عشوائية من مجتمع يتبع التوزيع $\Gamma(1,1/\theta)$، ويطلب إيجاد المقدّر الأمثل لكل من θ و $1/\theta$.

بما أن التوزيع $\Gamma(1,1/\theta)$ ينتمي لعائلة النماذج الأسية فإن الإحصاء $T = \sum_{i=1}^{n} X_i$ يعتبر إحصاءً كافياً لـ θ [راجع المثال (1.4.4)]، ويمكن بسهولة إثبات أنه تام أيضاً.

لإيجاد المقدّر الأكفأ لـ θ (إن وجد) نبحث عن دالة ما في الإحصاء الكافي التام $T = \sum_{i=1}^{n} X_i$ توقعها الرياضي θ، ونلاحظ بسهولة أن ذلك محقق في الدالة:

$$H_1(T) = \frac{n-1}{\sum_{i=1}^{n} X_i} = \frac{n-1}{T}$$

وبالتالي $H(T)$ المقدّر الأكفأ للمعلمة θ، على الرغم من أنه ليس المقدّر الأمثل [راجع المثال (6.7.5)]، وهذا يعني أن المقدّر الأكفأ ليس بالضرورة أمثل لـ $\tau(\theta)$.

وبشكل مشابه نجد أن الإحصاء:

$$H_2(T)=\frac{T}{n}=\overline{X}$$

هو المقدّر الأكفأ لـ $1/\theta$.

تمارين

1. بفرض $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $B(1,\theta)$:

   1. عين متوسط مربع الخطأ، ومن ثم مقدار التحيز (إن وجد) لكل من المقدّرات الآتية لتقدير المعلمة θ.

   $$T_1=\overline{X}\quad,\quad T_2=\frac{X_1+X_2}{2n}\quad,\quad T_3=\frac{X_1+X_2}{2}$$

   2. ما هو المقدّر الأفضل من بين المقدّرات غير المتحيزة في (1).

2. لتكن $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $N(\theta,4)$:

   1. عين متوسط مربع الخطأ، ومن ثم مقدار التحيز للمقدّر $T_1=2\overline{X}$ لتقدير المعلمة θ.

   2. هل $T_2=\frac{n}{n+1}\overline{X}^2+X_1$ مقدّر غير متحيز لـ $\tau(\theta)=\theta(1+\theta)$؟

3. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $f(x;\theta)$ بمتوسط θ وتباين σ^2 محدود، فأثبت أن المقدّر:

$$T=\frac{2}{n(n+1)}\sum_{i=1}^{n} iX_i$$

غير متحيز ومتسق لتقدير المعلمة θ.

4. إذا كانت $\boldsymbol{X}=(X_1,\dots,X_n)$ عينة عشوائية من توزيع $B(1,\theta)$، فبين أي من المقدّرات الآتية لتقدير المعلمة θ متسق:

$$T_1=2\overline{X}-X_1 \quad , \quad T_2=\frac{X_1+X_2}{2} \quad , \quad T_3=\overline{X}$$

5. إذا كانت $\boldsymbol{X}$ عينة عشوائية بحجم $n=2$ من توزيع $N(\theta,1)$، وكانت:

$$T_1=\frac{2}{3}X_1+\frac{1}{3}X_2 \quad , \quad T_2=\frac{1}{5}X_1+\frac{4}{5}X_2 \quad , \quad T_3=\overline{X}$$

فبين المقدّر الأفضل منهم.

6. إذا كانت $\boldsymbol{X}=(X_1,\dots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\frac{1}{\theta+2}e^{-x/(\theta+2)} \quad ; \quad x<0$$

فأوجد الحد الأدنى لتباين المقدّرات غير المتحيزة لتقدير θ.

7. لتكن $\boldsymbol{X}=(X_1,\dots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\theta(1+x)^{-(1+\theta)} \quad ; \quad x>0\ ,\ \theta>0$$

1. أوجد إحصاء كافي تام لـ θ (إن وجد).
2. أوجد الحد الأدنى لمتباينة كرامر وراو من أجل المقدّرات غير المتحيزة لتقدير $\tau(\theta)=1/\theta$.
3. أوجد المقدّر الأكفأ لتقدير $\tau(\theta)=1/\theta$ (إن وجد).

8. لتكن $\boldsymbol{X}=(X_1,\dots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\frac{\ln\theta}{\theta-1}\theta^{x} \quad ; \quad 0<x<1\ ,\ \theta>1$$

1. أوجد إحصاء كافي تام لـ θ (إن وجد).
2. أوجد دالة ما في θ يوجد لها المقدّر الأكفأ.

9. لتكن X عينة عشوائية بحجم n من توزيع:

$$f(x;\theta)=e^{-(x-\theta)} \quad ; \quad x\geq\theta\ ,\ -\infty<\theta<+\infty$$

1. هل يوجد إحصاء كافي تام لـ θ؟ إن كان الجواب بنعم فأوجده.
2. أوجد المقدّر الأكفأ لتقدير θ (إن وجد).
3. أوجد دالة ما $\tau(\theta)$ يوجد لها المقدّر الأكفأ.

10. لتكن $X=(X_1,...,X_n)$ عينة عشوائية من توزيع $f(x;\theta)$ بمتوسط θ وتباين σ^2.

1. أثبت أن $\sum_{i=1}^{n}a_i=1$; $T(X)=\sum_{i=1}^{n}a_iX_i$ مقدّرات غير متحيزة لتقدير المعلمة θ.
2. عين قيمة الثوابت $a_1,...,a_n$ بحيث يكون المقدّر $T(X)$ ذو تباين أصغري ضمن مجموعة المقدّرات (1).

11. لتكن $X=(X_1,...,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\frac{2x}{\theta^2} \quad ; \quad 0<x<\theta\ ,\ \theta>0$$

1. هل $Y_n=X_{(n)}$ إحصاء تام لـ θ؟
2. هل يوجد المقدّر الأكفأ لـ θ؟ إن كان الجواب بنعم فأوجده.

12. لتكن $X=(X_1,...,X_n)$ عينة عشوائية من توزيع $\Pi(\theta)$.

1. هل يوجد مقدّر غير متحيز لتقدير $\tau_1(\theta)=\frac{1}{\theta}$.
2. أوجد مقدّر غير متحيز لتقدير $\tau_2(\theta)=(1+\theta)e^{-\theta}$.
3. أوجد المقدّر الأكفأ لتقدير $\tau_2(\theta)$ (إن وجد).

13. إذا كانت $X=(X_1,\dots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\theta x^{-2} \quad ; \quad x\ge 0\ ,\ \theta>0$$

فأوجد دالة ما $\tau(\theta)$ يوجد لها المقدّر الأكفأ.

14. لنفترض بناءً على ملاحظة واحدة على متغير عشوائي ع يتبع توزيع:

$$f(x;\theta)=e^{-\theta}\ \theta^{x}/x! \Big/ \left(1-e^{\theta}\right) \quad ; \quad x=1,2,\dots$$

يطلب تقدير الدالة $\tau(\theta)=1-e^{-\theta}$. أوجد الإحصاء $T(X_1)$، الذي يحقق شرط عدم التحيز $E_\theta T(X_1)=1-e^{-\theta}$، وأثبت أنه غير مفيد تطبيقياً.

15. إذا كانت $X=(X_1,\dots,X_n)$ عينة عشوائية من توزيع $\Gamma(\alpha,\theta)$، فأثبت أن الإحصاء $T(X)=\frac{1}{n}\sum_{i=1}^{n}\ln X_i-\ln\alpha$ يعتبر المقدّر الأكفأ لتقدير المعلمة $\tau(\theta)=\frac{\Gamma'(\theta)}{\Gamma(\theta)}$، وأن تباينه $V_\theta T=\frac{\tau'(\theta)}{n}$.

16. لتكن $X=(X_1,\dots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=e^{-x+\theta}\left(1+e^{-x+\theta}\right) \quad ; \quad -\infty<x<\infty$$

1. أثبت أن الإحصاء $T(X)=\overline{X}$ مقدّر غير متحيز لتقدير θ، وأن تباينه $V_\theta T=\frac{\pi^2}{3n}$.
2. أثبت أن دالة المعلومات $i(\theta)=\frac{1}{3}$.
3. أثبت أن $\overline{X}$ لا يعتبر المقدّر الأكفأ لتقدير المعلمة θ.

الفصل السادس:

طرق إيجاد مقدرات النقطة

1.6 مقدمة

تطرقنا في الفصل السابق لمفهوم التقدير النقطي، وبينا الخواص التي نريد لمقدِّر أن يتصف بها. وتعرضنا بشيء من التفصيل لتلك الصفات وأظهرنا أن المقدِّر الجيد بشكل عام هو الذي تتوفر فيه الصفات الآتية:

1. الكفاية .
2. عدم التحيز .
3. الإتساق .
4. الكفاءة .

وهناك عدة طرق تعطي مقدِّرات نقطية تتصف بكل أو بعض تلك الصفات. وفي هذا الفصل سندرس بعض الطرق الشائعة الاستخدام في الإحصاء للحصول على المقدِّرات.

2.6 طريقة المعقولية العظمى
MAXIMUM LIKELIHOOD METHOD

في مقالتين نشرتا عام 1920م، عرض فيهما العالم فيشر (R. A. Fisher) أحد رواد الإحصاء في عصرنا طريقة عامة للتقدير أطلق عليها اسم "طريقة المعقولية العظمى" أو "الإمكان الأكبر"، كما بين مميزات هذه الطريقة. تعتبر طريقة

المعقولية العظمى إحدى أهم وأكثر الطرق انتشاراً في الإحصاء لتقدير معالم التوزيع الاحتمالي الإحصائي (النموذج الإحصائي) المقترح.

إذا كانت $X = (X_1,\ldots,X_n)$ عينة عشوائية من مجتمع توزيعه الاحتمالي $f(x;\theta)$، فإننا سنرمز لمقدّر المعلمة θ الذي نحصل عليه باستخدام طريقة المعقولية العظمى بـ $\hat{\theta}(X)$، كما سنرمز لمقدّر دالة معلمية $\tau(\theta)$ بـ $\hat{\tau} = \hat{\tau}(X)$، ومن ثم للتقدير بنقطة الموافق لعينة مشاهدة $x = (x_1,\ldots,x_n)$ بـ $\hat{\theta} = \hat{\theta}(x)$ و $\hat{\tau} = \hat{\tau}(x)$ على الترتيب.

لتكن $X = (X_1,\ldots,X_n)$ عينة عشوائية من التوزيع:

$$\mathcal{L}(\xi) \in \boldsymbol{F}\{F(x;\theta)\,;\,\theta \in \Theta\}$$

و $L(x;\theta)$ دالة المعقولية من أجل قيمة ملاحظة $x = (x_1,\ldots,x_n)$ للعينة X.

تعريف 1.2.6

تدعى القيمة (النقطة) $\hat{\theta}(x) \in \Theta$ بتقدير المعقولية العظمى (maximum likelihood estimating) للمعلمة θ، إذا كانت دالة المعقولية $L(x;\theta)$ تبلغ نهايتها العظمى عندها، أي أن:

$$L(x;\hat{\theta}) \geq L(x;\theta) \quad ; \quad \forall\, \theta \in \Theta \tag{1.2.6}$$

أو:

$$L(x;\hat{\theta}) = \sup_{\theta \in \Theta} L(x;\theta)$$

وهذا يعني أن التقدير $\hat{\theta}$ في حالة التوزيعات المنقطعة عبارة عن قيمة θ التي تجعل احتمال سحب العينة المشاهدة x (قيمة لـ X) أكبر ما يمكن. أما بالنسبة للتوزيعات المستمرة فإن قيمة θ التي تجعل احتمال الحصول على مفردات العينة X قريبة جداً من القيم التي حصلنا عليها (القيم $x_1,\ldots,x_n$) أكبر ما يمكن.

ويمكن تفسير ما سبق على النحو الآتي:

إذا فرضنا أن θ معلومة ولتكن θ_0، فإن قيمة X التي يكون وقوعها أو تحققها أكثر احتمالاً هي تلك القيمة ولنرمز لها بـ $x' = (x_1',\ldots,x_n')$ التي تجعل الدالة $L(x;\theta_0)$ أعظم ما يمكن. وعندما تأخذ θ قيمها الممكنة في فضاء المعلمة Θ فإنها تعرف من أجل كل قيمة $\theta \in \Theta$ دالة كثافة (أو دالة احتمال) معينة وبالتالي مجتمعاً معيناً. وبعد الحصول فعلاً على قيمة مشاهدة $x = (x_1,\ldots,x_n)$ للعينة $X = (X_1,\ldots,X_n)$، فإننا نرغب في معرفة ذلك المجتمع الذي يمكن أن يعطي مثل تلك العينة المشاهدة x بأكبر احتمال ممكن، أي نريد إيجاد قيمة لـ θ من Θ (نرمز لها بـ $\hat{\theta}$) تجعل دالة المعقولية $L(x;\theta)$ عندها في نهايتها العظمى. وقيمة $\hat{\theta}$ هذه ستكون بشكل عام دالة في $x = (x_1,\ldots,x_n)$ تسمى تقدير المعقولية العظمى (maximum likelihood estimate) و $\hat{\theta}(X)$ بمقدّر المعقولية العظمى (maximum likelihood estimator).

إذا كان من أجل كل قيمة $x \in G$ دالة المعقولية $L(x;\theta)$ تبلغ نهايتها العظمى عند نقطة داخلية من فضاء المعلمة Θ وقابلة للاشتقاق بالنسبة لـ θ، فإن تقدير المعقولية العظمى $\hat{\theta} = \hat{\theta}(x)$ يحقق المعادلة:

$$\frac{\partial L(x;\theta)}{\partial\theta} = 0 \quad \text{أو} \quad \frac{\partial \ln L(x;\theta)}{\partial\theta} = 0 \tag{2.2.6}$$

بالإضافة إلى المتباينة:

$$\frac{\partial^2 L(x;\theta)}{\partial\theta^2} < 0 \quad \text{أو} \quad \frac{\partial^2 \ln L(x;\theta)}{\partial\theta^2} < 0 \tag{3.2.6}$$

وهذا يعني، أننا نحصل على التقدير $\hat{\theta}$ بحل المعادلة (2.2.6) بعد التأكد من تحقق المتباينة (3.2.6).

أما إذا كانت المعلمة θ متعددة الأبعاد، أي $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_r)$، فإن التقدير $\hat{\boldsymbol{\theta}}$ نحصل عليه بحل منظومة المعادلات:

$$\frac{\partial L(x;\theta)}{\partial \theta_i} = 0 \quad ; \quad i = 1,2,\dots,r \tag{4.2.6}$$

أو

$$\frac{\partial \ln L(x;\theta)}{\partial \theta_i} = 0 \quad ; \quad i = 1,2,\dots,r$$

التي تدعى بمعادلات المعقولية.

لنرى الآن بعض الأمثلة لإيجاد مقدّرات المعقولية العظمى.

مثال 1.2.6

إذا كانت $X = (X_1,\dots,X_n)$ عينة عشوائية من مجتمع كثافة توزيعه الاحتمالية:

$$f(x;\theta) = \theta e^{-\theta x} \quad ; \quad x > 0 \ , \ \theta > 0$$

فأوجد مقدّر المعقولية العظمى للمعلمة θ.

بما أن:

$$L(x;\theta) = \theta^n e^{-\theta \sum x_i}$$

فإن:

$$\ln L(x;\theta) = n \ln \theta - \theta \sum x_i$$

$$\frac{\partial \ln L(x;\theta)}{\partial \theta} = \frac{n}{\theta} - \sum x_i$$

وبوضع $\frac{\partial \ln L(x;\theta)}{\partial \theta} = 0$ نجد:

$$\frac{n}{\theta} - \sum x_i = 0$$

وبحل هذه المعادلة بالنسبة لـ θ نحصل على تقدير المعقولية العظمى:

$$\hat{\theta}=\hat{\theta}(x)=\frac{n}{\sum x_i}=\frac{1}{\bar{x}}$$

ومن ثم فإن مقدّر المعقولية العظمى للمعلمة θ هو:

$$\hat{\theta}(X)=\frac{1}{\overline{X}}$$

مثال 2.2.6

إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع بواسون $\Pi(\theta)$، فأوجد مقدّر المعقولية العظمى للمعلمة θ.

بما أن:

$$f(x;\theta)=\frac{\theta^{x}}{x_i\,!}e^{-\theta}\qquad ;\qquad x=0,1,2,\ldots$$

فإن:

$$L(x;\theta)=\frac{\theta^{\sum x_i}}{\prod x_i\,!}e^{-n\theta}$$

$$\ln L(x;\theta)=-n\theta+\sum x_i\ln\theta-\sum\ln(x_i\,!)$$

وبالتالي:

$$\frac{\partial\ln L(x;\theta)}{\partial\theta}=-n+\frac{\sum x_i}{\theta}$$

بوضع $\frac{\partial\ln L(x;\theta)}{\partial\theta}=0$ وحلها بالنسبة لـ θ نجد:

$$-n+\frac{\sum x_i}{\theta}=0$$

$$\hat{\theta}=\hat{\theta}(x)=\frac{\sum x_i}{n}=\bar{x}$$

ومقدّر المعقولية العظمى لـ θ يكون:

$$\hat{\theta}(\boldsymbol{X}) = \overline{X}$$

ونلاحظ بسهولة أن $\overline{X}$ مقدّر غير متحيز لـ θ.

مثال 3.2.6

لتكن $\boldsymbol{X} = (X_1, \ldots, X_n)$ عينة عشوائية من توزيع $N(\mu, \theta^2)$ ونرغب في إيجاد مقدّر المعقولية العظمى لـ θ^2.

كما نعلم:

$$f(x;\theta) = \frac{1}{\theta\sqrt{2\pi}} e^{-\frac{1}{2\theta^2}(x-\mu)^2}$$

وعلى ذلك:

$$L(\boldsymbol{x};\theta) = \left(2\pi\theta^2\right)^{-\frac{n}{2}} e^{-\frac{1}{2\theta^2}\sum(x_i-\mu)^2}$$

$$\ln L(\boldsymbol{x};\theta) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\theta^2 - \frac{1}{2\theta^2}\sum\left(x_i - \mu\right)^2$$

$$\frac{\partial \ln L(\boldsymbol{x};\theta)}{\partial\theta} = -\frac{n}{2\theta^2} + \frac{1}{2\theta^4}\sum_{i=1}^{n}\left(x_i - \mu\right)^2 = 0$$

وبحل هذه المعادلة بالنسبة لـ θ^2 نجد:

$$\hat{\theta}^2(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \mu\right)^2$$

ومن ثم فمقدّر المعقولية العظمى:

$$\hat{\theta}^2(\boldsymbol{X}) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \mu\right)^2$$

وهو مقدّر غير متحيز لـ θ^2.

مثـــــال 4.2.6

إذا كـــانت $X=(X_1,\dots,X_n)$ عينة عشــوائية مـــأخوذة من التوزيع $N(\theta_1,\theta_2^2)$، فأوجد مقدّر المعقولية العظمى للمعلمة $\theta=(\theta_1,\theta_2^2)$.

بما أن:

$$f(x;\theta)=\frac{1}{\theta_2\sqrt{2\pi}}e^{-\frac{1}{2\theta_2^2}(x-\theta_1)^2} \quad ; \quad \theta=(\theta_1,\theta_2)$$

فإن:

$$L(x;\theta)=(2\pi)^{-\frac{n}{2}}\left(\theta_2^2\right)^{-\frac{n}{2}}e^{-\frac{1}{2\theta_2^2}\sum(x_i-\theta_1)^2}$$

$$\ln L(x;\theta)=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\theta_2^2-\frac{1}{2\theta_2^2}\sum_{i=1}^{n}(x_i-\theta_1)^2$$

$$=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\theta_2^2-\frac{n}{2\theta_2^2}s^2+\frac{n(\bar{x}-\theta_1)^2}{2\theta_2^2}$$

حيث أن $s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ القيمة الملاحظة لتباين العينة S^2.

وعلى ذلك نجـــد:

$$\frac{\partial\ln L(x;\theta)}{\partial\theta_1}=\frac{n(\bar{x}-\theta_1)}{\theta_2^2}=0$$

$$\frac{\partial\ln L(x;\theta)}{\partial\theta_2^2}=-\frac{n}{2\theta_2^2}+\frac{n}{2\theta_2^4}s^2+\frac{n(\bar{x}-\theta_1)^2}{2\theta_2^4}=0$$

وبحل جملة المعادلتين نحصل على تقديري المعقوليـــة العظمـــى:

$$\hat{\theta}_1=(x)=\bar{x} \quad , \quad \hat{\theta}_2^2(x)=s^2$$

ومن ثم مقدّري المعقولية العظمــــى:

$$\hat{\theta}_1=(X)=\bar{X} \quad , \quad \hat{\theta}_2^2(X)=S^2$$

إن المقدّر $\overline{X}$ لـ θ_1 غير متحيز بينما الإحصاء S^2 مقدّر لـ θ_2^2 متحيز. وبالتالي فإن مقدّر المعقولية العظمى لـ θ موجود ووحيد $\hat{\theta} = \left(\hat{\theta}_1, \hat{\theta}_1^2\right) = \left(\overline{X}, S^2\right)$. وهذا المقدّر دالة في الإحصاء الكافي $\boldsymbol{T} = \left(T_1, T_2\right)$، حيث أن $T_1 = \overline{X}$ و $T_2 = \sum_{i=1}^{n}\left(X_i - \overline{X}\right)$.

هكذا، فمقدّر المعقولية العظمى لدالة $\tau(\theta)$ يمكن أن يكون متحيزاً، أي ليس بالضرورة أن يكون مقدّر المعقولية العظمى غير متحيز.

مثال 5.2.6

بفرض $\boldsymbol{X} = \left(X_1, \ldots, X_n\right)$ عينة عشوائية من التوزيع المنتظم $R(0,\theta)$، فأوجد مقدّر المعقولية العظمى للمعلمة θ.

بما أن:

$$f(x;\theta) = \frac{1}{\theta} \quad ; \quad 0 < x \le \theta$$

فيمكن كتابتها على النحو الآتي [راجع مثال (11.2.4)].

$$f(x;\theta) = \frac{1}{\theta} \quad ; \quad \theta \ge x_{(n)} \;\; , \;\; x_{(n)} = \max x_i$$

وبالتالي فدالة المعقولية:

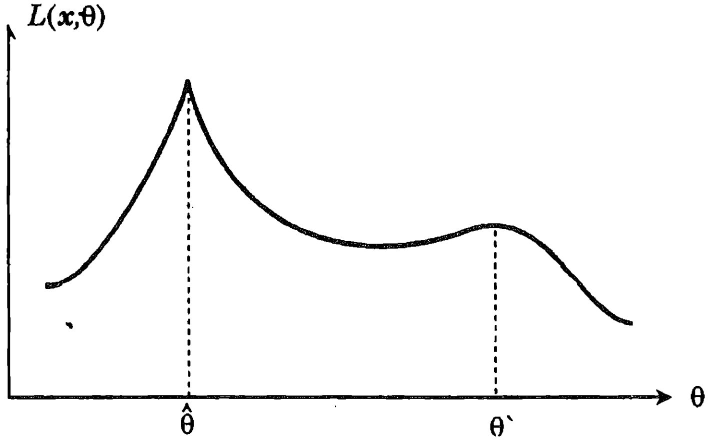
$$L(\boldsymbol{x};\theta) = \frac{1}{\theta^n} \quad ; \quad \theta \ge x_{(n)}$$

يعني هذا أن دالة المعقولية L متناقصة تماماً في θ. وبما أن θ لا يمكن أن تكون أصغر من $x_{(n)}$، فإن $L(\boldsymbol{x};\theta)$ تبلغ نهايتها العظمى عندما $\theta = x_{(n)}$. وبالتالي فمقدّر المعقولية العظمى لـ θ هو $\hat{\theta}(\boldsymbol{X}) = X_{(n)}$. ونلاحظ أن دالة المعقولية $L(\boldsymbol{x};\theta)$ غير مستمرة عند النقطة $\theta = x_{(n)}$، ومن ثم مشتقها عند هذه النقطة غير موجود، أي أن تقدير المعقولية العظمى لا يعتبر حلاً لمعادلة المعقولية:

$$\frac{\partial \ln L(x;\theta)}{\partial \theta} = 0$$

وهذه صفة مميزة للحالات التي لايكون فيها فضاء العينة G يعتمد على معلمة. لذا في مثل هذه الحالات يجب الحيطة وإتباع طرق أخرى (غير حل معادلة أو معادلات المعقولية) للحصول على مقدّر المعقولية العظمى، وسنتطرق لبعض هذه الطرق لاحقاً.

ولإيضاح تلك الميزة بيانياً، نفترض أن دالة المعقولية $L(x;\theta)$ يمكن تمثيلها بيانياً كما هي مبينة على الشكل (1.2.6)، حيث تعطي طريقة الاشتقاق النقطة θ' بينما التقدير المطلوب (التقدير الذي يجعل دالة المعقولية $L(x;\theta)$ تبلغ نهاية عظمى عنده) هو $\hat{\theta}$.



شكل (1.2.6)

مثال 6.2.6

إذا كانت $X = (X_1,...,X_n)$ عينة عشوائية من مجتمع توزيعه $R(\alpha,\beta)$، فأوجد مقدّر المعقولية العظمى للمعلمة $\theta = (\alpha,\beta)$.

بما أن:

$$f(x;\theta)=\begin{cases}1/(\beta-\alpha) & ; \quad \alpha<x<\beta \\ 0 & ; \quad x\le\alpha \text{ or } x\ge\beta\end{cases}$$

فإن:

$$L(x;\theta)=\frac{1}{(\beta-\alpha)^n} \quad ; \quad \alpha<x_i<\beta \ , \ i=1,\ldots,n$$

وباشتقاق L جزئياً بالنسبة لكل من α,β والمساواة بالصفر نحصل على معادلتين، وبحلهما نجد أحد المعلمتين α,β على الأقل غير محدود. وهذه النتيجة (الحل) غير مقبولة (غير منطقية)، أي تخفق طريقة الاشتقاق في إعطاء التقدير المطلوب. لكن نلاحظ بوضوح من صيغة دالة المعقولية $L(x;\theta)$ أنها تبلغ أكبر قيمة لها عندما يكون الفرق $(\beta-\alpha)$ أصغر ما يمكن، وهذا يتحقق عندما $\beta-\alpha=x_{(n)}-x_{(1)}$، وبالتالي مقدِّر المعقولية العظمى:

$$\hat{\theta}(X)=\left(X_{(1)},X_{(n)}\right)$$

مثال 7.2.6

إذا كانت $X=\left(X_1,\ldots,X_n\right)$ عينة عشوائية من توزيع:

$$f(x;\alpha,\beta)=\frac{1}{2\beta} \quad ; \quad \alpha-\beta\le x\le\alpha+\beta$$

فأوجد مقدِّر المعقولية العظمى للمعلمة $\theta=(\alpha,\beta)$.

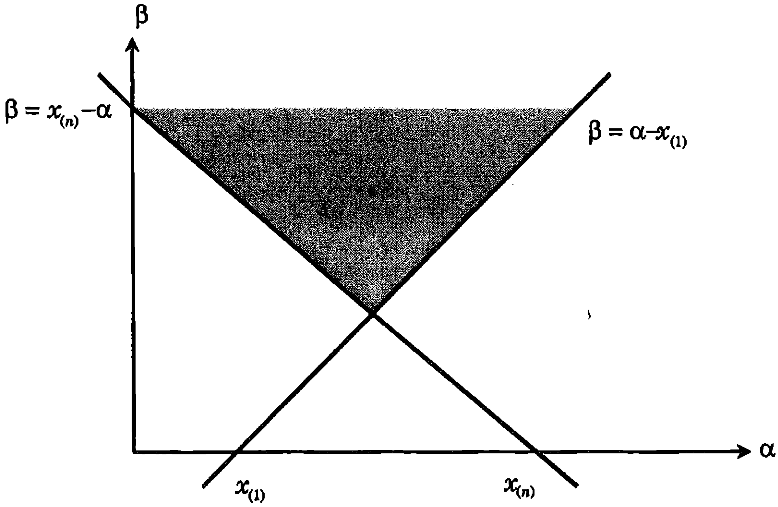
بما أن:

$$L(x;\theta)=\frac{1}{(2\beta)^n} \quad ; \quad \alpha-\beta\le x_i\le\alpha+\beta \ , \ i=1,\ldots,n$$

أي أن:

$$L(x;\theta)=\frac{1}{(2\beta)^n} \quad ; \quad \alpha-\beta\le x_{(1)}<x_{(n)}\le\alpha+\beta \ , \ i=1,\ldots,n$$

لإيجاد مقدّر المعقولية العظمى نستخدم الطريقة البيانية، كما هي مبينة على الشكل (2.2.6)



شكل (2.2.6)

نلاحظ أن دالة المعقولية $L(x;\theta)$ تساوي $(1/2\beta)^n$ في المنطقة المظللة على الشكل (2.2.6)، وأن أكبر قيمة ممكنة لها عندما تكون β أصغر ما يمكن، وتكون β أصغر ما يمكن عند تقاطع الخطين:

$$\beta = \alpha - x_{(1)} \quad , \quad \beta = x_{(n)} - \alpha$$

وبالتالي بحلهما المشترك نجد:

$$\hat{\alpha} = \frac{x_{(1)} + x_{(n)}}{2} \quad , \quad \hat{\beta} = \frac{x_{(n)} - x_{(1)}}{2}$$

ومن ثم فإن مقدّر المعقولية العظمى لـ $\theta = (\alpha,\beta)$ هو:

$$\hat{\theta}(\boldsymbol{X}) = \left(\hat{\alpha},\hat{\beta}\right) \quad ; \quad \hat{\alpha}(\boldsymbol{X}) = \frac{X_{(1)} + X_{(n)}}{2} \ , \ \hat{\beta}(\boldsymbol{X}) = \frac{X_{(n)} - X_{(1)}}{2}$$

مثـــال 8.2.6

إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشـــوائية من التـــوزيع $\mathcal{L}(\xi)\in R(\theta,\theta+1)$ ، فأوجد مقدِّر المعقولية العظمى للمعلمة θ.

بما أن:

$$f(x;\theta)=1 \quad , \quad \theta\le x\le\theta+1$$

فإن دالة المعقولية $L(x;\theta)$ يمكن كتابتها بدلالة دالة خوفســـيد علـــى الصـــورة:

$$L(x;\theta)=e\left(\theta+1-x_{(n)}\right)e\left(x_{(1)}-\theta\right)$$

ونلاحظ بوضوح من هذه الصيغة أن $L(x;\theta)$ تأخذ أكـــبر قيمـــة ممكنـــة لهـــا عندمـــا تكـون:

$$\left.\begin{array}{r}\theta+1\ge x_{(n)}\\ \\ \theta\le x_{(1)}\end{array}\right\} \Rightarrow \quad x_{(n)}-1\le\theta\le x_{(1)}$$

وعلى ذلك فإن تقدير المعقولية العظمى ليس وحيداً، حيـــث أن أيـــة قيمـــة:

$$\hat{\theta}(\mathrm{x})\in\left[x_{(n)}-1;x_{(1)}\right]$$

هي تقدير المعقولية العظمى. فمثلاً، كمقدِّر معقولية عظمى يمكـــن أخـــذ:

$$\hat{\theta}(\mathrm{X})=\frac{X_{(1)}+X_{(n)}-1}{2}$$

وفي هذه الحالة $\hat{\theta}(\mathrm{X})$ دالة في الإحصاء الكافي $T=\left(X_{(1)},X_{(2)}\right)$. ونخلص من هذا المثـــال إلى أن مقدِّر المعقولية العظمى لمعلمة θ ليس بالضرورة وحيداً.

مثـــال 9.2.6

(نموذج طبيعي متعدد الأبعاد، تقدير المعقولية العظمـــى لمعالمـــه).

لنفترض أن المتغير العشوائي الملاحظ ξ متعدد الأبعاد (k بعد مثلاً) يخضع للتوزيع الطبيعي $N(\mu,\Sigma)$، حيث أن $\mu=(\mu_1,\dots,\mu_k)$ متجهة القيم المتوسطة و $\Sigma=\|\sigma_{ij}\|_1^k$ $(|\Sigma|\neq 0)$ مصفوفة العزوم المركزية من المرتبة الثانية. إن عدد المعالم غير المعلومة (آخذين بالاعتبار تناظر المصفوفة Σ) يساوي $k+k(k+1)/2$.

إذا كانت $X=(X_1,\dots,X_n)$ عينة عشوائية مأخوذة من التوزيع $\mathcal{L}(\xi)$، فالمطلوب إيجاد تقديرات المعقولية العظمى لمعالم هذا التوزيع.

بما أن $X_i\ ;\ i=1,\dots,n$ متغيرات عشوائية مستقلة وكل منها ذات k بعد، بكثافة:

$$f(x_i;\theta)=\frac{1}{\sqrt{(2\pi)^k|\Sigma|}}\exp\left[-\frac{1}{2}(x_i-\mu)^t\Sigma^{-1}(x_i-\mu)\right]\quad;\quad\theta=(\mu,\Sigma)$$

فإن:

$$L(x,\theta)=\prod_{i=1}^{n}f(x_i;\theta)=(2\pi)^{-\frac{kn}{2}}|\Sigma|^{-\frac{n}{2}}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^t\Sigma^{-1}(x_i-\mu)\right]$$
(5.2.6)

لنرمز بـ $\bar{x}=(x_1+\dots+x_n)/n$، عندها:

$$\sum_{i=1}^{n}(x_i-\mu)^t\Sigma^{-1}(x_i-\mu)=$$
$$=\sum_{i=1}^{n}(x_i-\bar{x})^t\Sigma^{-1}(x_i-\bar{x})+n(\bar{x}-\mu)^t\Sigma^{-1}(\bar{x}-\mu)\tag{6.2.6}$$

إذا رمزنا بـ $\hat{\Sigma}(x)$ لمصفوفة العزوم من المرتبة الثانية للعينة، فإن:

$$\hat{\Sigma}(\mathrm{x})=\frac{1}{n}\sum_{i=1}^{n}(\mathrm{x}_i-\bar{\mathrm{x}})(\mathrm{x}_i-\bar{\mathrm{x}})^t=\|\hat{\sigma}_{ij}\|_1^k$$

حيث أن $\hat{\sigma}_{ij}$ العزم المركزي من المرتبة الثانية للعينة ويحسب من العلاقة:

$$\hat{\sigma}_{ij}=\frac{1}{n}\sum_{l=1}^{n}(x_{li}-\bar{x}_i)(x_{lj}-\bar{x}_j)\quad;\quad i,j=1,\dots,k$$

(هـــــــــنا $\bar{x}_i = \frac{1}{n}(x_{1i}+\ldots+x_{ni})$ و $x_l = (x_{l1},\ldots,x_{lk})\ ;\ l = 1,\ldots,n$
.($\bar{x} = (\bar{x}_1,\ldots,\bar{x}_k)$)

وباستخدام المساواة $y'By = tr(BY)$, $Y = y'y$ والمؤثر الخطي tr نجد:

$$\sum_{l=1}^{n}(x_l - \bar{x})'\Sigma^{-1}(x_l - \bar{x}) = ntr\left(\Sigma^{-1}\hat{\Sigma}(x)\right)$$

وبناءً على المساواتين (6.2.6) و (7.2.6) يمكن كتابة دالة المعقولية $L(x;\theta)$ على الصورة:

$$L(x;\theta) = (2\pi)^{-\frac{k\pi}{2}}\exp\left[-\frac{n}{2}(\bar{x}-\mu)'\Sigma^{-1}(\bar{x}-\mu) - \frac{n}{2}tr\Sigma^{-1}\hat{\Sigma}(x) - \frac{n}{2}\ln|\Sigma|\right]$$

نلاحظ أن أكبر قيمة ممكنة لدالة المعقولية L بالنسبة لـ θ توافق أصغر قيمة ممكنة للدالة:

$$\psi(x,\mu,\Sigma) = (\bar{x}-\mu)'\Sigma^{-1}(\bar{x}-\mu) + \left[tr\Sigma^{-1}\hat{\Sigma}(x) - k - \ln\left|\Sigma^{-1}\hat{\Sigma}(x)\right|\right]$$

بالنسبة لـ μ,Σ. (إدخال الثابتين k و $\ln|\hat{\Sigma}|$ من أجل تبسيط الصياغات اللاحقة).

إذا رمزنا بـ $\lambda_1,\ldots,\lambda_k$ لجذور المعادلة المميزة:

$$\left|\Sigma^{-1}\hat{\Sigma}(x) - \lambda E_n\right| = 0 \quad \text{أو} \quad \left|\hat{\Sigma}(x) - \lambda\Sigma\right| = 0$$

فإن:

$$tr\Sigma^{-1}\hat{\Sigma}(x) - k - \ln\left|\Sigma^{-1}\hat{\Sigma}(x)\right| = \lambda_1+\ldots+\lambda_k - k - \ln(\lambda_1+\ldots+\lambda_k)$$

بما أن Σ^{-1} مصفوفة محددة وموجبة و $a \geq 0\ ;\ a - 1 - \ln a \geq 0$ ، فإن:

$$\psi(x,\mu,\Sigma) \geq 0$$

وتتحقق المساواة فقط عندما $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 1$ و $\mu = \bar{x}$ أي عندما $\mu = \bar{x}$ و $\Sigma = \hat{\Sigma}(x)$، وبالتالي فإن مقدّري المعقولية العظمى لـ μ و $\Sigma = \left\|\sigma_{ij}\right\|_1^k$ هما على الترتيب $\bar{X}$ و $\hat{\Sigma}(x) = \left\|\hat{\sigma}_{ij}\right\|_1^k$.

يتمتع مقدّر المعقولية العظمى لمعلمة θ ببعض الخواص الهامة، والتي سنذكرها بصيغة مبرهنات.

مبرهنة 1.2.6

لتكن $X = (X_1, \ldots, X_n)$ عينة عشوائية من مجتمع توزيعه $f(x;\theta)$ يعتمد على معلمة θ (وحيدة البعد أو متعددة الأبعاد) و $\hat{\theta}(x)$ تقدير المعقولية العظمى لـ θ.

إذا كانت $q = q(\theta)$ دالة تناظر أحادية ما في θ معرفة على Θ وتأخذ قيمها في $Q = \{q = (q_1, \ldots, q_r)\} \subset R^r$، فإن تقدير المعقولية العظمى لـ $q = q(\theta)$ هو $\hat{q} = q(\hat{\theta})$.

الإثبات

بما أن $q = q(\theta)$ دالة تناظر أحادية في θ (حسب الفرض)، فإن الدالة العكسية موجودة وهي $\theta = \theta(q)$، عندها:

$$\sup_{\theta \in \Theta} L(x;\theta) = \sup_{q \in Q} L(x;\theta(q))$$

وإذا كانت الدالة $L(x;\theta)$ تبلغ أكبر قيمة ممكنة لها عند النقطة $\hat{\theta} = \hat{\theta}(x)$، فإن الدالة $L(x;\theta(q))$ في q تبلغ أكبر قيمة لها عند النقطة $\hat{q}$، المحققة للمعادلة $\hat{\theta} = \theta(\hat{q})$ أي عند النقطة $\hat{q} = q(\hat{\theta})$.

تدعى هذه الخاصة لمقدّر المعقولية العظمى بخاصة الثبات أو عدم الاختلاف (invariance property of maximum likelihood estimator). وتعطي هذه الخاصة إمكانية الحصول على مقدّرات المعقولية العظمى لأسرة هامة من

الدوال المعلمية $\tau(\theta)$ ذات التناظر الأحادي في θ. وهذا ما سنوضحه من خلال الأمثلة الآتية.

مثال 10.2.6

إذا كانت $X = (X_1,\dots,X_n)$ عينة عشوائية من مجتمع كثافة توزيعه الاحتمالي:

$$f(x;\theta) = \theta^x (1-\theta)^{1-x} \quad ; \quad x = 0,1$$

فأوجد مقدِّر المعقولية العظمى للدالة $\tau(\theta) = \dfrac{2\theta}{1+\theta}$.

نبحث أولاً عن مقدِّر المعقولية العظمى للمعلمة θ:

$$L(x;\theta) = \theta^{n\bar{x}} (1-\theta)^{n(1-\bar{x})}$$
$$\ln L(x;\theta) = n\bar{x}\ln\theta + n(1-\bar{x})\ln(1-\theta)$$
$$\frac{\partial \ln L}{\partial\theta} = \frac{n\bar{x}}{\theta} - \frac{n(1-\bar{x})}{1-\theta} = 0$$

ومنها

$$\hat{\theta}(X) = \bar{X}$$

بما أن $\tau(\theta) = \dfrac{2\theta}{1+\theta}$ دالة تناظر أحادية في θ، فحسب المبرهنة (1.2.6)، نجد أن مقدِّر المعقولية العظمى لـ $\tau(\theta)$ هو:

$$\hat{g} = \tau(\hat{\theta}) = \frac{2\bar{X}}{1+\bar{X}}$$

مثال 11.2.6

إذا كانت لدينا معطيات المثال (1.2.6)، فأوجد مقدِّر المعقولية العظمى لكل من الدوال المعلمية الآتية:

$$\tau_1(\theta)=\frac{1}{\theta} \quad , \quad \tau_2(\theta)=\ln\theta+1 \quad , \quad \tau_3(\theta)=\frac{\ln\theta}{\theta}$$

بما أن مقدّر المعقولية للمعلمة θ هو $\hat{\theta}(X)=\frac{1}{\overline{X}}$ وكل من الدوال المعلمية τ_1,τ_2,τ_3 ذات تناظر أحادي في θ، فحسب المبرهنة (1.2.6) نجد:

$$\hat{g}_1=\tau_1(\hat{\theta})=\overline{X} \quad , \quad \hat{g}_2=\tau_2(\hat{\theta})=\ln\frac{1}{\overline{X}}+1 \quad , \quad \hat{g}_3=\tau_3(\hat{\theta})=\overline{X}\ln\frac{1}{\overline{X}}$$

مبرهنة 2.2.6

إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من مجتمع توزيعه $f(x;\theta)$، وكان هناك إحصاء كافي للمعلمة θ، فإن مقدّر المعقولية العظمى للمعلمة θ يكون دالة في هذا الإحصاء الكافي.

الإثبات

نفترض أن $T=T(X)$ إحصاء كافي للمعلمة θ، ومن ثم يمكن كتابة دالة المعقولية (حسب المبرهنة 1.2.4) على الصورة:

$$L(x;\theta)=g(t;\theta)h(x)$$

وعلى ذلك:

$$\ln L(x;\theta)=\ln g(t;\theta)+\ln h(x)$$

$$\frac{\partial \ln L(x;\theta)}{\partial\theta}=\frac{\partial \ln g(t;\theta)}{\partial\theta}=K(t;\theta)$$

لإيجاد مقدّر المعقولية العظمى نضع:

$$K(T;\theta)=0$$

وبحل هذه المعادلة بالنسبة لـ θ نجد:

$$\hat{\theta}(X)=\varphi(T)$$

مـــــبرهنة 3.2.6

إذا وجد المقدِّر الأمثل للمعلمة θ في توزيع $f(x;\theta)$، فإنه مقدِّر المعقولية العظمى لهذه المعلمة.

الإثبـــات

لتكن $X=(X_1,\ldots,X_n)$ عينة عشوائية من مجتمع توزيعه $f(x;\theta)$، ولنفـــترض أن $T^*=T^*(X)$ المقدِّر الأمثل لتقدير المعلمة θ.

بما أن T^* المقدِّر الأمثل للمعلمة θ، فـــإن:

$$E_\theta T^*=\theta \quad ; \quad V_\theta T^*=\frac{1}{ni(\theta)}$$

وبالتالي حسب العلاقتـــين (2.7.5) و (8.7.5) نجـــد:

$$U(X;\theta)=\frac{\partial \ln L(X;\theta)}{\partial\theta}=\frac{1}{ni(\theta)}\left(T^*-\theta\right)$$

بوضـع:

$$\frac{1}{ni(\theta)}\left(T^*-\theta\right)=0$$

وبحلها بالنسبة لــ θ نجـــد أن:

$$\hat{\theta}(X)=T^*$$

مثـــال 12.2.6

إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من التوزيع $\mathcal{L}(\xi)\in N(\theta_1,\theta_2^2)$، فـأوجد مقدِّر المعقولية العظمى للدالة:

$$\tau(\theta)=\Phi\left(\frac{x_0-\theta_1}{\theta_2}\right) \quad ; \quad \theta=(\theta_1,\theta_2)$$

المثلة للاحتمال $P_\theta(\xi < x_0)$.

يمكن في هذه الحالة وضع $q(\theta) = (\tau(\theta), \theta_2)$ ، وهي دالة تناظر أحادية في $\theta = (\theta_1, \theta_2)$ ، وبالتالي حسب المبرهنة (1.2.6) فتقدير المعقولية العظمى لـ $q(\theta)$ يكون:

$$\hat{q} = \left(\tau(\hat{\theta}), \hat{\theta}_2\right)$$

لكن كما نعلم [راجع المثال (4.2.6)] فتقدير المعقولية العظمى لكل من θ_1، θ_2 هو $\bar{x}$، s على الترتيب، إذن:

$$\hat{q} = \left(\Phi\left(\frac{x_0 - \bar{x}}{s}\right), s\right)$$

أي أن تقدير المعقولية العظمى لـ $\tau(\theta)$ هو:

$$\hat{\tau} = \tau(\hat{\theta}) = \Phi\left(\frac{x_0 - \bar{x}}{s}\right)$$

ومن ثم فمقدّر المعقولية العظمى لـ $\tau(\theta)$:

$$\hat{\tau}(\boldsymbol{X}) = \Phi\left(\frac{x_0 - \bar{X}}{S}\right)$$

مثال 13.2.6

إذا كانت $\boldsymbol{X} = (X_1, \ldots, X_n)$ عينة عشوائية من التوزيع الطبيعي ثنائي البعد $N(\boldsymbol{0}, \Sigma)$ ، حيث أن:

$$\boldsymbol{0} = (0,0) \quad , \quad \Sigma = \begin{Vmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{Vmatrix}$$

و $\sigma^2 > 0$, $|\rho| < 1$ مجهولين.

إن $\mathcal{L}(X_i) \in N(0,\Sigma)\,;\, i=1,\ldots,n$ ، أي متغير عشوائي ذو بعدين كثافة توزيعه الاحتمالية:

$$f(x,y;\theta)=\frac{1}{2\pi}\exp\left\{-\frac{x^2+y^2}{2\sigma^2\left(1-\rho^2\right)}+\frac{\rho(x+y)}{\sigma^2\left(1-\rho^2\right)}-\frac{1}{2}\ln\left[\sigma^2\left(1-\rho^2\right)\right]\right\}\ ;\ \theta=\left(\sigma^2,\rho\right)$$

نلاحظ هنا أن صيغة دالة المعقولية معقدة وهي تعتمد على معلمتين σ^2,ρ وبالتالي لتبسيط عملية إيجاد تقديري المعقولية العظمى $\hat{\rho}$ و $\hat{\sigma}^2$ من المفيد الانتقال إلى معلمة جديدة $q=\left(q_1,q_2\right)$ بدلاً من $\theta=\left(\sigma^2,\rho\right)$ ، بوضع:

$$\begin{aligned} q_1 &= q_1(\theta)=-\frac{1}{2\sigma^2\left(1-\rho^2\right)} \\ q_2 &= q_2(\theta)=\frac{\rho}{\sigma^2\left(1-\rho^2\right)} \end{aligned} \qquad \ldots\ldots(1)$$

وعندئذٍ يمكن كتابة دالة الكثافة الاحتمالية على الصورة:

$$f(x,y;\theta)=\frac{1}{2\pi}\exp\left[q_1\left(x^2+y^2\right)+q_2xy+\tau(q)\right]$$

حيث أن $\tau(q)=\left(1/2\right)\ln\left(4q_1^2-q_2^2\right)$.

وعلى ذلك يمكن بسهولة إيجاد معادلتي المعقولية لتقدير $q_1; q_2$ ، وهما:

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n}\left(x_i^2+y_i^2\right) &= -\frac{\partial\tau(q)}{\partial q_1}=-\frac{4q_1}{4q_1^2-q_2^2} \\ \frac{1}{n}\sum_{i=1}^{n}x_iy_i &= -\frac{\partial\tau(q)}{\partial q_2}=-\frac{q_2}{4q_1^2-q_2^2} \end{aligned} \qquad \ldots\ldots(2)$$

لكن من العلاقتين في (1) نجد:

$$\sigma^2=\frac{2q_1}{4q_1^2-q_2^2}\qquad,\qquad \rho=-\frac{q_2}{2q_1}$$

وبالتالي من العلاقتين في (2) نحصل على تقدير المعقولية العظمى:

$$\hat{\sigma}^2 = \frac{1}{2n}\sum_{i=1}^{n}\left(x_i^2 + y_i^2\right)$$

$$\hat{\rho} = 2\sum_{i=1}^{n} x_i y_i \Big/ \sum_{i=1}^{n}\left(x_i^2 + y_i^2\right)$$

حيث أن (x_i, y_i) القيمة الملاحظة لـ $X_i\ ;\ i = 1,\dots,n$.

1.2.6 طريقة التراكم لحساب تقدير المعقولية العظمى بالتقريب

لا يتاح دائماً الحصول على تقديرات المعقولية العظمى بصورة صريحة محددة. في مثل تلك الحالات نلجأ إلى طريقة التقريب العددية لحل معادلة أو معادلات المعقولية (عندما يمكن صياغة مثل تلك المعادلات)، أي الحصول على تقدير المعقولية العظمى بالتقريب. وإحدى طرق التقريب هذه تدعى بطريقة التراكم، التي قدمت من قبل العالم فيشر، وتتمثل بما يلي:

لتكن θ معلمة حقيقية ودالة المعقولية $L(x;\theta)$ قابلة للاشتقاق مرتين بالنسبة لـ θ. ننشر دالة المساهمة $U(\theta) = U(x;\theta)$ [راجع الصيغة (1.6.4)] وفق متسلسلة تيلور في جوار النقطة θ_0، المتخذة كتقريب ابتدائي لـ $\hat{\theta} = \hat{\theta}(x)$، ونحسب هذا المنشور عند $\hat{\theta} = \theta$. عندئذٍ، بما أن $U(\hat{\theta}) = 0$، فإن:

$$0 = U(\theta_0) + \left(\hat{\theta} - \theta_0\right)U'\left(\theta^*\right)$$

حيث θ^* نقطة ما بين $\hat{\theta}$، θ_0 وبالتالي:

$$\hat{\theta} = \theta_0 - \left(\frac{U(\theta_0)}{U'\left(\theta^*\right)}\right)$$

إذا استبدلنا الآن في هذه المساواة θ^* بـ θ_0 و $U'(\theta_0)$ بـ $-ni(\theta_0) = EU'(X;\theta_0)$ [راجع البند (7.5)] نحصل على التقريب الأول:

$$\theta_1 = \theta_0 + \frac{U(\theta_0)}{ni(\theta_0)}$$

يمكن تكرار هذه العملية، بأخذ θ_1 كتقريب ابتدائي جديد، وهكذا فالتقريب $(k+1)$ وفق طريقة التراكم يحسب وفق العلاقة:

$$\theta_{k+1} = \theta_k + \frac{U(\theta_k)}{ni(\theta_k)} \quad , \quad k = 0,1,2,\ldots \qquad (12.2.6)$$

تستمر عملية التكرار حتى بلوغ الدقة المطلوبة:

$$|\theta_{k+1} - \theta_k| < \varepsilon$$

إن اختيار النقطة الأولية θ_0 يعتبر هاماً. وعادة يتم أخذ قيمة ما بحيث تسهل حساب تقدير متسق للمعلمة θ، عندها النقاط الثلاث θ_0، θ^* و $\hat{\theta}$، في حالة عينة كبيرة الحجم n، تقع قريبة من القيمة الحقيقية للمعلمة المجهولة. وفي حالات عدة يمكن استخدام تقديرات طريقة العزوم، التي سنتطرق لها لاحقاً، كنقطة أولية θ_0.

تجدر الإشارة، إلى أن طريقة التراكم بشكل عام تؤدي إلى الهدف وبشكل سريع، مع أنها في بعض الحالات لا تنتهي.

مثال 14.2.6

(نموذج كوشي تقدير المعلمة بطريقة التراكم).

لتكن $X = (X_1,\ldots,X_n)$ عينة من توزيع كوشي $K(\theta)$، إن دالة المساهمة:

$$U(\theta) = 2\sum_{i=1}^{n} \frac{(x_i - \theta)}{1 + (x_i - \theta)^2}$$

إن الحصول على حل دقيق لمعادلة المعقولية $U(\theta) = 0$ غير ممكن. وبما أن دالة المعلومات من أجل نموذج كوشي تساوي $1/2$ [راجع جدول (1.7.4)]، فالمتتالية (12.2.6) تأخذ الشكل:

$$\theta_{k+1} = \theta_k + (2/n)U(\theta_k) \quad , \quad k = 0,1,2,\ldots$$

وكتقريب ابتدائي يمكن أخذ θ_0 تساوي وسيط العينة $\bar{X}$، الذي يتقارب بالاحتمال، عندما $n \to \infty$، من القيمة الحقيقية للمعلمة θ، التي تعتبر في الحالة المفروضة الوسيط النظري (وسيط المجتمع $\tilde{\mu}$).

2.2.6 الخواص التقاربية لمقدّر المعقولية العظمى

إن طريقة المعقولية العظمى لا تعطي دوماً مقدّر غير متحيز. فمثلاً، نجد في المثال (4.2.6)، أن تباين العينة $S^2(X)$ هو مقدّر المعقولية العظمى للتباين θ_2^2 في النموذج $N(\theta_1, \theta_2^2)$، وهذا المقدّر متحيز (biased). إلا أن مقدار التحيز $E_\theta S^2 - \theta_2^2 = -\theta_2^2/n$ يتناقص بازدياد حجم العينة، وينتهي إلى الصفر عندما $n \to \infty$، أي أن مقدّر المعقولية العظمى يصبح غير متحيز. تدعى المقدّرات المتحيزة التي تتمتع بمثل هذه الخاصة بغير المتحيزة تقاربياً أو غير المتحيزة بالتقارب (asymptotically unbiased).

إن خاصة تلك المقدّرات المرتبطة بعدم التحيز تتحسن بازدياد حجم العينة. وهذه الحالة لها طبيعة عامة إلى حد ما. وبشكل خاص إن أهم خاصة لمقدّرات المعقولية العظمى تتمثل في أن لها طبيعة تقاربية، أي صحيحة من أجل عينات كبيرة. لذا فإن الاستخدام الواسع لمقدّرات المعقولية العظمى مرتبط أساساً بخواصها التقاربية. إن عرض مبرهنات التقارب لمقدّرات المعقولية العظمى تشكل محتوى هذه الفقرة، وللإشارة إلى ارتباط المقدّرات بحجم العينة سنشير إليها بالدليل n.

عند التكلم عن الخواص التقاربية للمقدّرات (أو الخواص من أجل عينات كبيرة)، فقبل كل شيء نهتم باتساقها، وبعبارة أخرى، بالتقارب بالاحتمال للمقدّرات المفروضة من المميزات النظرية الموافقة لها، لأنه كما بينا في الفصل الثالث، أن غالبية مميزات الغينة (عزوم العينة، القيم الحرجة، قيمة دالة التوزيع

التجريبي في كل نقطة، ...الخ) تتقارب بالاحتمال من المميزات النظرية الموافقة لها، عندما $n \to \infty$، وبالتالي تعتبر مقدِّرات متسقة. لنقدم الآن التعريف الدقيق للاتساق.

لتكن $X = (X_1, ..., X_n)$ عينة عشوائية من التوزيع $\mathcal{L}(\xi) \in \boldsymbol{F}\{F(x;\theta) \ ;\theta \in \Theta\}$ حيث في الحالة العامة فضاء المعلمة Θ مجال مفتوح من الفضاء R^r.

يُدعى، حسب التعريف، المقدِّر $T_n = T_n(\boldsymbol{X})$ من أجل دالة معلمية $\tau(\theta)$ بالمتسق، إذا حقق الشرط التالي:

$$T_n \xrightarrow{P_\theta} \tau(\theta) \quad ; \quad \forall\, \theta \in \Theta$$

أي أنه مهما تكن القيمة الحقيقية للمعلمة θ من فضاء المعلمة Θ، فإن T_n يتقارب احتمالياً (بالنسبة للتوزيع P_θ) من القيمة الحقيقية للدالة المقدَّرة $\tau(\theta)$.

إن خاصة الاتساق ضرورة من أجل أي قاعدة تقدير، إلا أنها في حقيقة الأمر، تعتبر خاصة تقاربية ومستقلة عن خواص المقدِّر عند ثبات حجم العينة (خلافاً لخاصتي عدم التحيز والكفاءة).

لنرى الآن المعيار الهام والبسيط للاتساق.

مبرهنة 4.2.6

إذا كان:

$$E_\theta T_n = \tau(\theta) + \varepsilon_n \quad , \quad V_\theta T_n = \delta_n$$

و

$$\varepsilon_n = \varepsilon_n(\theta) \to 0 \quad , \quad \delta_n = \delta_n(\theta) \to 0 \quad ; \quad \forall\, \theta \in \Theta$$

عندما $n \to \infty$، فإن T_n مقدِّر متسق للدالة $\tau(\theta)$.

الإثبـــات

بما أن:

$$|T_n - E_\theta T_n| = |(T_n - \tau) - (E_\theta T_n - \tau)| \geq |T_n - \tau| - |E_\theta T_n - \tau|$$

والحادث $\{|T_n - \tau| \geq \varepsilon\}$ محتوى في الحادث $\{|T_n - E_\theta T_n| \geq \varepsilon - |\varepsilon_n|\}$ فحسب متباينـــة تشيبيشف نجد:

$$P_\theta(|T_n - \tau| \geq \varepsilon) \leq P_\theta(|T_n - E_\theta T_n| \geq \varepsilon - |\varepsilon_n|) \leq \frac{\delta_n}{(\varepsilon - |\varepsilon_n|)^2} \to 0$$

عندما $n \to \infty$ مهما تكـــن $\theta \in \Theta$.

نورد فيما يلي بعض الخواص لتقارب دوال في متغيرات عشوائية، المفيدة لاحقاً.

1. إذا كان $\eta_n \xrightarrow{P} \eta$ و φ دالة مســـتمرة، فـــإن: $\varphi(\eta_n) \xrightarrow{P} \varphi(\eta)$ وأيضـــاً $\mathcal{L}(\varphi(\eta_n)) \longrightarrow \mathcal{L}(\varphi(\eta))$.

2. لتكن $n = 1,2,\ldots$ ، $\{\eta_n, \zeta_n\}$ متتالية أزواج من المتغـــيرات العشـــوائية. عندئـــذٍ:

1. $\eta_n - \zeta_n \xrightarrow{P} 0 \ , \ \zeta_n \xrightarrow{P} \zeta \ \Rightarrow \ \eta_n \xrightarrow{P} \zeta$

2. $\eta_n - \zeta_n \xrightarrow{P} 0 \ , \ \mathcal{L}(\zeta_n) \to \mathcal{L}(\zeta) \Rightarrow \mathcal{L}(\eta_n) \to \mathcal{L}(\zeta)$

3. $\mathcal{L}(\eta_n) \to \mathcal{L}(\eta) \ , \ \zeta_n \xrightarrow{P} 0 \Rightarrow \eta_n \zeta_n \xrightarrow{P} 0$

4. $\mathcal{L}(\eta_n) \to \mathcal{L}(\eta) \ , \ \zeta_n \xrightarrow{P} c \Rightarrow \mathcal{L}(\eta_n + \zeta_n) \to \mathcal{L}(\eta + c)$
$\mathcal{L}(\eta_n \zeta_n) \to \mathcal{L}(c\eta) \ , \ \mathcal{L}(\eta_n / \zeta_n) \to \mathcal{L}(\eta / c) \ ; \ c \neq 0$

5. $\eta_n - \zeta_n \xrightarrow{P} 0 \ , \ \mathcal{L}(\zeta_n) \longrightarrow \mathcal{L}(\zeta) \ \Rightarrow \ \varphi(\eta_n) - \varphi(\zeta_n) \xrightarrow{P} 0$

حيث φ دالة مستمرة. نترك برهان هـــذه الخصـــائص البســـيطة للقـــارئ علـــى سبيل المثـــل.

3. إذا كانت $X=(X_1,\dots,X_n)$ و $T_n=T_n(X)$ مقدِّر المعلمة وحيدة البعد θ في النموذج $\theta \in \Theta$; $F(x;\theta)$، بحيث أن:

$$\mathcal{L}_\theta\left(\sqrt{n}(T_n-\theta)\right)\to N\left(0,\sigma^2(\theta)\right) \quad ; \quad \forall\,\theta\in\Theta$$

عندما $n\to\infty$. وإذا كانت الدالة φ قابلة للاشتقاق و $\varphi'\neq 0$، فإن:

$$\mathcal{L}_\theta\left(\sqrt{n}(\varphi(T_n)-\varphi(\theta))\right)\to N\left\{0,[\varphi'(\theta)]^2\sigma^2(\theta)\right\} \tag{13.2.6}$$

بالإضافة إلى ذلك، إذا كانت الدالة φ' مستمرة، فإن:

$$\mathcal{L}_\theta\left(\sqrt{n}(\varphi(T_n)-\varphi(\theta))/\varphi'(T_n)\right)\to N\left(0,\sigma^2(\theta)\right) \tag{14.2.6}$$

وإذا كان $\sigma(\theta)$ مستمراً فإن:

$$\mathcal{L}_\theta\left(\sqrt{n}\,\frac{\varphi(T_n)-\varphi(\theta)}{\varphi'(T_n)\sigma(T_n)}\right)\to N(0,1) \tag{15.2.6}$$

يعتمد الإثبات على منشور تيلور:

$$\varphi(T_n)-\varphi(\theta)=(T_n-\theta)(\varphi'(\theta)+\zeta_n)$$

حيث عندما $|T_n-\theta|<\delta=\delta_{(\varepsilon)}$ فإن $|\zeta_n|<\varepsilon\ ;\ \forall\,\varepsilon>0$.

من ذلك:

$$P_\theta(|\zeta_n|<\varepsilon)\geq P_\theta(|T_n-\theta|<\delta)$$

لكن الطرف الأيمن حسب الفرض، ينتهي إلى 1 عندما $n\to\infty$ وبالتالي $\zeta_n \xrightarrow{P_\theta} 0$. وبناءً على الخاصة [(3) من 2]:

$$\sqrt{n}[\varphi(T_n)-\varphi(\theta)]-\sqrt{n}[(T_n-\theta)-\varphi'(\theta)]=\sqrt{n}(T_n-\theta)\zeta_n \xrightarrow{P_\theta} 0$$

وهذا يعني حسب الخاصة [(2) من 2] أن المتغير العشوائي $\sqrt{n}\left(\varphi(T_n)-\varphi(\theta)\right)$ له أيضـــاً نفس التوزيع المقارب، كما للمتغير العشوائي $\sqrt{n}(T_n-\theta)\varphi'(\theta)$ أي التوزيـــع الطبيعـــي $N\left(0,[\varphi'(\theta)]^2\sigma^2(\theta)\right)$، وهذا ما تؤكده العلاقة (13.2.6). وإذا كانت φ' مســـتمرة، فحسب الخاصة 1:

$$\varphi'(T_n)\xrightarrow{P_\theta}\varphi'(\theta)$$

من ذلك ومما سبق وحسب الخاصة [(4) من 2] نحصل على العلاقة (14.2.6). وبـــإجراء مناقشة مشابهة، يمكن إثبات صحة الصيغة (15.2.6).

نقـــدم بـــدون إثبـــات تعميمـــاً للخاصـــة 3 علـــى حالـــة معلمـــة متجـــه $\theta=(\theta_1,\ldots,\theta_r)$.

4. ليكـــن $T_n=(T_{1n},\ldots,T_{rn})$ مقـــــدِّراً للمعـــــلمة θ، المحقـــق للشـــرط $\mathcal{L}_\theta\left(\sqrt{n}(T_n-\mu)\right)\to N(0,\Sigma(\mu))$ عندمـــا $n\to\infty$ ومـــهما تكـــن $\theta\in\Theta$. عندئـــذٍ من أجل أي دالة φ في r متغير وقابلـــة للاشـــتقاق:

$$\mathcal{L}_\theta\left(\sqrt{n}(T_n-\theta)\right)\to N\left(0,\upsilon^2(\theta)\right) \qquad (14.2.6)$$

بفرض $\upsilon(\theta)\neq 0$، حيث $\upsilon^2(\theta)=b'(\theta)\Sigma(\theta)b(\theta)$ و $b(\theta)=\left(\dfrac{\partial\varphi(\theta)}{\partial\theta_1},\ldots,\dfrac{\partial\varphi(\theta)}{\partial\theta_r}\right)$.

بالإضافة لذلك، إذا كانت الدالة φ مستمرة وقابلة للاشتقاق وكل عناصر مصفوفة العـزوم من المرتبة الثانية $\Sigma(\theta)$ مستمرة بالنسبة لـــ θ فإن:

$$\mathcal{L}_\theta\left(\sqrt{n}\left(\varphi(T_n)-\varphi(\theta)\right)/\upsilon(T_n)\right)\to N(0,1) \qquad (17.2.6)$$

لنصيغ الآن الخواص المقاربة الأساسية لمقـــدِّرات المعقوليـــة العظمـــى. لنفـــترض أن النمـــوذج $F(x;\theta)$ نظـــامي، ودالة المعقولية $L_n(x;\theta)$ لهـــا نهايـــة عظمـــى واحـــدة

بالنسبة لـ $\theta \in \Theta$ من أجل $n \geq 1$ و $x \in G$، أي أن مقدِّر المعقولية العظمى موجود $\hat{\theta}_n(X)$ عندئذٍ:

1. مقدِّر المعقولية $\hat{\theta}(X)$ للمعلمة θ متسق.

2. إذا كانت الدالة $f(x;\theta)$ قابلة للاشتقاق ثلاث مرات بالنسبة لـ θ وتوجد دالة $M(x)$ مستقلة عن θ، بحيث أن:

$$\left|\frac{\partial^3 \ln f(x;\theta)}{\partial\theta_i \partial\theta_j \partial\theta_k}\right| \leq M(x) \quad , \quad E_\theta M(X_1) < \infty \quad ; \quad i,j,k = 1,\ldots,r$$

فعندما $n \to \infty$، ومهما تكن $\theta \in \Theta$، فإن:

$$\mathcal{L}_\theta\left(\sqrt{n}\left(\hat{\theta}_n - \theta\right)\right) \to N\left(O, I^{-1}(\theta)\right) \tag{18.2.6}$$

حيث أن $I(\theta)$ مصفوفة المعلومات، المعرفة بالعلاقة (6.7.4) ، والتي تعتبر حسب الفرض نظامية، أي غير شاذة $\left(|I(\theta)| \neq 0\right)$. بالإضافة لذلك، إذا كانت الدالة $\tau(\theta)$ مستمر وقابلة للاشتقاق بالنسبة لـ θ و $\hat{\tau}_n = \tau\left(\hat{\theta}_n\right)$ مقدِّر المعقولية العظمى لها، فإن:

$$\mathcal{L}_\theta\left(\sqrt{n}\left(\hat{\tau}_n - \tau(\theta)\right)\right) \to N\left(0, \sigma_\tau^2(\theta)\right) \tag{19.2.6}$$

حيث أن:

$$\sigma_\tau^2(\theta) = b'(\theta) I^{-1}(\theta) b(\theta) \quad , \quad b(\theta) = \left(\frac{\partial\tau(\theta)}{\partial\theta_1}, \ldots, \frac{\partial\tau(\theta)}{\partial\theta_r}\right)$$

هكذا، من أجل صف واسع من النماذج فإن مقدِّرات المعقولية العظمى متسقة وطبيعية بالتقارب.

في حالة معلمة θ وحيدة البعد تأخذ العلاقتان (18.2.6) و (19.2.6) الشكل:

$$\mathcal{L}_\theta\left(\sqrt{n}\left(\hat{\theta}_n-\theta\right)\right)\to N\left(0,1/i(\theta)\right) \tag{20.2.6}$$

$$\mathcal{L}_\theta\left(\sqrt{n}\left(\hat{\tau}_n-\tau(\theta)\right)\right)\to N\left(0,[\tau'(\theta)]^2/i(\theta)\right) \tag{21.2.6}$$

حيث $i(\theta)$ دالة المعلومات المعرفة بالعلاقة (13.5).

إثبات خاصة أن مقدِّر المعقولية العظمى طبيعي بالتقارب (في حالة معلمة وحيدة البعد) يعتمد على نشر دالة المساهمة $U_n(\theta)=U_n(X;\theta)$ وفق متسلسلة تيلور بالنسبة للقيمة الحقيقية للمعلمة θ وملاحظة هذا المنشور في النقطة $\hat{\theta}_n$، نجد:

$$0=U_n\left(\hat{\theta}_n\right)=U_n(\theta)+\left(\hat{\theta}_n-\theta\right)U'_n(\theta)+\frac{1}{2}\left(\hat{\theta}_n-\theta\right)^2U''\left(\theta^*\right)$$

حيث θ^* نقطة ما بين θ ، $\hat{\theta}_n$. وهذه المساواة يمكن كتابتها على الشكل:

$$\sqrt{n}\left(\hat{\theta}_n-\theta\right)=\frac{U_n(\theta)}{\sqrt{n}\,i(\theta)}\left[-\frac{U'_n(\theta)}{ni(\theta)}+\varepsilon_n\right]^{-1} \tag{22.2.6}$$

حيث:

$$|\varepsilon_n|=\frac{\left|\hat{\theta}_n-\theta\right|\left|U''_n\left(\theta^*\right)\right|}{2ni(\theta)}\leq\frac{\left|\hat{\theta}_n-\theta\right|}{2i(\theta)}\frac{1}{n}\sum_{j=1}^{n}M\left(X_j\right)$$

بما أن $\hat{\theta}_n \xrightarrow{P_\theta} \theta$، من الشرط المفروض على الدالة $M(x)$، وبناءً على الخاصة [(2) من 2] ينتج أن $\varepsilon_n \xrightarrow{P_\theta} 0$. وبتطبيق قانون الأعداد الكبيرة على المقدار:

$$\frac{1}{n}U'_n(\theta)=\frac{1}{n}\sum_{j=1}^{n}\frac{\partial^2\ln f\left(X_j;\theta\right)}{\partial\theta^2}$$

آخذين بالاعتبار العلاقة (4.7.4) نجد:

$$-\frac{1}{ni(\theta)}U_n'(\theta)\xrightarrow{P_\theta}-\frac{1}{i(\theta)}E_\theta\left(\frac{\partial^2 \ln f(X_1;\theta)}{\partial\theta^2}\right)=1$$

وبتطبيق مبرهنة النهاية المركزية على المتغير العشوائي:

$$\frac{1}{\sqrt{n}\,i(\theta)}U_n(\theta)=\frac{1}{\sqrt{n}\,i(\theta)}\sum_{j=1}^{n}\frac{\partial \ln f(X_j;\theta)}{\partial\theta}$$

آخذين بالاعتبار العلاقتين (3.6.4) و (4.7.4) نجد:

$$\mathcal{L}_\theta\left(\frac{1}{\sqrt{n}\,i(\theta)}U_n(\theta)\right)\to N\left(0,\frac{1}{i(\theta)}\right)$$

عندما $n\to\infty$. من ذلك ومن العلاقة (22.2.6) والخاصة [(4) من **2**] ينتج أن المتغير العشوائي $\sqrt{n}(\theta_n-\theta)$ له نفس التوزيع الحدي أيضاً، أي أن العلاقة (20.2.6) صحيحة.

تعتبر العلاقة (21.2.6) نتيجة مباشرة للعلاقة (20.2.6) والخاصة 3.

نسمي المقدار $\sigma^2(\theta)/n$ بالتباين المقارب للإحصاء T_n، المحقق للشرط:

$$\mathcal{L}_\theta\left[\sqrt{n}\left(T_n-\tau(\theta)\right)\right]\to N\left(0,\sigma^2(\theta)\right)$$

عندما $n\to\infty$. عندئذٍ من العلاقتين (20.2.6) و (21.2.6) ينتج أن التباين المقارب لمقدِّر المعقولية العظمى $\hat{\theta}_n$ (مقدِّر المعقولية العظمى $\hat{\tau}_n$) ينطبق على الحد الأدنى في متباينة كرامر وراو (1.7.5)، من أجل تباينات كل المقدِّرات غير المتحيزة للدالة المعلمية المفروضة.

تعريف 2.2.6: المقدِّر الأكفأ تقاربياً

Asymptotically Efficient Estimator

إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $f(x;\theta)$، وكان المقدِّر T_n للمعلمة θ طبيعياً بالتقارب $N(\theta,1/ni(\theta))$، فإن هذا المقدِّر يكون

الأمثل تقاربيا، ومن ثم الأكفأ تقاربياً (asymptotically efficient)، ومن أجل أي مقدّر T_n محقق للشرط:

$$\mathcal{L}_\theta(T_n) \to N\left(\theta, \sigma_T^2(\theta)/n\right) \tag{23.2.6}$$

فإن كفاءته التقريبية $eff(T_n,\theta)$ تعين كنسبة الحد الأدنى لمتباينة كرامر وراو إلى التباين المقارب للمقدّر T_n:

$$eff(T_n;\theta) = \left[i(\theta)\sigma_T^2(\theta)\right]^{-1} \tag{24.2.6}$$

ينتج من العلاقتين (*20.2.6*) و (*24.2.6*) الخاصة الثالثة لمقدّرات المعقولية العظمى.

3. يعتبر المقدّر $\hat{\theta}_n(X)$ للمعلمة θ أكفأ تقاربياً، أي أن كفاءته المقاربة:

$$e = eff\left(\hat{\theta}_n, \theta\right) \equiv 1 \quad ; \quad \forall\, \theta \in \Theta$$

مثال 15.2.6

إذا كانت $X = (X_1,\ldots,X_n)$ عينة عشوائية كبيرة من توزيع:

$$f(x;\theta) = \theta e^{-\theta x} \quad ; \quad x > 0$$

فأوجد مقدّر المعقولية العظمى لكل من θ ، $\tau(\theta) = \dfrac{1}{\theta}$، ثم أوجد التوزيع التقريبي لكل منهما.

بما أن:

$$L(x;\theta) = \theta^n e^{-n\theta\bar{x}}$$

فإن:

$$\ln L(x;\theta) = n\ln\theta - n\theta\bar{x}$$

$$\frac{\partial \ln L(x;\theta)}{\partial\theta} = \frac{n}{\theta} - n\bar{x} = 0 \Rightarrow \theta = \frac{1}{\bar{x}}$$

أي أن مقدِّر المعقولية العظمى للمعلمة θ هو $\hat{\theta}_n(X) = \frac{1}{\overline{X}}$، وبما أن $\tau(\theta) = \frac{1}{\theta}$ دالة تناظر أحادية في θ، فإن مقدِّر المعقولية العظمى لـ $\frac{1}{\theta}$ هو $\tau(\hat{\theta}) = \overline{X}$.

لنحسب الآن معلومات فيشر حول المعلمة θ.

$$i_n(\theta) = -E_\theta\left(\frac{\partial^2 \ln L(X;\theta)}{\partial\theta^2}\right) = -E_\theta\left(-\frac{n}{\theta^2}\right) = \frac{n}{\theta^2} = ni(\theta)$$

وعلى ذلك فالتوزيع التقريبي للمقدِّر $\hat{\theta}_n(X) = \frac{1}{\overline{X}}$ حسب العلاقة (20.2.6) هو:

$$\mathcal{L}_\theta\left(\frac{1}{\overline{X}}\right) \approx N(\theta, 1/ni(\theta)) = N(\theta, \theta^2/n)$$

وبشكل مشابه فالتوزيع التقريبي للمقدِّر $\overline{X}$ لـ $\tau(\theta) = \frac{1}{\theta}$، حسب العلاقة (21.2.6) هو:

$$\mathcal{L}_\theta(\overline{X}) \approx N\left(\tau(\theta), \frac{[\tau'(\theta)]^2}{ni(\theta)}\right) = N\left(\frac{1}{\theta}, \frac{(-1/\theta^2)^2}{n/\theta^2}\right) = N\left(\frac{1}{\theta}, \frac{1}{n\theta^2}\right)$$

إن مقدِّرات المعقولية العظمى ليست دائماً طبيعية بالتقارب. وكمثال على ذلك، نذكر مقدِّر المعقولية العظمى للمعلمة θ في توزيع $R(0,\theta)$، هو كما نعلم $\hat{\theta}_n(X) = X_{(n)}$. وأن:

$$\mathcal{L}_\theta\left(\frac{n}{\theta}\left(\theta - \hat{\theta}_n(X)\right)\right) \to \Gamma(1,1)$$

عندما $n \to \infty$. أي أن توزيع $\hat{\theta}_n(X)$ المقارب هو توزيع أسي. هكذا، التوزيع المقارب للمقدِّر $\hat{\theta}(X) = X_{(n)}$ ليس طبيعي، والسبب في ذلك يكمن في عدم نظامية النموذج $R(0,\theta)$.

3.6 طريقة العزوم THE METHOD OF MOMENTS

تعتبر طريقة العزوم تاريخياً إحدى أقدم طرق التقدير، والتي قدمت من قبل العالم الإحصائي كارل بيرسون عام 1894م. ويتمثل جوهر هذه الطريقة في المساواة بين بعض عزوم المجتمع وعزوم العينة المناظرة لها، فنحصل بذلك على جملة من المعادلات بحلها بالنسبة لمعالم المجتمع نحصل على التقديرات المطلوبة، التي تدعى بتقديرات طريقة العزوم. وبعبارة أخرى تؤخذ العزوم التجريبية (عزوم العينة) كتقديرات للعزوم النظرية (عزوم المجتمع) الموافقة لها، ومنها نستخلص تقديرات معالم المجتمع بدلالة العزوم التجريبية. لنوضح ذلك على النحو الآتي: لتكن $X = (X_1,\dots,X_n)$ عينة عشوائية من التوزيع:

$$\mathcal{L}(\xi) \in \boldsymbol{F} = \{F(x;\mu)\ \mu \in \Theta\}$$

حيث أن $\theta = (\theta_1,\dots,\theta_r) \in \Theta \subseteq R^r$، ولنفترض أن العزوم الابتدائية الـ r الأولى للمتغير العشوائي الملاحظ ξ موجودة:

$$\alpha_k = E\xi^k \qquad ; \qquad k = 1,2,\dots,r$$

وتعتبر هذه العزوم، بصورة عامة، دوال في المعلمة المجهولة θ، أي أن:

$$\alpha_k = \alpha_k(\theta)$$

والعزوم الابتدائية التجريبية الموافقة لها:

$$A_{nk}(X) = \frac{1}{n}\sum X_i^k$$

و $a_k = A_{nk}(x)$ قيم تلك العزوم عند العينة المشاهدة $x = (x_1,\dots,x_n)$.

بمساواة العزوم الابتدائية للمجتمع بما يناظرها من عزوم العينة فإننا نحصل على المعادلات:

$$\alpha_k(\theta) = a_k \quad ; \quad k = 1,2,\ldots,r \tag{1.3.6}$$

وبحل هذه المعادلات بالنسبة لـــ $\theta_1,\ldots,\theta_r$ نصل إلى التقديرات $\tilde{\theta}_1,\ldots,\tilde{\theta}_r$. ومن الواضح أن هذه التقديرات دوال في عزوم العينة.

يمكن بسهولة إثبات أن عزوم العينة $A_{nk}(X)$ تعتبر مقدرات غير متحــــيزة ومتســـقة للعزوم النظرية $\alpha_k(\theta)$.

لنفرض أن التوافق بين $\theta_1,\ldots,\theta_r$ و $\alpha_1,\ldots,\alpha_r$ يمكن تمثيله بدوال مستمرة وذات تناظر أحادي، أي توجد دوال مستمرة $\varphi_1,\ldots,\varphi_r$ بحيث أن:

$$\theta_i = \varphi_i(\alpha_1,\ldots,\alpha_r) \quad ; \quad i = 1,\ldots,r$$

والمقدرات الموافقـــة:

$$\tilde{\theta}_i(X) = \varphi_i(A_{1n}(X),\ldots,A_{nr}(X))$$

حسب المبرهنة (2.14.1) متقاربة احتمالياً، عند كل $\theta \in \Theta$، من θ_i. وهــــذا يعـــني أن الإحصاءات $\tilde{\theta}_i(X)$ تعتبر مقدرات متسقة للمعالم $\theta_i \; ; \; i = 1,\ldots,r$.

هكذا، طريقة العزوم عند شروط معينة، تعطي مقدرات متسقة. وعندئـــذٍ المعـــادلات (1.3.6) في حالات عدة بسيطة وحلها (خلافاً لطريقة المعقوليـــة العظمـــى) لا يتطلـــب عمليات حسابية معقدة. ولكن الكفاءة التقريبية للمقدرات التي نحصل عليها بهذه الطريقـــة أقل من واحد.

مثـــال 1.3.6

إذا كانت $X = (X_1,\ldots,X_n)$ عينة عشوائية من توزيع $\Gamma(\theta_1,\theta_2)$، فأوجد مقدر كـل من المعلمتين θ_1,θ_2 باستخدام طريقة العزوم.

بما أن:

$$f(x;\theta_1,\theta_2)=\frac{1}{\theta_2^{\theta_1}\Gamma(\theta_1)}x^{\theta_1-1}e^{-x/\theta_2}\qquad;\qquad \theta_1,\theta_2>0\ ,\ x>0$$

وعلى ذلـك:

$$\alpha_k=\frac{1}{\theta_2^{\theta_1}\Gamma(\theta_1)}\int_0^\infty x^k x^{\theta_1-1}e^{-x/\theta_2}dx=\frac{1}{\theta_2^{\theta_1}\Gamma(\theta_1)}\int_0^\infty x^{\theta_1+k-1}e^{-x/\theta_2}dx$$

وبوضـع $y=\dfrac{x}{\theta_2}$ نجـد:

$$\alpha_k=\frac{\theta_2^{\theta_1+k}}{\theta_2^{\theta_1}\Gamma(\theta_1)}\int_0^\infty y^{\theta_1+k-1}e^{-y}dy=\frac{\theta_2^k\Gamma(\theta_1+k)}{\Gamma(\theta_1)}$$

وبإعطاء $k=1,2$ نحصل علـى:

$$\alpha_1=\theta_1\theta_2\qquad,\qquad \alpha_2=\theta_1\theta_2^2(1+\theta_1)$$

لكن:

$$a_1=\frac{1}{n}\sum_1^n x_i=\bar{x}\qquad,\qquad a_2=\frac{1}{n}\sum_1^n x_i^2=\overline{x^2}$$

وبالتـالي فمعادلتـي التقدير همـا:

$$\theta_1\theta_2=\bar{x}$$
$$\theta_1\theta_2^2(1+\theta_1)=\overline{x^2}$$

بحلهمـا بالنسـبة لـ θ_1، θ_2 نجـد:

$$\tilde{\theta}_1=\frac{\bar{x}^2}{\overline{x^2}-\bar{x}^2}=\frac{\bar{x}^2}{s^2}$$
$$\tilde{\theta}_2=\frac{\overline{x^2}-\bar{x}^2}{\bar{x}^2}=\frac{s^2}{\bar{x}^2}$$

أي أن مقـدري العـزوم:

$$\tilde{\theta}_1(\boldsymbol{X}) = \frac{\overline{X}^2}{S^2} \quad , \quad \tilde{\theta}_2(\boldsymbol{X}) = \frac{S^2}{\overline{X}^2}$$

مثـــــال 2.4.6

إذا كانت $\boldsymbol{X} = (X_1,\dots,X_n)$ عينة عشوائية من التوزيع $\mathcal{L}(\xi) \in N(\theta_1,\theta_2^2)$، فأوجد مقدر العزوم للمعلمة $\boldsymbol{\theta} = (\theta_1,\theta_2^2)$.

كما نعلـم:

$$\alpha_1 = E_\theta\xi = \theta_1 \quad , \quad \alpha_2 = V_\theta\xi + (E_\theta\xi)^2 = \theta_2^2 + \theta_1^2$$
$$a_1 = \overline{x} \quad , \quad a_2 = \overline{x^2}$$

وبالمســاواة نحصـل على المعـادلتين:

$$\theta_1 = \overline{x}$$
$$\theta_2^2 + \theta_1^2 = \overline{x^2}$$

بحلهمـا بالنسـبة لـ θ_1، θ_2^2 نجـد:

$$\tilde{\theta}_1 = \overline{x}$$
$$\tilde{\theta}_2^2 = \overline{x^2} - \overline{x}^2 = s^2$$

أي أن مقـدر العـزوم للمعلمـــة $\boldsymbol{\theta} = (\theta_1,\theta_2^2)$ هــو:

$$\theta(\boldsymbol{X}) = (\overline{X}, S^2)$$

وهـذا نفـس مقـدر المعقوليـة العظمـى للمعلمـة $\boldsymbol{\theta}$.

مثـــــال 3.3.6

إذا كـانت $\boldsymbol{X} = (X_1,\dots,X_n)$ عينـة عشـوائية مـــن التوزيـــع:

$$f(x;\theta) = \theta x^{\theta-1} \quad , \quad 0 \le x \le 1$$

فأوجد مقدر العزوم للمعلمة θ.

$$\alpha_1 = \int_0^1 \theta x^{\theta} dx = \frac{\theta}{1+\theta}$$

وعلى ذلــك:

$$\frac{\theta}{1+\theta} = \bar{x} \Rightarrow \tilde{\theta} = \frac{\bar{x}}{1-\bar{x}}$$

وبالتالي مقدر العزوم للمعلمة θ:

$$\tilde{\theta}(X) = \frac{\bar{X}}{1-\bar{X}}$$

وهو مختلف عن مقدر المعقولية العظمــــى:

$$\hat{\theta}(X) = -\frac{n}{\sum_1^n \ln X_i}$$

مثـــــال 4.3.6

إذا كانت $X = (X_1,\ldots,X_n)$ عينة عشوائية من التوزيع $\mathcal{L}(\xi) \in \Gamma(\alpha, 1/\beta)$ فــأوجد مقدر العزوم لكل من المعلمتين α, β.

نعلم أن [راجع العلاقـــة (23.3.2)]

$$E_{\theta}\xi = \frac{\alpha}{\beta} \quad , \quad E_{\theta}\xi^2 = \frac{\alpha(\alpha+1)}{\beta^2} \quad , \quad \theta = (\alpha, \beta)$$

وعلى ذلــك:

$$\frac{\alpha}{\beta} = \bar{x}$$

$$\frac{\alpha(\alpha+1)}{\beta^2} = \overline{x^2}$$

بحلهما بالنسـبة لـ α,β نجد:

$$\tilde{\alpha} = \frac{\bar{x}^2}{s^2} \quad , \quad \tilde{\beta} = \frac{\bar{x}}{s^2}$$

أي أن:

$$\tilde{\alpha}(\boldsymbol{X}) = \frac{\bar{X}^2}{S^2} \quad , \quad \tilde{\beta}(\boldsymbol{X}) = \frac{\bar{X}}{S^2}$$

نلاحظ أن طريقة العزوم لا تستخدم عندما تكون العزوم من المراتب المطلوبة غير موجودة (فمثلاً، في حالة توزيع كوشي). بالإضافة لذلك، يقال بشكل عام أن مقدرات طريقة العزوم ليست أكفأ. لهذا تستخدم غالباً كتقريب أولي لإيجاد مقدرات أكثر كفاءة نحصل عليها بطرق أخرى (فمثلاً، طريقة التراكم).

4.6 طرق أخرى OTHER METHODS

هنالك طرق أخرى مختلفة للحصول على مقدرات نقطية للمعالم. نذكر منها:

1. طريقة χ^2 الصغرى.
2. طريقة المسافات الصغرى.
3. طريقة بتمان.
4. طريقة بيز.

سندرس بشكل موجز في هذا البند الطرق الثلاث الأولى.

1.4.6 طريقة χ^2 الصغرى Minimum-Chi-Square Method

يمكن الحصول على مجموعة من الطرق لإيجاد تقديرات معالم توزيع $f(x;\theta)$، وذلك بناءً على صياغة مقياس D بهذه الطريقة أوتلك لقياس إنحراف دالة التوزيع التجريبي $F_n^*(x)$ عن $F(x;\theta)$. ويعتبر مثل هذا المقياس في كل الحالات دالة في العينة $\boldsymbol{X} = (X_1,\dots,X_n)$ والمعلمة θ، أي أن $D = D(\boldsymbol{X};\theta)$.

إذا وجد مثل هذا المقياس، فعندئذٍ كتقدير للمعلمة θ نـأخذ القيمـة الـتي تجعـل D تبلغ أصغر قيمة ممكنة لهـا.

وإحدى أهم وأكثر مثل هذه المقاييس استخداماً هو ما يدعى بمقياس χ^2، الذي قدم من قبل العالم كارل بيرسون. ويتعين هذا المقياس (مقياس χ^2) على النحو الآتي: نقسم مجموعة القيم الممكنة Ω_ξ للمتغير العشوائي الملاحظ ξ (نطاق ξ range of) إلى k بمجموعة جزئيـة (خلية cell) منفصلة $\zeta_1,\zeta_2,\ldots,\zeta_k$ أي $\left(\bigcup_{i=1}^{k}\zeta_i\ ,\ \zeta_i \cap \zeta_j = \varnothing\ ;\ i \neq j\right)$. وليكـن المتغـير العشوائي γ_j يشير إلى عدد عناصر العينة $X=(X_1,\ldots,X_n)$ التي تقع في الخليـة ζ_i، أي أن:

$$\gamma_j = \left|\{i\,;\,X_i \in \zeta_j\}\right|;\ j=1,\ldots,k\,,\left(\sum_{1}^{k}\gamma_j = n\right)$$

لنرمز بـ $p_j(\theta)$ لاحتمال وقوع المتغـير العشـوائي الملاحـظ ξ في الخليـة ζ_j، وهذا يمكن إيجاده، بحيث إذا كانت $F(x;\theta)$ دالة توزيـع ξ فـإن:

$$p_j(\theta) = P_\theta\left(\xi \in \zeta_j\right) = \int_{\zeta_j} dF(x;\theta)\quad ;\quad j=1,\ldots,k$$

حيث أن $\sum_{j=1}^{k} p_j(\theta) = 1$.

إن التكرار النسبي $\frac{\gamma_j}{n}$ لوقوع عناصر العينة X في الخلية ζ_j يعتـبر مقـدر متسـق للاحتمال $p_j(\theta)$ [راجع العلاقة (5.3.1)]، لهذا مقيـاس انحـراف معطيـات العينـة عن القيم النظرية الموافقة لها يمكن إعطاءه بالمقياس المعـرف علـى الصـورة:

$$D = \sum_{j=1}^{k} c_j\left(\gamma_j/n - p_j(\theta)\right)^2$$

إذا وضعنا هنا $c_j = n/p_j(\theta)$، فنحصل على مقياس χ^2:

$$\chi^2 = \sum_{j=1}^{k}\frac{n}{p_j(\theta)}\left(\frac{\gamma_j}{n} - p_j(\theta)\right)^2 = \sum_{j=1}^{k}\frac{\left(\gamma_j - np_j(\theta)\right)^2}{np_j(\theta)} \qquad (1.4.6)$$

يمثـل المقـدار $\sum_{j=1}^{n}\left(\gamma_j - np_j(\theta)\right)^2$ مجمـوع مربعـات الفروقـات بـين العـدد الملاحظ والمتوقع لوقـوع الملاحظـات في الخليـة $\zeta_j\ ;\ j=\overline{1,k}$. ويمكـن كتابـة العلاقـة (1.4.6) علـى النحـو الآتـي:

$$\chi^2 = \sum_{j=1}^{k}\frac{\gamma_j^2}{np_j(\theta)} - n$$

سنرمز للقيمـة الملاحظـة للمتغـير العشـوائي γ_j بـ n_j. وتجـدر الإشـارة هنـا إلى أن مقـدر χ^2 الصغـرى يعتمـد علـى التجـزأة المقترحـة $\zeta_1,\zeta_2,\ldots,\zeta_k$.

لاحقاً سنستخدم أحياناً رمزاً آخـراً للدلالـة علـى هـذا المقيـاس يدعـى بمقيـاس X^2، ولـه توزيـع χ^2.

إن مقـدر المعلمـة θ، الـذي نحصـل عليـه مـن شـرط بلـوغ المقيـاس χ^2 قيمتـه الصغـرى، يدعـى بمقـدر طريقـة χ^2 الصغـرى. وأن هـذا المقـدر، عنـد شـروط عامـة، يتمتـع بالخصـائص التاليـة: متسـق، طبيعـي بالتقـارب، والأكفـأ تقاربيـاً (كمقـدر المعقوليـة العظمـى).

لإيجاد تقديـرات طريقـة χ^2 الصغـرى يجـب حـل منظومـة المعـادلات:

$$\sum_{j=1}^{k}\frac{\gamma_j^2}{p_j^2(\theta)}\frac{\partial p_j(\theta)}{\partial \theta_i} = 0 \quad ; \quad i=1,2,\ldots,r\ ,\ \left(\theta = \left(\theta_1,\ldots,\theta_r\right)\right) \tag{2.4.6}$$

الـتي نحصـل عليهـا مـن الشـرط $\frac{\partial \chi^2}{\partial \theta} = 0$. لكـن حـل هـذه المنظومـة ليـس سـهلاً حتـى في أبسـط الحـالات، لهـذا تستبدل عـادة بمنظومـة المعـادلات:

$$\sum_{j=1}^{k}\frac{\gamma_j}{p_j(\theta)}\frac{\partial p_j(\theta)}{\partial \theta_j} = 0 \quad ; \quad i=1,2,\ldots,r \tag{3.4.6}$$

الـتي حلهـا أبسـط بكثـير. والمقـدرات الـتي نحصـل عليهـا بنـاءً علـى حـل منظومـة المعـادلات (3.4.6)، عند شـروط معينـة، تتمتـع في حالـة عينـات كبـيرة الحجـم بنفـس

الخواص المقاربة التي تتمتع بها مقدرات طريقة χ^2 الصغـــرى باســـتخدام حـــل منظومـــة المعـــادلات (2.4.6).

لهذه الطريقة في تقدير معالم توزيع أهمية عند اختبـــــار جـــودة التـــلاؤم باســـتخدام اختبار χ^2.

مثـــــال 1.4.6

إذا كانت $X=(X_1,\dots,X_n)$ عينة عشوائية من التوزيع $\mathcal{L}(\xi)\in\Pi(\theta)$ فأوجد مقدر طريقة χ^2 الصغرى:

لنـأخذ:

$$\zeta_j=\{j-1\}\quad ;\quad j=1,\dots,k-2\quad ,\quad \zeta_k=\{k-1,k,\dots\}$$

عندئذٍ:

$$p_j(\theta)=f(j-1;\theta)=\frac{\theta^{j-1}}{(j-1)!}e^{-\theta}\quad ;\quad j=1,\dots,k-2$$

$$p_k(\theta)=\sum_{l=k-1}^{\infty}f(l;\theta)=\sum_{l=k-1}^{\infty}\frac{\theta^l}{l!}e^{-\theta}$$

حيـــث أن:

$$f(x;\theta)=\frac{\theta^x}{x!}e^{-\theta}\quad ;\quad x=0,1,\dots.$$

وعلى ذلـــك:

$$\frac{dp_j(\theta)}{d(\theta)}=\left[\frac{(j-1)\theta^{j-2}}{(j-1)!}-\frac{\theta^{j-1}}{(j-1)!}\right]e^{-\theta}=\left(\frac{j-1}{\theta}-1\right)\frac{\theta^{j-1}}{(j-1)!}e^{-\theta}$$

$$\frac{dp_k(\theta)}{d\theta}=\sum_{l=k-1}^{\infty}\left(\frac{l\theta^{l-1}}{l!}e^{-\theta}-\frac{\theta^l}{l!}e^{-\theta}\right)=$$

$$=\sum_{l=k-1}^{\infty}\left(\frac{l}{\theta}-1\right)\frac{\theta^l}{l!}e^{-\theta}=\sum_{l=k-1}^{\infty}\left(\frac{l}{\theta}-1\right)f(l;\theta)$$

بالتعويض في (3.4.6) نجد:

$$\sum_{j=1}^{k-1}\frac{\gamma_j}{p_j(\theta)}\frac{dp_j(\theta)}{d\theta}+\frac{\gamma_k}{p_k(\theta)}\frac{dxp_k(\theta)}{d\theta}=$$
$$=\sum_{j=0}^{k-2}\left(\frac{j}{\theta}-1\right)\gamma_{j+1}+\gamma_k\sum_{l=k-1}^{\infty}\left(\frac{1}{\theta}-1\right)f(l;\theta)\Big/\sum_{l=k-1}^{\infty}f(l;\theta)=0$$

بحل هذه المعادلة بالنسبة لـ θ نحصل على مقدر طريقة χ^2 الصغرى للمعلمة θ:

$$\theta^*(\boldsymbol{X})=\frac{1}{n}\left[\sum_{j=0}^{k-2}j\gamma_{j+1}+\gamma_k\sum_{l=k-1}^{\infty}lf(l;\theta)\Big/\sum_{l=k-1}^{\infty}f(l;\theta)\right]$$

نلاحظ أن الحد الأول (ضمن القوسين) يمثل مجموع X_i حيث أن $X_i \le k-2$، والحد الثاني يساوي تقريبا مجموع X_i، حيث $X_i \ge k-2$ (حيث أن $X_1,...,X_n$ القيم الملاحظة للمتغير ξ)، أي أن $\theta^*(\boldsymbol{X})\approx\overline{X}$.

مثال 2.4.6

لتكن $\boldsymbol{X}=(X_1,...,X_n)$ عينة عشوائية من توزيع بيرنولي $B(1,\theta)$، ونريد إيجاد مقدر θ باستخدام طريقة χ^2 الصغرى.

لنأخذ $\gamma_j=$ عدد الملاحظات في العينة $\boldsymbol{X}$ المساوية لـ $j\ ;\ j=0,1$. أي أن مدى المتغير العشوائي الملاحظ ξ جزء إلى خليتين:

$$\zeta_0=\{0\}\quad,\quad\zeta_1=\{1\}$$

ومن ثم:

$$\chi^2=\sum_{j=0}^{1}\frac{\left[n_j-np_j(\theta)\right]^2}{np_j(\theta)}=\frac{(n_1-n\theta)^2}{n}\frac{1}{\theta(1-\theta)}$$

أو:

$$\chi^2 = \sum_{j=0}^{1} \frac{n_j^2}{np_j(\theta)} - n = \frac{n_0^2}{n(1-\theta)} + \frac{n_1^2}{n\theta} - n$$

حيث $n_j \; ; \; j = 0,1$ القيمة الملاحظة للمتغير العشوائي γ_j:

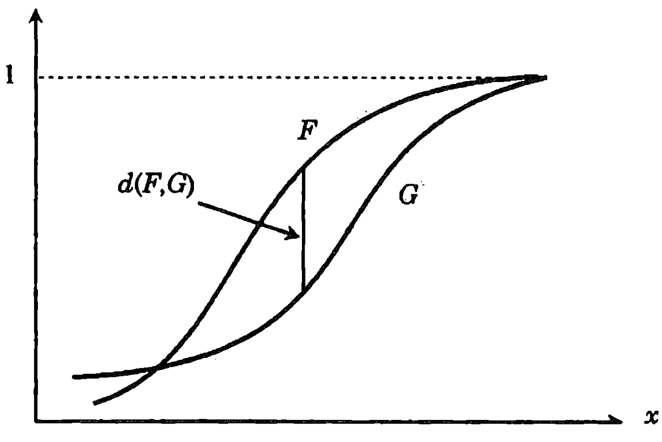
نلاحظ بسهولة أن القيمة الصغري لـ χ^2 كدالة في θ هي عبارة عن قيمة χ^2 وذلك عندما $\theta = \theta^* = \frac{n_1}{n}$، ومن ثم مقدر χ^2 الصغرى للمعلمة θ هو $\theta^*(X) = \frac{\gamma_1}{n}$.

2.4.6 طريقة المسافة الصغرى Minumum-Distance method

لتكن $X = (X_1, \ldots, X_n)$ عينة عشوائية من توزيع $F(x;\theta)$، ولتكن $d(F,G)$ دالة المسافة، التي تعين التباعد الاقصى بين دالتي التوزيع F و G. وكمثال لدالة المسافة نذكر:

$$d(F,G) = \sup_x |F(x) - G(x)|$$

التي تمثل أكبر مسافة شاقولية (رأسية) بين F و G. ويبدو ذلك بوضوح على الشكل (1.4.6).



شكل (1.4.6)

إن تقدير طريقة المسافة الصغرى للمعلمة θ وليكن θ^* هو عبارة عن القيمة $\theta \in \Theta$ التي تجعل $d\left[F(x;\theta),F_n^*(x)\right]$ تبلغ نهاية صغرى، حيث $F_n^*(x)$ دالة التوزيع التجريبي. هكذا، يتم اختيار θ^* بحيث تكون $F\left(x;\theta^*\right)$ الأقرب إلى $F_n(x)$ ضمن عائلة النموذج $\boldsymbol{F} = \left\{F(x;\theta)\ ,\ \theta \in \Theta\right\}$. ومن الطبيعي دائماً الرغبة في الحصول على مقدر المسافة الصغرى، لكن غالباً إيجاد ذلك أمر صعب. والمثال الآتي يعتبر استثناءً.

مثال 3.4.6

لتكن $X = \left(X_1,\dots,X_n\right)$ عينة عشوائية من توزيع بيرنولي:

$$F(x;\theta) = \begin{cases} 0 & ; \quad x \le 0 \\ 1-\theta & ; \quad 0 < x \le 1 \\ 1 & ; \quad x > 1 \end{cases}$$

حيث $0 \le \theta \le 1$.

لنرمز بـ n_j لعدد الملاحظات المساوية لـ $j\ ;\ j = 0,1$. عندئذٍ:

$$F_n^*(x) = \begin{cases} 0 & ; \quad x \le 0 \\ n_0/n & ; \quad 0 < x \le 1 \\ 1 & ; \quad x > 1 \end{cases}$$

والآن إذا استخدمنا دالة المسافة $d(F,G) = \sup_x \left|F(x) - G(x)\right|$، فإن:

$$d\left[F(x;\theta) - F_n^*(x)\right]$$

تبلغ نهاية صغرى إذا أخذنا:

$$1-\theta = \frac{n_0}{n}$$

أو:

$$\theta = \frac{n_1}{n} = \frac{1}{n}\sum_1^n x_i = \bar{x}$$

أي أن مقدر المسافة الصغـــرى:

$$\theta^*(X) = \overline{X}$$

3.4.6 مقدر بتمان لمعلمة الموضـــع والمقيـــاس
Pitman Estimator of Location and Scale Parametrec

لنعرف في البداية معلمة الموضع ومعلمـــة المقيـــاس.

تعريف 1.4.6: معلمـــة الموضـــع Location Parameter

إذا كان التوزيع الاحتمالي $f(x;\theta)$ للمتغير العشوائي ξ يعتمد على معلمـــة مـــا θ وحيـــدة البعـــد، فيقال أن θ معلمة موضع إذا أمكن كتابة $f(x;\theta)$ كدالـة في $(x-\theta)$، أي أن:

$$f(x;\theta) = g(x-\theta)$$

حيث $g(.)$ دالة ما. وبعبارة أخرى نقول عن معلمـــة θ أنهـــا معلمـــة موضـــع إذا كان التوزيع الاحتمالي للمتغير العشوائي $\zeta = \xi - \theta$ لا يعتمـــد علـــى θ، وتوزيـــع ζ يكـون:

$$g(y) = f(y+\theta;0)$$

مثـــال 4.4.6

إذا كان ξ متغير عشوائي دالة كثافتـــه الاحتماليـــة:

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2} \quad ; \quad -\infty < x < +\infty$$

فإن θ معلمة موضع لأن $f(x;\theta)$ دالة في $(x-\theta)$ وتوزيع $\zeta=\xi-\theta$ هو:

$$g(y)=f(y+\theta;0)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2} \qquad ; \qquad -\infty<x<+\infty$$

أي إذا كان $\mathcal{L}(\xi)\in N(\theta,1)$ فإن $\mathcal{L}(\zeta)=N(0,1)$.

وكمثال آخر لمعلمة الموضع نذكر المعلمة θ في التوزيع $R(\theta-1/2,\theta+1/2)$.

تعريف 2.4.6: معلمة مقياس Scale Parameter

إذا كان التوزيع الاحتمالي $f(x;\theta)$ للمتغير العشوائي ξ يعتمد على معلمة θ وحيدة البعد، فيقال أن θ معلمة مقياس إذا أمكن كتابة $f(x;\theta)$ على الصورة:

$$f(x;\theta)=\frac{1}{\theta}g\left(\frac{x}{\theta}\right)$$

حيث $g(.)$ دالة ما. وعندئذٍ يكون توزيع المتغير العشوائي $\zeta=\frac{\xi}{\theta}$:

$$g(y)=f(\theta y,1)$$

وهو لا يعتمد على θ.

مثال 5.4.6

إذا كان ξ متغير عشوائي كثافة توزيعه:

$$f(x;\theta)=\frac{1}{\theta}e^{-x/\theta} \qquad ; \qquad x<0$$

فإن θ معلمة مقياس، لأنه إذا وضعنا $\zeta=\frac{\xi}{\theta}$ فإن كثافة توزيع ζ هي:

$$g(y)=f(y\theta,1)=e^{-y} \qquad ; \qquad y>0$$

لا يعتمد على θ.

وكأمثلة أخرى على معلمة المقياس نذكر المعلمة θ في النموذج الطبيعي $N(0,\theta^2)$، وكذلك θ في النموذج المنتظم $R(0,\theta)$.

مقدر بتمــان لمعلمــة الموضــع

Pitman Estimator for Location Parameter

تعطي المبرهنة الآتية، التي سنقبلها بدون إثبات، مقدر بتمان لمعلمة موضع θ (وحيدة البعد) في التوزيع $f(x;\theta)$، وتبين أن له أقل متوسط مربع خطأ بانتظام بين مجموعة مقدرات الموضع (uniformly smallest mean square error).

مـــبرهنة 5.4.6

إذا كـانت $X=(X_1,\ldots,X_n)$ عينـة عشـوائية مـأخوذة مـن التوزيـع $f(x;\theta)$، وكـانت θ معلمـة موضـع وحيـدة البعـد، فـإن الإحصـاء $T(X)$ المعـرف بالعلاقـة:

$$T(X)=\frac{\int \theta L(X;\theta)d\theta}{\int L(X;\theta)d\theta}$$

هو مقـدر للمعلمـة θ بـأقل متوسـط مربـع خطـأ بانتظـام بـين مجموعـة مقـدرات الموضـع، ويدعـى هـذا المقـدر لمعلمـة الموضـع θ بمقـدر بتمـان (Pitman etimator)، ومـن ثـم فإن أيـة قيمـة $t=T(x)$ تدعـى بتقدير بتمـان لـ θ.

مثـــال 6.4.6

إذا كـانت $X=(X_1,\ldots,X_n)$ عينـة عشـوائية مـأخوذة مـن توزيـع $N(\theta,4)$ فـأوجد مقـدر بتمـان للمعلمـة θ.

يمكـن التـأكد بسـهولة أن θ معلمـة موضـع. وبمـا أن:

$$f(x;\theta)=\frac{1}{2\sqrt{2\pi}}e^{-\frac{1}{8}(\bar{x}-\theta)^2}$$

فإن:

$$L(x;\theta)=(8\pi)^{-\frac{n}{2}}\exp\left\{-\frac{1}{8}\left[\sum_{i=1}^{n}(x_i-\theta)^2\right]\right\}=$$

$$= (8\pi)^{-\frac{n}{2}} \exp\left\{-\frac{1}{8}\left[\sum_{1}^{n}(x_i - \bar{x})^2 + n(\theta - \bar{x})^2\right]\right\}$$

وعلى ذلك فإن تقدير بتمان:

$$t = T(x) = \frac{\int \theta L(x;\theta)d\theta}{\int L(x;\theta)d\theta} = \frac{\int \theta \exp\left[-\frac{n}{8}(\theta - \bar{x})^2\right]d\theta}{\int \exp\left[-\frac{n}{8}(\theta - \bar{x})^2\right]d\theta}$$

$$= \frac{\int \frac{\sqrt{n}}{2\sqrt{2\pi}}\theta e^{-\frac{n}{8}(\theta-\bar{x})^2}d\theta}{\int \frac{\sqrt{n}}{2\sqrt{2\pi}}e^{-\frac{n}{8}(\theta-\bar{x})^2}d\theta}$$

نلاحظ أن $\frac{\sqrt{n}}{2\sqrt{2\pi}}e^{-\frac{n}{8}(\theta-\bar{x})^2}$ ما هي إلا دالة كثافة التوزيع الطبيعي للمتغير العشوائي θ بمتوسط $\bar{x}$ وتباين $\sigma^2 = \frac{4}{n}$، وبالتالي فإن التكامل في البسط عبارة عن $E\theta = \bar{x}$ بينما التكامل في المقام يساوي الواحد. أي أن $t = \bar{x}$، ومن ثم فمقدر بتمان:

$$T = T(X) = \bar{X}$$

يمكن التأكد من أن $T = \bar{X}$ مقدٌ غير متحيز وتباينه يساوي الحد الأدنى في متباينة كرامر ورامر، أي أنه المقدر الأمثل (ومن ثم الأكفأ) للمعلمة θ.

مثال 7.4.6

إذا كانت $X = (X_1,...,X_n)$ عينة عشوائية مأخوذة من توزيع منتظم $R\left(\theta - \frac{1}{2},\theta + \frac{1}{2}\right)$، فأوجد مقدّر بتمان للمعلمة θ.

بما أن θ معلمة موضع في التوزيع $R\left(\theta - \frac{1}{2},\theta + \frac{1}{2}\right)$، و

$$f(x;\theta)=1 \quad ; \quad \theta-\frac{1}{2}\le x\le\theta+\frac{1}{2}$$

فإن:

$$L(x;\theta)=1 \quad ; \quad \theta-\frac{1}{2}\le x_{(1)}\le x_{(n)}\le\theta+\frac{1}{2}$$

إذن:

$$L(x;\theta)=1 \quad ; \quad y_n-\frac{1}{2}\le\theta\le y_1+\frac{1}{2}$$

حيث أن:

$$y_1=x_{(1)} \quad , \quad y_n=x_{(1)}$$

وعلى ذلك:

$$t=T(x)=\int_{y_n-\frac{1}{2}}^{y_1+\frac{1}{2}}\theta L(x;\theta)d\theta \Bigg/ \int_{y_n-\frac{1}{2}}^{y_1+\frac{1}{2}}L(x;\theta)d\theta=\int_{y_n-\frac{1}{2}}^{y_1+\frac{1}{2}}\theta d\theta \Bigg/ \int_{y_n-\frac{1}{2}}^{y_1+\frac{1}{2}}d\theta$$

$$=\left(\frac{\theta^2}{2}\bigg|_{y_n-\frac{1}{2}}^{y_1+\frac{1}{2}}\right)\Bigg/\left(\theta\bigg|_{y_n-\frac{1}{2}}^{y_1+\frac{1}{2}}\right)=\frac{1}{2}\frac{\left(y_1+\frac{1}{2}\right)^2-\left(y_n-\frac{1}{2}\right)^2}{\left(y_1+\frac{1}{2}\right)-\left(y_n+\frac{1}{2}\right)}=\frac{y_1+y_2}{2}$$

وبالتالي فمقدر بتمان:

$$T(X)=\frac{X_{(1)}+X_{(n)}}{2}$$

مقدر بتمان لمعلمة مقياس Pitman Estimator for Scal Parameter

إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية مأخوذة من توزيع متغير عشوائي ع يفترض قيماً موجبة ويعتمد على معلمة مقياس وحيدة البعد $\theta>0$، فإن المبرهنة الآتية تعطي مقدر بتمان لـ θ:

مــــبرهنة 2.4.6

إن مقــدر بتمــان لمعلمــة المقيــاس θ في التوزيـــع $f(x;\theta)$ للمتغـــير العشـــوائي ع يعطى بالعلاقـة:

$$T = T(X) = \frac{\int_0^\infty \frac{1}{\theta^2} L(x;\theta)d\theta}{\int_0^\infty \frac{1}{\theta^3} L(x;\theta)d\theta}$$

وهذا المقدر هـو دالـة في إحصـاء كـافي.

مثـــال 8.4.6

إذا كـانت $X = (X_1,\ldots,X_n)$ عينـة عشـوائية مـن توزيـــع منتظـــم:

$$f(x;\theta) = \frac{1}{\theta} \quad ; \quad 0 < x \le \theta$$

فـأوجد مقـدر بتمـان للمعلمـة θ.

نلاحـظ أن θ معلمـة مقيـاس، لأنـه بوضـع $y = \frac{x}{\theta}$ نجـد:

$$g(y) = f(\theta y, 1) = 1 \quad ; \quad 0 < y \le 1$$

وهـي لا تعتمـد علـى θ.

بمـا أن:

$$L(x;\theta) = \frac{1}{\theta^n} \quad ; \quad \theta \ge x_{(n)} = y_n$$

فإن تقديـر بتمـان للمعلمـة θ يكـون:

$$T(x) = \int_{y_n}^\infty \frac{1}{\theta^2}\frac{1}{\theta^n}d\theta \Big/ \int_{y_n}^\infty \frac{1}{\theta^3}\frac{1}{\theta^n}d\theta = \int_{y_n}^\infty \frac{1}{\theta^{n+2}}d\theta \Big/ \int_{y_n}^\infty \frac{1}{\theta^{n+3}}d\theta = \frac{n+2}{n+1}y_n$$

وبالتـالي مقـدر بتمـان للمعلمـة θ:

$$T(\boldsymbol{X}) = \frac{n+2}{n+1} Y_n = \frac{n+2}{n+1} X_{(n)}$$

بما أن:

$$E_\theta Y_n = \frac{n}{n+1}\theta$$

لأن كثافة توزيع Y_n هو:

$$h(y) = \frac{n}{\theta^n} y^{n-1} \quad ; \quad 0 < y \le \theta$$

وعلى ذلك فإن:

$$E_\theta T(\boldsymbol{X}) = \frac{n+2}{n+1} EY_n = \frac{n+2}{n+1}\frac{n}{n+1}\theta = \frac{n(n+2)}{(n+1)^2}\theta = \theta - \frac{\theta}{(n+1)^2} \ne \theta$$

إذن فمقدّر بتمان للمعلمة θ متحيز، ومقدار التحيز هذا:

$$E_\theta T - \theta = -\frac{\theta}{(n+1)^2}$$

وهو مقدار صغير عندما تكون n كبيرة.

يمكن إثبات أن المقدر $\frac{n+1}{n} Y_n$ لـ θ غير متحيز وبأقل تباين. ونترك إثبات ذلك للقارئ على سبيل المثال.

مثال 9.4.6

إذا كانت $\boldsymbol{X} = (X_1, \ldots, X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta) = \frac{1}{\theta} e^{-x/\theta} \quad ; \quad x > 0 \ , \ \theta > 0$$

فأوجد مقدّر بتمان للمعلمة θ.

نعلم من المثال (5.4.6) أن θ معلمة مقياس. كما أن:

$$L(x;\theta) = \frac{1}{\theta^n} e^{-\sum x_i/\theta}$$

وبالتـالي فتقدير بتمـان للمعلمـة θ هـو:

$$T(x) = \int_0^\infty \frac{1}{\theta^{n+2}} e^{-\sum x_i/\theta} d\theta \Big/ \int_0^\infty \frac{1}{\theta^{n+3}} e^{-\sum x_i/\theta} d\theta$$

$$= \frac{\Gamma(n+1)}{\left(\sum x_i\right)^{n+1}} \frac{\left(\sum x_i\right)^{n+2}}{\Gamma(n+2)} = \frac{\sum x_i}{n+1}$$

ومن ثم فمقـدّر بتمـان:

$$.\ T(X) = \frac{n}{n+1}\overline{X}$$

وهو مقدّر غـير متحـيز بـأقل تبـاين.

تمـــارين

1. بفـرض $X = (X_1,...,X_n)$ عينـة عشـوائية مـن توزيـع:

$$f(x;\theta) = \theta x^{-2} \quad ; \quad x \geq \theta > 0$$

   1. أوجد مقـدر المعقوليـة لـ θ، ثـم بـين إن كـان متحـيزاً أم لا.
   2. أوجـد مقـدر العـزوم للمعلمـة θ، ثـم بـين إن كـان متحـيزاً أم لا.

2. لتكن $X = (X_1,...,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta) = \frac{1}{2} e^{-|x-\theta|} \quad ; \quad -\infty < \theta < +\infty$$

   1. أوجد مقـدر المعقوليـة العظمـى لـ θ.
   2. أوجد مقـدر العـزوم لـ θ.
   3. أوجد مقـدر بتمـان لـ θ.
   4. هـل ينتمـي التوزيـع $f(x;\theta)$ لعائلـة النمـاذج الأسـية؟

.3 لتكن $X=(X_1,...,X_n)$ عينة عشوائية من التوزيع:

$$f(x;\theta)=\frac{1}{\theta} \quad ; \quad 0\le x\le\theta \ , \ \theta>0$$

1. أوجد مقدر θ باستخدام طريقة العزوم، وإذا رمزنا له بـ T_1 فأوجد متوسطه ومتوسط مربع الخطأ له.
2. أوجد مقدر المعقولية العظمى لـ θ وإذا رمزنا له بـ T_2، فأوجد متوسطه ومتوسط مربع الخطأ له.
3. أوجد المقدر الأكفأ لـ θ، وإذا رمزنا له بـ T_4، فأوجد متوسطه ومتوسط مربع الخطأ له.
4. إذا رمزنا بـ $T_5=X_{(1)}+X_{(n)}$، فأوجد متوسط ومتوسط مربع الخطأ له.

.4 لتكن $X=(X_1,...,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\theta x^{\theta-1} \quad ; \quad 0<x<1 \ , \ \theta>0$$

1. أوجد مقدر المعقولية العظمى لـ $\tau(\theta)=\dfrac{4\theta+1}{1+\theta}$.
2. هل توجد دالة في θ لها مقدر غير متحيز بأقل تباين؟.

.5 لتكن $X=(X_1,...,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\theta(1+x)^{-(1+\theta)} \quad ; \quad x<0 \ , \ \theta>0$$

1. قدر θ باستخدام طريقة العزوم.
2. أوجد مقدر المعقولية العظمى لـ $\tau(\theta)=\dfrac{1}{\theta}$.
3. أوجد مقدر المعقولية لكل من الدوال:

$$\frac{2}{3+\theta} \quad , \quad \ln\theta \quad , \quad \frac{2\theta}{5\theta+4}$$

.6 لتكن $X=(X_1,...,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\frac{\ln\theta}{\theta-1}\theta^{x} \quad ; \quad 0<x<1 \ , \ \theta>1$$

1. أوجد مقدر المعقولية العظمى لـ θ.
2. هل يوجد المقدر الأمثل لـ θ، إن وجد فما هو؟.
3. أوجد توزيع مقدر المعقولية لـ θ عندما $n\to\infty$.
4. أوجد مقدر المعقولية العظمى لـ $\tau(\theta)=1/(1+\theta)$، ثم أوجد توزيعه عندما $n\to\infty$.

7. بفرض $\boldsymbol{X}=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=c_2^x\theta(1-\theta)^{2-x} \quad ; \quad x=0,1,2 \ , \ \theta>0$$

1. أوجد مقدر المعقولية العظمى لـ θ^2، هل هو غير متحيز؟
2. هل يوجد المقدر الأمثل لـ θ^2.
3. أوجد توزيع مقدر المعقولية العظمى لـ θ^2 عندما $n\to\infty$.
4. هل مقدر المعقولية العظمى لـ θ^2 متسق.

8. لتكن $\boldsymbol{X}=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\begin{cases}\dfrac{2x}{\theta} & ; \quad 0<x\le\theta\\[2mm] \dfrac{2(1-x)}{1-\theta} & ; \quad \theta<x\le 1\end{cases}$$

حيث $0\le\theta\le1$.

1. قدر θ بطريقة العزوم.
2. أوجد مقدّر المعقولية العظمى لـ θ من أجل $n=1$ و $n=2$.
3. أوجد المقدّر الأمثل لـ θ عندما $n=1$ (إن وجد).
4. أوجد مقدّر المعقولية العظمى لـ θ.

9. بفرض $\boldsymbol{X}=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=e^{-(x-\theta)} \quad ; \quad x\geq\theta \ , \ -\infty<\theta<\infty$$

1. أوجد مقدّر المعقولية العظمى لـ θ.
2. أوجد مقدّر المعقولية العظمى لـ $\tau(\theta)=\frac{1}{\theta}$.
3. أوجد مقدّر العزوم للمعلمة θ.
4. أوجد توزيع مقدّر المعقولية العظمى لـ θ.

10. بفرض $\boldsymbol{X}=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\frac{2x}{\theta^2} \quad ; \quad 0<x<\theta \ , \ \theta>0$$

فأوجد:

1. مقدّر المعقولية العظمى للمعلمة θ.
2. مقدّر المعقولية العظمى لـ $\tau_1(\theta)=\frac{1}{\theta}$.
3. مقدّر المعقولية العظمى لـ $\tau_2(\theta)=(1+\theta)e^{-\theta}$.

11. بفرض $\boldsymbol{X}=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $\Gamma(\theta,\lambda)$:

1. أثبت أن مقدر المعقولية العظمى لـ $\tau(\theta)=\frac{1}{\theta}$ هو $\hat{\tau}_n=\frac{\lambda}{\overline{X}}$.
2. هل المقدر $\hat{\tau}_n$ متسق.
3. أوجد توزيع $\hat{\tau}_n$ عندما $n\to\infty$.

12. أثبت أن:

$$\hat{\theta}(\boldsymbol{X})=\frac{\overline{X}}{r+\overline{X}} \quad , \quad \overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

مقدّر المعقولية العظمى لـ θ في النموذج $\overline{B}=(r,\theta)$، ثم احسب التباين التقريبي له، بفرض n كبيرة.

13. ليكن ξ متغير عشوائي يخضع للتوزيع:

$$f(x;\theta)=(2x/\theta)e^{-x^2/\theta} \quad ; \quad x\geq 0$$

أوجد مقدر المعقولية العظمى لـ θ بناءً على عينة عشوائية $\boldsymbol{X}=(X_1,\ldots,X_n)$.

14. لدينا عينة عشوائية $((X_1,Y_1),\ldots,(X_n,Y_n))$ من توزيع طبيعي ثنائي البعد:

$$N\left((0,0)\ .\ \begin{Vmatrix}1 & \theta\\ \theta & 1\end{Vmatrix}\right) \quad ; \quad \theta\in(-1,1)$$

أثبت أن التباين التقريبي للمقدر $\hat{\theta}_n$ يساوي $(1-\theta^2)^2/n(1+\theta^2)$.

15. لتكن $\boldsymbol{X}=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\frac{1}{\theta} \quad ; \quad \theta\leq x\leq 2\theta$$

1. أوجد مقدر المعقولية العظمى لـ θ.
2. أوجد مقدر بتمان لـ θ.

16. إذا كانت $\boldsymbol{X}=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $Be(\theta_1,\theta_2)$ فأوجد مقدر المعقولية العظمى ومقدر العزوم لـ $\theta=(\theta_1,\theta_2)$.

17. إذا كانت $((X_1,Y_1),\ldots,(X_n,Y_n))$ عينة عشوائية من توزيع طبيعي ثنائي:

$$N((0,0),\Sigma);S=\begin{Vmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{Vmatrix}$$

فأوجد مقدر المعقولية العظمى للمعالم المجهولة $\sigma_1^2,\sigma_2^2,\rho$.

الفصل السابع:

التقدير بفترة

1.7 مقدمة

تعرضنا في الفصلين السابقين إلى خواص المقدِّرات النقطية وطرق الحصول على أفضل مقدِّر نقطي من أجل معلمة θ (أو دالة لها) في نموذج معطى $\boldsymbol{F} = \{F(x;\theta)\ ,\ \theta \in \Theta\}$. لكن لم نستعرض مقياساً للثقة التي يتمتع بها مثل هذا المقدِّر.

ورأينا عند بحث المقدَّرات النقطية في الفصل السادس، إن أي مقدِّر نقطي يمثل بدالة $T = T(X)$ في العينة $X = (X_1,...X_n)$. وتأخذ هذه الدالة عند كل ملاحظة x للعينة X قيمة وحيدة، أي نقطة واحدة $t = T(x)$، التي تتخذ كقيمة تقريبية (تقدير) للمعلمة المقدَّرة θ. هذا بالإضافة إلى أن قيم المقدِّر $T = T(X)$، عند ملاحظات متكررة للعينة X، بشكل عام مختلفة فيما بينها وتختلف عن القيمة الحقيقية للمعلمة θ، ومن ثم من المفيد أيضاً معرفة الخطأ الممكن ارتكابه نتيجة لاستخدام التقدير $t = T(x)$ بدلاً من القيمة الحقيقية للمعلمة غير المعلومة θ، أي إيجاد مقياس لمثل ذلك الخطأ. لذا في حالات عدة من المفيد إيجاد فترة (أو منطقة في حالة معلمة متجه)، التي تشتمل على القيمة الحقيقية للمعلمة θ باحتمال كبير γ. في هذه الحالة يتم الحديث عن التقدير بفترة (interval estimation)، ويتم اختيار قيمة α مسبقاً، بحيث تعكس "درجة الاستعداد لقبول الخطأ". يصف طول الفترة دقة تقدير المعلمة، لذا يفضل اختيار أقصر فترة. وإذا أمكن بناء مثل هذه

الفترة، فإنها تعكس دقة أكبر لتقدير المعلمة. هكذا، مما سبق نجد أنه من المفيد إيجاد تقدير بفترة نتوقع أن تقع فيها القيمة الحقيقية للمعلمة، بالإضافة إلى مقياس للتأكد من أن المعلمة تقع داخل هذه الفترة. وهذا يعني بدلاً من تقدير المعلمة بنقطة واحدة فإننا نقدرها بفترة، وطرق الحصول على مثل هذه الفترات للمعالم المجهولة لتوزيع ستكون موضوع هذا الفصل.

2.7 فترات الثقة CONFIDENCE INTERVALS

لتكن $X=(X_1,\dots,X_n)$ عينة عشوائية مأخوذة من التوزيع $f(x;\theta)$، الذي يعتمد على معلمة وحيدة البعد θ غير معلومة، ونرغب في تقدير θ بفترة.

لتقدير المعلمة θ بفترة يتم البحث عن إحصائين (مقدّرين)

$$T_1=T_1(X) \quad , \quad T_2=T_2(X) \quad ; \quad T_1<T_2$$

يحققان، عند قيمة معطاة $0<\gamma<1$. الشرط الآتي:

$$P_\theta\big(T_1(X)<\theta<T_2(X)\big)=\gamma \quad ; \quad \theta\in\Theta \tag{1.2.7}$$

تعريف 1.2.7: فترة الثقة Confidence Interval

إن الفترة العشوائية $\big(T_1(X),T_2(X)\big)$ تسمى $100\gamma\%$ فترة ثقة للمعلمة θ، أو $100\gamma\%$ مقدّر فترة ثقة لـ θ، ويسمى العدد γ بدرجة أو معامل الثقة (degree or coefficient confidence)، كما تسمى T_1,T_2 بحدي الثقة الأدنى والأعلى (lower and upper confidence limits) على الترتيب للمعلمة θ.

هكذا، $100\gamma\%$ فترة ثقة للمعلمة θ عبارة عن فترة عشوائية محتواة في فضاء المعلمة Θ، أي أن $(T_1,T_2)\subset\Theta$ ومرتبطة بالعينة X (لكن ليس بـ θ) وتشتمل على القيمة الحقيقية للمعلمة المجهولة θ باحتمال يساوي γ، حيث γ لا تعتمد على θ.

إذا كان الشرط (1.2.7) محققاً، فهذا يعني عملياً الآتي: لنفترض أجرينا عدداً كبيراً من التجارب المستقلة، وكل واحدة تتألف من n ملاحظة على متغير عشوائي ع توزيعه $f(x;\theta)$، وكانت نتائج كل تجربة موصوفة بالعينة $\boldsymbol{X}$. وحسبنا عند كل مشاهدة $\boldsymbol{x}=(x_1,\dots,x_n)$ للعينة $\boldsymbol{X}$ الفترة $\left(T_1(\boldsymbol{x}),T_2(\boldsymbol{x})\right)$، فإن $100\gamma\%$ من هذه الفترات تشتمل على القيمة الحقيقية للمعلمة المجهولة θ. أي أن القيمة الحقيقية لـ θ تنتمي لفترة الثقة $\left(T_1(\boldsymbol{X}),T_2(\boldsymbol{X})\right)$، وهذه القاعدة يمكن أن تكون خاطئة في بعض الحالات، لكن عدد مثل هذه الحالات، عندما تكون γ قريبة من الواحد (أو $\alpha=1-\gamma$ قريبة من الصفر)، صغير ويشكل $100\alpha\%$ من مجمل حالات تطبيق هذه القاعدة. وتدعى أيضاً القيمة $\left(T_1(\boldsymbol{x}),T_2(\boldsymbol{x})\right)$ للفترة العشوائية $\left(T_1(\boldsymbol{X}),T_2(\boldsymbol{X})\right)$ $100\gamma\%$ فترة الثقة لـ θ.

يتم البحث أحياناً عن فترات ثقة ذات جانب واحد (one-sided confidence interval)، وعلى الترتيب من الأعلى (الشكل $\theta<T_2(X)$) ومن الأدنى (الشكل $\theta>T_1(X)$)، أي أن أحد الحدين T_1,T_2 (وليس كلاهما) محدد، بحيث:

$$P_\theta\left(\theta<T_2(\boldsymbol{X})\right)=\gamma \qquad (2.2.7)$$
$$P_\theta\left(\theta>T_1(\boldsymbol{X})\right)=\gamma$$

وبشكل مشابه تعرف فترة الثقة من أجل مركبات معلمة $\boldsymbol{\theta}$ متعددة الأبعاد، فمثلاً من أجل المركبة θ_1 للمعلمة $\boldsymbol{\theta}=(\theta_1,\theta_2)$:

$$P_{\boldsymbol{\theta}}\left(T_1(\boldsymbol{X})<\theta_1<T_2(\boldsymbol{X})\right)=\gamma \quad ; \quad \boldsymbol{\theta}\in\Theta \qquad (3.2.7)$$

وأيضاً فترة الثقة من أجل دالة معلمية حقيقية $\tau(\theta)$ (θ يمكن أن تكون وحيدة البعد أو متعددة الأبعاد):

$$P_\theta\left(T_1(\boldsymbol{X})<\tau(\theta)<T_2(\boldsymbol{X})\right)=\gamma \quad ; \quad \theta\in\Theta \qquad (4.2.7)$$

مـــلاحظة

إذا كـانت $\tau(\theta)$ دالـة مضطـردة تماماً في θ، وكـانت $(T_1(X),T_2(X))$ هـي $100\gamma\%$ فـترة ثقـة لـ θ، فإن $(\tau(T_1),\tau(T_2))$ هـي $100\gamma\%$ فـترة ثقـة للدالـة $\tau(\theta)$ (إذا كـانت τ مـتزايدة تماماً) و $(\tau(T_2),\tau(T_1))$ هـي $100\gamma\%$ فـترة لـ $\tau(\theta)$ (إذا كـانت τ متناقصة تماماً).

كما في التقدير بنقطة أن مسألة التقدير بفترة تتمثل في:

1. طريقة بناء أو إيجاد فترة الثقة.
2. معايير تقييم جودة فترة الثقة، أي معايير المفاضلة بين فترات الثقة.

3.7 طريقة الكمية المحورية *Pivotal Quantity Method*

إحدى الطرق التي تمكننا في حالات عدة من بناء فترات الثقة تسمى طريقة الكمية المحورية. وقبل عرض هذه الطريقة لابد من تعريف الكمية المحورية (pivotal quantity) أولاً.

تعريف 1.3.7: الكمية المحورية Pivotal Quantity

إذا كـانت $X=(X_1,...,X_n)$ عينة عشوائية من توزيع مستمر $F(x,\theta)$، ووجد متغير عشوائي $Q=Q(X;\theta)$ يتصف بما يلي:

1. دالة في العينة X والمعلمة θ، التي نريد بناء فترة الثقة من أجلها دون أي معلمة أخرى، ومعرفة عند كل نقطة $\theta\in\Theta$ وعند كل نقطة $x\in G$.
2. دالة مستمرة ومضطردة تماما (متزايدة تماماً أو متناقصة تماماً) في θ من أجل كل $x\in G$.
3. التوزيع الاحتمالي لـ $Q(X;\theta)$ معلوم ولا يعتمد على المعلمة θ.

فمثل هذا المتغير العشوائي يسمى كمية محورية (pivotal quantity) من أجل المعلمة θ في $F(x;\theta)$.

وبشكل مشابه نعرف الكمية المحورية من أجل المركبات الفردية (فمثلاً θ_1) في حالة معلمة θ متعددة الأبعاد ونرمز لها بـ $Q(\boldsymbol{X};\theta_1)$، وأيضاً من أجل دالة معلمية حقيقية $\tau=\tau(\theta)$ (θ وحيدة البعد أو متعددة الأبعاد) ونرمز لها بـ $Q(\boldsymbol{X};\tau)$. وسنتطرق فيما يلي، ومن أجل الاختصار والتبسيط لحالة معلمة وحيدة البعد θ أو دالة حقيقية لها.

مثال 1.3.7

إذا كانت $\boldsymbol{X}=(X_1,\ldots,X_n)$ عينة عشوائية مأخوذة من توزيع طبيعي $N(\theta,\sigma^2)$، فإن:

$$Q(\boldsymbol{X};\theta)=\frac{\overline{X}-\theta}{\sigma/\sqrt{n}}$$

كمية محورية للتوزيع $N(\theta,\sigma^2)$ من أجل المعلمة θ وتوزيعها:

$$f_Q(q)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}q} \quad ; \quad -\infty<q<\infty$$

لا يعتمد على المعلمة θ، بينما $\overline{X}/\theta$ ليس كمية محورية لأن توزيعه $N(1,\sigma^2/n\theta)$ يعتمد على المعلمة θ.

مثال 2.3.7

إذا كانت $\boldsymbol{X}=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\theta e^{-\theta x} \quad ; \quad x>0$$

وكانت $Q(\boldsymbol{X};\theta)=\theta\sum_{i=1}^{n}X_i$ فإن توزيع Q:

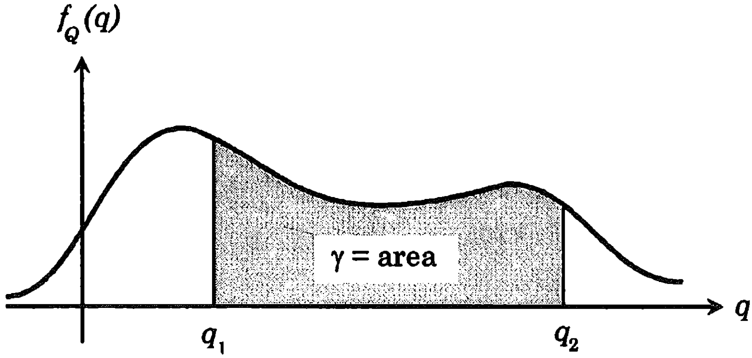
$$f_Q(q)=\frac{1}{\Gamma(n)}q^{n-1}e^{-q} \quad ; \quad q>0$$

وبذلك تكون Q كمية محورية من أجل θ لأنها تحقق الشروط الثلاث المعرفة للكمية المحورية (يمكن للقارئ التأكد منها بسهولة).

لنفترض من أجل نموذج مستمر $F(x;\theta)$ يعتمد على معلمة وحيدة البعد θ أمكن إيجاد كمية محورية $Q(\boldsymbol{X};\theta)$، كثافة توزيعها الاحتمالية $f_Q(q)$، عندئذٍ نستطيع من أجل قيمة معينة معطاة $0<\gamma<1$ إيجاد عددين $q_1<q_2$; q_1,q_2 (وهذا ممكن بطرق مختلفة طبعاً)، بحيث أن:

$$P\left(q_1<Q(\boldsymbol{X};\theta)<q_2\right)=\int_{q_1}^{q_2} f_Q(q)dq=\gamma \tag{1.3.7}$$

أنظر شكل (1.3.7).



شكل (1.3.7)

لنعين الآن من أجل $\boldsymbol{x}\in G$ العددين (التقديرين) $T_i(\boldsymbol{x})$; $i=1,2$، حيث $T_1(\boldsymbol{x})<T_2(\boldsymbol{x})$، وذلك كحلين للمعادلتين:

$$Q(\boldsymbol{x};\theta)=q_1 \quad , \quad Q(\boldsymbol{x};\theta)=q_2 \tag{2.3.7}$$

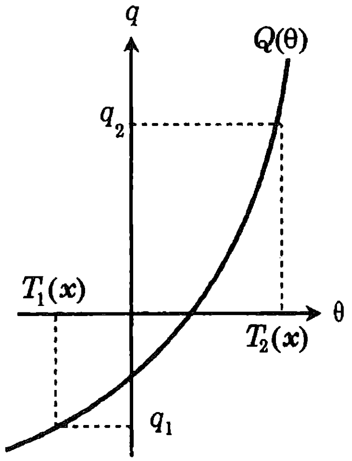
بالنسبة لـ θ [يتعين هذان العددان تماماً، حسب الشرط (2) المفروض على $Q(\boldsymbol{X};\theta)$]، عندئذٍ المتباينة المزدوجة:

$$q_1<Q(\boldsymbol{x};\theta)<q_2$$

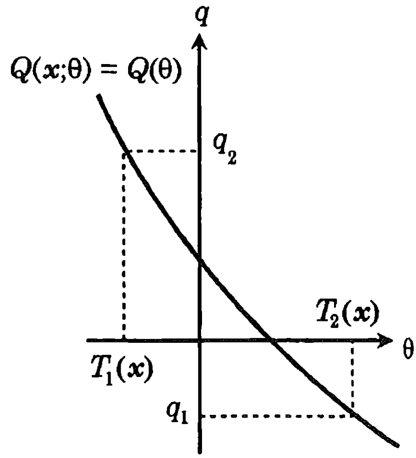
مكافئة للمتباينة المزدوجة:

$$T_1(x) < \theta < T_2(x)$$

كما يبدو ذلك بوضوح على الشكلين (2.3.7) و (3.3.7).



شكل (3.3.7)
$Q(\theta)$ دالة متزايدة تماماً في θ



شكل (2.3.7)
$Q(\theta)$ دالة متناقصة تماماً في θ

وبالتالي يمكن إعادة صياغة العلاقة (1.3.7) على الصورة:

$$P_\theta\left(T_1(X) < \theta < T_2(X)\right) = \gamma \quad ; \quad \theta \in \Theta \tag{3.3.7}$$

هكذا، الفترة $\left(T_1(X), T_2(X)\right)$ هي عبارة عن $100\gamma\%$ فترة ثقة للمعلمة θ وطولها (length of the confidence) هو $L_\gamma = T_2 - T_1$.

تجدر الإشارة هنا إلى الملاحظات الآتية:

1. q_1, q_2 مستقلان عن θ، لأن توزيع Q لا يعتمد على θ.
2. من أجل كل قيمة معينة $0 < \gamma < 1$ هنالك أزواج عدة ممكنة (q_1, q_2) تحقق العلاقة:

$$P_\theta(q_1 < Q(X;\theta) < q_2) = \gamma \quad ; \quad \theta \in \Theta \tag{4.3.7}$$

وأن كل زوج من الأعداد (q_1, q_2) ينتج عنه فترة ثقة (t_1, t_2)، وبالتالي يجب اختيار الزوج الذي يجعل القيمة المتوقعة لطول الفترة العشوائية (T_1, T_2)، أي L_γ أصغر ما يمكن.

3. الميزة الرئيسية لطريقة الكمية المحورية تتمثل في أن المتباينة:

$$q_1 < Q(x;\theta) < q_2$$

يمكن إعادة صياغتها أو عكسها (inverted) أو تحويرها (pivoted) إلى متباينة من الشكل:

$$T_1(x) < \theta < T_2(x)$$

وذلك من أجل كل قيمة $x \in G$.

إن طريقة الكمية المحورية في بناء فترة الثقة للمعالم المجهولة في توزيع تشتمل على مرحلتين:

المرحلة الأولى: إيجاد كمية محورية.

المرحلة الثانية: إعادة صياغة المتباينة $q_1 < Q(x;\theta) < q_2$ وكتابتها على الشكل $T_1(x) < \theta < T_2(x)$.

4.7 بناء فترات الثقة باستخدام كمية محورية في حال المعاينة من توزيع طبيعي

سنتطرق في هذا البند لبناء فترات الثقة لبعض الحالات الهامة والشائعة تطبيقياً والمتعلقة بالمعاينة من توزيع طبيعي (sampling from the normal distribution).

1.4.7 فترة الثقة لمتوسط النموذج الطبيعي I

لتكن $\boldsymbol{X}=(X_1,\ldots,X_n)$ عينة عشوائية مأخوذة من توزيع $N(\theta,\sigma^2)$، ونريد بناء $100\gamma\%$ فترة ثقة للمعلمة θ.

بما أن:

$$Q(\boldsymbol{X};\theta)=\frac{\overline{X}-\theta}{\sigma/\sqrt{n}}$$

كمية محورية متناقصة تماماً في θ وتوزيعها:

$$\mathcal{L}\left(\frac{\overline{X}-\theta}{\sigma/\sqrt{n}}\right)=N(0,1)$$

فالمعادلتان (2.3.7) تكتبان على الصورة:

$$\frac{\overline{x}-\theta}{\sigma/\sqrt{n}}=q_1 \quad , \quad \frac{\overline{x}-\theta}{\sigma/\sqrt{n}}=q_2$$

وبحل كل منهما بالنسبة لـ θ نجد:

$$T_1(\boldsymbol{x})=\overline{x}-\frac{\sigma}{\sqrt{n}}q_2 \quad , \quad T_2(\boldsymbol{x})=\overline{x}-\frac{\sigma}{\sqrt{n}}q_1$$

وعلى ذلك فإن $100\gamma\%$ فترة ثقة للمعلمة θ هي عبارة عن أي فترة من الشكل:

$$\Delta_\gamma(\boldsymbol{X})=\left(\overline{X}-\frac{\sigma}{\sqrt{n}}q_2,\overline{X}-\frac{\sigma}{\sqrt{n}}q_1\right) \tag{1.4.7}$$

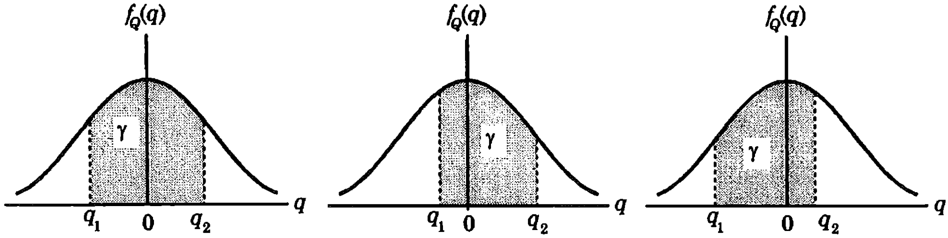
حيث $q_1<q_2$ أي عددين يحققان الشرط:

$$P(q_1<Q<q_2)=\gamma$$

لكن $\mathcal{L}(Q)=N(0,1)$، إذن:

$$P(q_1 < Q < q_2) = \Phi(q_2) - \Phi(q_1) = \gamma \tag{2.4.7}$$

وفي الحقيقة هنالك أزواج عدة (q_1,q_2) تحقق العلاقة (2.4.7)، يمكن إيضاح بعضها من خلال الأشكال الآتية:



طول الفترة العشوائية $\Delta_\gamma(X)$ المعرفة بالعلاقة (1.4.7):

$$L_\gamma = L_\gamma(q_1,q_2) = \frac{\sigma}{\sqrt{n}}(q_2 - q_1) \tag{3.4.7}$$

يعتمد على q_1,q_2 وحجم العينة n. وبازدياد حجم العينة n يتناقص طول فترة الثقة، بينما بازدياد الفرق $(q_2 - q_1)$ يزداد طول فترة الثقة هذه، وهذا الفرق يتعلق بقيمة γ واختيار q_1,q_2. لذا للحصول على أصغر فترة ثقة ضمن الفترات من الشكل (1.4.7) بافتراض n و γ معطاة، يجب البحث عن القيمة الصغرى لـ $L_\gamma(q_1,q_2)$ عند الشرط (2.4.7).

باشتقاق العلاقة (2.4.7) بالنسبة لـ q_1 نحصل على:

$$\varphi(q_2)\frac{dq_2}{dq_1} - \varphi(q_1) = 0 \Rightarrow \frac{dq_2}{dq_1} = \frac{\varphi(q_1)}{\varphi(q_2)}$$

حيث أن:

$$\varphi(x) = \Phi'(x)$$

وللحصول على القيمة الصغرى لـ L_γ نضع:

$$\frac{\partial L_\gamma}{\partial q_1} = \frac{\sigma}{\sqrt{n}}\left(\frac{dq_2}{dq_1} - 1\right) = \frac{\sigma}{\sqrt{n}}\left(\frac{\varphi(q_1)}{\varphi(q_2)} - 1\right) = 0 \Rightarrow \varphi(q_1) = \varphi(q_2)$$

ويتحقق ذلك عندما تكون $q_1 = q_2$ وهذا حل مرفوض [يتعارض مع الشرط (2.4.7)]. أو عندما تكون $q_1 = -q_2$ لأن الدالة $\varphi(q)$ زوجية. وبناءً على هذه النتيجة وأيضاً المعادلة (2.4.7) والعلاقة $\Phi(-x) = 1 - \Phi(x)$ نحصل على المساواة

$$\Phi(q_2) = \frac{1+\gamma}{2} = 1 - \frac{\alpha}{2} \quad ; \quad \gamma = 1 - \alpha$$

وباستخدام جدول التوزيع الطبيعي المعياري نجد $q_2 = Z_{\alpha/2}$ [$Z_{\alpha/2}$ القيمة الحرجة للتوزيع $N(0,1)$ الموافقة لاحتمال $(1-\alpha/2)$].

هكذا، $100\gamma\%$ فترة ثقة من الشكل $100\gamma\%$ [أصغر فترة ثقة المثلى $\Delta_\gamma(X)$] من أجل المعلمة θ في النموذج $N(\theta,\sigma^2)$ هي الفترة:

$$\Delta^*_\gamma(X) = \left(\bar{X} - \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \, , \, \bar{X} + \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}\right) \tag{4.4.7}$$

المتناظرة بالنسبة للنقطة العشوائية $\bar{X}$ وطولها $2\frac{\sigma}{\sqrt{n}} Z_{\alpha/2}$. فمثلاً، إذا كنا نرغب في تقدير متوسط مجتمع طبيعي بتباين معلوم $\sigma^2 = 225$، ولدينا عينة عشوائية بحجم $n = 20$ ومتوسط $\bar{x} = 64.3$، عندئذٍ من أجل $\gamma = 0.95$ نجد $Z_{(1+\gamma)/2} = Z_{0.975} = 1.96$. وبالتالي 97.5% فترة ثقة لمتوسط المجتمع θ بناءً على العينة المشاهدة:

$$64.3 - 1.96\frac{15}{\sqrt{20}} < \theta < 64.3 + 1.96\frac{15}{\sqrt{20}}$$

$$57.7 < \theta < 70.9$$

2.4.7 فترة الثقة لتباين النموذج الطبيعي II

لتكن $X = (X_1,\ldots,X_n)$ عينة عشوائية من التوزيع $N(\mu,\theta^2)$، ونريد بناء $100\gamma\%$ فترة ثقة للتباين غير المعلوم θ^2.

يمكن بسهولة في هذه الحالة إيجاد كمية محورية تعتمد على $\tau = \tau(\theta) = \theta^2$، وهي:

$$Q = Q(X;\tau) = \frac{1}{\tau}\sum_{i=1}^{n}\left(X_i - \mu\right)^2$$

بما أن:

$$\mathcal{L}\left(\frac{X_i - \mu}{\theta}\right) = N(0,1)$$

فحسب المبرهنة (7.3.2):

$$\mathcal{L}\left(\left(\frac{X_i - \mu}{\theta}\right)^2\right) = \chi^2_{(1)}$$

وبالتالي فإن:

$$\mathcal{L}\left(Q(X;\tau)\right) = \chi^2_{(n)}$$

بما أن Q دالة متناقصة تماماً في τ ، فبحل المعادلتين:

$$G(x,\tau_1) = \frac{1}{\tau_1}\sum_{i=1}^{n}(x_i - \mu)^2 = q_2$$

$$G(x,\tau_2) = \frac{1}{\tau_2}\sum_{i=1}^{n}(x_i - \mu)^2 = q_1$$

بالنسبة لـ τ_1،τ_2 نجد:

$$\tau_1 = T_1(x) = \frac{1}{q_2}\sum_{i=1}^{n}\left(x_i - \mu\right)^2$$

$$\tau_2 = T_2(x) = \frac{1}{q_1}\sum_{i=1}^{n}\left(x_i - \mu\right)^2 \quad ; \quad q_1 < q_2$$

وعلى ذلك فإن $100\gamma\%$ فترة ثقة لـ $\tau = \theta^2$ هي أي فترة من الشكل:

$$\Delta_\gamma(X) = \left(\frac{1}{q_2}\sum_{i=1}^{n}(X_i - \mu)^2, \frac{1}{q_1}\sum_{i=1}^{n}\left(X_i - \mu\right)^2\right) \tag{5.4.7}$$

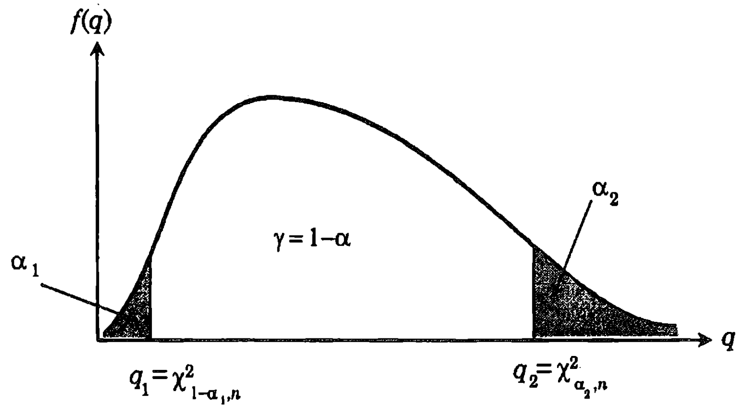
حيث أن q_1, q_2 عددان يحققان العلاقة:

$$\int_{q_1}^{q_2} f_Q(q)dq = \gamma \tag{6.4.7}$$

و $f_Q(q)$ دالـة كثافـة توزيـع $\chi^2_{(n)}$ المعرفـة بالعلاقـة (29.3.2)، ويمكـن الحصـول علـى q_1,q_2 بالمحاولـة والخطأ، ولنرمز لهمـا بـ:

$$q_1 = \chi^2_{1-\alpha_1,n} \quad , \quad q_2 = \chi^2_{\alpha_2,n} \quad ; \quad \gamma = 1-\alpha \quad , \quad \alpha = \alpha_1 + \alpha_2$$

[يمكـن إيجـاد أكـثر مـن زوج (q_1,q_2) طبعـاً، وهـذا يتوقـف علـى α_1,α_2]. ويبـدوان بوضـوح على الشـكل (1.4.7).



شكل (1.4.7)

ومن ثـم تـأخذ فـترة الثقـة (5.4.7) الصـورة الآتيـة:

$$\Delta_\gamma(\boldsymbol{X}) = \left(\frac{1}{\chi^2_{\alpha_2,n}} \sum_{i=1}^{n} (X_i - \mu)^2 , \frac{1}{\chi^2_{1-\alpha_1,n}} \sum_{i=1}^{n} (X_i - \mu)^2 \right) \tag{7.4.7}$$

للحصـول علـى أصغر فـترة ثقـة مـن الشـكل (4.4.7) يجـب اختيــار q_1,q_2 بحيـث تجعل طـول الفـترة:

$$L_\gamma = \left(\frac{1}{q_1} - \frac{1}{q_2} \right) \sum_{i=1}^{n} (X_i - \mu)^2$$

أصغر ما يمكن عند الشرط (6.4.7)، أي البحث عن أصغر قيمة ممكنة للنسبة q_2/q_1. وباشتقاق العلاقة (6.4.7) بالنسبة لـ q_1 نجد:

$$\frac{dq_2}{dq_1}f_Q(q_2)-f_Q(q_1)=0\Rightarrow\frac{dq_2}{dq_1}=\frac{f_Q(q_1)}{f_Q(q_2)}$$

وعلى ذلك وباشتقاق L_γ بالنسبة لـ q_1 نجد:

$$\frac{\partial L_\gamma}{\partial q_1}=\left(-\frac{1}{q_1^2}+\frac{1}{q_2^2}\frac{dq_2}{dq_1}\right)\sum_{i=1}^{n}(X_i-\mu)^2=$$

$$=\left(-\frac{1}{q_1^2}+\frac{1}{q_2^2}\frac{f_Q(q_1)}{f_Q(q_2)}\right)\sum_{i=1}^{n}(X_i-\mu)^2=0\ \Rightarrow$$

$$q_1^2f_Q(q_1)=q_2^2f_Q(q_2)\ \Rightarrow\ q_1f_Q(q_1)=q_2f_Q(q_2)$$

وهذا يعني أن طول فترة الثقة المعرف في (5.4.7) يكون أصغر ما يمكن إذا اخترنا q_1,q_2 بحيث يحققان المساواة:

$$q_1f_Q(q_1)=q_2f_Q(q_2)$$

وبوضع:

$$q_1=\chi^2_{1-\alpha_1,n}\quad,\quad q_2=\chi^2_{\alpha_2,n}\quad;\quad \alpha_1+\alpha_2=1-\gamma$$

نحصل على المعادلة:

$$\left(\chi^2_{\alpha_2,n}/\chi^2_{1-\alpha_1,n}\right)^{n/2}=\exp\left[\left(\chi^2_{\alpha_2,n}-\chi^2_{1-\alpha_1,n}\right)/2\right]\tag{8.4.7}$$

ويعني هذا أن طول فترة الثقة من الشكل (5.4.7) يكون أقل ما يمكن إذا اخترنا α_1,α_2 بحيث يحققان المعادلة (8.4.7).

هكذا، إن الصعوبة الأساسية في الحصول على أصغر فترة ثقة وفق هذه الطريقة تكمن في اختيار α_1,α_2 بحيث يحققان المعادلة (8.4.7)، وهذا بالمحاولة والخطأ. لذا عادة يتم اختيار $\alpha_1=\alpha_2=\frac{1-\gamma}{2}=\frac{\alpha}{2}$ وذلك عندما يكون حجم

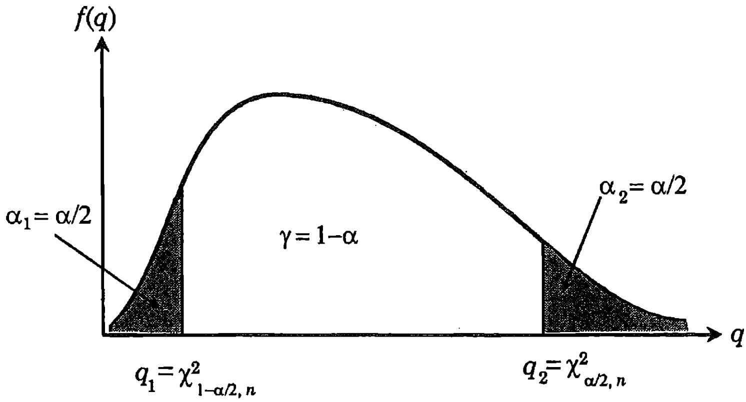
العينـــة $n \geq 20$، لأن النســـبة $\chi^2_{\alpha_2,n}/\chi^2_{1-\alpha_1,n}$ تقـــترب مـــن النســـبة $\chi^2_{\alpha/2,n}/\chi^2_{1-\alpha/2,n}$ بازدياد حجم العينة n، بحيث يصبح الفـــرق بـــين هـــاتين النســـبتين لا يتجـــاوز 0.02 تقريباً عندما تكون $n \geq 20$، وهذا ما يبينه الجدول (1.4.7) مـــن أجـــل درجـــة ثقـــة $\gamma = 0.95$ وبعض القيم لـــ n. وتسمى فترة الثقـــة في هـــذه الحالـــة بفـــترة ثقـــة ذات ذيلين متساويين (شكل 2.4.7) وتـــأخذ الشـــكل:

$$\Delta^*_\gamma = \left(\frac{1}{\chi^2_{\alpha/2,n}} \sum_{i=1}^{n} (X_i - \mu)^2 \; , \; \frac{1}{\chi^2_{1-\alpha/2,n}} \sum_{i=1}^{n} (X_i - \mu)^2 \right) \qquad (9.4.7)$$

وهي كما ذكرنا صيغة تقريبية لأصغر فترة ثقة مـــن الشـــكل (5.4.7).

جـــــدول (1.4.7):

n	$\chi^2_{1-\alpha/2,n}$	$\chi^2_{\alpha/2,n}$	$\chi^2_{1-\alpha_1,n}$	$\chi^2_{\alpha_2,n}$	$\chi^2_{\alpha/2,n}/\chi^2_{1-\alpha/2,n}$	$\chi^2_{\alpha_2,n}/\chi^2_{1-\alpha_1,n}$
2	0.05	7.38	0.08	8.09	147.6	101.13
5	0.83	13.83	0.99	14.37	15.46	14.52
10	3.25	20.48	3.52	21.73	6.30	6.17
20	9.59	34.17	9.96	35.23	3.56	3.54



شكل (2.4.7)

فمثلاً، إذا كنا نرغب في تقدير فترة لتباين مجتمع طبيعي متوسطه $\mu = 65$، بناءً على عينة مشاهدة بحجم $n = 20$. وافترضنا $\sum_{i=1}^{20}(x_i - 65)^2 = 300$. فعندئذٍ من أجل درجة ثقة $\gamma = 0.95$، نجد:

$$q_1 = \chi^2_{0.975,20} = 9.59 \quad , \quad q_2 = \chi^2_{0.025,20} = 34.17$$

وبالتالي 95% فترة ثقة لـ θ^2 (أصغر فترة ثقة) حسب (9.4.7) تكون:

$$\Delta^*_\gamma(x) = \left(\frac{1}{34.17}(300)\ ,\ \frac{1}{9.59}(300)\right) = (8.79\ ,\ 31.28)$$

وباستخدام الصيغة (8.4.7) نجد:

$$\Delta^*_\gamma(x) = \left(\frac{1}{35.23}(300)\ ,\ \frac{1}{9.96}(300)\right) = (8.52\ ,\ 30.12)$$

ويتقلص الفرق بين ناتج الصيغتين (8.4.7) و (9.4.7) بازدياد حجم العينة n.

3.4.7 فترة الثقة للمتوسط والتباين في النموذج الطبيعي العام

لتكن $X = (X_1, ..., X_n)$ عينة عشوائية من توزيع طبيعي $N(\theta_1, \theta_2^2)$، ونريد بناء $100\gamma\%$ فترة ثقة لكل من $\theta = (\theta_1, \theta_2^2)\ ;\ \tau_1(\theta) = \theta_1\ ,\ \tau_2(\theta) = \theta_2^2$.

نلاحظ أن:

$$Q(X;\theta_1) = \sqrt{n-1}\,\frac{\overline{X} - \theta_1}{S(X)}$$

دالة مستمرة ومتناقصة تماماً في θ_1 وتتبع توزيع ستيودنت بـ $(n-1)$ درجة حرية، أي أن:

$$\mathcal{L}(Q) = S(n-1)$$

وذلك حسب المبرهنة (6.9.3)، وبالتالي فإن $Q(X;\theta_1)$ كمية محورية.

لإيجاد $100\gamma\%$ فترة ثقة لـــ θ_1 نتبع نفس الخطوات الـــواردة في الفقـــرة الســـابقة (1.4.7) نجد:

$$\sqrt{n-1}\,\frac{\bar{x}-\theta_1}{S(x)}=q_2 \;\Rightarrow\; T_1(x)=\bar{x}-\frac{S(x)}{\sqrt{n-1}}q_2$$

$$\sqrt{n-1}\,\frac{\bar{x}-\theta_1}{S(x)}=q_1 \;\Rightarrow\; T_2(x)=\bar{x}-\frac{S(x)}{\sqrt{n-1}}q_1$$

وعلى ذلك فإن $100\gamma\%$ فترة ثقـــة لــــــ θ_1 هـــي عبـــارة عـــن أي فـــترة مـــن الشـــكل:

$$\Delta_\gamma(X)=\left(\overline{X}-\frac{S(x)}{\sqrt{n-1}}q_2\;,\;\overline{X}-\frac{S(x)}{\sqrt{n-1}}q_1\right) \tag{10.4.7}$$

حيث أن $q_1<q_2$ عددان يحققان الشرط:

$$P(q_1<Q<q_2)=\int_{q_1}^{q_2}f(t)dt=\gamma$$

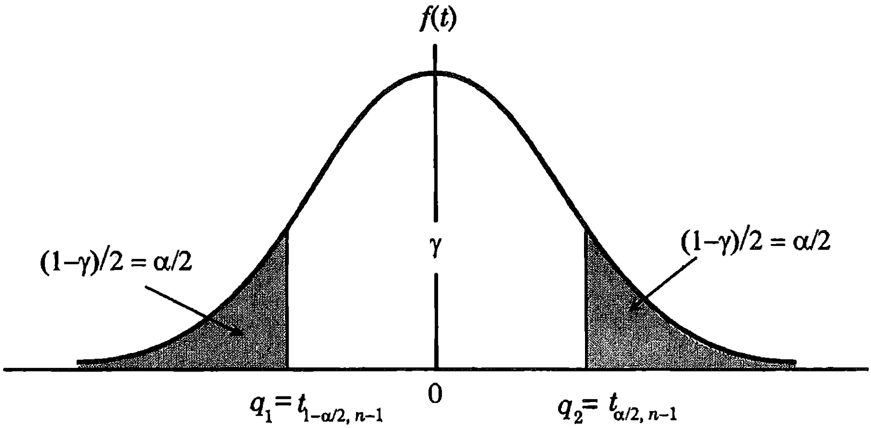
وأن $f(t)=f_Q(t)$ هي دالة كثافة توزيع ستيودنت بـــ $(n-1)$ درجة حرية. وطول هـــذه الفترة:

$$L_\gamma=\frac{S(X)}{\sqrt{n-1}}(q_2-q_1) \tag{11.4.7}$$

ويكون أقل ما يمكن عندما يكون الفرق (q_2-q_1) أصغر ما يمكن، ونلاحظ بســـهولة بناءً على تناظر توزيع t حول الصفر، أن ذلك يحدث عندما $q_1=-q_2$. ويمكن إثبات ذلك جبرياً بنفس الأسلوب الذي اتبعناه في الفقرة (1.4.7)، أي أن:

$$q_2=t_{(1-\gamma)/2,n-1}=t_{\alpha/2,n-1}\qquad,\qquad q_1=t_{(1+\gamma)/2,n-1}=-t_{\alpha/2,n-1}$$

حيث $t_{\alpha/2,n-1}$ القيمة الحرجة لتوزيع $S(n-1)$ الموافق لاحتمال $\alpha/2$، ويمكن الحصـــــول عليها من جدول توزيع t (الملحق جدول رقم 10).



شكل (3.4.7)

وعلى ذلك فإن $100\gamma\%$ فترة ثقة لـــ θ_1 بأقل طول هي:

$$\Delta_\gamma(X) = \left(\overline{X} - \frac{S(x)}{\sqrt{n-1}} t_{\alpha/2,n-1} \ , \ \overline{X} + \frac{S(x)}{\sqrt{n-1}} t_{\alpha/2,n-1}\right) \tag{12.4.9}$$

وطولها:

$$L_\gamma = \frac{2S(x)}{\sqrt{n-1}} t_{\alpha/2,n-1} \tag{13.4.7}$$

والقيمة المتوقعة لـــه:

$$E\left(L_\gamma\right) = \frac{2}{\sqrt{n-1}} t_{\alpha/2,n-1} E(S)$$

فمثلاً، من أجل عينة مشاهدة حجمها $n = 10$، $\overline{x} = 25$ و $S = 4$، فإن 95% فـترة ثقة لـــ θ_1 تكون:

$$\left(25 - \frac{4}{\sqrt{9}}(2.262) \ , \ 25 + \frac{4}{\sqrt{9}}(2.262),\right)$$

أي أن:

$$-21.198 < \theta_1 < 28.016$$

أما لإيجاد $100\gamma\%$ فترة ثقة لـ θ_2^2 بطول أصغري، نتبع ذات الخطوات التي اتبعناها في الفقرة (2.4.7)، بعد تعيين كمية محورية $G\left(\boldsymbol{X};\theta_2^2\right)$. نلاحظ أن:

$$Q = \frac{nS^2(\boldsymbol{X})}{\theta_2^2} \quad ; \quad S^2(\boldsymbol{X}) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

تتبع توزيع $\chi^2_{(n-1)}$ [حسب المبرهنة (5.9.3)]، وعلى ذلك فإن Q كمية محورية لتقدير $\tau_2 = \theta_2^2$، ومن ثم:

$$\frac{nS^2(\boldsymbol{x})}{\theta_2^2} = q_2 \;\Rightarrow\; T_1(\boldsymbol{x}) = \frac{nS^2(\boldsymbol{x})}{q_2}$$

$$\frac{nS^2(\boldsymbol{x})}{\theta_2^2} = q_1 \;\Rightarrow\; T_2(\boldsymbol{x}) = \frac{nS^2(\boldsymbol{x})}{q_1}$$

وبالتالي $100\gamma\%$ فترة ثقة لـ θ_2^2 هي عبارة عن أي فترة من الشكل:

$$\Delta_\gamma(\boldsymbol{X}) = \left(\frac{nS^2(\boldsymbol{X})}{q_2}, \frac{nS^2(\boldsymbol{X})}{q_1}\right) \tag{14.4.7}$$

حيث $q_1 < q_2$ عددان يحققان الشرط:

$$P\left(q_1 < Q < q_2\right) = \int_{q_1}^{q_2} f(q)dq = \gamma$$

وأن $f_Q(q) = f(q)$ دالة كثافة توزيع $\chi^2_{(n-1)}$، أي أن: $\alpha_1 + \alpha_2 = \alpha = 1 - \gamma$; $q_2 = \chi^2_{\alpha_2, n-1}$, $q_1 = \chi^2_{1-\alpha_1, n-1}$

وطول هذه الفترة:

$$L_\gamma = nS^2\left(\frac{1}{q_2} - \frac{1}{q_1}\right) \tag{15.4.7}$$

وعندما $n \geq 20$ فأصغر قيمة للطول L_γ هي:

$$L_\gamma^* = nS^2\left(\frac{1}{\chi^2_{\alpha/2,n-1}} - \frac{1}{\chi^2_{1-\alpha/2,n-1}}\right)$$

وعلى ذلك %100γ فترة ثقة لـ θ_2^2، والتي تتمتع بطول أصغري ضمن %100γ فترات ثقة من الشكل (14.4.7) هي:

$$\Delta_\gamma^*(X) = \left(\frac{nS^2(X)}{\chi^2_{\alpha/2,n-1}}\,,\,\frac{nS^2(X)}{\chi^2_{1-\alpha/2,n-1}}\right) \tag{16.4.7}$$

أما إذا كانت $n < 20$ فإن %100γ فترة ثقة لـ θ_2^2 ذات طول أصغري هي من الشكل:

$$\left(\frac{nS^2(X)}{\chi^2_{\alpha_2,n-1}}\,,\,\frac{nS^2(X)}{\chi^2_{1-\alpha_1,n-1}}\right) \quad ; \quad \alpha_1 + \alpha_2 = \alpha \tag{17.4.7}$$

حيث أن $\chi^2_{\alpha_2,n-1}$ ، $\chi^2_{1-\alpha_1,n-1}$ يعينان من العلاقة (8.4.7) باستبدال n بـ $(n-1)$.

فمثلاً، من أجل عينة عشوائية بحجم $n = 20$، $S^2 = 9$ فإن %90 فترة ثقة لـ θ_2^2 هي:

$$\left(\frac{180}{16.919}\,,\,\frac{180}{3.325}\right)$$

أي أن:

$$10.645 < \theta_2^2 < 54.135$$

إن طريقة الكمية المحورية يمكن تطبيقها على حالات أخرى غير إيجاد فترات الثقة لمعالم توزيع، ونورد فيما يلي ثلاث حالات هامة تطبيقياً.

4.4.7 فترة الثقة للفرق بين متوسطي نموذجين طبيعيين

لتكن $X = (X_1,\dots,X_m)$ عينة عشوائية مأخوذة من توزيع $N(\theta_1^{(1)},\theta_2^2)$ و $Y = (Y_1,\dots,Y_n)$ عينة عشوائية من توزيع $N(\theta_1^{(2)},\theta_2^2)$، وكانت العينتان مستقلتين، ونريد بناء $100\gamma\%$ فترة ثقة للفرق $(\theta_1^{(1)} - \theta_1^{(2)})$.

نلاحظ أن النموذجين مختلفين فقط بمتوسطيهما، وكما نعلم:

$$\mathcal{L}_{\theta_1}(\overline{X}) = N\left(\theta_1^{(1)},\frac{\theta_2^2}{m}\right) \quad , \quad \theta_1 = (\theta_1^{(1)},\theta_2^2)$$

$$\mathcal{L}_{\theta_2}(\overline{Y}) = N\left(\theta_1^{(2)},\frac{\theta_2^2}{n}\right) \quad , \quad \theta_2 = (\theta_1^{(2)},\theta_2^2)$$

وعلى ذلك:

$$\mathcal{L}_{\theta}(\overline{X}-\overline{Y}) = N\left((\theta_1^{(1)}-\theta_1^{(2)}),\frac{m+n}{mn}\theta_2^2\right) \quad ; \quad \theta = (\theta_1^{(1)},\theta_1^{(2)},\theta_2^2)$$

$$\mathcal{L}_{\theta}\left(\frac{(\overline{X}-\overline{Y})-(\theta_1^{(1)}-\theta_1^{(2)})}{\sqrt{\frac{m+n}{mn}}\,\theta_2}\right) = N(0,1)$$

وكما نعلم أن [راجع مبرهنة (5.9.3)]:

$$\mathcal{L}_{\theta_1}\left(mS^2(X)/\theta_2^2\right) = \chi^2_{(m-1)}$$
$$\mathcal{L}_{\theta_2}\left(nS^2(Y)/\theta_2^2\right) = \chi^2_{(n-1)}$$

ومن ثم:

$$\mathcal{L}_{\theta}\left(\frac{mS^2(X)+nS^2(Y)}{\theta_2^2}\right) = \chi^2_{m+n-2}$$

بأخذ:

$$Q = \left(\frac{\left(\overline{X} - \overline{Y}\right) - \left(\theta_1^{(1)} - \theta_1^{(2)}\right)}{\sqrt{\frac{m+n}{mn}}\sqrt{\frac{mS^2(X) + nS^2(Y)}{m+n-2}}} \right) =$$

$$= \sqrt{\frac{mn(m+n-2)}{m+n}} \frac{\left(\overline{X} - \overline{Y}\right) - \left(\theta_1^{(1)} - \theta_1^{(2)}\right)}{\sqrt{mS^2(X) + nS^2(Y)}}$$

نجد أن:

$$\mathcal{L}(Q) = S(m+n-2)$$

أي أن Q تتبع توزيع ستيودنت (توزيع t) بـ $(m+n-2)$ درجة حرية، لذا يمكن أخذ Q ككمية محورية لبناء فترة ثقة للفرق $\left(\theta_1^{(1)} - \theta_1^{(2)}\right)$. وبإتباع نفس الخطوات الواردة في الفقرة (3.4.7) نحصل على $100\gamma\%$ فترة ثقة لـ $\left(\theta_1^{(1)} - \theta_1^{(2)}\right)$ المطلوبة وهي من الشكل:

$$\left(\left(\overline{X} - \overline{Y}\right) \pm t_{\alpha/2, m+n-2} \left[\frac{m+n}{mn(m+n-2)} \left(mS^2(X) + nS^2(Y)\right) \right]^{1/2} \right) \qquad (18.4.7)$$

إذا كان تباينا النموذجين معلومين ويساويان على الترتيب σ_1^2, σ_2^2 فإن الكمية المحورية لتقدير فرق المتوسطين $\left(\theta^{(1)} - \theta^{(2)}\right)$ يمكن بناءها بسهولة على النحو الآتي:

بما أن $\overline{X}, \overline{Y}$ مستقلان، وأن:

$$\mathcal{L}_{\theta^{(1)}}\left(\overline{X}\right) = N\left(\theta^{(1)}, \sigma_1^2/m\right) \quad , \quad \mathcal{L}_{\theta^{(2)}}\left(\overline{Y}\right) = N\left(\theta^{(2)}, \sigma_2^2/n\right)$$

فإن:

$$\mathcal{L}\left(\left(\overline{X} - \overline{Y}\right) - \left(\theta^{(1)} - \theta^{(2)}\right) \Big/ \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} \right) = N(0,1)$$

مهما تكن $\theta^{(1)}$ و $\theta^{(2)}$.

وعلى ذلك يمكن أخـــذ:

$$Q = \frac{\left(\overline{X} - \overline{Y}\right) - \left(\theta^{(1)} - \theta^{(2)}\right)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

ككمية محورية لبناء $100\gamma\%$ فترة ثقة لـ $\left(\theta^{(1)}, \theta^{(2)}\right)$. وبناء $100\gamma\%$ فترة ثقـــة يتـــم بشكل مشابه لما ورد في الفقرة (1.4.7). وفترة الثقة المطلوبة تأخذ الشكل:

$$\left(\overline{X} - \overline{Y} \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}\right) \quad ; \quad \frac{\alpha}{2} = \frac{1-\gamma}{2} \tag{19.4.7}$$

5.4.7 فترة الثقة للفـــرق $(\mu_1 - \mu_2)$ فـــي حالـــة المعاينـــة مـــن توزيع طبيعي ثنـــائي

لتكن $\left((X_1, Y_1), \ldots, (X_n, Y_n)\right)$ عينـــة عشـــوائية مـــن مجتمـــع يخضـــع للتوزيـــع الطبيعي الثنائي بمعالم مجهولـــة:

$$\mu_x = \mu_1 \quad , \quad \mu_y = \mu_2 \quad , \quad \sigma_x^2 = \sigma_1^2 \quad , \quad \sigma_y^2 = \sigma_2^2 \quad , \quad \rho_{xy} = \rho$$

ونريد بناء $100\gamma\%$ فترة ثقة للفرق $(\mu_1 - \mu_2)$.

إذا وضعنـــا

$$Z_i = X_i - Y_i \quad ; \quad i = 1, \ldots, n$$

فإن $Z = (Z_1, Z_2, \ldots, Z_n)$ تعتبر عينة عشوائية مأخوذة من توزيع طبيعـــي بمتوســـط $\mu = (\mu_1 - \mu_2)$ وتباين $\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$، أي من توزيع $N(\mu, \sigma^2)$، وعلـــى ذلك فإن:

$$Q = \frac{\overline{Z} - \mu}{S(Z)/\sqrt{n-1}} \quad ; \quad S^2(Z) = \frac{1}{n}\sum_{i=1}^{n}\left(Z_i - \overline{Z}\right)^2 \;,\; \overline{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i$$

وتوزيعـها:

$$\mathcal{L}(Q) = S(n-1)$$

وبالتـالي Q تعتبر كمية محـورية، وأن $100\gamma\%$ فـترة ثقـة لـ $\mu = \mu_1 - \mu_2$ المطلوبـة:

$$\left(\overline{Z} - \frac{S(Z)}{\sqrt{n-1}} t_{\alpha/2,n-1} \ , \ \overline{Z} + \frac{S(Z)}{\sqrt{n-1}} t_{\alpha/2,n-1}\right) \qquad (20.4.7)$$

تستخدم هذه الطريقـة في حالـة المشـاهدات الثنائيـة (paired observation)، أي عندما تكون كل ملاحظة x_i مترافقة مع ملاحظة y_i.

6.4.7 فترة الثقة للنسبة بين تبـايني نموذجيـن طبيعييـن

لنرى الآن الحالة التي تكون فيها $X = (X_1, \ldots X_m)$ عينـة عشـوائية مـن التوزيـع $N\left(\theta_1^{(1)}, \theta_2^{(2)^2}\right)$، بينمـا $Y = (Y_1, \ldots, Y_n)$ عينـة عشـوائية مـن توزيـع $N\left(\theta_1^{(2)}, \theta_2^{(2)^2}\right)$، والعينتـان مسـتقلتان، ويطلـب تقديـر النسـبة $\theta_2^{(1)^2} / \theta_2^{(2)^2}$ غـير المعلومـة.

يمكن بناء كمية محورية في هذه الحالـة علـى أسـاس المبرهنـة (9.3.2)، والـتي تعطى بالمتغير العشـوائي:

$$F = \frac{m(n-1)S^2(X)}{n(m-1)S^2(Y)} \Bigg/ \frac{\theta_2^{(1)^2}}{\theta_2^{(2)^2}}$$

وهذا الأخير يتبع توزيـع F بـ $\gamma_1 = m-1 \ , \ \gamma_2 = n-1$ درجـة حريـة مـن أجل أي قيم للمعالم $\theta_1^{(1)}, \theta_1^{(2)}, \theta_2^{(1)}, \theta_2^{(2)}$. وبالتـالي يعتـبر F كميـة محوريـة لتقديـر النسبـة $\theta_2^{(1)^2} / \theta_2^{(2)^2}$، وعندئذٍ تبنى $100\gamma\%$ فـترة ثقـة للنسـبة $\theta_2^{(1)^2} / \theta_2^{(2)^2}$ بشـكل مشـابه لبنـاء $100\gamma\%$ فترة الثقة (16.4.7) الـواردة في الفقـرة (3.4.7) وتـأخذ الشـكل:

$$\left(\frac{m(n-1)S^2(X)}{n(m-1)S^2(Y)}\Big/F_{\alpha/2,m-1,n-1}\ ,\ \frac{m(n-1)S^2(X)}{n(m-1)S^2(Y)}\Big/F_{1-\alpha/2,m-1,n-1}\right) \qquad (21.4.7)$$

حيـــث إن $F_{P,m-1,n-1}$ القيمـــة الحرجـــة لتوزيـــع F الموافقـــة لاحتمـــال p و $\gamma_1 = m-1\ ,\ \gamma_2 = n-1$ درجة حرية [الملحق جدول رقم 12].

فمثلاً، من أجل عينتان بحجم $m = n = 9$ و $S(y) = 8\ ,\ S(x) = 10$ ، فـإن 0.90 فترة ثقة لـ $\theta_2^{(1)^2}/\theta_2^{(2)^2}$ هي:

$$(0.454,5.375)$$

ومن أجـل $n = m = 13$:

$$(0.578,4.203)$$

حيـث إن:

$$F_{0.95,8,8} = F^{-1}_{0.05,8,8} = 3.44 \quad , \quad F_{0.95,12,12} = F^{-1}_{0.05,12,12} = 0.69$$

5.7 طرق إيجاد كمية محورية

METHODS OF FINDING PIVOTAL QUANTITY

نلاحظ من خلال الأمثلة السابقة أن الخطوة الأساسية في مســألة بنــاء فــترة ثقــة لمعلمة θ (أو دالة $\tau(\theta)$) بإتبــاع طريقــة الكميــة المحوريــة تتمثــل في إيجــاد الدالــة $Q(X;\theta)$. وإيجــاد هذه الدالة يعتمد على خصوصيـــة النمـــوذج المفـــروض $f(x;\theta)$. وفي حالات عدة لا توجد مثل هذه الدالــة أو مـــن الصعوبــة بمكــان إيجادهــا، لكــن يمكن تمييز مجموعة من النماذج المســتمرة $F(x;\theta)$، الــتي توجــد مــن أجلــها دائمــاً كمية محورية لها شكل بسيط، وخاصــة إذا كــانت دالــة التوزيــع $F(x;\theta)$ مســتمرة ومضطردة تماماً بالنسبة لـ θ من أجل كل نقطة في فضـــاء x، فــإن:

$$Q(X;\theta) = -2\sum_{i=1}^{n} \ln F(X_i;\theta) \qquad (1.5.7)$$

تعتبر كمية محورية، وهذا ما تؤكده المبرهنــــة الآتيــة.

مـــــبرهنة 1.5.7

إذا كـــانت $X = (X_1,...,X_n)$ عينة عشـــوائية مـــن توزيـــع مســـتمر $F(x;\theta)$، وكانت دالة التوزيع $F(x;\theta)$ مستمرة ومضطردة تماماً (مـــتزايدة تمامـــاً أو متناقصـــة تماماً) في θ من أجل كل نقطة في فضاء x، فتوجـــد كميـــة محوريـــة لــــ θ تعطـــى بالعلاقـة

$$Q(X;\theta) = -2\sum_{i=1}^{n} \ln F(X_i;\theta)$$

وتخضع لتوزيـــع $\chi^2_{(2n)}$.

الإثبـــات

نلاحظ بسهولة أن الدالة $Q(X;\theta)$ مستمرة ومضطردة تماماً في θ من أجـــل كـــل $x \in G$، لأن $F(X_i;\theta)$ دالة مستمرة ومضطردة تماماً في θ من أجـــل كـــل نقطـــة في فضاء x (حسب الفرض).

وكما نعلم أن المتغيرات العشوائية $F(X_i;\theta)\ ;\ i=\overline{1,n}$ مستقلة، ولكل منها توزيـــع منتظم $R(0,1)$، أي أن:

$$\mathcal{L}_\theta(F(X_i;\theta)) = R(0,1) \quad ; \quad \theta \in \Theta$$

ومن ثم $-\ln F(X_i;\theta)$ متغـــيرات عشـــوائية مســـتقلة ولكـــل منـــها توزيـــع جامـــا $\Gamma(1,1)$. وبنـــاءً علـــى ذلـــك وخـــواص توزيـــع جامـــا فـــإن المتغـــير العشـــوائي $-\sum \ln F(X_i;\theta)$ يتبع توزيع جامـــا $\Gamma(n,1)$، وبالتـــالي:

$$Q(X;\theta) = -2\sum_{i=1}^{n} \ln F(X_i;\theta)$$

تتبع توزيع $\chi^2_{(2n)}$، وهذا الأخير لا يعتمد على θ. إذن $Q(X;\theta)$ كمية محورية لـ θ.

يمكننا باستخدام الكمية المحورية (1.5.7) بناء $100\gamma\%$ فترة ثقة لـ θ على النحو الآتي:

نختار، عند درجة ثقة معينة γ، عددين $q_1 < q_2$ يحققان العلاقة:

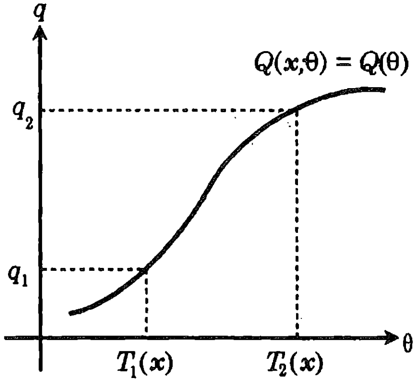
$$\int_{q_1}^{q_2} f_Q(q)dq = \frac{1}{2^n \Gamma(n)} \int_{q_1}^{q_2} q^{n-1} e^{-q/2} dq = \gamma \tag{2.5.7}$$

وبحل المعادلتين:

$$\begin{aligned} -2\sum_{i=1}^{n} \ln F(x_i;\theta) &= q_1 \\ -2\sum_{i=1}^{n} \ln F(x_i;\theta) &= q_2 \end{aligned} \tag{3.5.7}$$

بالنسبة لـ θ، نحصل على العددين $T_1(x) < T_2(x)$. وبالتالي $\left(T_1(X), T_2(X)\right)$ تكون $100\gamma\%$ فترة الثقة المطلوبة للمعلمة θ.

نلاحظ بوضوح أن الصعوبة الأساسية في تطبيق هذه الطريقة على مسألة معينة تكمن في إيجاد حلي المعادلتين (3.5.7)، أي في الحصول على $T_1(x), T_2(x)$. فمثلاً، إذا كانت $Q(X;\theta)$ متزايدة تماماً في θ، فالشكل (1.5.7) يبين فترة الثقة $\left(T_1(x), T_2(x)\right)$.



شكل (1.5.7)

مثـــــال 1.5.7

إذا كــانت $X=\left(X_1,\dots,X_n\right)$ عينــة عشــوائية مــن توزيــع منتظــم $R(0,\theta)$ ، فـأوجد $100\gamma\%$ فـترة ثقـة للمعلمــة θ .

بمـا أن:

$$F(x;\theta)=\frac{x}{\theta} \quad ; \quad 0\le x\le\theta \ , \ \theta>0$$

دالـة مسـتمرة ومتناقصـة تمامـاً في θ ، فحســب المبرهنــة (1.5.7):

$$Q(X;\theta)=-2\sum_{1}^{n}\ln F\left(X_i;\theta\right)=-2\sum_{1}^{n}\ln X_i+2n\ln\theta$$

كميـة محوريـة للمعلمـة θ تخضـع لتوزيـع $\chi^2_{(2n)}$ ، أي أن كثافـة توزيعهــا:

$$f_Q(q)=\frac{1}{2^n\Gamma(n)}q^{n-1}e^{-q/2} \quad ; \quad q>0$$

وباختيــار عدديــن $q_1<q_2$ يحققــان العلاقــة (2.5.7)، بإتبــاع طريقــة التجربــة والخطأ، وعـادة يتـم اختيــار q_1,q_2 بحيــث تكــون هنــالك احتمــالات متســاوية علــى ذيلـي توزيـع $\chi^2_{(2n)}$ ، أي يتـم اختيــار:

$$q_1=\chi^2_{1-\alpha/2,2n} \quad , \quad q_2=\chi^2_{\alpha/2,2n} \quad ; \quad \gamma=1-\alpha$$

ومن ثم بحـل المعـادلتين:

$$-2\sum_{1}^{n}\ln x_i+2\ln\theta=q_1$$

$$-2\sum_{1}^{n}\ln x_i+2\ln\theta=q_2$$

بالنسبة لـ θ نجد:

$$T_1(x)=e^{q_1/2n}\left(\prod_{1}^{n}x_i\right)^{1/n}$$

$$T_2(x)=e^{q_2/2n}\left(\prod_{1}^{n}x_i\right)^{1/n}$$

وبالتالي $100\gamma\%$ فترة ثقة للمعلمة θ عبارة عن أي فترة من الشكل:

$$\left(T_1(\boldsymbol{X}),T_2(\boldsymbol{X})\right)$$

حيث أن:

$$T_j(\boldsymbol{X})=e^{q_j/2n}\left(\prod_1^n X_i\right)^{1/n} \quad ; \quad j=1,2$$

يمكن، في هذه المسألة، أخذ كمية محورية أخرى وهي:

$$Q(\boldsymbol{X};\theta)=\left(X_{(n)}/\theta\right)^n$$

تتبع التوزيع المنتظم $R(0,1)$، وعندئذٍ $100\gamma\%$ فترة ثقة لـ θ تأخذ الشكل:

$$\left(X_{(n)}/q^{1/n} \ , \ X_{(n)}/(q-\gamma)^{1/n}\right)$$

حيث أن $\gamma \le q \le 1$. ومن الواضح أن أصغر $100\gamma\%$ فترة ثقة لـ θ توافق $q=1$:

$$\left(X_{(n)} \ , \ X_{(n)}/\alpha^{1/n}\right) \quad ; \quad \gamma=1-\alpha$$

مثال 2.5.7

إذا كانت $\boldsymbol{X}=\left(X_1,\ldots,X_n\right)$ عينة عشوائية من مجتمع كثافة توزيعه:

$$f(x;\theta)=\theta x^{\theta-1} \quad ; \quad 0<x<1 \ , \ \theta>0$$

فأوجد $100\gamma\%$ فترة ثقة للمعلمة θ.

بما أن:

$$F(x;\theta)=\theta\int_0^x x^{\theta-1}dx=x^\theta$$

وهي دالة مستمرة ومتزايدة تماماً في θ مهما تكن $0<x<1$، فحسب المبرهنة (1.5.7) توجد كمية محورية لـ θ تعطى بالعلاقة (1.5.7):

$$Q(X;\theta) = -2\sum_{1}^{n} \ln F(X_i;\theta) = -2\theta\sum_{1}^{n} \ln X_i$$

وتوزيعـها:

$$\mathcal{L}_\theta(Q) = \chi^2_{(2n)} \quad ; \quad \theta \in \theta$$

وباختيار عددين q_1, q_2 يحققان العلاقة:

$$P_\theta\left(q_1 < -2\theta\sum_{1}^{n} \ln X_i < q_2\right) = \gamma$$

من أجل قيمة معطاة γ، وبحل المعادلتين:

$$-2\theta\sum_{1}^{n} \ln x_i = q_1$$

$$-2\theta\sum_{1}^{n} \ln x_i = q_2$$

بالنسبة لـــ θ عند عينة مشاهدة $x = (x_1,\dots,x_n)$ نجد:

$$T_1(x) = \frac{q_2}{2\sum_{1}^{n} \ln x_i} \quad , \quad T_2(x) = \frac{-q_1}{2\sum_{1}^{n} \ln x_i}$$

وبالتالي $100\gamma\%$ فترة ثقة للمعلمة θ هي من الشكل:

$$\Delta_\gamma = \left(-q_2\Big/2\sum_{1}^{n} \ln X_i \ , \ -q_1\Big/2\sum_{1}^{n} \ln X_i\right)$$

وأن طول هذه الفــترة:

$$L_\gamma = (q_2 - q_1)\frac{1}{2\sum_{1}^{n} \ln X_i}$$

والقيمة المتوقعة لهذه الفــترة:

$$E_\theta\left(L_\gamma\right) = \left(q_2 - q_1\right) E\left(1 \Big/ 2\sum_1^n \ln X_i\right)$$

وعادةً، كما أشرنا سابقاً، نختار:

$$q_1 = \chi^2_{1-\alpha/2,2n} \quad , \quad q_2 = q_1 = \chi^2_{\alpha/2,2n}$$

نلاحظ من تعريف الكمية المحورية أنها تعتمد على المعلمة المراد تقديرها بالإضافة إلى العينة $\boldsymbol{X}$، ومن ثم فهي تعتمد على إحصاء من خلال اعتمادها على العينة $\boldsymbol{X}$. ويفضل أن يكون هذا الإحصاء كافياً (إن وجد) للمعلمة المطلوب إيجاد فترة ثقة لها. وإذا رمزنا بـ C لإحصاء كافي للمعلمة θ، فيمكن كتابة الكمية المحورية Q على الصورة:

$$Q(\boldsymbol{X};\theta) = Q(C,\theta)$$

وعندئذٍ الملاحظتان التاليتان تساعدان، في حالات عدة، في إيجاد كمية محورية للمعلمة θ المراد تقديرها.

1. إذا كانت θ معلمة موضع (location parameter)، أي إذا كان التوزيع الاحتمالي للمجتمع الذي سحبت منه العينة العشوائية $\boldsymbol{X} = \left(X_1,\ldots,X_n\right)$ يمكن كتابته على الصورة:

$$f(x;\theta) = f(x - \theta)$$

فنبحث عن Q كدالة في $(C - \theta)$، أي من الشكل:

$$Q(\boldsymbol{X};\theta) = Q(C - \theta)$$

2. إذا كانت θ معلمة مقياس (scale parameter)، أي إذا كان التوزيع الاحتمالي للمجتمع الذي سحبت منه العينة $\boldsymbol{X} = \left(X_1,\ldots,X_n\right)$ يمكن كتابته على الصورة:

$$f(x;\theta) = f\left(\frac{x}{\theta}\right)$$

فنبحث عن Q كدالة في $\frac{C}{\theta}$، أي من الشكل:

$$Q(X;\theta) = Q\left(\frac{C}{\theta}\right)$$

لنوضح كيفية الاستفادة من تلك الملاحظتين للحصول على كمية محورية $Q(X;\theta)$ مفيدة لبناء فترة ثقة لمعلمة θ من خلال الأمثلة الآتية.

مثال 3.5.7

إذا كانت $X = (X_1,\ldots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta) = e^{-(x-\theta)} \quad ; \quad x > \theta$$

فأوجد $100\gamma\%$ فترة ثقة لـ θ.

بما أن θ معلمة موضع و $C = X_{(n)}$ إحصاء كافي لها، فنأخذ:

$$Q(X;\theta) = n(C - \theta)$$

وهي دالة مستمرة ومتناقصة تماماً في θ مهما تكن $x \in G$.

كما نعلم [راجع العلاقة (3.10.3)] أن توزيع $C = X_{(n)}$ هو:

$$h(c;\theta) = ne^{-n(C-\theta)} \quad ; \quad c > \theta$$

ومن ثم توزيع Q هو:

$$f_Q(q) = e^{-q} \quad ; \quad q > 0$$

ولا يعتمد على θ. إذن فإن $Q(X;\theta) = n(C - \theta)$ كمية محورية للمعلمة θ ويمكن الاستفادة منها لبناء $100\gamma\%$ فترة ثقة لـ θ على النحو الآتي:

نختار عددين $q_1 < q_2$، عند قيمة معينة معطاة γ، بحيث يحققان العلاقة:

$$P\left(q_1 < Q(X;\theta) < q_2\right) = \int_{q_1}^{q_2} e^{-q} dq = e^{-q_1} - e^{-q_2} = \gamma \qquad . . . (1)$$

وبحل المعادلتين بالنسبة لـ θ:

$$n(c-\theta) = q_1 \quad , \quad n(c-\theta) = q_2 .$$

عند عينة مشاهدة $x = (x_1,\dots,x_2)$ نحصل على العددين:

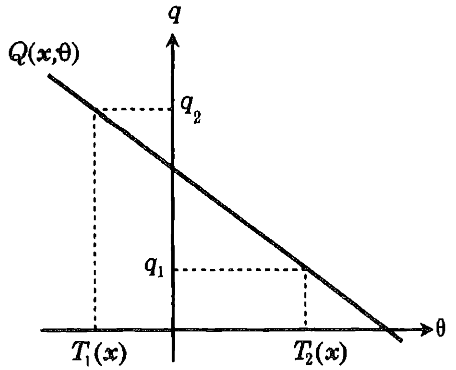
$$\theta_1 = T_1(x) = c - \frac{q_2}{n} \quad , \quad \theta_2 = T_2(x) = c - \frac{q_1}{n}$$

يحققان العلاقة:

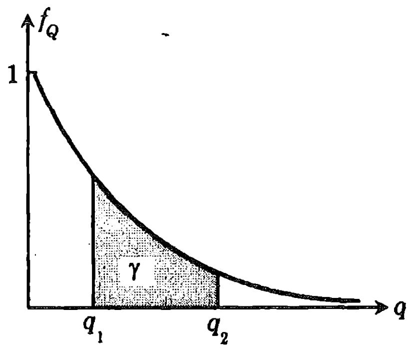
$$P\left(T_1(x) < \theta < T_2(x)\right) = \gamma$$

ومن ثم فإن $100\gamma\%$ فترة ثقة للمعلمة θ هي من الشكل:

$$\left(C - \frac{q_2}{n} \; , \; C - \frac{q_1}{n}\right)$$



شكل (3.5.7)



شكل (2.5.7)

وطول هذه الفترة:

$$L_\gamma = \frac{1}{n}(q_2 - q_1) \qquad . . . (2)$$

وأصغر قيمة له توافق $q_1 = 0$. ويمكن إثبات ذلك جبرياً كالآتي:

باشتقاق العلاقة (1) بالنسبة لـ q_1 نجد:

$$-e^{-q_1} + e^{-q_2}\frac{dq_2}{dq_1} = 0 \Rightarrow \frac{dq_2}{dq_1} = e^{q_2 - q_1} \qquad . . . (3)$$

وباشتقاق العلاقة (2) بالنسبة لـ q_1 والاستفادة من (1) و (3) نجد:

$$\frac{\partial L_\gamma}{\partial q_1} = \frac{1}{n}\left(\frac{dq_2}{dq_1} - 1\right) = \frac{1}{n}\left(e^{q_2 - q_1} - 1\right)$$
$$= \frac{1}{n}\left[e^{q_2}\left(\gamma + e^{-q_2}\right) - 1\right] = \frac{\gamma}{n}e^{q_2} > 0$$

وهذا يعني أن L_γ دالة متزايدة في q_1 ، ومن ثم فأقل قيمة لها عندما $q_1 = 0$ (لأن $q > 0$)، وبوضع $q_1 = 0$ في العلاقة (1) نجد:

$$1 - e^{-q_2} = \gamma \Rightarrow q_2 = -\ln(1-\gamma) = -\ln\alpha \quad ; \quad \alpha = 1 - \gamma$$

وبالتالي تكون $100\gamma\%$ فترة الثقة المطلوبة (أصغر فترة ثقة) للمعلمة θ هي:

$$\left(C + \frac{\ln\alpha}{n} \;,\; C\right)$$

وأن طولها:

$$L_\gamma = -\frac{\ln\alpha}{n}$$

وكذلك اتبعنا نفس الأسلوب في إيجاد كمية محورية لبناء $100\gamma\%$ فترة ثقة للمعلمة θ في النموذج الطبيعي I (بند 1.4.7)، حيث أن θ معلمة موضع.

مثال 4.5.7

إذا كانت $X = (X_1, \ldots, X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\frac{1}{\theta} \quad ; \quad 0<x<\theta$$

فأوجد $100\gamma\%$ فترة ثقة للمعلمة θ.

بما أن θ معلمة مقياس و $C=X_{(n)}$ إحصاء كافي لها، لذلك نبحث عن كمية محورية كدالة في $\frac{C}{\theta}$ ولنأخذ:

$$Q(X;\theta)=\frac{C}{\theta} \quad ; \quad \theta>0$$

وهي دالة مستمرة ومتناقصة تماماً في θ من أجل كل $x\in G$. ويمكن معرفة توزيعها بسهولة على النحو الآتي:

بما أن توزيع $C=X_{(n)}$ [راجع العلاقة (13.3)] هو:

$$h(c;\theta)=\frac{nc^{n-1}}{\theta^{n}} \quad ; \quad 0<c<\theta$$

فإن توزيع Q:

$$f_Q(q)=nq^{n-1} \quad ; \quad 0<q<1$$

وهو لا يعتمد على θ، إذن Q كمية محورية:

لبناء $100\gamma\%$ فترة ثقة لـ θ باستخدام الكمية المحورية Q نتبع ما يلي:

نختار عددين $q_1<q_2$ بحيث أن:

$$P(q_1<Q<q_2)=\int_{q_1}^{q_2} f_Q(q)dq=n\int_{q_1}^{q_2} q^{n-1}dq=q_2^n-q_1^n=\gamma \quad ..... (1)$$

وبحل المعادلتين بالنسبة لـ θ:

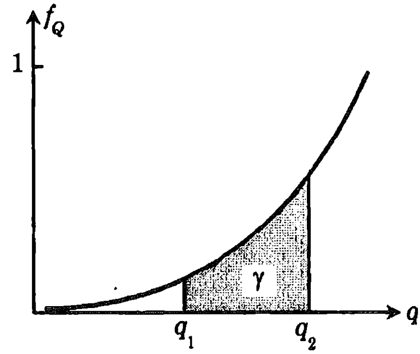
$$\frac{c}{\theta}=q_1 \quad , \quad \frac{c}{\theta}=q_2$$

عند عينة مشاهدة $x=(x_1,\ldots,x_n)$ ، نجد:

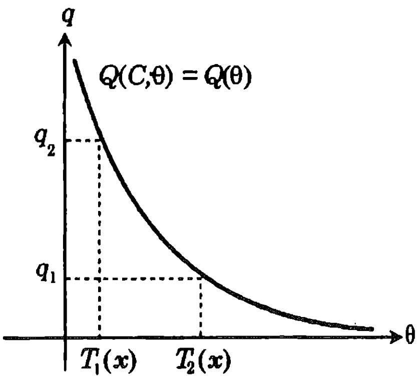
$$T_1(x)=\theta_1=\frac{c}{q_2}\quad,\quad T_2(x)=\theta_2=\frac{c}{q_1}$$

أي أن $100\gamma\%$ فترة ثقة لـ θ هي من الشكل:

$$\Delta_\gamma=\left(\frac{C}{q_2}\ ,\ \frac{C}{q_1}\right)$$



شكل (4.5.7)



شكل (5.5.7)

وطول الفترة:

$$L_\gamma=C\left(\frac{1}{q_1}-\frac{1}{q_2}\right)\qquad\ldots(2)$$

وأن أقصر فترة توافق $q_2=1$ ، ويمكن إثبات ذلك كالآتي:

نشتق (1) بالنسبة لـ q_2 :

$$nq_2^{n-1}-nq_1^{n-1}\frac{dq_1}{dq_2}=0\ \Rightarrow\ \frac{dq_1}{dq_2}-\left(\frac{q_2}{q_1}\right)^{n-1}\qquad\ldots(3)$$

وباشتقاق (1) بالنسبة لـ q_2 والاستفادة من (3) نجد:

$$\frac{\partial L_\gamma}{\partial q_2} = C\left(-\frac{1}{q_1^2}\frac{dq_1}{dq_2} + \frac{1}{q_2^2}\right)$$

$$= C\left[-\frac{1}{q_1^2}\left(\frac{q_2}{q_1}\right)^{n-1} + \frac{1}{q_2^2}\right] = -C\left(\frac{q_2^{n+1} - q_1^{n+1}}{q_2^2 q_1^{n+1}}\right) < 0.$$

وهذا يعني أن L_γ دالة متناقصة في q_2 وتكون أقل قيمة لها عندما تكون q_2 أكبر ما يمكن أي $q_2 = 1$ (لأن $0 < q < 1$). ومن ثم بالتعويض في (1) نجد:

$$1 - q_1^n = \gamma \Rightarrow q_1 = \sqrt[n]{1-\gamma} = \alpha^{1/n} \quad ; \quad \gamma = 1 - \alpha$$

وتكون $100\gamma\%$ فترة الثقة المطلوبة لـ θ هي من الشكل:

$$\left(C, C\alpha^{-1/n}\right)$$

وأتبعنا نفس الأسلوب عند بناء فترة الثقة للمعلمة θ^2 في توزيع $N(\mu,\theta^2)$ الوارد في الفقرة (2.4.7)، حيث θ^2 معلمة مقياس.

6.7 طريقة الإحصاء STATISTICAL METHOD

إن طريقة الكمية المحورية لبناء فترة ثقة للمعلمة وحيدة البعد θ قابلة للتطبيق في حالات عدة، وغير ممكنة في حالات أخرى ومنها حالة توزيع منقطع، وذلك إما لعدم إمكانية إيجاد مثل هذه الكمية أو لصعوبة بناء فترة الثقة بهذه الطريقة والمتمثلة بتعيين حدي الثقة. وعندئذٍ يمكن استخدام طريقة أخرى تعتمد على إحصاء ما $T = T(X)$ لبناء فترة ثقة لـ θ، لذا تسمى هذه الطريقة بـ "طريقة الإحصاء".

نفترض كالعادة لدينا عينة عشوائية $X = (X_1,\ldots,X_n)$ من توزيع $f(x;\theta)$ يعتمد على معلمة وحيدة البعد. ونرغب في بناء فترة ثقة للمعلمة θ. نختار إحصاء ما $T = T(X)$، وهذا ممكن بطرق مختلفة. فمثلاً، إذا وجد إحصاء كافي للمعلمة

θ، فيؤخذ T إحصاء كافي، أو إذا لم يوجد إحصاء كافي لـ θ فيؤخذ T مقدر نقطي لها، ومن الممكن أن يكون مقدر المعقولية العظمى. ويتم اختيار T عملياً بحيث تكون العمليات التي نحتاجها لبناء فترة الثقة لـ θ ممكنة وسهلة قدر الإمكان، ومنها سهولة معرفة توزيعه الاحتمالي $F_T(t;\theta)$.

سوف نوضح هذه الطريقة في الفقرتين الآتيتين باعتبار T متغير مستمر ومن ثم باعتبار T متغير منقطع.

1.6.7 حالة T متغير مستمر

لتكن $X = (X_1,\dots,X_n)$ عينة عشوائية من توزيع مستمر $F(x;\theta)$ يعتمد على معلمة وحيدة البعد θ. إذا كان $T = T(X)$ إحصاء ما مستمر دالة توزيعه معلومة $F_T(t;\theta)$، مستمرة ومضطردة تماماً في θ، فيمكن بناء $100\gamma\%$ فترة ثقة للمعلمة θ بناءً على هذا الإحصاء.

لنعرف عند كل $\theta \in \Theta$ عددين $t_i = \tau_i(\theta)\ ;\ i = 1,2$، بحيث $t_1 < t_2$ ويحققان العلاقة:

$$P_\theta(t_1 < T(X) < t_2) = F_T(t_2;\theta) - F_T(t_1;\theta) = \gamma \tag{1.6.7}$$

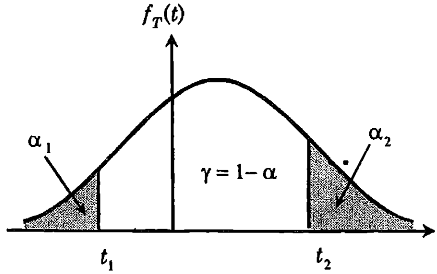
أو:

$$F_T(t_1;\theta) = \alpha_1 \quad , \quad F_T(t_2;\theta) = 1 - \alpha_2 \quad ; \quad \alpha_1 + \alpha_2 = \alpha \ , \ \gamma = 1 - \alpha \tag{2.6.7}$$

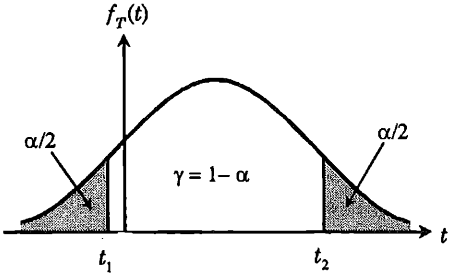
ويتم عادة اختيار هذين العددين بحيث يحققان:

$$F_T(t_1;\theta) = \alpha/2 \quad ; \quad F_T(t_2;\theta) = 1 - \alpha/2 \tag{3.6.7}$$

أي يتم البحث عن فترة ثقة ذات ذيلين متساويي الاحتمال. ويمكن إيضاح الحالتين على الشكلين (1.6.7) و (2.6.7).



شكل (1.6.7)

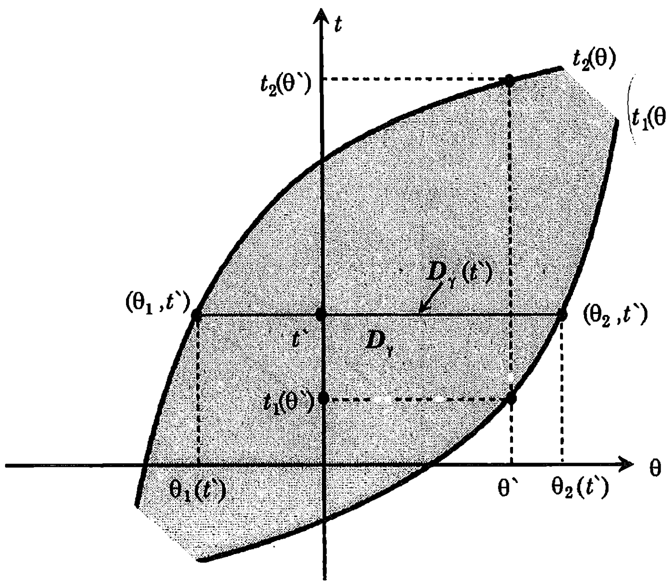


شكل (2.6.7)

لنعرف الفئة الجزئية D_γ من $\Im \times \Theta$، حيث $\Im$ فضاء القيم الممكنة للإحصاء T، على النحو الآتي:

$$D_\gamma = \{(\theta,t) \ ; \ t_1(\theta) < t < t_2(\theta)\} \tag{4.6.7}$$

ويمكن تمثيلها بيانيا (بافتراض $t_1(\theta), t_2(\theta)$ متزايدتان تماماً في θ) كما هو مبين بالمنطقة المظللة على الشكل (3.6.7).



شكل (3.6.7)

عندئذٍ، حسب العلاقة (1.6.7)، فإن:

$$P_\theta\left((\theta, T(X)) \in D_\gamma\right) = \gamma \quad ; \quad \theta \in \Theta$$

وعند قيمة معينة $t' \in \mathfrak{I}$، فإن:

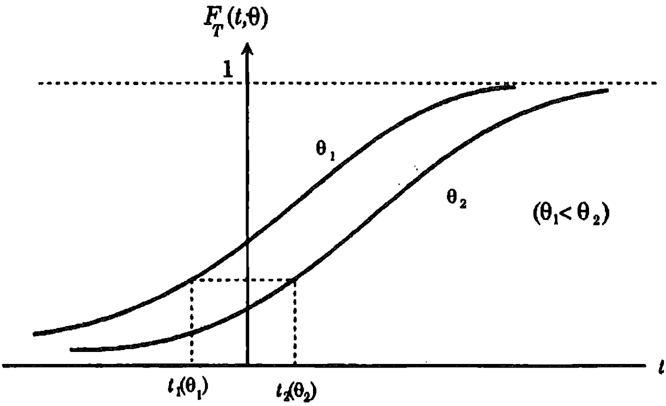
$$D_\gamma(t') = \left\{\theta : (\theta, t') \in D_\gamma\right\}$$

أنظر شكل (3.6.7).

لنرى الآن الفئة العشوائية $D_\gamma(T(X)) \subset \Theta$. يقع الحادث $\theta \in D_\gamma(T(X))$ فقط عندما $T(X) \in (t_1(\theta), t_2(\theta))$، وبالتالي، عند كل قيمة لـ θ، فاحتماله يساوي γ، أي أن:

$$P_\theta\left(\theta \in D_\gamma(T(X))\right) = \gamma \quad ; \quad \theta \in \Theta$$

وبالتالي، تم تعيين فئة عشوائية $D_\gamma(T(X))$ تشتمل على القيمة الحقيقية للمعلمة θ باحتمال يساوي γ. وإذا كانت هذه الفئة تمثل فترة، نكون قد توصلنا إلى بناء $100\gamma\%$ فترة ثقة للمعلمة θ، وهي محققة، إذا كانت الدالتان $t_i(\theta)\ ;\ i = 1,2$ من نفس النوع (متزايدتان تماماً أو متناقصتان تماماً في θ). وهذا صحيح دوماً لكون الدالة $F_T(t;\theta)$ مستمرة ومضطردة تماماً بالنسبة لـ θ (حسب الفرض). ويبدو ذلك بوضوح على الشكل (4.6.7) في حالة $F_T(t;\theta)$ دالة متناقصة تماماً في θ.



شكل (4.6.7)

هكذا، حسب الفروض المقدمة تمثل الفئة $D_\gamma(t)$، عند كل قيمة t لـ $T(X)$، فترة حداها $\theta_1(t)<\theta_2(t)$، وبالتالي $100\gamma\%$ فترة ثقة للمعلمة θ هي $\left(T_1(X),T_2(X)\right)$، حيث أن $T_i(X)=\theta_i\left(T(X)\right)\ ;\ i=1,2$.

يعطي التمثيل البياني المدلول الواضح لهذه الطريقة، فالرسم المبنى وفق الخط الرأسي (من أجل أي قيمة $\theta'\in\Theta$ نحدد من الشرط (3.6.7) القيمتين المناظرتين لها $t_1(\theta'),t_2(\theta')$). وبشكل مشابه الرسم البياني وفق الخط الأفقي من أجل أي قيمة ملاحظة $t'=T(x)$ (x قيمة ملاحظة للعينة X) نحصل على القيمتين المناظرتين لها $\theta_1(t'),\theta_2(t')$، بحيث أن $\theta\in\left(\theta_1,\theta_2\right)$. هكذا، إذا اعتبرنا الرسم وفق الخط الرأسي، فحدي المنطقة D_γ يوصفان بالنقطتين المتغيرتين $(\theta,t_1),(\theta,t_2)$ بحيث يحققان الشرط (3.6.7). أما إذا كان الرسم وفق الخط الأفقي، فحدي المنطقة D_γ يمكن وصفهما بشكل مكافئ بالنقطتين المتغيرتين $(t,\theta_1),(t,\theta_2)$ بأخذ القيمة الملاحظة t للإحصاء $T(X)$ كمتغير مستقل. عندئذٍ يمكن إعادة صياغة العلاقة (3.6.7) على النحو الآتي:

1. إذا كانت الدالتان $t_i(\theta)\ ;i=1,2$ متزايدتان تماماً في θ [شكل (3.6.7)] ويقابل ذلك تناقص الدالة $F_T(t;\theta)$ بالنسبة لـ θ، فإن:

$$F_T\left(t;\theta_1\right)=1-\alpha_2 \quad , \quad F_T\left(t;\theta_2\right)=\alpha_1 \tag{5.6.7}$$

وعندما $\alpha_1=\alpha_2=\alpha/2$:

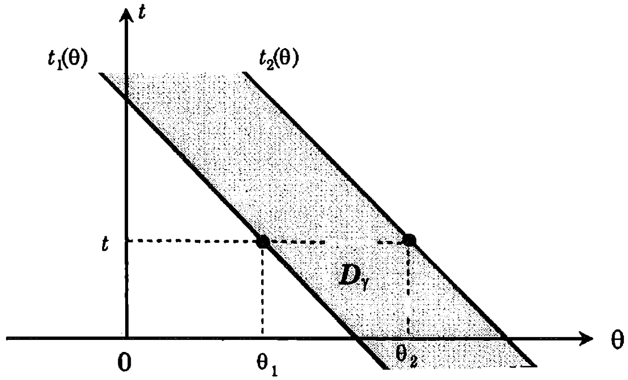
$$F_T\left(t;\theta_1\right)=1-\alpha/2 \quad , \quad F_T\left(t;\theta_2\right)=\alpha/2 \tag{6.6.7}$$

2. إذا كانت الدالتان $t_i(\theta)\ ;\ i=1,2$ متناقصتان تماماً في θ [شكل (5.6.7)]، ويقابل ذلك تزايد الدالة $F_T(t;\theta)$ بالنسبة لـ θ، فإن:

$$F_T\left(t;\theta_1\right)=\alpha_1 \quad , \quad F_T\left(t;\theta_2\right)=1-\alpha_2 \tag{7.6.7}$$

وعندما $\alpha_1=\alpha_2=\alpha/2$:

$$F_T(t;\theta_1) = \alpha/2 \quad , \quad F_T(t;\theta_2) = 1 - \alpha/2 \tag{8.6.7}$$



شكل (5.6.7)

هكذا، خوارزمية بناء $100\gamma\%$ فترة ثقة للمعلمة θ، في حالة دالة التوزيع $F_T(t;\theta)$ للإحصاء T مستمرة ومضطردة تماماً بالنسبة لـ θ، يمكن إيجازها كالآتي:

لتكن $t = T(x)$ قيمة ملاحظة للإحصاء $T = T(X)$، بحل المعادلتين:

$$F_T(t,\theta) = \alpha/2 \quad , \quad F_T(t,\theta) = 1 - \alpha/2 \quad ; \quad \gamma = 1 - \alpha$$

بالنسبة لـ θ، نحصل على عددين $\theta_1 < \theta_2$، حيث تشتمل الفترة (θ_1, θ_2) على القيمة الحقيقية للمعلمة θ باحتمال يساوي γ. وإذا كانت γ قريبة من الواحد (أو $\alpha = 1 - \gamma$ قريبة من الصفر)، فاحتمال الخطأ في تطبيق هذه القاعدة يكون صغيراً ويساوي α، أي نحصل بالنتيجة على $100\gamma\%$ فترة ثقة للمعلمة θ وهي من الشكل:

$$\left(T_1(X), T_2(X)\right) \quad ; \quad T_i(X) = \theta_i\left(T(X)\right) \quad ; \quad i = 1,2$$

لنوضح ما سبق من خلال الأمثلة الآتية:

مثـــال 1.6.7

لتكن $X=(X_1,...,X_n)$ عينة عشوائية من توزيع $N(\theta,\sigma^2)$، ونريد بنـــاء $100\gamma\%$ فترة ثقة للمعلمة θ.

كما نعلم $T=T(X)=\overline{X}$ إحصاء كافي وفي نفس الوقت مقدر المعقولية العظمى لـــ θ، وتوزيعه:

$$\mathcal{L}_\theta(\overline{X})=N\left(\theta,\frac{\sigma^2}{n}\right)$$

أي أن:

$$F_T(\overline{x};\theta)=\Phi\left(\frac{\overline{x}-\theta}{\sigma/\sqrt{n}}\right)$$

وهذه دالة مستمرة ومضطردة في θ، وبالتالي:

$$F_T(t_1(\theta);\theta)=\Phi\left(\frac{t_1(\theta)-\theta}{\sigma/\sqrt{n}}\right)=\alpha_1$$

ومن جدول التوزيع الطبيعي [الملحق جدول رقـــم 9] نجـــد:

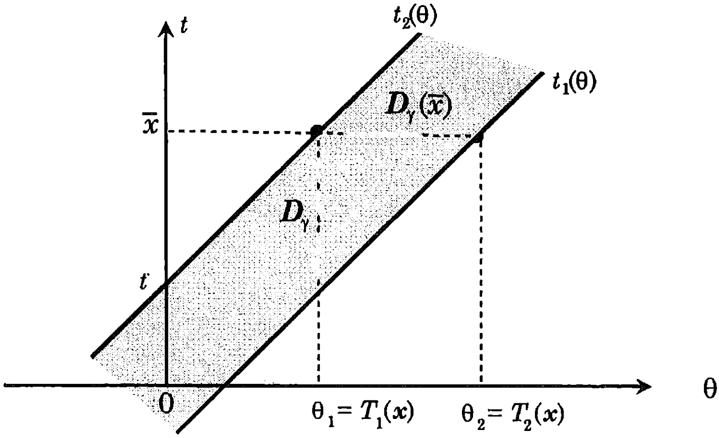
$$\frac{t_1(\theta)-\theta}{\sigma/\sqrt{n}}=Z_{\alpha_1}\ \Rightarrow\ t_1(\theta)=\theta+\frac{\sigma}{\sqrt{n}}Z_{\alpha_1}\qquad ...(1)$$

وبـــالمثل:

$$F_T(t_2(\theta);\theta)=\Phi\left(\frac{t_2(\theta)-\theta}{\sigma/\sqrt{n}}\right)=1-\alpha_2$$

$$\frac{t_2(\theta)-\theta}{\sigma/\sqrt{n}}=Z_{\alpha_2}\ \Rightarrow\ t_2(\theta)=\theta+\frac{\sigma}{\sqrt{n}}Z_{\alpha_2}\qquad ...(2)$$

يمكن تمثيل (1) و (2) بيانياً كمـــا في الشـــكل (6.6.7).



شكل (6.6.7)

نلاحظ أن $t_i(\theta)$; $i = 1,2$ دالتان متزايدتان تماماً في θ. وعلى ذلك فإن:

$$\bar{x} = \theta_1 + \frac{\sigma}{\sqrt{n}} Z_{1-\alpha_2} \Rightarrow \theta_1 = \bar{x} - \frac{\sigma}{\sqrt{n}} Z_{1-\alpha_2}$$

$$\bar{x} = \theta_2 + \frac{\sigma}{\sqrt{n}} Z_{\alpha_1} \Rightarrow \theta_2 = \bar{x} - \frac{\sigma}{\sqrt{n}} Z_{\alpha_1}$$

ويمكن الحصول على θ_1, θ_2 بطريقة أخرى على النحو الآتي:

بما أن $t_i(\theta); i = 1,2$ دالتان متزايدتان تماماً في θ و $F_T(t;\theta)$ دالة متناقصة في θ، فنحصل على θ_1 من حل المعادلة بالنسبة لـ θ_1:

$$F_T(\bar{x};\theta_1) = \Phi\left(\frac{\bar{x} - \theta_1}{\sigma/\sqrt{n}}\right) = 1 - \alpha_2$$

$$\frac{\bar{x} - \theta_1}{\sigma/\sqrt{n}} = Z_{1-\alpha_2} \Rightarrow \theta_1 = \bar{x} - \frac{\sigma}{\sqrt{n}} Z_{1-\alpha_2}$$

ونحصل على θ_2 بحل المعادلة بالنسبة لـ θ_2:

$$F_T\left(\overline{x};\theta_2\right)=\phi\left(\frac{\overline{x}-\theta_2}{\sigma/\sqrt{n}}\right)=\alpha_1$$

$$\frac{\overline{x}-\theta_2}{\sigma/\sqrt{n}}=Z_{\alpha_1}\ \Rightarrow\ \theta_1=\overline{x}-\frac{\sigma}{\sqrt{n}}Z_{\alpha_1}$$

وتكون $\left(\overline{x}-\frac{\sigma}{\sqrt{n}}Z_{1-\alpha_2}\ ,\ \overline{x}-\frac{\sigma}{\sqrt{n}}Z_{\alpha_1}\right)$ هي $100\gamma\%$ فترة ثقة للمعلمة θ، حيث أن $\gamma=1-\left(\alpha_1+\alpha_2\right)=1-\alpha$.

إذا أخذنا $\alpha_1=\alpha_2=\alpha/2$ فنحصل على $100\gamma\%$ فترة ثقة للمعلمة θ ذات ذيلين متساويي الاحتمال:

$$\left(\overline{x}-\frac{\sigma}{\sqrt{n}}Z_{\alpha/2}\ ,\ \overline{x}+\frac{\sigma}{\sqrt{n}}Z_{\alpha/2}\right)$$

وهي نفس فترة الثقة التي حصلنا عليها في الفقرة (1.4.7) علاقة (4.4.7).

مثال 2.6.7

إذا كانت $X=\left(X_1,\ldots,X_n\right)$ عينة عشوائية من توزيع منتظم $R(0,\theta)$، فأوجد $100\gamma\%$ فترة ثقة للمعلمة θ.

نعلم أن $T=T(X)=X_{(n)}$ إحصاء كافي للمعلمة θ وأن توزيعه:

$$f_T(t;\theta)=\frac{nt^{n-1}}{\theta^n}\qquad;\qquad 0\le t\le\theta$$

ومن ثم دالة توزيعه:

$$F_T(t;\theta)=\int_0^t\frac{nt^{n-1}}{\theta^n}dt=\left(\frac{t}{\theta}\right)^n\qquad;\qquad 0\le t\le\theta$$

بحل المعادلتين:

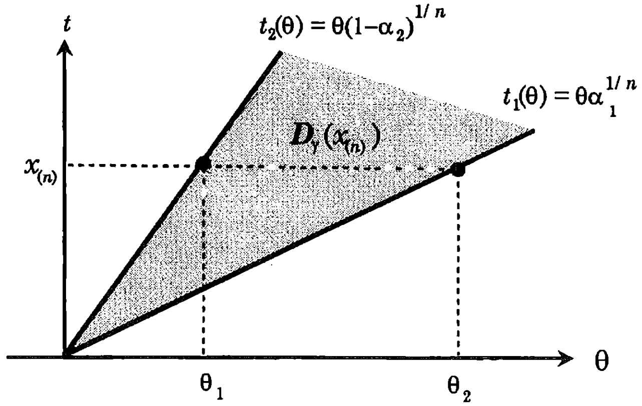
$$F_T\left(t_1(\theta),\theta\right)=\left(\frac{t_1(\theta)}{\theta}\right)^n=\alpha_1$$

$$F_T\left(t_2(\theta),\theta\right)=\left(\frac{t_2(\theta)}{\theta}\right)^n=1-\alpha_2 \quad ; \quad \gamma=1-\alpha \ , \ \alpha=\alpha_1+\alpha_2$$

بالنسبة لـ t_1 و t_2 نجد:

$$t_1(\theta)=\theta\alpha_1^{1/n} \quad , \quad t_2(\theta)=\theta\left(1-\alpha_2\right)^{1/n}$$

ويمكن تمثيلهما بيانياً كما هو مبين على الشكل (7.6.7):



شكل (7.5.7)

وبما أن دالة التوزيع $F_T(t;\theta)$ مستمرة ومتناقصة تماماً في θ، فنحصل على حدي الثقة $\theta_1<\theta_2$ بحل المعادلتين، عند قيمة ملاحظة $t=x_{(n)}$ لـ $T=X_{(n)}$:

$$F_T\left(x_{(n)};\theta_1\right)=\left(\frac{x_{(n)}}{\theta_1}\right)^n=1-\alpha_2$$

$$F_T\left(x_{(n)};\theta_2\right)=\left(\frac{x_{(n)}}{\theta_2}\right)^n=\alpha_1$$

بالنسبة لـ θ_1 و θ_2 نجد:

$$T_1(x)=\theta_1=x_{(n)}(1-\alpha_2)^{-1/n} \quad , \quad T_2(x)=\theta_2=x_{(n)}\alpha_1^{-1/n}.$$

وتكـون:

$$\Delta_\gamma(X)=(T_1(X),T_2(X))=\left(X_{(n)}(1-\alpha_2)^{-1/n},X_{(n)}\alpha_1^{-1/n}\right)$$

هي $100\gamma\%$ فـترة ثقـة للمعلمـة θ. وطـول هـذه الفـترة هـو:

$$L_\gamma=X_{(n)}\left(\frac{1}{\sqrt[n]{\alpha_1}}-\frac{1}{\sqrt[n]{1-\alpha_2}}\right)$$

وتكون هـذه الفـترة أقصـر مـا يمكـن إذا اخترنـا α_1,α_2 بحيث يكـون المقـدار بـين القوسين أقـل مـا يمكـن، ويحـدث ذلـك عندمـا تكـون $\alpha_1=\alpha\ ,\ \alpha_2=0$، وفي هـذه الحالـة تكـون $\left(X_{(n)},X_{(n)}\alpha^{-1/n}\right)$ هـي $100\gamma\%$ فـترة ثقـة للمعلمـة θ. وهـي نفـس الفـترة الـتي حصلنـا عليهـا باسـتخدام طريقـة الكميـة المحورية في المثـال (1.5.7).

مثـــال 3.6.7

إذا كـانت $X=(X_1,\ldots,X_n)$ عينـة عشـوائية مـن توزيـع:

$$f(x;\theta)=\theta e^{-\theta x} \quad ; \quad x>0\ ,\ \theta>0$$

أوجـد $100\gamma\%$ فـترة ثقـة للمعلمـة θ.

نعلـم أن $T=\sum_{1}^{n}X_i$ إحصـاء كـافي للمعلمــة θ وأن توزيعـه:

$$f_T(t;\theta)=\frac{\theta^n}{\Gamma(n)}t^{n-1}e^{-\theta t} \quad ; \quad t>0$$

[راجـع مثـال (1.4.4)].

ويمكـن التـأكد بسـهولة أن المتغـير العشـوائي $Y=2\theta T$ يخضـع لتوزيـع $\chi^2_{(2n)}$، وبالتـالي نحصـل علـى $t_1(\theta)$ بحـل المعادلـة بالنسـبة لـ t_1:

$$F_T\left(t_1(\theta);\theta\right) = F_Y\left(2\theta t_1(\theta);\theta\right) = \alpha_1$$

ومن جدول توزيع χ^2 بـ $2n$ درجة حرية [الملحق جدول رقم 11] نجد:

$$\chi^2_{1-\alpha_1,2n} = 2\theta t_1(\theta) \Rightarrow t_1(\theta) = \chi^2_{1-\alpha_1,2n}/2\theta$$

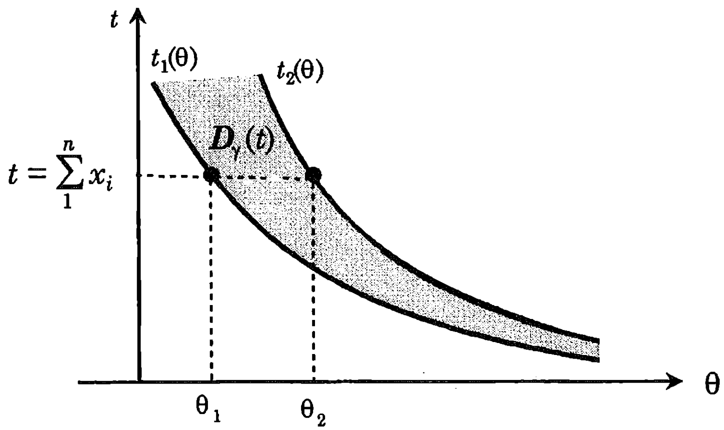
أما للحصول على $t_2(\theta)$ نقوم بحل المعادلة الآتية بالنسبة لـ t_2:

$$F_T\left(t_2(\theta);\theta\right) = F_Y\left(2\theta t_2(\theta);\theta\right) = 1-\alpha_2$$

ومن جدول χ^2 بـ $2n$ درجة حرية نجد:

$$\chi^2_{\alpha_2,2n} = 2\theta t_2(\theta) \Rightarrow t_2(\theta) = \chi^2_{\alpha_2,2n}/2\theta$$

ويمكن تمثيلها بيانياً كما هو مبين على الشكل (8.6.7).



شكل (8.6.7)

نلاحظ من الشكل أن الدالتين $t_1(\theta),t_2(\theta)$ متناقصتان تماماً في θ، وهو مختلف عن المثالين السابقين. وهذا يقابل تزايد دالة التوزيع $F_T(t;\theta)$ بالنسبة لـ θ، وبالتالي نحصل على θ_1 بحل المعادلة:

$$F_T(t;\theta_1) = F_Y(2\theta_1 t;\theta_1) = \alpha_1 \;\Rightarrow$$
$$\chi_{1-\alpha_1,2n} = 2\theta_1 t \;\Rightarrow\; \theta_1 = \chi^2_{1-\alpha_1,2n}/2t$$

وكما نحصل على θ_2 بحل المعادلة:

$$F_T(t;\theta_2) = F_Y(2\theta_2 t;\theta_2) = 1-\alpha_2 \;\Rightarrow$$
$$\chi_{\alpha_2,2n} = 2\theta_2 t \;\Rightarrow\; \theta_2 = \chi^2_{\alpha_2,2n}/2t$$

وبذلك تكون:

$$\left(\frac{\chi^2_{1-\alpha_1,2n}}{2\sum X_i}\,,\,\frac{\chi^2_{\alpha_2,2n}}{2\sum X_i}\right)$$

هي $100\gamma\%$ فترة ثقة للمعلمة θ، وبأخذ $\alpha_1=\alpha_2=\alpha/2$ نحصل على:

$$\left(\frac{\chi^2_{1-\alpha/2}}{2\sum X_i}\,,\,\frac{\chi^2_{1-\alpha/2}}{2\sum X_i}\right)$$

نلاحظ من الأمثلة السابقة، بأننا لا نحتاج عملياً لتعيين $t_1(\theta),t(\theta)$ عند بناء فترات الثقة، ومن أجل قيمة ملاحظة t لـ $T=T(X)$، نحتاج فقط لإيجاد $\theta_1<\theta_2$.

2.6.7 حالة T متغير منقطع

لتكن $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع منقطع $F(x;\theta)$ يعتمد على معلمة وحيدة البعد θ، إذا كان $T=T(X)$ إحصاء ما منقطع دالة توزيعه $F_T(t;\theta)$ مستمرة ومضطردة تماماً في θ، فيمكن بناء $100\gamma\%$ فترة ثقة للمعلمة θ باستخدام هذا الإحصاء.

يمكن إجراء مناقشة مشابهة لحالة T متغير مستمر من أجل T متغير منقطع، والاختلاف الوحيد يتمثل بما يلي:

نتيجة لكون دالة التوزيع $F_T(t;\theta)$ درجية، فإن:

$$P_\theta\left(t_1 < T(X) < t_2\right) = F\left(t_2;\theta\right) - F_T\left(t_1 + 0;\theta\right)$$

حيث أن $F_T(t;\theta) = P(T < t)$. ولا يمكن دائماً، من أجل قيمة معطاة γ، إيجاد فترة ثقة للمعلمة θ لها درجة ثقة مساوية تماماً لـ γ. لهذا يتم البحث عن فترات ثقة للمعلمة θ ذات درجة ثقة لا تقل عن القيمة المختارة γ، أي يتم اختيار عددين $t_1 < t_2$ يحققان الشرط:

$$P_\theta\left(t_1 < T(X) < t_2\right) = F_T\left(t_2;\theta\right) - F_T\left(t_1 + 0;\theta\right) \geq \gamma$$

بدلاً من الشرط (1.6.7) في حالة T متغير مستمر.

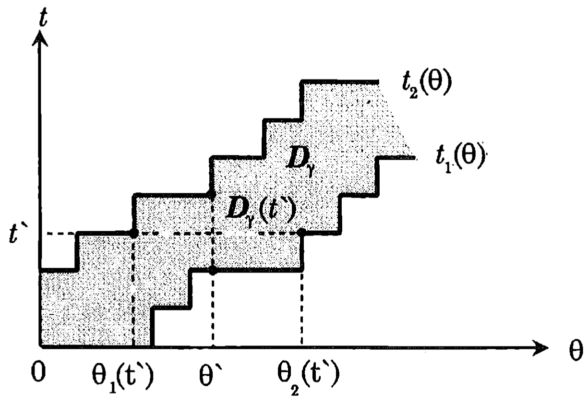
ومن ثم بدلاً من الشرطين (2.6.7) نأخذ:

$$F\left(t_1 + 0;\theta\right) \leq \alpha_1 \quad , \quad F\left(t_2;\theta\right) \geq 1 - \alpha_2 \tag{10.6.7}$$

وعندما $\alpha_1 = \alpha_2 = \alpha/2$، نجد:

$$F\left(t_1 + 0;\theta\right) \leq \alpha/2 \quad , \quad F\left(t_2;\theta\right) \geq 1 - \alpha/2 \tag{11.6.7}$$

تعتبر الدالتان $t_i(\theta)\ ;\ i = 1,2$، في هذه الحالة، درجيتان، ويبدو ذلك بوضوح على الشكل (9.6.7).



شكل (9.6.7)

إن خوارزمية بناء فترة ثقة للمعلمة θ بدرجة ثقة لا تقل عن γ، في حالة نموذج منقطع، هي ذات خوارزمية بناء $100\gamma\%$ فترة ثقة للمعلمة θ في حالة النموذج المستمر، على أن نعين $\theta_1 < \theta_2$ من العلاقتين:

$$F(t+0;\theta) \le \alpha_1 \quad , \quad F(t;\theta) \ge 1-\alpha_2 \tag{12.6.7}$$

أو من العلاقتين:

$$F(t+0;\theta) \le \alpha/2 \quad , \quad F(t;\theta) \ge 1-\alpha/2 \tag{13.6.7}$$

عندما نأخذ $\alpha_1 = \alpha_2 = \alpha/2$.

مثال 4.5.7

لتكن $X = (X_1,\ldots,X_n)$ عينة عشوائية من توزيع بيرنولي $B(1;\theta)$، ونريد بناء فترة ثقة للمعلمة θ.

نعلم أن $T(X) = \overline{X}$ إحصاء كافي للمعلمة θ، والمتغير العشوائي T يفترض القيم $0,1/n,\ldots,n/n$ ويخضع لتوزيع ذي الحدين $B(n;\theta)$:

$$F_T\left(T = \frac{k}{n};\theta\right) = \sum_{r=0}^{k-1} C_n^r \theta^r (1-\theta)^{n-r} \quad ; \quad k = 1,\ldots,n$$

ومشتق هذه الدالة بالنسبة لـ θ:

$$\frac{d}{d\theta} F_T\left(\frac{k}{n};\theta\right) = -nC_{n-1}^{k-1}\theta^{k-1}(1-\theta)^{n-k} < 0$$

وهذا يعني أن الدالة $F_T(t;\theta)$ متناقصة تماماً بالنسبة لـ θ. وبالتالي، بتطبيق الطريقة الواردة أعلاه، نحصل على عددين $\theta_1 < \theta_2$ من العلاقتين:

$$F_T\left(\frac{k}{n};\theta_1\right) = \sum_{r=0}^{k-1} C_n^r \theta_1^r (1-\theta_1)^{n-r} \ge 1-\alpha_2$$

$$F_T\left(\frac{k+1}{n};\theta_2\right) = \sum_{r=0}^{k} C_n^r \theta_2^r (1-\theta_2)^{n-r} \le \alpha_1 \quad ; \quad \gamma = 1-\alpha, \alpha = \alpha_1 + \alpha_2$$

وذلك باتباع طريقة التجربة والخطـــأ.

فمثلاً، إذا كانت $x=(x_1,\dots,x_{20})$ عينة مشاهدة و $T(x)=\bar{x}=\frac{4}{20}$. ونريد بنـــاء فترة ثقة بمعامل ثقة لا يقل عن 0.93.

تأخذ العلاقتان السابقتان الصورة الآتيـــة:

$$F_T\left(\frac{4}{20};\theta_1\right)=\sum_{r=0}^{3}C_{20}^r\theta_1^r(1-\theta_1)^{20-r}\geq 1-\alpha_2$$

$$F_T\left(\frac{5}{20};\theta_2\right)=\sum_{r=0}^{4}C_{20}^r\theta_2^r(1-\theta_2)^{20-r}\leq \alpha_1 \quad ; \quad \gamma=1-(\alpha_1+\alpha_2)$$

باستخدام جدول توزيع ذي الحدين واتباع أسلوب التجربـــة والخطـــأ، نجـــد:

عندما: $\alpha_1'=0.0509$, $1-\alpha_2'=0.9841$ $(\gamma\leq 1-\alpha_1'-\alpha_2'=0.9332)$ فـــإن:

$\theta_1=0.05$ ؛ $\theta_2=0.40$، وبالتالي 93.32% فترة ثقة للمعلمة θ هي $(0.05,0.40)$.

مثـــال 5.6.7

لتكن $X=(X_1,\dots,X_n)$ عينة عشوائية من توزيع بواسون $\Pi(\theta)$، ونريد بناء فـــترة ثقة للمعلمة θ بمعامل ثقة لا يقل عن γ.

كما نعلم $T=T(X)=\overline{X}$ إحصاء كافي ومقدر المعقولية العظمى لـــ θ، ويفـــترض القيم $k=0,1,\dots\ ;\ \frac{k}{n}$.

بما أن $\mathcal{L}_\theta\left(\sum_{1}^{n}X_i\right)=\Pi(n\theta)$، فإن:

$$F_T\left(\frac{k}{n};\theta\right)=\sum_{r=0}^{k-1}e^{-n\theta}\frac{(n\theta)^r}{r!}$$

ومشتقه بالنسبة لـــ θ:

$$\frac{d}{d\theta}F_T\left(\frac{k}{n};\theta\right)=-ne^{-n\theta}\frac{(n\theta)^{k-1}}{(l-1)!}<0$$

وهذا يعني أن دالة التوزيع $F_T(t;\theta)$ متناقصة تماماً في θ. وبالتالي فإن فترة الثقة بدرجة ثقة لا تقل عن γ هي (θ_1,θ_2)، حيث θ_1,θ_2 يعينان من العلاقتين:

$$F_T\left(\frac{k}{n};\theta_1\right)=\sum_{r=0}^{k-1}e^{-n\theta_1}\frac{(n\theta_1)^r}{r!}\geq 1-\alpha_2$$

$$F_T\left(\frac{k+1}{n};\theta_2\right)=\sum_{r=0}^{k}e^{-n\theta_2}\frac{(n\theta_2)^r}{r!}\leq \alpha_1$$

فمثلاً، إذا كانت $x=(x_1,\dots,x_{10})$ عينة مشاهدة لـ X و $T(x)=\bar{x}=\frac{3}{10}$ ونريد بناء فترة ثقة لـ (θ_1,θ_2) بمعامل ثقة لا يقل عن $\gamma=0.95$.

$$\left.\begin{aligned}F_T\left(\frac{3}{10};\theta_1\right)&=\sum_{r=0}^{2}e^{-10\theta_1}\frac{(10\theta_1)^r}{r!}\geq 1-\alpha_2\\F_T\left(\frac{4}{10};\theta_2\right)&=\sum_{r=0}^{3}e^{-10\theta_2}\frac{(10\theta_2)^r}{r!}\leq \alpha_1\end{aligned}\right\}\quad\ldots.(1)$$

بالعودة إلى جدول توزيع بواسون وباتباع أسلوب التجربة والخطأ، بحيث $\alpha_1+\alpha_2\leq 1-\gamma=0.05$ نجد:

$$\left.\begin{aligned}&\sum_{r=0}^{2}e^{-0.6}\frac{(0.6)^r}{r!}=0.9779=1-\alpha_2' \Rightarrow \alpha_2'=0.0221\\&\sum_{r=0}^{3}e^{-8.7}\frac{(8.7)^r}{r!}=0.0262=\alpha_1'\end{aligned}\right\}\quad\ldots.(2)$$

وبمقارنة (1) و (2) نجد:

$$10\theta_1=0.6 \Rightarrow \theta_1=0.06$$
$$10\theta_2=8.7 \Rightarrow \theta_2=0.87$$

وبالتالي 95.17% فترة ثقة للمعلمة θ هي $(0.06,0.87)$.

7.7 فترات الثقة عندما يكون حجم العينة كبيراً
LARGE SAMPLE CONFIDENCE INTERVAL

رأينا من خلال الأمثلة السابقة بأن فترة الثقة لمعلمة θ في نموذج $f(x;\theta)$ يعتمد على اختيار الإحصاء $T = T(X)$. أي أن اختيار إحصاءات مختلفة T يؤدي إلى فترات ثقة مختلفة بشكل عام. إلا أن الهدف النهائي يتمثل في الحصول على أقصر فترة ثقة عند قيمة معينة مسبقا γ . لنفترض استخدمنا مقدِّرات غير متحيزة وطبيعية تقريباً. عندئذٍ طول فترة الثقة متناسبة طرداً مع تباين المقدر، وهذا يعني كلما كان التباين أقل فإن طول فترة الثقة يكون اقصر. وبالتالي، إذا وجد المقدر الأكفأ أو الأكفأ تقاربياً، فيمكننا في حالات عدة وبسهولة تامة بناء فترات فترة تقريبية وهي الأقصر أو الأقصر تقاربياً، عند معامل ثقة معينة γ.

لقد لاحظنا عند دراستنا للخواص المقاربة لمقدر المعقولية العظمى $\hat{\theta}_n$ للمعلمة θ في توزيع $f(x;\theta)$ أنه الأكفأ تقاربياً وتوزيعه المقارب طبيعي بمتوسط $E_\theta\left(\hat{\theta}_n(X)\right) = \theta$ وتباين $\sigma_n^2 = \dfrac{1}{ni(\theta)}$ ، أي أن:

$$\mathcal{L}_\theta\left(\hat{\theta}(X)\right) = N\left(\theta, \frac{1}{ni(\theta)}\right) \quad ; \quad \theta \in \Theta$$

عندما $n \to \infty$.

ومن ثم عندما يكون حجم العينة كبيراً فإن $\hat{\theta}_n(X)$ يتبع تقريباً توزيع طبيعي $N\left(\theta, \dfrac{1}{ni(\theta)}\right)$ ، حيث أن:

$$ni(\theta) = -E\left(\frac{\partial^2 \ln L}{\partial \theta^2}\right)$$

معلومات فيشر. وعلى ذلك فإن:

$$\mathcal{L}\left(\left(\hat{\theta}_n(\boldsymbol{X})-\theta\right)\sqrt{ni(\theta)}\right)\approx N(0,1)$$

هكذا، إذا وجد مقدر المعقولية العظمى $\hat{\theta}_n(\boldsymbol{X})$ لمعلمة θ في نموذج $f(x;\theta)$ يمكن استخدامه بسهولة تامة للحصول على فترة ثقة تقريبية لهذه المعلمة.

وعندئذٍ يمكن اختيار $\left(\hat{\theta}_n(\boldsymbol{X})-\theta\right)/\sqrt{ni(\theta)}$ ككمية محورية تقريبية. وبإتباع خوارزمية الطريقة المحورية [راجع بند (3.7)] نحصل على فترة الثقة $\left(T_1(\boldsymbol{X}),T_2(\boldsymbol{X})\right)$ بمعامل ثقة يساوي تقريباً γ.

نفترض $\mathrm{X}=\left(X_1,\dots,X_n\right)$ عينة عشوائية كبيرة الحجم من توزيع $f(x;\theta)$ (مستمر أو منقطع)، و $\hat{\theta}_n(\boldsymbol{X})$ مقدر المعقولية العظمى للمعلمة θ، وإذا كان γ معامل الثقة المطلوب، فإنه يمكن إيجاد عددين $q_1<q_2$ بحيث يكون:

$$P_\theta\left(q_1<\left(\hat{\theta}_n(X)-\theta\right)\sqrt{ni(\theta)}<q_2\right)=\gamma$$

وفي هذه الحالة نعلم أن فترة الثقة تكون أقصر ما يمكن [راجع فقرة (1.4.7)] عندما تكون:

$$q_2=-q_1=Z_{\alpha/2}\qquad;\qquad\gamma=1-\alpha$$

وبتحوير المتباينة:

$$-Z_{\alpha/2}<\left(\hat{\theta}_n(x)-\theta\right)\sqrt{i(\theta)}<Z_{\alpha/2}$$

نحصل على:

$$\hat{\theta}_n(x)-\frac{Z_{\alpha/2}}{\sqrt{ni(\theta)}}<\theta<\hat{\theta}_n(x)+\frac{Z_{\alpha/2}}{\sqrt{ni(\theta)}}$$

وبالتالي فإن $100\gamma\%$ فترة ثقة للمعلمة θ هي:

$$\left(\hat{\theta}_n(\boldsymbol{X})-\frac{Z_{\alpha/2}}{\sqrt{ni(\theta)}},\hat{\theta}_n(\boldsymbol{X})+\frac{Z_{\alpha/2}}{\sqrt{ni(\theta)}}\right)$$

مثـــال 1.7.7

إذا كـــانت $X=(X_1,\dots,X_n)$ عينة عشـــوائية كبيرة الحجم من توزيـــع $N(0,\theta^2)$ فأوجد $100\gamma\%$ فترة ثقة للتباين θ^2.

نعلم أن مقدر المعقولية العظمى للمعلمة θ^2 [راجع مثال (3.2.6)] هو:

$$\hat{\theta}_n(X)=\frac{1}{n}\sum_1^n X_i^2$$

ويمكن التأكد بســـهولة أن:

$$i_n(\theta)=ni(\theta)=\frac{n}{2\theta^4}$$

وبالتـــالي:

$$Q(X;\theta)=\left(\hat{\theta}_n(X)-\theta^2\right)\sqrt{ni(\theta)}=\left(\frac{1}{n}\sum_1^n X_i^2-\theta^2\right)\sqrt{\frac{n}{2\theta^4}}$$
$$\mathcal{L}(Q)\approx N(0,1)$$

وبتحوير المتباينـــة:

$$-Z_{\alpha/2}<\left(\frac{1}{n}\sum_1^n X_i^2-\theta^2\right)\sqrt{\frac{n}{2\theta^4}}<Z_{\alpha/2}$$

إلى متباينة مزدوجة بالنسبة لـــ θ^2 نجد:

$$\frac{\sum X_i^2/n}{1+Z_{\alpha/2}\sqrt{n/2}}<\theta^2<\frac{\sum X_i^2/n}{1-Z_{\alpha/2}\sqrt{n/2}}$$

ومن ثم $100\gamma\%$ فترة الثقة التقريبية المطلوبة للمعلمة θ^2 هي من الشكل:

$$\left(\frac{\sum X_i^2/n}{1+Z_{\alpha/2}\sqrt{n/2}}\ ,\ \frac{\sum X_i^2/n}{1-Z_{\alpha/2}\sqrt{n/2}}\right)$$

مثـــال 2.7.7

إذا كانت $X=(X_1,\dots,X_n)$ عينة عشوائية من توزيع بواســون $\Pi(\theta)$ وكــانت n كبيرة، فأوجد $100\gamma\%$ فترة ثقة تقريبية للمعلمة θ.

يمكن التأكد بســهولة أن:

$$\hat{\theta}_n(X)=\overline{X} \quad ; \quad i_n(\theta)=\frac{n}{\theta}$$

ومــن ثم:

$$Q(X;\theta)=\left(\overline{X}-\theta\right)\sqrt{n/\theta}$$
$$\mathcal{L}(Q)\approx N(0,1)$$

إن المتباينة:

$$-Z_{\alpha/2}<\left(\overline{X}-\theta\right)\sqrt{n/\theta}<Z_{\alpha/2} \qquad ...(1)$$

لا يمكن تحويرها مباشرة كمتباينة مزدوجة في θ، ولكن بملاحظة أن:

$$P\left(\left(\overline{X}-\theta\right)^2<\frac{\theta}{n}Z^2_{\alpha/2}\right)=\gamma$$

لأن:

$$P\left(\left|\overline{X}-\theta\right|<\frac{\theta}{n}Z_{\alpha/2}\right)=\gamma$$

يمكن أخذ المتباينــة:

$$\left(\overline{X}-\theta\right)^2<\frac{\theta}{n}Z^2_{\alpha/2} \qquad ...(2)$$

كمتباينة مكافئة للمتباينــــة المزدوجـــة (1). وبإعـــادة كتابـــة (2) كصيغـــة مـــن الدرجة الثانية في θ وتحويرها نحصل على $100\gamma\%$ فـــترة الثقـــة المطلوبـــة للمعلمـــة θ وهي:

$$\overline{X} + \frac{Z^2_{\alpha/2}}{2n} \pm \left(\frac{\overline{X} Z_{\alpha/2}}{n} - \frac{Z^4_{\alpha/2}}{4n^2}\right)^{1/n}$$

يمكن تجاوز الصعوبة التي واجهتنا في تحوير المتباينة (1) في هذا المثال وفي كثير من المسائل وذلك باستبدال θ بمقدرها $\hat{\theta}_n(X)$ في عبارة $i_n(\theta)$.

بالعودة للمثال السابق نجد أن $Q(X;\theta)$ تصبح:

$$Q(X;\theta) = \left(\overline{X} - \theta\right)\sqrt{\frac{n}{\overline{X}}}$$

وعندئذٍ يمكن تحوير المتباينة المزدوجة:

$$-Z_{\alpha/2} < \left(\overline{X} - \theta\right)\sqrt{\frac{n}{\overline{X}}} < Z_{\alpha/2}$$

إلى:

$$\overline{X} - Z_{\alpha/2}\sqrt{\frac{\overline{X}}{n}} < \theta < \overline{X} + Z_{\alpha/2}\sqrt{\frac{\overline{X}}{n}}$$

وعملياً فإن هذه الفترة لا تختلف كثيراً عن الفترة السابقة. فمثلاً، إذا كانت $\gamma = 0.95$، $\overline{x} = 10$، فإن $Z_{\alpha/2} = 1.96$، ومن ثم الفترة السابقة هي $(9.40,10.63)$، بينما الفترة الأخيرة $(9.38,10.62)$.

مثال 3.7.7

إذا كانت $X = \left(X_1,\dots,X_n\right)$ عينة عشوائية من توزيع بيرنولي $B(1,\theta)$، وكانت n كبيرة، فأوجد $100\gamma\%$ فترة ثقة تقريبية لـ θ.

كما نعلم أن مقدِّر المعقولية العظمى لـ θ هو:

$$\hat{\theta}_n(X) = \overline{X}$$

ومعلومات فيشر:

$$i_n(\theta) = \frac{n}{\theta(1-\theta)} \approx \frac{n}{\overline{X}\left(1-\overline{X}\right)}$$

وبالتـالي:

$$Q(\boldsymbol{X};\theta) = \left(\overline{X} - \theta\right)\sqrt{\frac{n}{\overline{X}\left(1-\overline{X}\right)}}$$

وتكون فـترة الثقـة المطلوبـة:

$$\left(\overline{X} - Z_{\alpha/2}\sqrt{\frac{\overline{X}\left(1-\overline{X}\right)}{n}}\ ,\ \overline{X} + Z_{\alpha/2}\sqrt{\frac{\overline{X}\left(1-\overline{X}\right)}{n}}\right)$$

مثـــال 4.7.7

إذا كـانت $\boldsymbol{X} = \left(X_1,\dots,X_n\right)$ عينـة عشـوائية مــن توزيــع:

$$f(x;\theta) = \theta e^{-\theta x} \quad ; \quad x > 0\ ,\ \theta > 0$$

وكـانت n كبـيرة، فـأوجد $100\gamma\%$ فـترة ثقـة تقريبيـة لـ θ.

كما نعلم أن مقـدِّر المعقولية العظمـى لـ θ هـو:

$$\hat{\theta}(\boldsymbol{X}) = \frac{1}{\overline{X}}$$

[راجــع مثــال (1.2.6)]. وأن $i_n(\theta) = \frac{n}{\theta^2} \approx \frac{n}{\overline{X}^2}$، وبالتـــالي:

$$Q(\boldsymbol{X};\theta) = \left(\frac{1}{\overline{X}} - \theta\right)\sqrt{\frac{n}{\overline{X}^2}}$$

وتكون فـترة الثقـة المطلوبـة:

$$\left(\frac{1}{\overline{X}} - Z_{\alpha/2}\frac{\overline{X}}{\sqrt{n}}\ ,\ \frac{1}{\overline{X}} + Z_{\alpha/2}\frac{\overline{X}}{\sqrt{n}}\right)$$

يمكن استخدام طريقة تقريبية لبناء فترات ثقة. كما نعلم $U(\boldsymbol{X};\theta)=\dfrac{\partial \ln L}{\partial \theta}$ عبارة عن متغير عشوائي توقعه صفر وتباينه $i_n(\theta)$، ومن ثم إذا أخذنا:

$$Q(\boldsymbol{X};\theta)=\frac{U(\boldsymbol{X};\theta)}{\sqrt{i_n(\theta)}} \qquad .....(1)$$

فإن:

$$E_\theta(Q)=0 \quad , \quad \mathrm{var}(Q)=1$$

لكن:

$$U(\boldsymbol{X};\theta)=\frac{\partial \ln L}{\partial \theta}=\sum_1^n \frac{\partial}{\partial \theta}\ln f(X_i;\theta)$$

أي يمكن النظر إلى المتغير $Q(\boldsymbol{X};\theta)$، عند كل قيمة $\theta \in \Theta$ كمجموع n من المتغيرات $\frac{\partial}{\partial \theta}\ln f(X_i;\theta)$ مقسوماً على انحرافه المعياري، وبالتالي حسب مبرهنة النهاية المركزية فإن:

$$\mathcal{L}(Q)=N(0,1)$$

عندما $n\to\infty$. ومن ثم يمكن استخدام (1) ككمية محورية تقريبية، عندما تكون n كبيرة، لإيجاد فترات الثقة.

فمثلاً، بالعودة إلى المثال السابق نجد:

$$U(\boldsymbol{X};\theta)=\frac{n}{\theta}-\sum X_i \quad ; \quad i_n(\theta)=\frac{n}{\theta^2}$$

وبالتالي:

$$Q(\boldsymbol{X};\theta)=\frac{U(\boldsymbol{X};\theta)}{\sqrt{i_n(\theta)}}=\frac{\frac{n}{\theta}-\sum X_i}{\sqrt{n}/\theta}=\sqrt{n}\left(1-\theta\overline{X}\right)$$

ويمكن تحويـر المتباينـة:

$$-Z_{\alpha/2} < \sqrt{n}\left(1-\theta\overline{X}\right) < Z_{\alpha/2}$$

إلى:

$$\frac{1-Z_{\alpha/2}/\sqrt{n}}{\overline{X}} < \theta < \frac{1+Z_{\alpha/2}/\sqrt{n}}{\overline{X}}$$

ومـن ثـم $100\gamma\%$ فـترة الثقـة المطلوبـة للمعلمـة θ هـي:

$$\left(\frac{1-Z_{\alpha/2}/\sqrt{n}}{\overline{X}} < \theta < \frac{1+Z_{\alpha/2}/\sqrt{n}}{\overline{X}}\right)$$

8.7 مناطق الثقة من أجل معلمة متعددة الأبعاد

تطرقنـا حتـى الآن لتقدير فـترة ثقـة لمعلمـة وحيـدة البعـد (سـلمية)، أي لكيفيـة بنــاء فـترة عشـوائية $\left(T_1(X),T_2(X)\right)$ تغطي القيمـة الحقيقيـة للمعلمـة المجهولـة θ باحتمــال لا يقـل عــن γ. لكــن هنــالك حــالات يتعلـق فيهـا التقديـر بمعلمـة متعـددة الأبعـاد $\theta=\left(\theta_1,\ldots,\theta_r\right)$، فعندئـذٍ بـدلاً مـن فــترات ثقـة نعـين منــاطق ثقــة $\mathscr{y}=\mathscr{y}(X)\subset\Theta$ بنـاءً على شـرط مشـابه للشـرط (1.2.7):

$$P_\theta\left(\theta\in\mathscr{y}(X)\right)=\gamma \quad ; \quad \theta\in\Theta \tag{1.8.7}$$

هكـذا، $100\gamma\%$ منطقـة ثقـة هـي عبـارة عـن فئـة جزئيـة عشـوائية مــن فضــاء المعلمـة $\Theta\subset R^r$، المرتبطــة بالعينــة X (وليــس بـ θ)، والــتي تغطـي القيمـة الحقيقيــة للمعلمـة المجهولـة θ باحتمـال يســاوي γ.

الطريقـة العامـة لبنـاء منطقـة الثقـة تتمثـل بمـا يلي: نضـع مقــابل كــل قيمــة $\theta\in\Theta$ فئـة جزئيـة $H(\theta)$ مـن فضـاء العينـة G، أي أن $H(\theta)\subset G$، بحيـث يتحقــق الشــرط:

$$P_\theta\left(X\in H(\theta)\right)=\gamma \tag{2.8.7}$$

بهـذا الشـكل نحصـل على أسـر الفئـات الجزئيــة $\left\{H(\theta);\theta\in\Theta\right\}$.

لنعين الآن من أجل كل $x \in G$ فئة جزئية $\mathscr{Y}(x) \in \Theta$ وفق القاعدة التالية $\mathscr{Y}(x) = \{\theta : x \in H(\theta)\}$. بذلك، نحصل من فضاء المعلمة Θ على أسرة الفئات الجزئية $\{\mathscr{Y}(x); x \in G\}$. لنرى الفئة الجزئية العشوائية $\mathscr{Y}(X)$. إن الحادثين $\{X \in H(\theta)\}$ و $\{\theta \in \mathscr{Y}(X)\}$ متكافئان، لأنه حسب البناء كل منهما يقتضي الآخر، وبالتالي احتماليهما متساويان عند كل قيمة $\theta \in \Theta$. بأخذ العلاقة (2.8.7) بعين الاعتبار، نجد من أجل $\mathscr{Y}(X)$ تتحقق المساواة (1.8.7)، ومن ثم $\mathscr{Y}(X)$ تمثل $100\gamma\%$ منطقة ثقة من أجل المعلمة θ.

تجدر الإشارة هنا، إلى أن هذه الطريقة لبناء منطقة الثقة لا تعطي منطقة ثقة وحيدة (لأنه عند درجة ثقة معطاة γ، يمكن اختيار الفئة $H(\theta)$ بطرق مختلفة). وتهدف المسألة إلى بناء منطقة ثقة ذات قياس أصغري (أصغر ما يمكن)، التي تضمن أكبر دقة (عند γ معطاة) للمعلمة θ. في مسائل معينة الفئة $H(\theta)$، المحققة للشرط (1.8.7)، تبنى عادة بواسطة إحصاء ما $T(X)$ (بشكل عام يقال متجه) توزيعه معلوم. لنوضح هذه الطريقة من خلال الأمثلة الآتية.

مثال 1.8.7

إذا كانت $X = (X_1, \ldots, X_n)$ عينة عشوائية من توزيع طبيعي عام $N(\theta_1, \theta_2^2)$، ونريد بناء منطقة ثقة للمعلمة $\theta = (\theta_1, \theta_2^2)$، أي منطقة ثقة للمتوسط والتباين في آن واحد (simultaneous confidence regon for the mean and variance).

في الفقرة (3.4.7) تم بناء فترتي الثقة على انفراد لكل من المتوسط θ_1 والتباين θ_2^2 غير المعلومين في النموذج $N(\theta_1, \theta_2^2)$ [راجع العلاقتين (12.4.7) و (16.4.7)] وهما:

$$\left(\overline{X} - \frac{S}{\sqrt{n-1}} t_{\alpha/2,n-1} \,,\, \overline{X} + \frac{S}{\sqrt{n-1}} t_{\alpha/2,n-1}\right) \quad \ldots .(1)$$

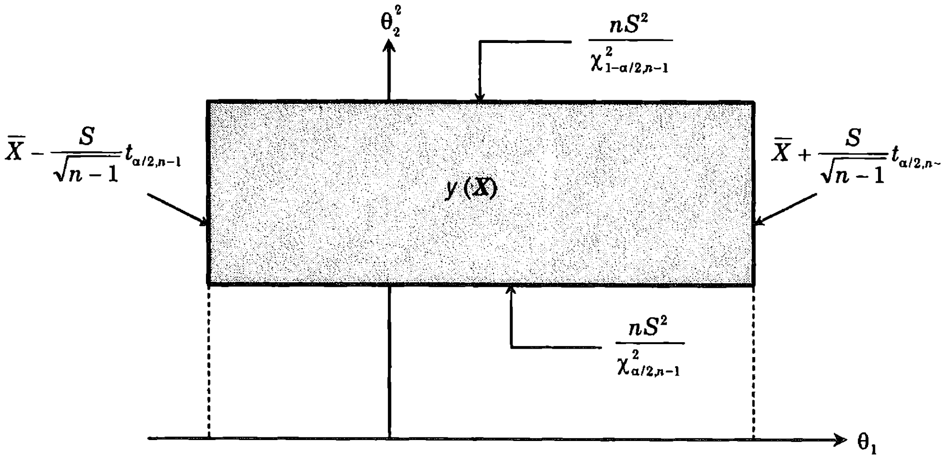
$$\left(\frac{nS^2}{\chi^2_{\alpha/2,n-1}} \,,\, \frac{nS^2}{\chi^2_{1-\alpha/2,n-1}}\right) \quad \ldots .(2)$$

بحيث أن:

$$P\left(\overline{X}-\frac{S}{\sqrt{n-1}}t_{\alpha/2,n-1}<\theta_1<\overline{X}+\frac{S}{\sqrt{n-1}}t_{\alpha/2,n-1}\right)=\gamma$$

$$P\left(\frac{nS^2}{\chi^2_{\alpha/2,n-1}}<\theta_2^2<\frac{nS^2}{\chi^2_{1-\alpha/2,n-1}}\right)=\gamma$$

قد نفكر للوهلة الأولى بدمج فترتي الثقة (1) و (2) للحصول على منطقة ثقة للمعلمة $\theta=\left(\theta_1,\theta_2^2\right)$ كما هي مبينة على الشكل (1.8.7) بالمنطقة المظللة.



شكل (1.8.7)

ومن الواضح أن هذه المنطقة $y(X)$ هي منطقة ثقة للمعلمة $\theta=\left(\theta_1,\theta_2^2\right)$ ولكن معامل الثقة غير محدد. وأن معامل الثقة ليس γ^2 لأن الحادثين (1) و (2) غير مستقلين (المتغيران العشوائيان $\left(\overline{X}-\theta_1\right)/S$ و S^2 اللذان على أساسهما بنيت فترتي الثقة مرتبطان).

يمكن بناء $100\gamma\%$ منطقة ثقة للمعلمة $\theta=\left(\theta_1,\theta_2^2\right)$ باستخدام الإحصاء ذو البعدين $T=\left(\overline{X},S^2\right)$، حيث أن الإحصائين $\overline{X},S^2$ مستقلان [راجع مبرهنة (5.9.3)]، وذلك على النحو الآتي:

بما أن:

$$\mathcal{L}_\theta\left(\frac{\overline{X}-\theta_1}{\theta_2/\sqrt{n}}\right)=N(0,1)$$
$$\mathcal{L}_\theta\left(\frac{nS^2}{\theta_2^2}\right)=\chi^2_{(n-1)}$$

نعين من أجل كل $\theta\in\Theta$ الفئة الجزئية $H(\theta)$ كالآتي:

$$H(\theta)=\left\{x:\left|\frac{\overline{x}-\theta_1}{\theta_2/\sqrt{n}}\right|<t=z_{\alpha_1/2},t`=\chi_{1-\alpha_2/2,n-1}<\frac{ns^2}{\theta_2^2}<t''=\chi^2_{\alpha_2/2,n-1}\right\}$$

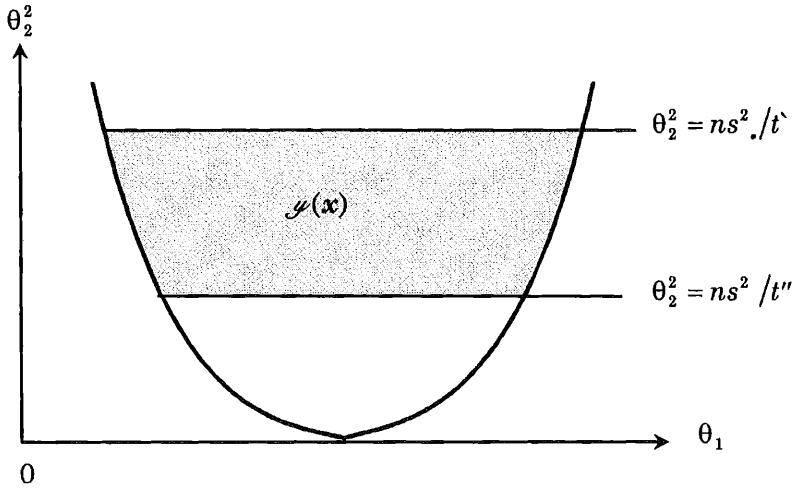
حيث أن $\Phi\left(-Z_{\alpha_1/2}\right)=\alpha_1/2$ ، $S^2=S^2(x)$ ، $\gamma=\gamma_1\gamma_2$ ، $\gamma_1=1-\alpha_1$ ، $\gamma_2=1-\alpha_2$ عندئذٍ، بناءً على استقلال $\overline{X}$ و S^2 نجد:

$$P_\theta\left(X\in H(\theta)\right)=P_\theta\left(\left|\frac{\overline{X}-\theta_1}{\theta_2/\sqrt{n}}\right|<z_{\alpha_1/2}\right)P_\theta\left(t`<\frac{nS^2}{\theta_2^2}<t''\right)=\gamma_1\gamma_2=\gamma$$

أي أن الفئات الجزئية $H(\theta)$ تحقق الشرط (2.8.7). وبحل المتباينتين المزدوجتين المعرفتين للفئات الجزئية $H(\theta)$ بالنسبة لـ θ_1,θ_2^2 نحصل على مناطق الثقة $y(x)$ المطلوبة المعينة بالشرطين:

$$\theta_2^2>\left(\overline{x}-\theta_1\right)^2 n/t \quad , \quad ns^2/t''<\theta_2^2<ns^2/t'$$

وتمثل المنطقة $y(x)$ بجزء من المستوي $\left(\theta_1,\theta_2^2\right)$ المحدد بالقطع المكافئ $\theta_2^2=\left(\overline{x}-\theta_1\right)^2 n/t$ والمستقيمين $\theta_2^2=ns^2/t`$ و $\theta_2^2=ns^2/t''$. وهذا ما يبدو بوضوح بالمنطقة المظللة على الشكل (2.8.7).



شكل (2.8.7)

تمــــارين

1. لتكن X ملاحظـة واحـدة مـن توزيـع:

$$f(x;\theta) = \theta x^{\theta-1} \quad ; \quad 0 < x < 1 \ , \ \theta > 0$$

والمطلـوب:

1. أوجد كميـة محورية واستخدمها لإيجاد مقـدر فـترة ثقـة للمعلمـة θ.
2. أثبـت أن $Y = -1/\ln X \ ; \ (Y/2, Y)$ فـترة ثقـة للمعلمـة θ، ثـم عـين معامل الثقة الموافـق لها.
3. أوجـد أفضـل $100\gamma\%$ فـترة ثقـة للمعلمـة θ.

2. إذا كـانت $X = (X_1, \ldots, X_n)$ عينـة عشـوائية مـن توزيـع $N(\theta,\theta)$، فـأوجد كمية محوريـة للمعلمـة θ، واستخدمها للحصـول على مقـدر فـترة ثقـة لها.

3. إذا كانت $X=(X_1,\dots,X_n)$ عينة عشوائية من توزيع $R\left(\theta-\frac{1}{2},\theta+\frac{1}{2}\right)$ وكانت $Y=(Y_1,\dots,Y_n)$ العينة المرتبة الموافقة للعينة X، فأثبت أن (Y_1,Y_n) فترة ثقة للمعلمة θ واحسب معامل الثقة لها.

4. لتكن $X=(X_1,\dots,X_n)$ عينة عشوائية من مجتمع توزيعه:

$$f(x;\theta)=\theta e^{-\theta x} \quad ; \quad x>0 \ , \ \theta>0$$

والمطلوب:

1. أوجد $100\gamma\%$ فترة ثقة لمتوسط المجتمع.
2. أوجد $100\gamma\%$ فترة ثقة لتباين المجتمع.
3. ما هو احتمال أن هاتين الفترتين تغطيان القيمة الحقيقية للمتوسط والقيمة الحقيقية للتباين في آن واحد؟
4. أوجد مقدر فترة ثقة لـ $\tau(\theta)=e^{-\theta}$.

تنبيه: استخدم نتائج المثال (3.6.7) والملاحظة في البند (2.7).

5. لتكن X ملاحظة واحدة من مجتمع توزيعه:

$$f(x;\theta)=\theta e^{-\theta x} \quad ; \quad x>0 \ , \ \theta>0$$

والمطلوب:

1. إذا كانت $(X,2X)$ فترة ثقة لـ $\tau(\theta)=1/\theta$ فأوجد معامل الثقة لها.
2. أوجد فترة ثقة أخرى لـ $\tau(\theta)=1/\theta$ لها نفس معامل الثقة، لكن توقع طولها أصغر.

6. لتكن $X=(X_1,X_2)$ عينة عشوائية من توزيع $N(\theta,1)$، ويطلب:

1. إذا كانت $(X_{(1)},X_{(2)})$ عبارة عن $100\gamma\%$ فترة ثقة للمعلمة θ، فأوجد الطول المتوقع لها.

2. أوجد مقدر فترة الثقة للمعلمة θ باستعمال $(\overline{X}-\theta)$ كمية محورية بمعامل ثقة γ، ثم قارن طول هذه الفترة بالطول المتوقع في (1).

7. إذا كانت $x=(3.3,-0.3,-0.6,-0.9)$ عينة عشوائية ملاحظة من توزيع طبيعي $N(\theta,\sigma^2=9)$، فأوجد 90% فترة ثقة للمتوسط θ، وإذا افترضنا التباين غير معلوم، فما هي 90% فترة ثقة لـ θ؟

8. إذا كانت $\boldsymbol{X}=(X_1,\dots,X_n)$ عينة عشوائية من مجتمع توزيعه $N(\theta_1,\theta_2^2)$، فما هو حجم العينة n اللازم ليكون 90% فترات ثقة للمعلمة θ_1 أطوالها أقل من $\theta_2/5$، وذلك باحتمال يساوي 0.95.

9. لدينا خمس مجتمعات طبيعية لها نفس التباين غير المعلوم. سحبنا عينة عشوائية من كل مجتمع، فكانت النتائج على النحو الآتي:

S^2 :	40	30	20	42	50
n	6	4	3	7	8

حيث $S^2=\frac{1}{n}\sum_1^n\left(X_i-\overline{X}\right)^2$. أوجد فترة ثقة للتباين المشترك.

10. إذا كانت $\boldsymbol{X}=(X_1,\dots,X_n)$ عينة عشوائية من توزيع منتظم $R(\theta/2;\theta)$، فأثبت أن $\left(X_{(n)},X_{(n)}/\sqrt[n]{1-\gamma}\right)$ هي $100\gamma\%$ فترة ثقة للمعلمة θ.

11. إذا كانت $\boldsymbol{X}=(X_1,\dots,X_n)$ عينة عشوائية من توزيع جاما $\Gamma(\theta,\gamma)$ وكانت n كبيرة، فبين أن أقصر $100\gamma\%$ فترة ثقة تقريبية للمعلمة θ هي:

$$\left(\frac{\overline{X}}{\gamma}\left(1-z_{\alpha/2}/\sqrt{n\gamma}\right),\ \frac{\overline{X}}{\gamma}\left(1+z_{\alpha/2}/\sqrt{n\gamma}\right)\right)$$

12. إذا كانت $\boldsymbol{X}=(X_1,\dots,X_n)$ عينة عشوائية من توزيع $\overline{B}(r,\theta)$ وكانت n كبيرة، فبين أن أقصر $100\gamma\%$ فترة ثقة تقريبية للمعلمة θ هي:

$$\left(\hat{\theta}_n(X) \pm z_{\alpha/2}\sqrt{\hat{\theta}_n(X)\left(1-\hat{\theta}_n(X)\right)^2 \Big/ (rn)}\right) = \frac{1}{r+\overline{X}}\left(\overline{X} \pm z_{\alpha/2}\sqrt{r\overline{X} \Big/ \left[n(r+\overline{X})\right]}\right)$$

13. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع أسي:

$$f(x;\theta) = e^{-(x-\theta)} \quad ; \quad 0 \le x < \theta$$

فبين أن $100\gamma\%$ فترة ثقة للمعلمة θ هي:

$$\left(X_{(1)} + \frac{\ln(1-\gamma)}{n}\ ,\ X_{(1)}\right) \quad ; \quad X_{(1)} = \min X_i$$

14. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta_1) = \theta_1 e^{-\theta_1 x} \quad ; \quad x > 0\ ,\ \theta_1 > 0$$

وكانت $Y=(Y_1,\ldots,Y_n)$ عينة عشوائية من توزيع:

$$f(y;\theta_2) = e^{-\theta_2 y} \quad ; \quad y > 0\ ,\ \theta_2 > 0$$

فالمطلوب:

1. أوجد $100\gamma\%$ فترة ثقة للمعلمة $\theta = \theta_1/\theta_2$.
2. أوجد منطقة ثقة للمعلمة $\theta = (\theta_1,\theta_2)$ على الصورة:

$$\left\{(\theta_1,\theta_2) : \theta_1\sum_1^n X_i + \theta_2\sum_1^n Y_i \le c\right\}$$

15. إذا كانت $X=(X_1,\ldots,X_n)$ عينة عشوائية من توزيع $N(\theta,\theta^2)$، فالمطلوب:

1. بين أن $100\gamma\%$ فترة ثقة للتباين θ^2 لها الشكل:

$$\left(\overline{X}\Big/\left(1+z_{\alpha/2}/\sqrt{n}\right)\ ,\ \overline{X}\Big/\left(1-z_{\alpha/2}/\sqrt{n}\right)\right)$$

2. إذا كانت n كبيرة، فأوجد $100\gamma\%$ فترة ثقة تقريبية لـ θ^2.
3. إذا كانت $x=(x_1,\ldots,x_{100})$ عينة ملاحظة من $N(\theta,\theta^2)$، وكانت $\bar{x}=3\ ,\ \alpha=0.05$ فاحسب الفترتين السابقتين.

الفصل الثامن:

طرق تقدير بييز

1.8 مقدمة

في دراستنا السابقة لطرق التقدير (بنقطة أو فترة) كنا نفترض أن لدينا مجتمع صيغة توزيعه معلومة $f(x;\theta)$ لكنها تحتوي على معلمة ثابتة غير معلومة θ من فضاء المعلمة Θ، ونرغب في تقدير θ بناءً على المعلومات المتوفرة في عينة عشوائية $x=(x_1,...,x_n)$ مأخوذة من هذا المجتمع. وبعبارة أخرى إن معلوماتنا المتوفرة حول المعلمة θ قبل سحب العينة العشوائية x تتمثل في أنها قيمة ثابتة، لكنها غير معلومة، وهي إحدى نقاط فضاء المعلمة Θ.

تتوفر في حالات عدة معلومات إضافية حول المعلمة غير المعلومة θ تبين أنها متغيرة (غير ثابتة)، ويمكن اعتبارها كقيمة لمتغير عشوائي نرمز له بـ Θ، ومن ثم له توزيع احتمالي. فمثلاً، لنفترض أننا نرغب في تقدير نسبة العطب في إنتاج مصنع ما لأجهزة التلفاز. وفي يوم ما أخذنا لأعلى التعيين 10 أجهزة ولنرمز لنتائج الملاحظة بـ $X_1,...,X_{10}$، حيث $X_i=1$ إذا كان التلفاز معطوب و $X_i=0$ إذا كان التلفاز غير معطوب (سليم)، أي يمكن اعتبار $\boldsymbol{X}=(X_1,...,X_n)$ عينة عشوائية من توزيع:

$$f(x;\theta)=\theta^x(1-\theta)^{1-x} \quad ; \quad x=0,1 \;,\; 0\le\theta\le 1$$

يمكن التأكد بسهولة من أن مقدِّر المعقولية العظمى لـ θ هو $\hat{\theta}(\boldsymbol{X})=\overline{X}$ وهو نفس مقدر العزوم. لنفترض الآن توفر بعض المعلومات الإضافية حول المعلمة

θ، وذلك قبل سحب عينة عشوائية، وهي أننا لاحظنا خلال أيام مختلفة أن نسبة العطب θ في إنتاج المصنع المذكور متغيرة وأن هذا التغير يمكن تمثيله بمتغير عشوائي Θ دالة احتماله:

$$g_\Theta(\theta) = 6\theta(1-\theta) \quad ; \quad 0 \le \theta \le 1$$

والسؤال المهم الآن، هو: كيف يمكن استخدام هذه المعلومات الإضافية حول المعلمة θ لتقدير θ_0 حيث θ_0 قيمة Θ يوم سحب عينة عشوائية x؟.

تدعى طرق التقدير الإحصائية التي تجيب على هذا السؤال بطرق تقدير بييز (Bayesian method of estimation).

هكذا، فطرق التقدير السابقة تفترض أن المعلمة θ التي نريد تقديرها عبارة عن قيمة ثابتة غير معلومة، لكن في طرق تقدير بييز ينظر إلى المعلمة θ كقيمة لمتغير عشوائي Θ له توزيع احتمالي.

إذا كان Θ متغير عشوائي له توزيع احتمالي، فسنرمز لدالة توزيع Θ بـ $G(\theta) = G_\Theta(\theta)$ ولدالة الكثافة (أو دالة الاحتمال) بـ $g(\theta) = g_\Theta(\theta)$، ونفترض أن هاتين الدالتين معرفتان على فضاء المعلمة θ، ولا تحتويان على معالم غير معلومة.

في مسائل كثيرة ليس واقعياً افتراض أن θ قيمة لمتغير عشوائي Θ، وفي مسائل أخرى يبدو ذلك ممكناً، ولكن توزيع Θ غير معلوم، وحتى إن كان معلوماً فهو يحتوي على معالم أخرى غير معلومة. ومع ذلك، يكون منطقياً في بعض المسائل الافتراض بأن توزيع Θ معلوم تماماً أو يمكن اشتقاقه، وسندرس فيما يلي هذه الحالة.

2.8 التوزيع القبلي والبعدي

PRIOR AND POSTERIOR DISTRIBUTION

نستعمل حتى الآن الرمز $f(x;\theta)$ للدلالة على توزيع متغير عشوائي ع من أجل كل قيمة $\theta \in \Theta$. وعندما نريد الإشارة إلى أن المعلمة θ قيمة لمتغير عشوائي

Θ، سنرمز لتوزيع ξ بـ $f_{\xi|\Theta}=f(x|\theta)$ للتمييز عن $f(x;\theta)$. ويجب ملاحظة أن $f(x|\theta)$ ترمز للتوزيع الشرطي (المشروط) للمتغير العشوائي ξ بافتراض $\Theta=\theta$، ومن ثم ترمز $f(x;\theta)$ للتوزيع المشترك للمتغيرين Θ و ξ.

لتكن $X=(X_1,...,X_n)$ عينة عشوائية بحجم n من توزيع $f(x|\theta)$، حيث أن θ قيمة لمتغير عشوائي Θ. ولنفترض أن توزيع Θ معلوم تماماً، أي أن $g(\theta)$ معلومة (لا تحتوي على معالم غير معلومة)، ونريد تقدير $\tau(\theta)$ (ومنها $\tau(\theta)=\theta$). كيف يمكننا الاستفادة من المعلومات الإضافية المتمثلة بمعرفة توزيع Θ في عمليةِ التقدير؟ في الفصول السابقة كنا نستخدم دالة المعقولية $L(x|\theta)$ كعبارة تشتمل على كل معلوماتنا حول المعلمة θ، أي أن دالة المعقولية تتضمن العينة الملاحظة $x=(x_1,...,x_n)$ وكذلك صيغة التوزيع $f(x;\theta)$. والآن نحتاج لعبارة تشتمل على كل المعلومات التي تحتويها دالة المعقولية بالإضافة إلى المعلومات الإضافية المتمثلة بمعرفة التوزيع $g(\theta)$ الذي يدعى بالتوزيع القبلي لـ Θ (prior distribution)، وهذا يلخص ما نعرفه حول θ قبل سحب عينة عشوائية. وبعبارة أخرى نبحث عن دالة تلخص ما نعرفه حول θ بعد أخذ عينة عشوائية، أي نبحث عن التوزيع البعدي لـ Θ بعد الحصول على عينة عشوائية $x=(x_1,...,x_n)$.

تعريف 1.2.8 التوزيع القبلي Prior Distribution

يدعى التوزيع $g_\Theta(\theta)=g(\theta)$ بالتوزيع القبلي لـ Θ وهو توزيع احتمالي يحتوي على كل معلوماتنا حول المعلمة θ قبل الحصول على عينة عشوائية $x=(x_1,...,x_n)$. نلاحظ أن الدالة $g(\theta)$ تصف درجة اعتقادنا (ثقتنا) في القيم الممكنة لـ Θ قبل المعاينة.

تعريف 2.2.8 التوزيع البعدي Posterior Distribution

يدعى التوزيع $g(\theta|x)$ بالتوزيع البعدي لـ Θ، وهو توزيع احتمالي شرطي لـ Θ، عند $X=x$، ويصف درجة اعتقادنا في القيم الممكنة للمعلمة Θ بعد

الحصول على x. أي يشتمل على كل المعلومات حول θ بعد الحصول على عينة عشوائية x.

نلاحظ من التعريفين أن الحصول على عينة عشوائية x تغير درجة اعتقادنا أو ثقتنا في القيم المختلفة لـ Θ، وذلك بتحويل التوزيع القبلي إلى التوزيع البعدي، وهذا ما تؤكده نظرية الاحتمالات بشكل عام.

تَستخدم تقنيات بيز التوزيع القبلي $g(\theta)$ والتوزيع المشترك للعينة العشوائية X و Θ، أي $f(x;\theta)$ لحساب التوزيع البعدي $g(\theta|x)$ لـ Θ.

كما نعلم من نظرية الاحتمالات أن التوزيع المشترك $f(x;\theta)$ لـ Θ و X يساوي إلى التوزيع الهامشي لأحدهما مضروباً بالتوزيع الشرطي للآخر، أي أن:

$$f(x;\theta) = g(\theta)f(x|\theta) = f(x)g(\theta|x)$$

ومن ثم:

$$g(\theta|x) = \frac{g(\theta)f(x|\theta)}{f(x)} \qquad (1.2.8)$$

حيث أن:

$$f(x) = \begin{cases} \int g(\theta)f(x|\theta)d\theta & \text{حالة } \Theta \text{ متغير مستمر} \\ \sum_{\theta} g(\theta)f(x|\theta) & \text{حالة } \Theta \text{ متغير منقطع} \end{cases}$$

$$f(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

نلاحظ من العلاقة (1.2.8) أن:

$$g(\theta|x) \propto g(\theta)f(x|\theta) = g(\theta)L(x|\theta) \qquad (2.2.8)$$

أي أن الدالة $f(x)$ يمكن أن تدخل في ثابت التناسب.

إذا وجد إحصاء كافي $T(X)$ للمعلمة θ، وكان توزيعه $h(t|\theta)$، فيمكن كتابة العلاقة (2.2.8) على الصورة:

$$g(\theta|x) \propto g(\theta)h(t|\theta)N(x)$$

ويمكن إدخال $N(x)$ في ثابت التناسب:

$$g(\theta|x) \propto g(\theta)h(t|\theta)$$

يتم تعيين ثابت التناسب باعتبار $g(\theta|x)$ دالة كثافة شرطية (أو دالة احتمال شرطية) لمتغير عشوائي Θ، أي أن:

$$\int g(\theta|x)d\theta = 1 \quad \text{في حالة } \Theta \text{ مستمرة}$$

أو:

$$\sum_{\theta} g(\theta|x) = 1 \quad \text{في حالة } \Theta \text{ منقطعة}$$

مثال 1.2.8

إذا كانت $x = (x_1,...,x_n)$ عينة عشوائية من توزيع:

$$f(x|\theta) = \theta^x (1-\theta)^{1-x} \quad ; \quad x = 0,1$$

وكان التوزيع القبلي لـ Θ:

$$g(\theta) = 6\theta(1-\theta) \quad ; \quad 0 \le \theta \le 1$$

فأوجد التوزيع البعدي لـ Θ.

بما أن:

$$f(x|\theta) = L(x|\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^t(1-\theta)^{n-t} \quad ; \quad t = \sum_{i=1}^{n} x_i$$

$$f(x) = \int g(\theta)L(x|\theta)d\theta = 6\int_0^1 \theta^{t+1}(1-\theta)^{n-t+1}d\theta = 6\beta(t+2, n-t+2)$$

وذلك بناءً على العلاقة (36.3.2). ويكون التوزيع البعدي لـ Θ:

$$g(x|\theta)=\frac{g(\theta)f(x|\theta)}{f(x)}=\frac{1}{\beta(t+2,n-t+2)}\theta^{t+1}(1-\theta)^{n-t+1} \quad ; \quad 0\le\theta\le 1$$

نلاحظ أن التوزيع القبلي ما هو إلا توزيع $Be(2,2)$ وكذلك التوزيع البعدي هو $Be(t+2,n-t+2)$، أي أن كلا التوزيعين من عائلة توزيعات بيتا.

مثال 2.2.8

إذا كانت $x=(x_1,\ldots,x_n)$ عينة عشوائية من توزيع بواسون:

$$f(x|\theta)=\frac{\theta^x}{x!}e^{-\theta} \quad ; \quad x=0,1,\ldots$$

وكان التوزيع القبلي لـ Θ هو $\Gamma(2,1/3)$، أي أن:

$$g(\theta)=9\theta e^{-3\theta} \quad ; \quad \theta>0$$

فأوجد التوزيع البعدي لـ Θ.

بما أن:

$$L(x|\theta)=f(x|\theta)=\prod_{i=1}^{n}\frac{\theta^{x_i}}{x_i!}e^{-\theta}=\frac{\theta^t}{\prod_1^n(x_i!)}e^{-n\theta} \quad ; \quad t=\sum_{i=1}^{n}x_i$$

فإن:

$$g(\theta|x)\propto g(\theta)L(x|\theta)=\frac{9}{\prod_1^n(x_i!)}\theta^{t+1}e^{-(n+3)\theta}\propto\theta^{t+1}e^{-(n+3)\theta} \quad ; \quad \theta>0$$

وهذا يعني أن التوزيع $g(\theta|x)$ ما هو إلا توزيع جاما $\Gamma\left(t+2,\frac{1}{n+3}\right)$، وحسب العلاقة (20.3.2)، فإن:

$$f(\theta|x)=\frac{(n+3)^{t+2}}{\Gamma(t+2)}\theta^{t+1}e^{-(n+3)\theta} \quad ; \quad \theta>0$$

نلاحظ أيضاً أن التوزيع البعدي ينتمي لنفس عائلة التوزيعات التي ينتمي إليها التوزيع القبلي وهي عائلة التوزيعات الجماوية $\Gamma(\alpha,\beta)$.

مثال 3.2.8

إذا كانت $x=(x_1,...,x_n)$ عينة عشوائية من توزيع:

$$f(x|\theta)=\theta e^{-\theta x} \quad ; \quad x>0 \;,\; \theta>0$$

وكان التوزيع القبلي لـ Θ هو $\Gamma(a+1,1/b)$، أي أن:

$$g(\theta)=\frac{b^{a+1}}{\Gamma(a+1)}\theta^{a}e^{-b\theta} \quad ; \quad \theta>0$$

وعلمت بأن $T(X)=-\sum_{i=1}^{n}\ln X_i$ إحصاء كافي للمعلمة θ، فأوجد باستخدام الإحصاء الكافي $T(X)$ التوزيع البعدي لـ Θ.

كما نعلم [راجع المثال (3.2.4)] أن توزيع الإحصاء الكافي $T(X)=-\sum_{i=1}^{n}\ln X_i$ هو عبارة عن توزيع $\Gamma(n,1/\theta)$، أي أن:

$$h(t|\theta)=\frac{\theta^{n}}{\Gamma(n)}t^{n-1}e^{-\theta t} \quad ; \quad t>\theta$$

وحيث أن:

$$g(\theta|x)\propto g(\theta)h(t|\theta)$$

فإن

$$g(t|\theta)\propto \theta^{a+n}\; e^{-(b+t)\theta}$$

وهذا يعني أن التوزيع البعدي لـ Θ هو $\Gamma\left(a+n+1,\frac{1}{b+t}\right)$، أي أن:

$$g(\theta|x)=\frac{(b+t)^{a+n+1}}{\Gamma(a+n+1)}\theta^{a+n}\ e^{-(b+t)\theta}\quad ;\quad \theta>0$$

نلاحظ أيضاً أن التوزيعين القبلي والبعدي لـ Θ ينتميان إلى عائلة التوزيعات الجماوية (أي نفس العائلة).

نلاحظ من خلال الأمثلة السابقة أن التوزيع البعدي لـ Θ هو توزيع مستمر، ومن ثم يتبادر لذهن القارئ بأن التوزيعات البعدية هي دوماً توزيعات مستمرة، وهذا غير صحيح. ويمكن إيضاح ذلك من خلال المثال الآتي:

مثال 4.2.8

إذا كانت X ملاحظة مفردة على متغير عشوائي ع توزيعه:

$$f(x|\theta)=\theta^{x}(1-\theta)^{1-x}\quad ;\quad x=0,1\ ,\ \theta=0.3,0.4$$

وكان التوزيع القبلي لـ Θ معطى بالجدول:

θ	$\theta_1=0.3$	$\theta_2=0.4$
$g(\theta)$	0.4	0.6

فأوجد التوزيع البعدي لـ Θ.

بما أن:

θ	θ_1	θ_2
$f(x_1=0\|\theta)$	0.7	0.6
$f(x_2=1\|\theta)$	0.3	0.4

فإن:

$$g(\theta_1|x_1)=\frac{g(\theta_1)f(x_1|\theta_1)}{\sum_{\theta}g(\theta)f(x_1|\theta)}=\frac{(0.4)(0.7)}{(0.4)(0.7)+(0.6)(0.6)}=\frac{28}{64}=\frac{7}{16}$$

$$g(\theta_2|x_1)=\frac{g(\theta_2)f(x_1|\theta_2)}{\sum_{\theta}g(\theta)f(x_1|\theta)}=\frac{(0.6)(0.6)}{0.64}=\frac{36}{64}=\frac{9}{16}$$

وبشكل مشابه:

$$g(\theta_1|x_2)=\frac{g(\theta_1)f(x_2|\theta_1)}{\sum_{\theta} g(\theta)f(x_2|\theta)}=\frac{(0.4)(0.3)}{(0.4)(0.3)+(0.6)(0.4)}=\frac{12}{36}=\frac{1}{3}$$

$$g(\theta_2|x_2)=\frac{g(\theta_2)f(x_2|\theta_1)}{\sum_{\theta} g(\theta)f(x_2|\theta)}=\frac{(0.4)(0.3)}{0.36}=\frac{24}{36}=\frac{2}{3}$$

وبذلك فإن التوزيع البعدي لـ Θ:

θ	$x_1=0$	$x_2=1$
$g(\theta_1\|x)$	7/16	1/3
$g(\theta_2\|x)$	9/16	2/3

3.8 التوزيعات القبلية المرافقة

CONJUGATE PRIOR DISTRIBUTION

رأينا من خلال الأمثلة السابقة أنه بالإمكان وبسهولة إيجاد التوزيع البعدي لـ Θ بناءً على معرفة التوزيع القبلي لـ Θ، وهذا ما كنا نفترضه دائماً. لكن المسألة ليست بهذه البساطة، حيث في مسائل كثيرة يكون التوزيع القبلي لـ Θ غير معلوم، وبالتالي يجب تعيينه.

توجد عدة طرق لاختيار التوزيعات القبلية لـ Θ أهمها اثنان:

1. التوزيعات القبلية المرافقة.
2. التوزيعات القبلية غير المُعلِمة.

وسندرس في هذا البند الطريقة الأولى.

رأينا في المثال (1.2.8) أن كلا التوزيعين القبلي والبعدي لـ Θ عضو في عائلة توزيعات بيتا $Be(\upsilon_1,\upsilon_2)$، حيث التوزيع القبلي $Be(2,2)$ والتوزيع البعدي

$Be(t+2,n-t+2)$. وكذلك رأينا في المثال (2.2.8) أن التوزيع القبلي والبعدي لـ Θ هما على الترتيب $\Gamma\left(2,\frac{1}{3}\right)$ و $\Gamma\left(t+2,\frac{1}{n-3}\right)$ ، أي كلاهما عضو في عائلة التوزيعات الجماوية $\Gamma(\alpha,\beta)$. ونلاحظ نفس الصفة في المثال (3.2.8). وبعبارة أخرى، نلاحظ في الأمثلة الثلاث أن التوزيع القبلي والبعدي لـ Θ لهما نفس النموذج الإحصائي. وفي مثل تلك الحالات يقال عن التوزيع القبلي $g(\theta)$ أنه مرافق للتوزيع الاحتمالي الذي سحبت منه العينة العشوائية (توزيع مجتمع الصفة المدروسة ع).

تعريف 1.3.8

يقال عن مجموعة التوزيعات القبلية لـ Θ أنها تشكل عائلة $\underline{G}$ مرافقة للتوزيع الاحتمالي الذي سحبت منه العينة العشوائية إذا كان التوزيع البعدي لـ Θ ، المشتق بناءً على توزيع قبلي $g(\theta)\in\underline{G}$ ، ينتمي إلى نفس عائلة التوزيعات التي ينتمي إليها التوزيع القبلي $g(\theta)$ مهما كانت العينة المشاهدة x. وتسمى هذه الخاصية أحياناً بخاصية الانغلاق تحت المعاينة (closed under sampling).

وتصبح المسألة الآن: كيف نحصل على التوزيع القبلي المرافق؟

لنفترض للتبسيط أن $x=(x_1,\ldots,x_n)$ عينة عشوائية من توزيع أسي:

$$f(x|\theta)=e^{A(\theta)B(x)+C(\theta)+D(x)}$$

عندئذٍ دالة المعقولية L (كدالة في θ عند عينة مشاهدة x):

$$L(x|\theta)=L(\theta)=e^{A(\theta)\sum B(x_i)+nC(\theta)+\sum D(x_i)}$$

ومن ثم يمكن كتابة:

$$L(\theta)=L\left(\theta,n,\sum B(x_i)\right)\propto e^{nC(\theta)+A(\theta)\sum B(x_i)} \qquad \ldots(1)$$

إذا اخترنا التوزيع القبلي لـ Θ متناسب مع دالة المعقولية، أي أن:

$$g(\theta;a_1,b_1) \propto L(\theta;a_1,b_1) \propto e^{a_1C(\theta)+b_1A(\theta)} \qquad \ldots(2)$$

وبدمج (ضربياً) الدالتين الواردتين في الطرف الأيمن في (1) و (2)، وبأخذ التوزيع البعدي لـ Θ متناسباً مع العبارة الناتجة عن الدمج، نجد:

$$g(\theta|x;a_2,b_2) \propto e^{a_2C(\theta)+b_2A(\theta)}$$

حيث أن $b_2 = b_1 + \sum B(x_i)$ و $a_2 = a_1 + n$. وهذا يعني أن التوزيع البعدي لـ Θ ينتمي لنفس عائلة التوزيعات، التي ينتمي إليها التوزيع القبلي لـ Θ، أي أن التوزيع القبلي $g(\theta;a_1,b_2)$ هو توزيع مرافق للتوزيع الاحتمالي $f(x|\theta)$.

مثال 1.3.8

إذا كانت $x=(x_1,\ldots,x_n)$ عينة عشوائية من توزيع بواسون:

$$f(x|\theta)=\frac{\theta^x}{x!}e^{-\theta} \quad ; \quad x=0,1,\ldots$$

فأوجد توزيع قبلي مرافق، ومن ثم أوجد التوزيع البعدي بناءً عليه.

إن دالة المعقولية:

$$L(x|\theta)=\frac{\theta^{\sum x_i}}{\prod_1^n (x_i!)}e^{-n\theta} \qquad x_i=0,1,\ldots$$

ويمكن كتابتها على الصورة:

$$L(\theta;n,t) \propto \theta^t e^{-n\theta} \quad ; \quad t=\sum_{i=1}^n x_i$$

لذلك نأخذ التوزيع القبلي:

$$g(\theta;a_1,b_1) \propto \theta^{a_1}e^{-b_1\theta} \quad ; \quad \theta>0$$

وهذا ما هو إلا توزيع $\Gamma\left(a_1+1,\frac{1}{b_1}\right)$، أي أن:

$$g(\theta;a_1,b_1)=\frac{b_1^{a+1}}{\Gamma(a_1+1)}\theta^{a_1}e^{-b_1\theta} \quad ; \quad \theta>0 \qquad \ldots(1)$$

حيث أن $(a_1>-1,b_1>0)$ a_1,b_1 ثوابت التوزيع. وبالتالي التوزيع البعدي لـ Θ يكون:

$$g(\theta|x)\propto\theta^{a_1+t}e^{-(n+b_1)\theta} \quad ; \quad \theta>0$$

وهذا عبارة عن توزيع $\Gamma\left(a_1+t+1,\frac{1}{n+b_1}\right)$، حيث أن a_1,b_1 ثوابت التوزيع، أي أن:

$$g(\theta|x)=\frac{(n+b_1)^{a_1+t+1}}{\Gamma(a_1+t+1)}\theta^{a_1+1}e^{-(n+b_1)\theta} \quad ; \quad \theta>0$$

بما أن التوزيع القبلي (1) والتوزيع البعدي (2) لـ Θ عضوان في عائلة التوزيعات الجماوية $\Gamma(\alpha,\beta)$، فإن التوزيع (1) هو توزيع قبلي لـ Θ مرافق للتوزيع الذي سحبت منه العينة العشوائية x.

وبشكل عام إذا كانت $x=(x_1,\ldots,x_n)$ عينة عشوائية من توزيع $f(x|\theta)$، وكانت $L(\theta)=L(x|\theta)$ دالة المعقولية. فيمكن كتابة دالة المعقولية على الصورة:

$$L(\theta;\alpha)$$

حيث أن α هي ثوابت دالة المعقولية باعتبارها دالة في θ. وإذا افترضنا التوزيع القبلي $g(\theta)$ لـ Θ على أنه متناسب مع دالة المعقولية، أي أن:

$$g(\theta;\alpha_1)\propto L(\theta;\alpha_1)$$

حيث α_1 هي ثوابت التوزيع القبلي. بدمج دالة الإمكان مع التوزيع القبلي (ضربياً)، وأخذ التوزيع القبلي متناسباً مع هذا الجداء نحصل على التوزيع البعدي لـ Θ، أي أن:

$$g(\theta|x;\alpha_2)\propto g(\theta;\alpha_1)L(\theta;\alpha)\propto L(\theta;\alpha_1)L(\theta;\alpha)\propto L(\theta;\alpha_2)$$

حيث α_2 هي ثوابت التوزيع البعدي لـ Θ وهي دالة في (α,α_1).

مثال 2.3.8

إذا كانت $x=(x_1,\ldots,x_n)$ عينة عشوائية من توزيع:

$$f(x|\theta)=\frac{1}{\theta} \quad ; \quad 0<x<\theta$$

فأوجد توزيع قبلي مرافق، ومن ثم أوجد التوزيع البعدي لـ Θ.

إن دالة المعقولية:

$$L(\theta;n)=\frac{1}{\theta^n} \quad ; \quad \theta>x_{(n)}$$

راجع علاقة (11.2.4). بأخذ التوزيع القبلي لـ Θ:

$$g(\theta;a,b)\propto\frac{1}{\theta^{a+1}} \quad ; \quad \theta>b$$

حيث $a,b>0$ ثوابت التوزيع. ويحسب ثابت التناسب c_1 على النحو الآتي:

$$\int_b^\infty g(\theta;a,b)d\theta=c_1\int_b^\infty\frac{1}{\theta^{a+1}}d\theta=1\Rightarrow c_1=ab^a$$

نحصل على التوزيع القبلي لـ Θ:

$$g(\theta)=g(\theta;a+b)=\frac{ab^a}{\theta^{a+1}} \quad ; \quad \theta>b$$

وبالتالي فالتوزيع البعدي لـ Θ:

$$g(\theta|x)\propto\frac{1}{\theta^{n+a+1}} \quad ; \quad \theta>t=\max\left(b,x_{(n)}\right)$$

ويعين ثابت التناسب c_2 على النحو الآتي:

$$\int_t^\infty g(\theta|x)d\theta = c_2\int_t^\infty \frac{d\theta}{\theta^{n+a+1}} = 1 \Rightarrow c_2 = (n+a)t^{n+a}$$

أي أن التوزيع البعدي لـ Θ هو:

$$g(\theta|x) = \frac{(n+a)t^{n+a}}{\theta^{n+a+1}} \quad ; \quad .\theta > t$$

مثـــــال 3.3.8

إذا كانت $x = (x_1,\dots,x_n)$ عينة عشوائية من توزيع:

$$f(x|\theta) = C_n^x \theta^x (1-\theta)^{n-x} \quad ; \quad x = 0,1,\dots,n$$

فأوجد التوزيع البعدي لـ Θ.

بما أن دالة المعقولية:

$$L(\theta) = L(x|\theta) = \prod_{i=1}^n C_n^{x_i}\theta^{x_i}(1-\theta)^{n-x_i} \propto \theta^t(1-\theta)^{n^2-t} \quad ; \quad 0<\theta<1 \;,\; t = \sum_1^n x_i$$

فإن التوزيع القبلي:

$$g(\theta) \propto \theta^a(1-\theta)^b$$

ومن ثم فالتوزيع البعدي:

$$g(\theta|x) \propto \theta^{a+t}(1-\theta)^{b+n^2-t}$$

وهذا ما هو إلا توزيع $Be(a+t-1, b+n^2-t-1)$، أي أن:

$$g(\theta|x) = \frac{1}{Be(a+t-1, b+n^2-t-1)}\theta^{a+t}(1-\theta)^{b+n^2-t} \quad ; \quad 0<\theta<1$$

وهو المطلوب.

4.8 التوزيعات القبلية غير المُعلِمة

NON-INFORMATIVE PRIOR DISTRIBUTIONS

في حالات عدة تكون معلوماتنا القبلية (قبل سحب عينة عشوائية) قليلة أو نادرة حول المعلمة θ، بل شبه معدومة بحيث لا تتجاوز معرفتنا كون θ متغيرة ويمكن اعتبارها قيمة لمتغير عشوائي Θ. ونرغب في تمثيل هذه المعرفة البسيطة بتوزيع احتمالي يعبر عنها. تدعى مثل هذه التوزيعات الاحتمالية القبلية بالتوزيعات القبلية غير المُعلِمة (non-informative priors distribution) وأحياناً تسمى بالتوزيعات القبلية الجاهلة (ignorance prior distribution)، وهي عبارة عن توزيعات تعبر عن المعرفة القليلة أو شبه المعدومة، وتوافق الحالة التي تكون المعلومات المتوفرة لدينا حول المعلمة θ قبل المعاينة ليست ذات أهمية بالمقارنة مع المعلومات المتوقع الحصول عليها من العينة العشوائية.

إن حالة المعرفة شبه المعدومة أو المعرفة البسيطة المحدودة حول المعلمة تفترض عدم التمييز بين القيم المختلفة الممكنة لهذه المعلمة، أي إعطائها نفس الوزن أو الأهمية. وتوجد عدة طرق للتعبير عن ذلك، نذكر منها:

1. التوزيع المنتظم

وهنا نميز حالتين:

الحالة الأولى: إذا كانت المعلمة θ تأخذ قيماً في مدى محدود (a,b) أو في المدى $(-\infty,+\infty)$، فنختار عندئذٍ توزيعاً قبلياً منتظماً لـ Θ على هذه الفترة.

إذا كانت $a<\theta<b$، فإن:

$$g(\theta)=\frac{1}{b-a} \tag{1.4.8}$$

وإذا كانت $-\infty<\theta<\infty$، فإن:

$$g(\theta)\propto d\theta \quad ; \quad -\infty<\theta<\infty \tag{2.4.8}$$

الحالة الثانية: إذا كانت المعلمة θ تأخذ قيماً في المدى $(0,\infty)$، فإننا نختار توزيعاً قبلياً منتظماً لـ $\tau(\theta)=\ln\theta$ على الفترة $(-\infty,+\infty)$، أي أن:

$$g(\theta)\propto\frac{d\theta}{\theta} \quad ; \quad 0<\theta<\infty \tag{3.4.8}$$

نلاحظ أن التوزيعات القبلية غير المُعلِمة لـ Θ في (2.4.8) و (3.4.8) هي توزيعات غير تامة أو ناقصة *(improper)* لأن تكامل هذه التوزيعات على فضاء المعلمة θ يعطي ما لانهاية (∞). وهذه ليست مشكلة عملياً لأنه يمكن استخدام التوزيع الناقص كدالة وزن تخصص أوزاناً متساوية على مسافة لانهائية.

2. قاعدة عدم الاختلاف أو الثبات (invariance)

هذه الطريقة للحصول على التوزيعات القبلية غير المُعلِمة مبنية على افتراض أن العنصر الاحتمالي في التوزيع القبلي لا يتغير تحت تأثير أي تحويلة. وهذا الافتراض يبرر اختيار التوزيع القبلي ليتناسب مع الجذر التربيعي الموجب لمعلومات فيشر، أي أن:

$$g(\theta)\propto\sqrt{i(\theta)} \tag{4.4.8}$$

مثال 1.4.8

إذا كان X متغير عشوائي توزيعه:

$$f(x|\theta)=\theta^{x}(1-\theta)^{1-x} \quad ; \quad x=0,1 \ , \ 0<\theta<1$$

أوجد توزيع قبلي غير مُعلِم للمتغير Θ.

بما أن $0<\theta<1$، فحسب العلاقة (1.4.8):

$$g(\theta)=1 \quad ; \quad 0<\theta<1$$

يعتبر توزيع قبلي غير مُعلِم.

ويمكن الحصول على توزيع قبلي غير مُعلِم لـ Θ باستخدام معلومات فيشر [العلاقة (4.4.8)]:

$$\ln f(x|\theta) = x\ln\theta + (1-x)\ln(1-\theta)$$

$$\frac{\partial \ln f(x|\theta)}{\partial\theta} = \frac{x}{\theta} - \frac{1-x}{1-\theta}$$

$$\frac{\partial^2 \ln f(x|\theta)}{\partial\theta^2} = -\frac{x}{\theta^2} - \frac{1-x}{(1-\theta)^2}$$

$$i(\theta) = -E\left[\frac{\partial^2 \ln f(x|\theta)}{\partial\theta^2}\right] = \frac{1}{\theta(1-\theta)}$$

وعلى ذلك فإن:

$$g(\theta) \propto \frac{1}{\sqrt{\theta(1-\theta)}}$$

وهناك توزيع قبلي آخر غير مُعلِم وغير تام ويستخدم في حالة توزيع ذي الحدين وتوزيع بيرنولي، وهو:

$$g(\theta) \propto \frac{1}{\theta(1-\theta)} \qquad (5.4.8)$$

مثال 2.4.8

إذا كان ع متغير عشوائي توزيعه:

$$f(x|\theta) = C_n^x\theta^x(1-\theta)^{n-x} \quad ; \quad x = 0,1,\ldots,n \;,\; 0<\theta<1$$

أوجد توزيع قبلي غير مُعلِم للمتغير Θ.

بما أن $0<\theta<1$، فإن التوزيع القبلي المنتظم:

$$g(\theta) = 1 \quad ; \quad 0<\theta<1$$

يعتبر توزيع قبلي غير مُعلِم لـ Θ.

ويمكن الحصول على توزيع قبلي لـ Θ باستخدام معلومات فيشر:

$$\ln f(x|\theta) = \ln C_n^x + x\ln\theta + (n-x)\ln(1-\theta)$$

$$\frac{\partial \ln f(x|\theta)}{\partial\theta} = \frac{x}{\theta} - \frac{n-x}{1-\theta}$$

$$\frac{\partial^2 \ln f(x|\theta)}{\partial\theta^2} = -\left[\frac{x}{\theta^2} + \frac{n-x}{(1-\theta)^2}\right]$$

$$i(\theta) = -E\left[\frac{\partial^2 \ln f(x|\theta)}{\partial\theta^2}\right] = E\left[\frac{X}{\theta^2} + \frac{n-X}{(1-\theta)^2}\right] = \frac{n}{\theta} - \frac{n}{1-\theta} = \frac{n}{\theta(1-\theta)}$$

ومن ثم:

$$g(\theta) \propto \sqrt{\frac{n}{\theta(1-\theta)}}$$

مثـال 3.4.8

بفرض X متغير عشوائي توزيعه:

$$f(x|\theta) = \theta e^{-\theta x} \quad ; \quad x>0 \ , \ \theta>0$$

اقترح توزيع قبلي غير مُعلِم لـ Θ.

بما أن $0<\theta<\infty$، فإن:

$$g(\theta) \propto \frac{1}{\theta}$$

وباستخدام معلومات فيشر:

$$\ln f(x|\theta) = \ln\theta - \theta x$$

$$\frac{\partial \ln f(x|\theta)}{\partial\theta} = \frac{1}{\theta} - x \quad , \quad \frac{\partial^2 \ln f(x|\theta)}{\partial\theta^2} = -\frac{1}{\theta^2}$$

$$i(\theta) = -E\left[\frac{\partial^2 \ln f(x|\theta)}{\partial\theta^2}\right] = \frac{1}{\theta^2}$$

وعلى ذلـك:

$$g(\theta) \propto \frac{1}{\theta}$$

5.8 حالة عدة معالم

اقتصر بحثنا فيما سبق من هذا الفصل على حالة مجتمع توزيعه $f(x|\theta)$ يحتوي على معلمة غير معلومة وحيدة البعد (حقيقية). سندرس الآن الحالة التي تكون فيها θ متعددة الأبعاد ولتكن $k > 1$; $\theta = (\theta_1,...,\theta_k)$.

إذا كان اهتمامنا يتركز على إحدى مركبات θ (معلمة وحيدة البعد)، فإن المركبات الأخرى تعتبر معالم مزعجة أو مقلقة لنا ولا نحتاجها، وبالتالي نرغب في التخلص منها، وهذا ممكن بإيجاد التوزيع الهامشي للمتغير الذي يعنينا. فمثلاً، بفرض $\theta = (\theta_1,\theta_2)$، فالتوزيع البعدي المشترك لـ Θ_1,Θ_2 عند $X = x$ نرمز له بـ $g(\theta_1,\theta_1|x)$. وإذا كان اهتمامنا يقتصر على المعلمة θ_1، فإن المعلمة θ_2 تعتبر مقلقة أو مزعجة، ويمكن التخلص منها بالحصول على التوزيع البعدي الهامشي للمتغير Θ_1 على النحو الآتي:

$$g(\theta_1|x) = \begin{cases} \int g(\theta_1,\theta_2|x)d\theta_2 \\ \sum_{\theta_2} g(\theta_1,\theta_2|x) \end{cases}$$

مثال 1.5.8

إذا كانت $X = (X_1,...,X_n)$ عينة عشوائية من توزيع $N(\theta,\phi)$، وكان المتغيران Θ,Φ مستقلان، فأوجد التوزيع البعدي لكل من Θ,Φ.

بما أن المتغيرين Θ,Φ مستقلان، فيمكن اقتراح التوزيع القبلي لـــ (Θ,Φ) :

$$g(\theta,\phi) \propto \frac{d\theta d\phi}{\phi}$$

وكما نعلم دالة المعقولية:

$$L(x|\theta,\phi) = \left(\frac{1}{2\pi\phi}\right)^{n/2} e^{-\frac{\sum(x_i-\theta)^2}{2\phi}}$$

وعلى ذلك، فإن التوزيع البعدي لـ (Θ,Φ) يكون:

$$L(x|\theta,\phi) \propto \frac{e^{-\frac{\sum(x_i-\theta)^2}{2\phi}}}{\phi^{\frac{n}{2}+1}}$$

وحيث أن:

$$\sum_{i=1}^{n}(x_i-\theta) = \sum_{1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\theta)^2 = nS^2 + n(\bar{x}-\theta)^2 \ ; \ S^2 = \frac{1}{n}\sum_{1}^{n}(x_i-\bar{x})^2$$

فإن:

$$g(\theta,\phi|x) \propto \frac{e^{-\frac{n\theta^2+n(\bar{x}-\theta)^2}{2\phi}}}{\phi^{\frac{n}{2}+1}}$$

بمكاملة الطرف الأيمن بالنسبة لـ Φ نحصل على التوزيع البعدي الهامشي لـ Θ :

$$g(\theta|x) \propto \left[\frac{1}{ns^2+n(\theta-\bar{x})^2}\right]^{n/2} \propto 1\Big/\left[1+\frac{(\theta-\bar{x})^2}{s^2}\right]^{n/2}$$

وبوضع:

$$t = \frac{(\theta-\bar{x})}{s}\sqrt{n-1}$$

حيث t يتبع توزيع $S(n-1)$ [راجع المبرهنة (5.9.3)]، نجد:

$$g(\theta|x) \propto 1 \Big/ \left[1+\frac{t^2}{n-1}\right]^{n/2}$$

وبشكل مشابه نجد التوزيع البعدي الهامشي لـ Φ:

$$g(\phi|x) \propto \frac{e^{-\frac{ns^2}{2\phi}}}{\phi^{\frac{n}{2}+1}} \int_{-\infty}^{+\infty} e^{-\frac{n(\theta-\bar{x})^2}{2\phi}} d\theta \propto \frac{\sqrt{2\pi}\, e^{-\frac{ns^2}{2\phi}}}{\phi^{\frac{n}{2}+1}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\phi}} e^{-\frac{n(\theta-\bar{x})^2}{2\phi}} d\theta$$

وبملاحظة أن التكامل الأخير يساوي 1، نجد:

$$g(\phi|x) \propto \frac{e^{-\frac{ns^2}{2\phi}}}{\phi^{\frac{n+1}{2}}} = \left(\frac{1}{\phi}\right)^{\frac{n+1}{2}} e^{-\frac{ns^2}{2\phi}}$$

وهذا يعني أن $\tau(\theta,\phi)=\frac{1}{\theta}$ يتبع توزيع جاما $\Gamma\left(\frac{n+3}{2},\frac{2}{ns^2}\right)$، ومن ثم فإن:

$$\mathcal{L}\left(\frac{ns^2}{\Phi}\right)=\chi^2_{(n-1)}$$

مثال 2.5.8

إذا كانت $x=(x_1,\ldots,x_m)$ عينة عشوائية من توزيع:

$$f(x|\theta)=\theta_1 e^{-\theta_1 x} \quad ; \quad x>0$$

وكانت $y=(y_1,\ldots,y_n)$ عينة عشوائية من توزيع:

$$f(y|\theta)=\theta_2 e^{-\theta_2 y} \quad ; \quad y>0$$

أوجد توزيع قبلي لـ $\Theta=(\Theta_1,\Theta_2)$، ثم أوجد التوزيع البعدي لـ $\tau(\Theta)=\frac{\Theta_1}{\Theta_2}$.

بفرض أن التوزيعين القبليين لـــ Θ_1,Θ_2 مستقلان، فإن:

$$g(\theta_1,\theta_2)=g(\theta_1)g(\theta_2)\propto\frac{1}{\theta_1\theta_2}$$

ودالة المعقولية:

$$L(\theta_1,\theta_2|x,y)=\theta_1^m\theta_2^n e^{-\theta_1 t-\theta_2 u}\quad;\quad t=\sum_1^m x_i\ ,\ u=\sum_1^n y_i$$

وعلى ذلك فالتوزيع البعدي لـــ $\Theta=(\Theta_1,\Theta_2)$ هو:

$$g(\theta|x,y)\propto\theta_1^{m-1}\theta_2^{n-1}e^{-\theta_1 t-\theta_2 u}$$
$$g(\theta|x,y)\propto\left(\theta_1^{m-1}e^{-t\theta_1}\right)\left(\theta_2^{n-1}e^{-u\theta_2}\right)$$

وباستعمال حقيقة أن تكامل دالة كثافة توزيع جاما (20.3.2) من 0 إلى ∞ يســاوي الواحد، أي أن:

$$\int_0^\infty x^{\alpha-1}e^{-x/\beta}dx=\beta^\alpha\Gamma(\alpha)$$

نجد:

$$L(\theta_1,\theta_2|x,y)=\frac{t^m u^m}{\Gamma(m)\Gamma(n)}\theta_1^{m-1}\theta_2^{n-1}e^{-t\theta_1-u\theta_2}=g(\theta_1|x)g(\theta_2|y)$$

وهذا يعني أن التوزيعين البعديين لـــ Θ_1 و Θ_2 مستقلان، وأن:

$$\mathcal{L}(\Theta_1)=\Gamma\left(m,\frac{1}{t}\right)$$
$$\mathcal{L}(\Theta_2)=\Gamma\left(n,\frac{1}{u}\right)$$

وحسب المبرهنة (4.3.2) فإن:

$$\mathcal{L}\left(T=\sum_{1}^{m} X_i\right) \in \Gamma\left(m, \frac{1}{\theta_1}\right)$$
$$\mathcal{L}\left(T=\sum_{1}^{n} Y_i\right) \in \Gamma\left(n, \frac{1}{\theta_2}\right)$$

ومن ثم:

$$\mathcal{L}(2\Theta_1 T) = \chi^2_{2m} \quad , \quad \mathcal{L}(2\Theta_2 T) = \chi^2_{2n}$$

وبالتالي حسب المبرهنة (9.3.2) فإن:

$$L\left(\frac{2\theta_1 T/2m}{2\theta_2 U/2n} = \frac{nT}{mU}\frac{\Theta_1}{\Theta_2}\right) = F_{2m,2n}$$

6.8 تقدير بييز بنقطة BAYES POINT ESTIMATE

لاحظنا مما سبق أن التوزيع البعدي يحل محل دالة المعقولية، حيث يشتمل على المعلومات المتوفرة حول المعلمة θ قبل الحصول على عينة عشوائية $x=(x_1,\ldots,x_n)$ من مجتمع الصفة المدروسة ξ، بالإضافة إلى المعلومات التي تقدمها العينة x حول هذه المعلمة. وبالتالي لتقدير المعلمة θ يمكن استخدام نفس الأسلوب الذي أتّبع باستخدام دالة المعقولية $L(x|\theta)$ لتقدير θ. وبعبارة أخرى، يمكن تقدير المعلمة θ بالقيمة التي تعظم التوزيع البعدي، أي بمنوال التوزيع البعدي لـ Θ. كما يمكن استخدام أحد مقاييس النزعة المركزية الأخرى للتوزيع البعدي لـ Θ كالمتوسط والوسيط لتقدير المعلمة θ، ويستخدم عادة متوسط (القيمة المتوقعة) التوزيع البعدي لـ Θ كتقدير نقطة لـ θ. كما يمكن تقدير أي دالة معلمية $\tau(\theta)$ بالقيمة المتوقعة لـ $\tau(\Theta)$.

إذا رمزنا بـ θ^* و $\tau^*(\theta)$ لتقديري بييز بنقطة لـ θ و $\tau(\theta)$ على الترتيب، الموافقين للتوزيع القبلي $g(\theta)$، فإن:

$$\theta^* = E(\Theta|\boldsymbol{x}) = \begin{cases} \int \theta g(\theta|\boldsymbol{x})d\theta \\ \sum\limits_{\theta} \theta g(\theta|\boldsymbol{x}) \end{cases} \tag{1.6.8}$$

$$\tau^*(\theta) = E[\tau(\Theta)|\boldsymbol{x}] = \begin{cases} \int \tau(\theta) g(\theta|\boldsymbol{x})d\theta \\ \sum\limits_{\theta} \tau(\theta) g(\theta|\boldsymbol{x}) \end{cases} \tag{2.6.8}$$

نلاحظ وجود تشابه بين تقدير بييز بنقطة لـ θ وتقدير بتمان لمعلمة الموضع θ.

مثـــــال 1.6.8

لتكـن $\boldsymbol{x} = (x_1,\dots,x_n)$ عينـة عشـوائية مـن توزيـــع:

$$f(x|\theta) = \theta^x(1-\theta)^{1-x} \quad ; \quad x = 0,1$$

وأن التوزيع القبلي لـ Θ هو $R(0,1)$. أوجد تقدير بييز لـ θ و $\tau(\theta) = \theta(1-\theta)$.

بمـا أن:

$$f(\boldsymbol{x}|\theta) = \theta^t(1-\theta)^{n-t} \quad ; \quad t = \sum_{i=1}^{n} x_i$$

$$f(\boldsymbol{x}) = \int_0^1 \theta^t(1-\theta)^{n-t}d\theta = \beta(t+1, n-t+1)$$

فإن:

$$g(\theta|\boldsymbol{x}) = \frac{g(\theta)f(\boldsymbol{x}|\theta)}{f(\boldsymbol{x})} = \frac{1}{\beta(t+1,n-t+1)}\theta^t(1-\theta)^{n-t} \quad ; \quad 0 \le \theta \le 1$$

وعلـى ذلـك فـإن تقديـر بيـيز بنقطـة للمعلمـة θ الموافـق للتوزيـع القبلـــي $R(0,1)$ هو:

$$E\left(\Theta|x\right)=\frac{1}{\beta(t+1,n-t+1)}\int_0^1\theta^{t+1}(1-\theta)^{n-t}\,d\theta$$

$$=\frac{\beta(t+2,n-t+1)}{\beta(t+1,n-t+1)}=\frac{\Gamma(t+2)\Gamma(n-t+1)\Gamma(n+2)}{\Gamma(t+1)\Gamma(n-t+1)\Gamma(n+3)}$$

$$=\frac{t+1}{n+2}=\frac{\sum x_i+1}{n+2}$$

ومن ثم فمقدّر بييز لـ θ الموافق لتوزيع قبلي منتظم $R(0,1)$:

$$\theta^*(X)=E\left(\Theta|X\right)=\frac{\sum X_i+1}{n+2}$$

نلاحظ أن مقدّر المعقولية لـ θ هو $\sum X_i/n$ وهو مقدّر غير متحيز، بينما مقدّر بييز لـ θ متحيز.

للحصول على مقدّر بييز لـ $\tau(\theta)=\theta(1-\theta)$ نحسب:

$$E\left[\tau(\Theta)|x\right]=\frac{1}{\beta(t+1,n-t+1)}\int_0^1\theta^{t+1}(1-\theta)^{n-t+1}\,d\theta$$

$$=\frac{\beta(t+2,n-t+2)}{\beta(t+1,n-t+1)}=\frac{(t+1)(n-t+1)}{(n+3)(n+2)}=\frac{\left(\sum x_i+1\right)\left(n-\sum x_i+1\right)}{(n+3)(n+2)}$$

ومن ثم مقدّر بييز لـ $\theta(1-\theta)$ الموافق للتوزيع القبلي $R(0,1)$ هو:

$$\frac{\left(\sum X_i+1\right)\left(n-\sum X_i+1\right)}{(n+3)(n+2)}$$

مثال 2.6.8

إذا كانت $X=\left(X_1,\dots,X_n\right)$ عينة عشوائية من توزيع بيرنولي $B(1,\theta)$، وكان التوزيع القبلي لـ Θ هو $R(0,1)$، فأوجد مقدّر بييز للمعلمة θ.

بما أن التوزيع القبلي لـ Θ:

$$g(\theta) = 1 \quad ; \quad 0 < \theta < 1$$

فهـذا يعـني أن معلوماتنـا السـابقة حـول المعلمـة θ (قبـل الحصـول علـى عينـة عشـوائية) قليلـة جـداً، لـذا نفـترض أن جميـع قيـم Θ متسـاوية الإمكانيـة، حيـث لا يوجـد مـبرر لإعطاء بعض قيـم Θ أهميـة أكـبر مـن غيرهـا.

ويمكـن التـأكد بسـهولة أن التوزيـع البعـدي لـ Θ:

$$g(\theta|x) = \frac{1}{b(t+1, n-t+1)} \theta^{t} (1-\theta)^{n-t} \quad ; \quad 0 < \theta < 1$$

وبالتالي تقديـر بيـز لـ θ:

$$\begin{aligned}
\theta^{*} = E(\Theta|x) &= \frac{1}{\beta(t+1, n-t+1)} \int_{0}^{1} \theta^{t+1} (1-\theta)^{n-t} dt \\
&= \frac{\beta(t+2, n-t+1)}{\beta(t+1, n-t+1)} = \frac{\Gamma(t+2)\Gamma(n-t+1)\Gamma(n+2)}{\Gamma(n+3)\Gamma(t+1)\Gamma(n-t+1)} \\
&= \frac{t+1}{n+2} \quad ; \quad t = \sum_{1}^{n} x_i
\end{aligned}$$

أي أن مقـدِّر بيـز لـ θ:

$$\theta^{*}(X) = E(\Theta|X) = \left(1 + \sum_{1}^{n} X_i\right) \Big/ (n+2)$$

مثـــــال 3.6.8

لتكـن لدينـا معطيـات ونتـائج المثـال (2.2.8)، ونريـد إيجـاد تقديـر بيـز بنقطـة للمعلمـة θ الموافـق للتوزيـع القبلـــي $\Gamma\left(2, \frac{1}{2}\right)$.

بمـا أن:

$$g(\theta|x) = \frac{(n+3)^{t+2}}{\Gamma(t+2)} \theta^{t+1} e^{-(n+1)\theta} \quad ; \quad \theta > 0$$

فإن تقدير بييز بنقطة للمعلمة θ:

$$\theta^* = E(\Theta|x) = \frac{(n+3)^{t+2}}{\Gamma(t+2)}\int_0^\infty \theta^{t+2}\, e^{-(n+1)\theta} d\theta$$

$$= \frac{(n+3)^{t+2}}{\Gamma(t+2)}\frac{\Gamma(t+3)}{(n+3)^{t+3}} = \frac{t+2}{n+3} \qquad ; \qquad t = \sum_{i=1}^{n} x_i$$

لأن:

$$\int_0^\infty \theta^{t+2} e^{-(n+1)\theta} d\theta = \frac{\Gamma(t+3)}{(n+3)^{t+3}}$$

وبناءً على تعريف دالة كثافة توزيع جاما وخواص دالة الكثافة [العلاقة (20.3.2)]. ومن ثم فمقدِّر بييز للمعلمة θ:

$$\theta^*(X) = \frac{T+2}{n+3} \qquad ; \qquad T = \sum_{i=1}^{n} X_i$$

مثال 4.6.8

إذا كانت لدينا معطيات ونتائج المثال (3.2.8)، فأوجد مقدِّر بييز للمعلمة θ، الموافق لتوزيع قبلي $\Gamma\left(a+1, \frac{1}{b}\right)$:

بما أن:

$$g(\theta|x) = \frac{(b+t)^{a+n+1}}{\Gamma(a+n+1)}\theta^{a+n} e^{-(b+t)\theta} \qquad ; \qquad \theta > 0$$

فإن:

$$\theta^* = E(\Theta|x) = \frac{(b+t)^{a+n+1}}{\Gamma(a+n+1)}\int_0^\infty \theta^{a+n+1} e^{-(b+t)\theta} d\theta$$

$$= \frac{(b+t)^{a+n+1}}{\Gamma(a+n+1)}\frac{\Gamma(a+n+2)}{(b+t)^{a+n+2}} = \frac{a+n+1}{b+t} \qquad ; \qquad t = -\sum_{i=1}^{n} \ln x_i$$

وبالتالي مقدر بييز لـ θ:

$$\theta^*(X)=\frac{a+n+1}{b+T} \quad ; \quad T=-\sum_{i=1}^{n}\ln X_i$$

مثــــال 5.6.8

إذا كانت $x=(x_1,\ldots,x_n)$ عينة عشوائية من توزيع $N(\theta,\sigma^2)$، وكـــان التوزيــع القبلي للمتغير Θ هو $N(a,b^2)$، حيث a,b^2 مقادير ثابتة ومعلومــة، فــأوجد تقديــر بييز للمعلمة θ.

كما نعلم $\overline{X}$ إحصاء كافي للمعلمة θ وأن $\mathcal{L}_\theta(\overline{X})\in N\left(\theta,\frac{\sigma^2}{n}\right)$. وعلــى ذلــك فإن:

$$\begin{aligned}
g(\theta|x)&=g(\theta)h(\overline{x}|\theta)\propto e^{-\frac{n(\overline{x}-\theta)^2}{2\sigma^2}-\frac{(\theta-a)^2}{2b^2}}\\
&=e^{-\frac{1}{2b^2\sigma^2}\left[nb^2(\overline{x}-\theta)^2+\sigma^2(\theta-a)^2\right]}\\
&=e^{-\frac{nb^2+\sigma^2}{2b^2\sigma^2}\left[\theta^2-2\frac{nb^2\overline{x}+a\sigma^2}{nb^2+\sigma^2}\theta+\frac{nb^2\overline{x}^2+\sigma^2a^2}{nb^2+\sigma^2}\right]}\\
&\propto e^{-\frac{nb^2+\sigma^2}{2b^2\sigma^2}\left(\theta-\frac{nb^2\overline{x}+a\sigma^2}{nb^2+\sigma^2}\right)^2}
\end{aligned}$$

وبالتالي:

$$\begin{aligned}
L(\Theta|x)&=N\left(\frac{nb^2\overline{x}+a\sigma^2}{nb^2+\sigma^2},\frac{b^2\sigma^2}{nb^2+\sigma^2}\right)\\
&=N\left(\frac{b^2\overline{x}+(\sigma^2/n)a}{b^2+(\sigma^2/n)},\frac{b^2(\sigma^2/n)}{b^2+(\sigma^2/n)}\right)\\
&=N(\theta^*,\sigma^{*2})
\end{aligned}$$

وعلى ذلك فإن تقدير بييز للمعلمة θ

$$\theta^* = \frac{b^2\overline{x} + \left(\sigma^2/n\right)a}{b^2 + \left(\sigma^2/n\right)}$$

الموافق للتوزيع القبلي $N(a,b^2)$.

إذا كانت:

1. $b^2 = \infty$، فإن $\theta^* = \overline{x}$ وهو نفس تقدير المعقولية العظمى. وهذا يعني أن المعلومات المتوفرة حول θ قبل سحب العينة قليلة جداً، وأن المعلومات مصدرها الأساسي معطيات العينة x.

2. $b^2 = 0$، فإن $\theta^* = a$. وهذا يعني مهما تكن العينة x فإن $\theta = a$ باحتمال يساوي الواحد.

مثال 6.6.8

إذا كانت x ملاحظة مفردة على متغير عشوائي ξ يخضع لتوزيع $B(n,\theta)$، وكان التوزيع القبلي لـ Θ هو $Be(\alpha,\beta)$، فأوجد تقدير بييز للمعلمة θ.

بما أن:

$$f\left(x|\theta\right) = C_n^x\theta^x(1-\theta)^{n-x} \quad ; \quad x = 0,1,\ldots,n$$

$$g(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} \quad ; \quad 0<\theta<1$$

ومن ثم:

$$f\left(x;\theta\right) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}C_n^x\theta^x(1-\theta)^{n-x}$$

$$= C_n^x\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}$$

$$f(x) = \int_0^1 f(x,\theta)d\theta = C_n^x\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1\theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}d\theta$$

$$= C_n^x \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+x)\Gamma(n-x+\beta)}{\Gamma(n+\alpha+\beta)} \quad ; \quad x = 0,1,\ldots,n$$

وعلى ذلـك:

$$g(\theta|x) = \frac{f(x;\theta)}{f(x)} = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(\alpha+x)\Gamma(n-x+\beta)} \theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1} \quad ; \quad 0<\theta<1$$

$$\theta^* = E(\theta|x) = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(\alpha+x)\Gamma(n-x+\beta)} \int_0^1 \theta^{x+\alpha}(1-\theta)^{n-x+\beta-1} d\theta$$

$$= \frac{\Gamma(n+\alpha+\beta)}{\Gamma(\alpha+x)\Gamma(n-x+\beta)} \frac{\Gamma(x+\alpha+1)\Gamma(n-x+\beta)}{\Gamma(n+\alpha+\beta+1)}$$

$$= \frac{x+\alpha}{\alpha+\beta+n}$$

7.8 تقدير بييز بفترة BAYES INTERVAL ESTIMATE

رأينا في البند السابق (6.8) أن تقدير بييز بنقطة لمعلمة θ هو عبارة عن قيمة وحيدة، وعادة تكون القيمة المتوقعة $E(\Theta|x)$ ، لكن لم نرفق مثل هذا التقدير بأي مقياس يعبر عن درجة الثقة التي يتمتع بها. لذا في حالات عدة من المفيد استبدال التقدير بنقطة للمعلمة θ بمجموعة من النقاط أي بفترة، التي تشتمل على القيمة الحقيقية للمعلمة θ باحتمال كبير γ. وكما أشرنا سابقاً فإن هذه الفترة تسمى بفترة ثقة لـ θ والاحتمال γ بمعامل الثقة (درجة الثقة).

إن طريقة بييز في بناء فترات ثقة تتمثل بما يلي:

لتكن $x=(x_1,\ldots,x_n)$ عينة عشوائية من توزيع $f(x|\theta)$ ، الذي يعتمد على معلمة مجهولة، و $g(\theta)$, $g(\theta|x)$ التوزيع القبلي والبعدي على الترتيب للمتغير Θ.

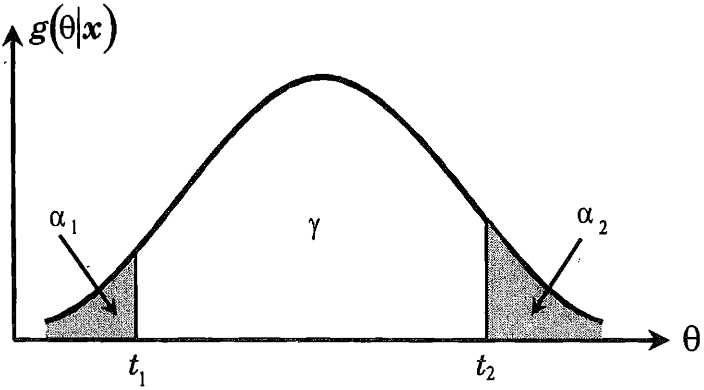
لتقدير المعلمة θ بفترة يتم البحث عن عددين:

$$t_1 = T_1(\boldsymbol{x}) \quad , \quad t_2 = T_2(\boldsymbol{x}) \quad ; \quad t_1 < t_2$$

يحققان، عند قيمة معينة مسبقاً γ، الشرط الآتي:

$$\int_{t_1}^{t_2} g(\theta|\boldsymbol{x})d\theta = \gamma \Leftrightarrow P(t_1 < \Theta < t_2) = \gamma \qquad (1.7.8)$$

تدعى الفترة (t_1,t_2) بـ $\gamma 100\%$ فترة ثقة بييز للمعلمة θ، وتبدو بوضوح على الشكل (1.7.8).



شكل (1.7.8)

نلاحظ أن الزوج (t_1,t_2) المحقق للعلاقة (1.7.8) ليس وحيداً، بل توجد أزواج عدة تحقق تلك العلاقة. ونفضل دائماً فترة الثقة (t_1,t_2) ذات أصغر طول، أي نبحث عن فترة الثقة (t_1,t_2) الموافقة لأصغر فرق $(t_2 - t_1)$ عند الشرط (1.7.8).

مثال 7.6.8

لتكن لدينا معطيات ونتائج المثال (5.6.8)، ونريد بناء فترة بييز بمعامل ثقة $\gamma = 1 - \alpha$.

بما أن:

$$g(\theta|x)=\frac{1}{\sigma^{*}\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\theta-\theta^{*}}{\sigma^{*}}\right)^{2}} \quad ; \quad -\infty<\theta<+\infty$$

حيث:

$$\theta^{*}=\frac{b^{2}\bar{x}+\left(\sigma^{2}/n\right)a}{b^{2}+\left(\sigma^{2}/n\right)} \quad , \quad \sigma^{*}=\frac{b^{2}\left(\sigma^{2}/n\right)}{b^{2}+\left(\sigma^{2}/n\right)}$$

وهذا يعني أن:

$$\mathcal{L}\left(\frac{\Theta-\theta^{*}}{\sigma^{*}}\right)=N(0,1)$$

ومن ثم البحث عن أقصر فترة (t_1,t_2)، بحيث:

$$P\left(t_1<\Theta<t_2\right)=\gamma$$

تكافئ البحث عن أقصر فترة (z_1,z_2) بحيث:

$$P\left(z_1<\frac{\Theta-\theta^{*}}{\sigma^{*}}<z_2\right)=\gamma$$

لكن كما نعلم [راجع العلاقة (4.4.7)] أقصر فترة من الشكل (z_1,z_2) هي التي تقابل ذيلين متساويي الاحتمال، أي أن:

$$z_1=-z_{\alpha/2} \quad , \quad z_2=z_{1-\alpha/2}=z_{\alpha/2}$$

وعلى ذلك فإن $\gamma 100\%$ فترة ثقة بييز للمعلمة θ هي:

$$\left(\theta^{*}-\sigma^{*}z_{\alpha/2}<\theta<\theta^{*}+\sigma^{*}z_{\alpha/2}\right)$$

مثـــال 8.6.8

مصنع لإنتاج المصابيح الكهربائية، عمر المصباح ع مقدر بساعة إضاءة يتبع تقريباً التوزيع الطبيعي بمتوسط غير معلوم θ وانحراف معياري يساوي 100 ساعة إضاءة. تشير الخبرة الماضية إلى أن θ قيمة لمتغير عشوائي طبيعي Θ بمتوسط 800 ساعة إضاءة وانحراف معياري 10 ساعات إضاءة. أخذت عينة عشوائية من 25 مصباحاً من إنتاج المصنع، فتبين أن متوسط عمر مصابيح هذه العينة يساوي 780 ساعة إضاءة. أوجد 95% فترة ثقة بييز للمعلمة θ.

بما أن، حسب الفرض:

$$f(x|\theta) = \frac{1}{(2\pi)^{25/2}100^{25}} \exp\left[-\frac{1}{2}\sum_{i=1}^{25}\left(\frac{x_i-\theta}{100}\right)^2\right] \quad ; \quad -\infty < x_i < \infty \ , \ i = 1,\ldots,25$$

$$g(\theta) = \frac{1}{10\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\theta-800}{10}\right)^2\right] \quad ; \quad -\infty < \theta < \infty$$

فإن:

$$\begin{aligned} f(x,\theta) &= g(\theta) f(x|\theta) = \\ &= \frac{1}{(2\pi)^{13}10^{51}} \exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{25}\left(\frac{x_i-\theta}{100}\right)^2 + \left(\frac{\theta-800}{10}\right)^2\right]\right\} \\ &= \frac{1}{(2\pi)^{13}10^{51}} \exp\left[-\frac{1}{2}\sum_{i=1}^{25}\left(\frac{x_i-780}{100}\right)^2\right] \times \\ &\quad \times \exp\left\{-\frac{1}{2}\left[(25)\left(\frac{780-\theta}{100}\right)^2\right] + \left(\frac{\theta-800}{10}\right)^2\right\} \end{aligned}$$

لأن:

$$\sum_{i=1}^{25}(x_i-\theta)^2 = \sum_{i=1}^{25}(x_i-780)^2 + 25(780-\theta)^2$$

وباستخدام طريقة الإتمام إلى مربع كامل نجد:

$$25\left(\frac{780-\theta}{100}\right)^2+\left(\frac{\theta-800}{10}\right)^2=\frac{\theta^2-1592\theta+635.28}{80}=\frac{(\theta-796)^2+1664}{80}$$

وبالتالي:

$$f(x;\theta)=k\exp\left[-\frac{1}{2}\left(\frac{\theta-796}{\sqrt{80}}\right)^2\right]$$

حيث k دالة في قيم العينة.

وعلى ذلك:

$$f(x)=\int_{-\infty}^{+\infty}f(x,\theta)d\theta=k\sqrt{2\pi}\sqrt{80}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sqrt{80}}e^{-\frac{1}{2}[(\theta-796)/\sqrt{80}]^2}d\theta=k\sqrt{2\pi}\sqrt{80}$$

ومن ثم التوزيع البعدي لـ Θ يكون:

$$g(\theta|x)=\frac{f(x,\theta)}{f(x)}=\frac{1}{\sqrt{2\pi}\sqrt{80}}\exp\left[-\frac{1}{2}\left(\frac{\theta-796}{80}\right)^2\right]\quad;\quad-\infty<\theta<\infty$$

وهذا يعني أن التوزيع البعدي لـ θ طبيعي بمتوسط $\theta^*=796$ وانحراف معياري $\sigma^*=\sqrt{80}$، و 95% فترة ثقة بييز تعطى على النحو الآتي:

$$\theta^*-1.96\sigma^*<\theta<\theta^*+1.96\sigma^*$$

أي أن:

$$796-1.96\sqrt{80}<\theta<796+1.96\sqrt{80}$$

أو:

$$778.5<\theta<813:5$$

نحصل على نفس النتيجة باستخدام نتيجة المثال السابق، ووضع:

$$a=\theta^{*}=796 \quad , \quad b^{2}=\sigma^{*2}=80 \quad , \quad \gamma=0.95 \quad , \quad z_{\alpha/2}=z_{0.025}=1.96$$

مثـــال 9.7.8

إذا كانت $x=(x_1,\ldots,x_n)$ عينة عشوائية من توزيع:

$$f(x|\theta)=\theta x^{\theta-1} \quad ; \quad 0\le x\le 1$$

وكان التوزيع القبلي لـ Θ هو $g(\theta)\propto\frac{1}{\theta}$، فأوجد $\gamma 100\%$ فترة ثقة بييز للمعلمة θ.

بما أن $T(X)=-\sum_{i=1}^{n}\ln X_i$ إحصاء كافي للمعلمة θ وأن:

$$h(t|\theta)=\frac{\theta^{n}}{\Gamma(n)}t^{n-1}e^{-\theta t} \quad ; \quad t>0$$

[راجع المثال (3.2.4)]، فإن:

$$g(\theta|t)\propto g(\theta)h(t|\theta)\propto\theta^{n-1}e^{-t\theta}$$

وعلى ذلك:

$$g(\theta|t)=\frac{t^{n}}{\Gamma(n)}\theta^{n-1}e^{-t\theta} \quad ; \quad \theta>0$$

ومن ثم يمكن التأكد بسهولة أن توزيع $\chi^{2}=2\Theta T$ يخضع لتوزيع $\chi^{2}_{(2n)}$.

الآن يمكن إيجاد قيمتين $\chi^{2}_{1-\alpha_1},\chi^{2}_{\alpha_2}\ ;\ \alpha_1+\alpha_2=1-\gamma$ بحيث يحققان الشرط:

$$P\left(\chi^{2}_{1-\alpha_1,2n}<2\theta t<\chi^{2}_{\alpha_2,2n}\right)=0.95$$

ومن ثم:

$$P\left(\frac{\chi^2_{1-\alpha_1,2n}}{2t} < \theta < \frac{\chi^2_{\alpha_2,2n}}{2t}\right)$$

وكما نعلم أقصر فترة هي الموافقة لـ $\alpha_1,\alpha_2 = 0.025$. وبذلك تكون 95% فترة ثقة بييز للمعلمة θ هي من الشكل:

$$\left(\frac{\chi^2_{1-\alpha/2,2n}}{2t},\frac{\chi^2_{\alpha/2,2n}}{2t}\right)$$

فمثلاً، إذا كان:

$$n = 10 \quad , \quad t = -\sum_{1}^{10} \ln x_i = 8 \quad , \quad \gamma = 0.95$$

فإن 95% فترة ثقة بييز للمعلمة θ هي:

$$\left(\frac{9.591}{16},\frac{34.17}{16}\right) = (0.6,2.1)$$

تمارين

1. إذا كانت $x = (x_1,\ldots,x_n)$ عينة عشوائية من توزيع:

$$f(x|\theta) = e^{-(x-\theta)} \quad ; \quad x \geq \theta\,,\ -\infty < \theta < \infty$$

وكان التوزيع القبلي لـ Θ هو:

$$g(\theta) = e^{-\theta} \quad ; \quad \theta > 0$$

فأوجد:

1. التوزيع البعدي لـ Θ.
2. تقدير بييز بنقطة لـ θ الموافق للتوزيع القبلي $g(\theta)$.

.2 لتكن $x=(x_1,\dots,x_n)$ عينة عشوائية من توزيع:

$$f(x|\theta)=\theta x^{\theta-1} \quad ; \quad 0<x<1 \ , \ \theta>0$$

بافتراض أن التوزيع القبلي لـ Θ هو:

$$g(\theta)=\frac{\lambda^r}{\Gamma(r)}\theta^{r-1}e^{-\theta\lambda} \quad ; \quad \theta>0$$

حيث أن r و λ مقدارين ثابتين معلومين.

1. ما هو التوزيع البعدي لـ Θ.
2. أوجد مقدِّر بييز للمعلمة θ الموافق للتوزيع القبلي المعطى $g(\theta)$.

.3 إذا كانت $x=(x_1,\dots,x_n)$ عينة عشوائية من توزيع $N(\theta,1)$ وكان التوزيع القبلي لـ Θ هو $N(\mu,1)$، فأوجد:

1. التوزيع البعدي لـ Θ.
2. مقدر بييز للمعلمة θ الموافق للتوزيع القبلي $N(\mu,1)$.

.4 بفرض أن $x=(x_1,\dots,x_n)$ عينة عشوائية من توزيع $R(1,\theta)$، والتوزيع القبلي لـ Θ هو $R(0,1)$، فأوجد التوزيع البعدي لـ Θ ومن ثم مقدر بييز للمعلمة الموافق للتوزيع القبلي $R(0,1)$.

.5 إذا كانت $x=(x_1,\dots,x_n)$ عينة عشوائية من توزيع :

$$f(x|\theta)=\frac{1}{\theta}e^{-\theta x} \quad ; \quad x>0$$

فأوجد توزيع قبلي مرافق لـ Θ، ومن ثم أوجد التوزيع البعدي لـ θ.

.6 إذا كانت $x=(x_1,\dots,x_n)$ عينة عشوائية من توزيع $N(\theta_1,\sigma_1^2)$ و $y=(y_1,\dots,y_n)$ عينة عشوائية من توزيع $N(\theta_2,\sigma_2^2)$، وكانت العينتان

مستقلتان، وكان التوزيع البعدي لكل من Θ_1 و Θ_2 منتظم على الفترة $(-\infty,\infty)$ فأوجد:

1. التوزيع البعدي لـ $\Theta_1 - \Theta_2$.
2. مقدر بييز للمعلمة $\theta_1 - \theta_2$.
3. فترة بييز للفرق $\theta_1 - \theta_2$ بمعامل ثقة γ.

7. إذا كانت لدينا معطيات ونتائج المثال (3.4.8) فأوجد $100\gamma\%$ فترة ثقة للمعلمة θ.

8. إذا كانت لدينا معطيات ونتائج المثال (2.5.8) فأوجد $100\gamma\%$ فترة ثقة بييز للمعلمة $\theta_1 - \theta_2$.

9. إذا كانت $x = (x_1,\dots,x_n)$ عينة عشوائية من توزيع:

$$f(x|\theta) = \theta(1-\theta)^{x-1} \quad , \quad x = 1,2,\dots.$$

وكان التوزيع القبلي لـ Θ هو $R(0,1)$، فأوجد 95% فترة ثقة بييز للمعلمة θ.

10. إذا كانت لدينا معطيات التمرين (4)، فأوجد 90% فترة ثقة بييز للمعلمة θ.

11. إذا كانت لدينا معطيات التمرين (5)، فأوجد 95% فترة ثقة بييز للمعلمة θ.

المراجــع العلميــة

أولاً: المراجع باللغة العربيــة

1. د. انيس كنجو، **الإحصاء الرياضي**، مطبعة زيـــد ثـابت، 1979.

2. د. أنيس كنجو، **الإحصاء وطرق تطبيقه في ميـــادين البحــث العلمــي**، الجــزء الأول والثاني، مؤسسة الرســالة، بــيروت، 1987.

3. د. جلال مصطفـــى الصيــاد، **الاســتدلال الإحصــائي**، دار المريــخ للنشــر، الرياض، المملكة العربية الســعودية، 1993.

4. د. عبدالحفيظ محمد فوزي مصطفى، **نظريـــة الاحتمــالات (I ، II)**، أكاديميــة الدراسات والعلوم الجوية، مصراتــه، ليبيــا، 1997.

ثانياً: المراجع باللغة الإنجليزيـــة

1. Alexander, M. M.; Franklin, A. G. and Duanc, C. B., *Introduction to The Theory of Statistics*, third edition, 1974.

2. Beaumout, G. P., *Intermediate Mathematical Statistics*, Champan and Hall, London and New York, 1980.

3. Hogg, R. V. and Craing, A. T., *Introduction to Mathematical Statistics*, Collier McMillan publisher, 1978.

4. Lymann, O., *An Introduction to Statistical Methods and Data Analysis*, Second edition, Duxbury press, Boston, USA, 1984.

5. Rohntgi, V. K., *An Introduction to Probability and Mathematical Statistics*, John Wiley and Sons, 1976.

6. Zacks, S., *Parametric Statistical Inference*, Pergamon press, 1981.

ثالثاً: المراجع باللغة الروسية

1. جرسيموفيتش، أ. ن.، **الإحصاء الرياضي**، الطبعة الثانية، العلوم، موسكو، 1983.

2. إيفتشينكو، ج. إ.، وميدفيديف، يو. إز، **الإحصاء الرياضي**، المدرسة العليا، موسكو، 1984

3. جمورمان، ب. إ.، **نظرية الاحتمالات والإحصاء الرياضي**، المدرسة العليا، موسكو، 1985.

4. كورووليوك، هـ. ن.، **نظرية الاحتمالات والإحصاء الرياضي**، العلوم، موسكو، 1985.

5. بوروفكوف، أ. أ.، **الإحصاء الرياضي: تقدير المعالم واختبار الفرضيات**، العلوم، موسكو، 1984.

6. كليموف، ج. ب.، **نظرية الاحتمالات والإحصاء الرياضي**، جامعة موسكو، 1980.

7. كنيدينكا، ب. ف.، **نظرية الاحتمالات**، العلوم، موسكو، 1988.

ملحق الجداول الإحصائية

The Binomial Distribution جدول 1: توزيع ذي الحدين

جدول قيم الدالة: $P_n(m) = C_n^m \; P^n \; q^{n-m}$

							p						
n	m	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	
1	0	.9900	.9500	.9000	.8500	.8000	.7500	.7000	.6500	.6000	.5500	.5000	1
	1	.0100	.0500	.1000	.1500	.2000	.2500	.3000	.3500	.4000	.4500	.5000	0
2	0	.9801	.9025	.8100	.7225	.6400	.5625	.4900	.4225	.3600	.3025	.2500	2
	1	.0198	.0950	.1800	.2550	.3200	.3750	.4200	.4550	.4800	.4950	.5000	1
	2	.0001	.0025	.0100	.0225	.0400	.0625	.0900	.1225	.1600	.2025	.2500	0
3	0	.9703	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664	.1250	3
	1	.0294	.1354	.2430	.3251	.3840	.4219	.4410	.4436	.4320	.4084	.3750	2
	2	.0003	.0071	.0270	.0574	.0960	.1406	.1890	.2389	.2880	.3341	.3750	1
	3	.0000	.0001	.0010	.0034	.0080	.0156	.0270	.0429	.0640	.0911	.1250	0
4	0	.9606	.8145	.6561	.5220	.4096	.3164	.2401	.1785	.1296	.0915	.0625	4
	1	.0388	.1715	.2916	.3685	.4096	.4219	.4116	.3845	.3456	.2995	.2500	3
	2	.0006	.0135	.0486	.0975	.1536	.2109	.2646	.3105	.3456	.3676	.3750	2
	3	.0000	.0005	.0638	.0115	.0256	.0469	.0756	.1115	.1536	.2005	.2500	1
	4	.0000	.0000	.0001	.0005	.0016	.0039	.0081	.0150	.0256	.0410	.0625	0
5	0	.9510	.7738	.5905	.4437	.3277	.2373	.1681	.1160	.0778	.0503	.0312	5
	1	.0480	.2036	.3280	.3915	.4096	.3955	.3602	.3124	.2592	.2069	.1562	4
	2	.0010	.0214	.0729	.1382	.2048	.2637	.3087	.3364	.3456	.3369	.3125	3
	3	.0000	.0011	.0081	.0244	.0512	.0879	.1323	.1811	.2304	.2757	.3126	2
	4	.0000	.0000	.0004	.0022	.0064	.0146	.0284	.0488	.0768	.1128	.1562	1
	5	.0000	.0000	.0000	.0001	.0003	.0010	.0024	.0053	.0102	.0185	.0312	0
6	0	.9415	.7351	.5314	.3771	.2621	.1780	.1176	.0754	.0467	.0277	.0156	6
	1	.0571	.2321	.3543	.3993	.3932	.3560	.3025	.2437	.1866	.1359	.0938	5
	2	.0014	.0305	.0984	.1762	.2458	.2966	.3241	.3280	.3110	.2780	.2344	4
	3	.0000	.0021	.0146	.0415	.0819	.1318	.1852	.2355	.2765	.3032	.3126	3
	4	.0000	.0001	.0012	.0055	.0154	.0330	.0595	.0951	.1382	.1861	.2344	2
	5	.0000	.0000	.0001	.0004	.0015	.0044	.0102	.0205	.0369	.0609	.0938	1
	6	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0018	.0041	.0083	.0156	0
7	0	.9321	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	.0078	7
	1	.0659	.2573	.3720	.3960	.3670	.3115	.2471	.1848	.1306	.0872	.0647	6
	2	.0020	.0406	.1240	.2097	.2753	.3115	.3177	.2985	.2613	.2140	.1641	5
	3	.0000	.0036	.0230	.0617	.1147	.1730	.2269	.2679	.2903	.2918	.2734	4
	4	.0000	.0002	.0026	.0109	.0287	.0577	.0972	.1542	.1935	.2388	.2734	3
	5	.0000	.0000	.0002	.0012	.0043	.0115	.0250	.0466	.0774	.1172	.1641	2
		.99	.95	.90	.85	.80	.75	.70	.65	.60	.55	.50	
							p						

تابع الجدول 1

n	m	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	
							p						
7	6	.0000	.0000	.0000	.0001	.0004	.0013	.0036	.0084	.0172	.0320	.0547	1
	7	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0016	.0037	.0078	0
8	0	.9227	.6634	.4305	.2725	.1678	.1002	.0576	.0319	.0168	.0084	.0039	8
	1	.0746	.2793	.3826	.3847	.3355	.2670	.1977	.1373	.0896	.0548	.0312	7
	2	.0026	.0515	.1488	.2376	.2936	.3115	.2965	.2587	.2090	.1569	.1094	6
	3	.0001	.0054	.0331	.0839	.1468	.2076	.2541	.2786	.2787	.2568	.2188	5
	4	.0000	.0004	.0046	.0185	.0459	.0865	.1361	.1875	.2322	.2627	.2734	4
	5	.0000	.0000	.0004	.0026	.0092	.0231	.0467	.0808	.1239	.1719	.2188	3
	6	.0000	.0000	.0000	.0002	.0011	.0038	.0100	.0217	.0413	.0403	.1094	2
	7	.0000	.0000	.0000	.0000	.0001	.0004	.0012	.0033	.0079	.0164	.0312	1
	8	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0017	.0039	0
9	0	.9135	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020	9
	1	.0830	.2985	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176	8
	2	.0034	.0629	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703	7
	3	.0001	.0077	.0446	.1069	.1762	.2336	.2668	.2716	.2508	.2119	.1641	6
	4	.0000	.0006	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461	5
	5	.0000	.0000	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2691	4
	6	.0000	.0000	.0001	.0006	.0028	.0087	.0210	.0424	.0743	.1160	.1641	3
	7	.0000	.0000	.0000	.0000	.0003	.0012	.0039	.0098	.0212	.0407	.0703	2
	8	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0013	.0035	.0083	.0176	1
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0020	0
10	0	.9044	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010	10
	1	.0914	.3151	3874	.3474	.2684	.1877	.1211	.0725	.0403	.0207	.0098	9
	2	.0042	.0746	.1937	.2759	.3020	.2816	.2336	.1757	.1209	.0763	.0439	8
	3	.0001	.0105	.0574	.1298	.2013	.2503	.2668	.2522	.2150	.1665	.1172	7
	4	.0000	.0010	.0112	.0401	.0881	.1460	.2001	.2377	.2508	.2384	.2051	6
	5	.0000	.0001	.0015	.0085	.0264	.0584	.1029	.1536	.2007	.2340	.2461	5
	6	.0000	.0000	.0001	.0012	.0055	.0162	.0368	.0689	.1115	.1596	.2051	4
	7	.0000	.0000	.0000	.0001	.0008	.0031	.0090	.0212	.0425	.0746	.1172	3
	8	.0000	.0000	.0000	.0000	.0001	.0004	.0014	.0043	.0106	.0229	.0439	2
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016	.0042	.0098	1
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	0
11	0	.8953	.5688	.3138	.1673	.0859	.0422	.0198	.0088	.0036	.0014	.0005	11
	1	.0995	.3293	.3836	.3248	.2362	.1549	.0932	.0518	.0266	.0125	.0064	10
	2	.0050	0876	.2131	.2866	.2953	.2581	.1998	.1395	.0887	.0513	.0269	9
	3	.0002	.0137	.0710	.1517	.2215	.2581	.2588	.2254	.1774	.1259	.0606	8
	4	.0000	.0014	.0158	.0536	.1107	.1721	.2201	.2428	.2365	.2060	.1611	7
		.99	.95	.90	.85	.80	.75	.70	.65	.60	.55	.50	
							p						

p

n	m	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	
11	5	.0000	.0001	.0025	.0132	.0388	.0803	.1321	.1830	.2207	.2360	.2256	6
	6	.0000	.0000	.0003	.0023	.0097	.0268	.0566	.0985	.1471	.1931	.2256	5
	7	.0000	.0000	.0000	.0003	.0017	.0064	.0173	.0379	.0701	.1128	.1611	4
	8	.0000	.0000	.0000	.0000	.0002	.0011	.0037	.0102	.0234	.0462	.0806	3
	9	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018	.0052	.0126	.0269	2
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0007	.0021	.0054	1
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0005	0
12	0	.8864	.5404	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008	.0002	12
	1	.1074	.3413	.3766	.3012	.2062	.1267	.0712	.0368	.0174	.0075	.0029	11
	2	.0060	.0988	.2301	.2924	.2835	.2323	.1678	.1088	.0639	.0339	.0161	10
	3	.0002	.0173	.0852	.1720	.2362	.2581	.2397	.1954	.1419	.0923	.0537	9
	4	.0000	.0021	.0213	.0683	.1329	.1936	.2311	.2367	.2128	.1700	.1208	8
	5	.0000	.0002	.0038	.0193	.0532	.1032	1585	.2039	.2270	.2225	.1934	7
	6	.0000	.0000	.0005	.0040	.0155	.0401	.0792	.1281	.1766	.2124	.2256	6
	7	.0000	.0000	.0000	.0006	.0033	.0115	.0291	.0591	.1009	.1489	.1934	5
	8	.0000	.0000	.0000	.0001	.0005	.0024	.0078	.0199	.0420	0.762	.1208	4
	9	.0000	.0000	.0000	.0000	.0001	.0004	.0015	.0048	.0125	.0277	.0637	3
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0025	.0068	.0161	2
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0029	1
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	0
13	0	.8775	.5133	.2542	.1209	.0550	.0238	.0097	.0037	.0013	.0004	.0001	13
	1	.1152	.3512	.3672	.2774	.1787	.1029	.0540	.0259	.0113	.0045	.0016	12
	2	.0070	.1109	.2448	.2937	.2680	.2059	.1388	.0836	.0453	.0220	.0095	11
	3	.0003	.0214	.0997	.1900	.2457	.2517	.2181	.1651	.1107	.0660	.0349	10
	4	.0000	.0028	.0277	.0838	1535	.2097	.2337	.2222	.1845	.1350	.0873	9
	5	.0000	.0003	.0055	.0266	.0691	.1258	.1803	.2154	.2214	.1989	.1571	8
	6	.0000	.0000	.0008	.0063	.0230	.0559	.1030	.1546	.1968	.2169	.2096	7
	7	.0000	.0000	.0001	.0011	.0058	.0186	.0442	.0833	.1312	.1775	.2095	6
	8	.0000	.0000	.0001	.0001	.0011	.0047	.0142	.0336	.0656	.1088	.1571	5
	9	.0000	.0000	.0000	.0000	.0001	.0009	.0034	.0101	.0243	.0495	.0873	4
	10	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0022	.0065	.0162	.0349	3
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0012	.0038	.0095	2
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0015	1
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	0
14	0	.8687	.4877	.2288	.1028	.0440	.0178	.0068	.0024	.0008	.0002	.0001	14
	1	.1229	.3593	.3559	.2539	.1539	.0832	.0467	.0181	.0073	.0027	.0009	13
	2	.0081	.1220	.2570	.2912	.2501	.1802	.1134	.0634	.0317	.0141	.0056	12
		.99	.95	.90	.85	.80	.75	.70	.65	.60	.55	.50	

p

تابع الجدول 1

n	m	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	
							p						
14	3	.0003	.0250	.1142	.2056	.2501	.2402	.1943	.1366	.0845	.0462	.0222	11
	4	.0000	.0037	.0349	.0998	.1720	.2202	.2290	.2022	.1549	.1040	.0611	10
	5	.0000	.0004	.0078	.0352	.0860	.1468	.1963	.2178	.2066	.1701	.1222	9
	6	.0000	.0000	.0013	.0093	.0322	.0734	.1262	.1750	.2066	.2088	.1833	8
	7	.0000	.0000	.0002	.0019	.0092	.0280	.0618	.1082	.1574	.1952	.2088	7
	8	.0000	.0000	.0000	.0003	.0020	.0082	.0232	.0510	.0918	.1398	.1833	6
	9	.0000	.0000	.0000	.0000	.0003	.0018	.0066	.0183	.0408	.0762	.1222	5
	10	.0000	.0000	.0000	.0000	.0000	.0003	.0014	.0049	.0136	.0312	.0611	4
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0010	.0033	.0083	.0222	3
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0019	.0058	2
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0009	1
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	0
15	0	.8601	.4633	.2059	.0874	.0352	.0134	.0047	.0016	.0005	.0001	.0000	15
	1	.1003	.3658	.2432	.2312	.1319	.0668	.0305	.0126	.0047	.0016	.0005	14
	2	.0092	.1348	.2869	.2856	.2309	.1559	.0916	.0476	.0219	.0090	.0032	13
	3	.0004	.0307	.1285	.2184	.2501	.2252	.1700	.1110	.0634	.0318	.0139	12
	4	.0000	.0049	.0428	.1156	.1876	.2252	.2186	.1792	.1268	.0760	.0417	11
	5	.0000	.0006	.0105	.0449	.1032	.1681	.2061	.2123	.1859	.1404	.0916	10
	6	.0000	.0000	.0019	.0132	.0430	.0917	.1472	.1906	.2066	.1914	.1527	9
	7	.0000	.0000	.0003	.0030	.0138	.0393	.0811	.1319	.1771	.2013	.1964	8
	8	.0000	.0000	.0000	.0005	.0035	.0131	.0348	.0710	.1181	.1647	.1964	7
	9	.0000	.0000	.0000	.0001	.0007	.0034	.0116	.0298	.0612	.1048	.1527	6
	10	.0000	.0000	.0000	.0000	.0001	.0007	.0030	.0096	.0245	.0515	.0916	5
	11	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0074	.0191	.0417	4
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052	.0139	3
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0032	2
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	1
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	0
16	0	.8515	.4401	.1853	.0743	.0281	.0100	.0033	.0010	.0003	.0001	.0000	16
	1	.1376	.3706	.3294	.2097	.1126	.0536	.0228	.0087	.0030	.0009	.0002	15
	2	.0104	.1463	.2745	.2775	.2111	.1336	.0732	.0353	.0150	.0056	.0018	14
	3	.0005	.0359	.1123	.2285	.2463	.2079	.1465	.0888	.0468	.0215	.0085	13
	4	.0000	.0061	.0514	.1311	.2001	.2252	.2040	.1553	.1014	.0572	.0278	12
	5	.0000	.0008	.0137	.0555	.1201	.1802	.2099	.2008	.1623	.1123	.0667	11
	6	.0000	.0001	.0028	.0180	.0550	.1101	.1649	.1982	.1983	.1684	.1222	10
	7	.0000	.0000	.0004	.0045	.0197	.0524	.1010	.1524	.1889	.1969	.1746	9
	8	.0000	.0000	.0001	.0009	.0055	.0197	.0487	.0923	.1417	.1812	.1964	8
		.99	.95	.90	.85	.80	.75	.70	.65	.60	.55	.50	
							p						

تابع الجدول 1

		p											
n	m	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	
16	9	.0000	.0000	.0000	.0001	.0012	.0058	.0185	.0442	.0840	.1318	.1746	7
	10	.0000	.0000	.0000	.0000	.0002	.0014	.0056	.0167	.0392	.0755	.1222	6
	11	.0000	.0000	.0000	.0000	.0000	.0002	.0013	.0049	.0142	.0337	.0667	5
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0040	.0115	.0278	4
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0025	.0085	3
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018	2
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	1
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	0
17	0	.8429	.4181	.1668	.0631	.0225	.0075	.0023	.0007	.0002	.0000	.0000	17
	1	.1447	.3741	.3150	.1893	.0957	.0426	.0169	.0060	.0019	.0005	.0001	16
	2	.0117	.1575	.2800	.2673	.1914	.1136	.0581	.0260	.0102	.0035	.0010	15
	3	.0006	.0415	.1556	.2359	.2393	.1893	.1245	.0701	.0341	.0144	.0052	14
	4	.0000	.0076	.0605	.1457	.2093	.2209	.1868	.1320	.0796	.0411	.0182	13
	5	.0000	.0010	.0175	.0668	.1361	.1914	.2081	.1849	.1379	.0875	.0472	12
	6	.0000	.0001	.0039	.0236	.0680	.1276	.1784	.1991	.1839	.1432	.0944	11
	7	.0000	.0000	.0007	.0065	.0267	.0668	.1201	.1685	.1827	.1841	.1484	10
	8	.0000	.0000	.0001	.0014	.0084	.0279	.0644	.1134	.1606	.1883	.1855	9
	9	.0000	.0000	.0000	.0003	.0021	.0093	.0276	.0511	.1070	.1540	.1855	8
	10	.0000	.0000	.0000	.0000	.0004	.0025	.0095	.0263	.0571	.1008	.1484	7
	11	.0000	.0000	.0000	.0000	.0001	.0005	.0026	.0090	.0242	.0525	.0944	6
	12	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0081	.0215	.0472	5
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0021	.0068	.0182	4
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052	3
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	2
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	1
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	0
18	0	.8345	.3972	.1501	.0536	.0180	.0056	.0016	.0004	.0001	.0000	.0000	18
	1	.1517	.3763	.3002	.1704	.0811	.0338	.0126	.0042	.0012	.0003	.0001	17
	2	.0130	.1683	.2835	.2566	.1723	.0958	.0458	.0180	.0069	.0022	.0006	16
	3	.0007	.0473	.1680	.2406	.2207	.1704	.1046	.0547	.0240	.0095	.0031	15
	4	.0000	.0093	.0700	.1592	.2153	.2130	.1681	.1104	.0614	.0291	.0117	14
	5	.0000	.0014	.0218	.0787	.1507	.1988	.2017	.1664	.1146	.0666	.0327	13
	6	.0002	.0002	.0052	.0301	.0816	.1436	.1873	.1941	.1656	.1181	.0708	12
	7	.0000	.0000	.0010	.0081	.0350	.0620	.1376	.1792	.1892	.1557	.1214	11
	8	.0000	.0000	.0002	.0022	.0120	.0376	.0811	.1327	.1734	.1884	.1669	10
	9	.0000	.0000	.0000	.0004	.0033	.0139	.0386	.0794	.1284	.1694	.1855	9
	10	.0000	.0000	.0000	.0001	.0008	.0042	.0149	.0385	.0771	.1248	.1669	8
		.99	.95	.90	.85	.80	.75	.70	.65	.60	.55	.50	
		p											

						p							
n	m	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	
18	11	.0000	.0000	.0000	.0000	.0001	.0010	.0046	.0151	.0374	.0742	.1214	7
	12	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0047	.0145	.0354	.0708	6
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0045	.0134	.0327	5
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0039	.0117	4
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0009	.0031	3
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006	2
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	1
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	0
19	0	.3262	.3774	.1351	.0456	.0144	.0042	.0011	.0003	.0001	.0000	.0000	19
	1	.1583	.3774	.2852	.1529	.0535	.0256	.0083	.0029	.0008	.0002	.0000	18
	2	.0144	.1787	.2852	.2428	.1540	.0803	.0358	.0138	.0046	.0013	.0003	17
	3	.0008	.0633	.1795	.2428	.2182	.1517	.0869	.0422	.0175	.0082	.0018	16
	4	.0000	.0112	.0798	.1714	.2182	.2023	.1491	.0909	.0467	.0203	.0074	15
	5	.0000	.0018	.0266	.0907	.1636	.2023	.1916	.1458	.0933	.0497	.0222	14
	6	.0000	.0002	.0069	.0374	.0955	.1574	.1916	.1844	.1451	.0949	.0518	13
	7	.0000	.0000	.0014	.0122	.0443	.0974	.1525	.1844	.1797	.1443	.0961	12
	8	.0000	.0000	.0002	.0032	.0166	.0487	.0981	.1489	.1797	.1771	.1442	11
	9	.0000	.0000	.0000	.0007	.0051	.0198	.0514	.0980	.1464	.1771	.1762	10
	10	.0000	.0000	.0000	.0001	.0013	.0066	.0220	.0628	.0976	.1449	.1762	9
	11	.0000	.0000	.0000	.0000	.0003	.0018	.0077	.0233	.0632	.0970	.1442	8
	12	.0000	.0000	.0000	.0000	.0000	.0004	.0022	.0083	.0237	.0529	.0961	7
	13	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0024	.0086	.0233	.0618	6
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0082	.0222	5
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0022	.0074	4
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0018	3
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	2
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	1
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	0
20	0	.8179	.3585	.1216	.0388	.0115	.0032	.0008	.0002	.0000	.0000	.0000	20
	1	.1652	.3774	.2702	.1368	.0576	.0211	.0058	.0020	.0005	.0001	.0000	19
	2	.0159	.1687	.2852	.2293	.1369	.0699	.0278	.0100	.0031	.0008	.0002	18
	3	.0010	.0595	.1901	.2428	.2054	.1339	.0716	.0323	.0123	.0040	.0011	17
	4	.0000	.133	.0898	.1821	.2182	.1897	.1304	.0738	.0350	.0138	.0045	16
	5	.0000	.0022	.0319	.1028	.1746	.2023	.1789	.1272	.0746	.0365	.0148	15
	6	.0000	.0003	.0089	.0454	.1081	.1686	.1916	.1712	.1244	.0746	.0370	14
	7	.0000	.0000	.0020	.0160	.0545	.1124	.1643	.1944	.1559	.1221	.0739	13
	8	.0000	.0000	.0004	.0046	.0222	.0609	.1144	.1614	.1797	.1623	.1201	12
		.99	.95	.90	.85	.80	.75	.70	.65	.60	.55	.50	
							p						

		p											
n	m	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	
20	9	.0000	.0000	.0001	.0011	.0074	.0271	.0654	.1158	.1597	.1771	.1602	11
	10	.0000	.0000	.0000	.0002	.0020	.0099	.0308	.0686	.1171	.1593	.1762	10
	11	.0000	.0000	.0000	.0000	.0005	.0030	.0120	.0336	.0710	.1185	.1602	9
	12	.0000	.0000	.0000	.0000	.0001	.0008	.0039	.0136	.0355	.0727	.1201	8
	13	.0000	.0000	.0000	.0000	.0000	.0002	.0010	.0045	.0148	.0366	.0738	7
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0049	.0150	.0370	6
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0048	.0148	5
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0046	4
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	3
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	2
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	1
	20	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	0
25	0	.7778	.2774	.0718	.0172	.0038	.0008	.0001	.0000	.0000	.0000	.0000	25
	1	.1964	.3650	.1994	.0759	.0236	.0063	.0014	.0003	.0000	.0000	.0000	24
	2	.0238	.2305	.2659	.1607	.0708	.0251	.0074	.0018	.0004	.0001	.0000	23
	3	.0018	.0930	.2265	.2174	.1358	.0641	.0243	.0076	.0018	.0004	.0001	22
	4	.0001	.0269	.1384	.2110	.1867	.1175	.0572	.0224	.0071	.0018	.0004	21
	5	.0000	.0060	.0646	.1564	.1960	.1645	.1030	.0506	.0199	.0063	.0018	20
	6	.0000	.0010	.0239	.0920	.1633	.1828	.1472	.0908	.0442	.0172	.0053	19
	7	.0000	.0001	.0072	.0441	.1108	.1654	.1712	.1327	.0800	.0381	.0143	18
	8	.0000	.0000	.0018	.0175	.0623	.1241	.1651	.1607	.1200	.0701	.0322	17
	9	.0000	.0000	.0004	.0058	.0294	.0781	.1336	.1635	.1511	.1084	.0609	16
	10	.0000	.0000	.000	.0016	.0118	.0417	.0916	.1409	.1612	.1419	.0974	15
	11	.0000	.0000	.0000	.0004	.0040	.0189	.0536	.1034	.1465	.1583	.1328	14
	12	.0000	.0000	.0000	.0000	.0012	.0074	.0268	.0650	.1140	.1511	.1550	13
	13	.0000	.0000	.0000	.0000	.0003	.0025	.0115	.0350	.0760	.1236	.1550	12
	14	.0000	.0000	.0000	.0000	.0000	.0007	.0042	.0161	.0434	.0867	.1328	11
	15	.0000	.0000	.0000	.0000	.0000	.0002	.0013	.0064	.0212	.0520	.0974	10
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0004	.0021	.0088	.0296	.0609	9
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0031	.0115	.0322	8
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0009	.0042	.0143	7
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0013	.0063	6
	20	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0016	5
	21	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0004	4
	22	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	3
		.99	.95	.90	.85	.80	.75	.70	.65	.60	.55	.50	
		p											

تابع الجدول 1

							p					
n	m	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	
30	0	.7397	.2146	.0424	.0076	.0012	.0002	.0000	.0000	.0000	.0000	.0000	30
	1	.2242	.3389	.1413	.0404	.0093	.0018	.0003	.0000	.0000	.0000	.0000	29
	2	.0328	.2586	.2277	.1034	.0337	.0086	.0018	.0003	.0000	.0000	.0000	28
	3	.0031	.1270	.2361	.1703	.0785	.0269	.0072	.0015	.0003	.0000	.0000	27
	4	.002	.0451	.1771	.2028	.1325	.0604	.0208	.0056	.0012	.0002	.0000	26
	5	.0000	.0124	.1023	.1861	.1723	.1047	.0464	.0157	.0041	.0008	.0001	25
	6	.0000	.0027	.0474	1368	.1795	.1455	.0829	.0363	.0115	.0029	.0006	24
	7	.0000	.0005	.0180	.0628	.1538	.1662	.1229	.0652	.0263	.0061	.0019	23
	8	.0000	.0001	.0068	.0420	.1106	.1593	.1501	.1009	.0505	.0181	.0055	22
	9	.0000	.0000	.0016	.0181	.0676	.1298	.1573	.1328	.823	.0382	.0133	21
	10	.0000	.0000	.0004	.0067	.0355	.0909	.1416	.1502	.1152	.656	.0280	20
	11	.0000	.0000	.0001	.0022	.0161	.0551	.1103	.1471	.1386	.0976	.0508	19
	12	.0000	.0000	.0000	.0006	.0064	.0291	.0749	.1254	.1474	.1265	.0806	18
	13	.0000	.0000	.0000	.0001	.0022	.0134	.0444	.0936	1380	.1433	.1115	17
	14	.0000	.0000	.0000	.0000	.0007	.0054	.0231	.0611	.1101	.1323	.1354	16
	15	.0000	.0000	.0000	.0000	.0002	.0019	.0106	.0351	.0783	.1242	.1445	15
	16	.0000	.0000	.0000	.0000	.0000	.0006	.0042	.0177	.0489	.0953	.1354	14
	17	.0000	.0000	.0000	.0000	.0000	.0002	.0015	.0079	.0289	.0642	.1115	13
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0005	.0031	.0129	.0379	.0806	12
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0010	.0064	.0196	.0508	11
	20	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0020	.0068	.0280	10
	21	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0034	.0133	9
	22	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0055	8
	23	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0019	7
	24	.0000	.0000	.0000	0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006	6
	25	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	5
		.99	.95	.90	.85	.80	.75	.70	.65	.60	.55	.50	
							p						

The Poisson Distribution جدول 2: توزيع بواسون

جدول قيم الدالة: $$P_m(t).10^6 = \frac{\lambda^m}{m!} e^{-\lambda} . 10^6$$

	λ										
m	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	m
0	904837	818731	740818	670320	606531	548812	496585	449329	406570	367879	0
1	90484	163746	222245	268128	303265	329287	347610	359463	365913	367879	1
2	4524	16375	33337	53626	75816	98786	121663	143785	164661	183940	2
3	151	1092	3334	7150	12636	19757	28388	38343	49398	61313	63
4	4	55	250	715	1580	2964	4968	7669	11115	15328	4
5		2	15	57	158	356	696	1227	2001	3066	5
6			1	4	13	36	81	164	300	511	6
7					1	3	8	19	39	73	7
8							1	2	4	9	8
										1	9

m	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	m
0	332871	301194	272532	246597	223130	201897	182684	165299	149569	135335	0
1	366158	361433	354291	345236	334695	323034	310562	297538	284180	270671	1
2	201387	216860	230289	241665	251021	258428	263978	267784	269971	270671	2
3	72842	86744	99792	112777	125510	137828	149587	160671	170982	180447	3
4	20307	26023	32432	39472	47067	55131	63575	72302	81216	90224	4
5	4467	6246	8432	11052	14120	17642	21615	26029	30862	36089	5
6	819	1249	1827	2579	3530	4705	6124	7809	9773	12030	6
7	129	214	339	516	756	1075	1487	2008	2653	3437	7
8	18	32	55	90	142	215	316	452	630	859	8
9	2	4	8	14	24	38	60	90	133	191	9
10		1	1	2	4	6	10	16	25	38	10
11						1	2	3	4	7	11
12									1	1	12

تابع الجدول 2

m	λ 2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0	m
0	122456	110803	100259	90718	82085	74274	67206	60810	55023	49787	0
1	257159	243767	230595	217723	205212	193111	181455	170268	159567	149361	1
2	270016	268144	265185	261268	256516	251045	244964	238375	231373	224042	2
3	189012	196639	203308	209014	213763	217572	220468	222484	223660	224042	3
4	99231	108151	116902	125409	133602	141422	148816	155739	162154	168013	4
5	41677	47587	53775	60196	66801	73539	80360	87214	94049	100819	5
6	14587	17448	20614	24078	27834	31867	36162	40700	45457	50409	6
7	4376	5484	6773	8255	9941	11836	13948	16280	18832	21604	7
8	1149	1508	1947	2477	3106	3847	4708	5698	6827	8102	8
9	268	369	498	660	863	1111	1412	1773	2200	2701	9
10	56	81	114	158	216	289	381	496	638	810	10
11	11	16	24	35	49	68	94	126	168	221	11
12	2	3	5	7	10	15	21	29	41	55	12
13		1	1	1	2	3	4	6	9	13	13
14						1	1	1	2	3	14
15										1	15
	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0	
0	45049	40762	36883	33373	30197	27324	24724	22371	20242	18316	0
1	139653	130439	121714	113469	105691	98365	91477	85009	78943	73263	1
2	216461	208702	200829	192898	184959	177058	169233	161517	153940	146525	2
3	223677	222616	220912	218617	215785	212469	208720	204588	200122	195367	3
4	173350	178093	182252	185825	188812	191222	193066	194359	195119	195367	4
5	107477	113979	120286	126361	132169	137680	142869	147713	152193	156293	5
6	55530	60789	66158	71604	77098	82608	88102	93551	98925	104196	6
7	24592	27789	31189	34779	38549	42484	46568	50785	55115	59540	7
8	9529	11116	12865	14781	16865	19118	21538	24123	26869	29770	8
9	3282	3952	4717	5584	6559	7647	8854	10185	11643	13231	9
10	1018	1265	1557	1899	2296	2753	3276	3870	4541	5292	10
11	287	368	567	587	730	901	1102	1337	1610	1925	11
12	74	98	128	166	213	270	340	423	523	642	12
13	18	24	33	43	57	75	97	124	157	197	13

تابع الجدول 2

m	λ										m
	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0	
14	4	6	8	11	14	19	26	34	44	56	14
15	1	1	2	2	3	5	6	9	11	15	15
16				1	1	1	1	2	3	4	16
17									1	1	17
	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	5.0	
0	16573	14996	13569	12277	11109	10052	9095	8230	7447	6738	0
1	67948	62981	58345	54020	49990	46238	42748	39503	36488	33690	1
2	139293	132261	125441	118845	112479	106348	100457	94807	89396	84224	2
3	190368	185165	179799	174305	168718	163068	157383	151691	146014	140374	3
4	195127	194424	193184	191736	189808	187528	194925	182029	178867	175467	4
5	160004	163316	166224	168728	170827	172525	173830	174748	175290	175467	5
6	109336	114321	119127	123734	128120	132270	136167	139798	143153	146223	6
7	64040	68593	73178	77775	82363	86920	91426	95862	100207	104445	7
8	32820	36011	39333	42776	46329	49979	53713	57517	61377	65278	8
9	14951	16805	18793	20913	23165	25545	28050	30676	33416	36266	9
10	6130	7058	8081	9202	10424	11751	13184	14724	16374	18138	10
11	2285	2695	3159	3681	4264	4914	5633	6425	7294	8242	11
12	781	943	1132	1350	1599	1884	2206	2570	2978	3434	12
13	246	305	374	457	554	667	798	949	1123	1321	13
14	72	91	115	144	178	219	268	325	393	472	14
15	20	26	33	42	53	67	84	104	128	157	15
16	5	7	9	12	15	19	25	31	39	49	16
17	1	2	2	3	4	5	7	9	11	14	17
18			1	1	1	1	2	2	3	4	18
19									1	1	19
	5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0	
0	6097	5517	4992	4517	4087	3698	3346	3028	2739	2479	0
1	31093	28686	26455	24390	22477	20477	19072	17560	16163	14873	1
2	79288	74584	70107	65852	61812	57982	54355	50923	47680	44618	2
3	134790	129279	123856	118533	113323	108234	103275	98452	93771	89235	3

تابع الجدول 2

m	λ										m
m	5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0	m
4	171857	168063	164109	160020	155819	151528	147167	142775	138312	133853	4
5	175294	174785	173955	172821	171401	169711	167770	165596	163208	160623	5
6	149000	151480	153660	155539	157117	158397	159382	160076	160488	160623	6
7	108557	112528	116343	119987	123449	126717	129782	132635	135268	137677	7
8	69205	73143	77077	80991	84871	88702	92470	96160	99760	103258	8
9	39216	42261	45390	48595	51866	55192	58564	61970	65398	68838	9
10	20000	21976	24057	26241	28526	30908	33382	35943	38585	41303	10
11	9273	10388	11591	12882	14263	15735	17298	18952	20696	22529	11
12	3941	4502	5119	5797	6537	7343	8216	9160	10175	11264	12
13	1546	1801	2087	2408	2766	3163	3603	4087	4618	5199	13
14	563	669	790	929	1087	1265	1467	1693	1946	2228	14
15	191	232	279	334	398	472	557	655	766	891	15
16	61	75	92	113	137	165	199	237	282	334	16
17	18	23	29	36	44	54	67	81	98	118	17
18	5	7	8	11	14	17	21	26	32	39	18
19	1	2	2	3	4	5	6	8	10	12	19
20			1	1	1	1	2	2	3	4	20
21								1	1	1	21
	6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0	
0	2243	2029	1836	1662	1503	1303	1360	1231	1114	1008	0
1	13682	12582	11569	10634	9772	8978	8247	7574	6954	6383	1
2	41729	39006	36441	34029	31760	29629	27628	25751	23990	22341	2
3	84848	80612	76527	72595	68814	65183	61702	58368	55178	52129	3
4	129393	124948	120530	116151	111822	107553	103351	99225	95182	91226	4
5	157860	154936	151868	148674	145369	141969	138490	134946	131351	127717	5
6	160491	160100	159561	158585	157483	156166	154648	152939	151053	149003	6
7	139856	141803	143515	144992	146234	147243	148020	148569	148895	149003	7
8	106640	109897	113018	115994	118815	121475	123967	126284	128422	130377	8
9	72278	75707	79113	82484	85811	89082	92286	95415	98457	101405	9
10	44090	46938	49841	52790	55777	58794	61832	64882	67935	70983	10
11	14450	26456	28545	30714	32959	35276	37661	40109	42614	45171	11

m	λ										m
	6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0	
12	12429	13669	14986	16381	17853	19402	21028	22728	24503	26350	12
13	5832	6519	7263	8064	8926	9850	10837	11889	13005	14188	13
14	2541	2887	3268	3687	4144	4644	5186	5774	6410	7094	14
15	1033	1193	1373	1573	1796	2043	2317	2618	2949	3311	15
16	394	462	540	629	730	843	970	1113	1272	1448	16
17	141	169	200	237	279	327	382	445	516	596	17
18	48	58	70	84	101	120	142	168	198	232	18
19	15	19	23	28	34	42	50	60	72	85	19
20	5	6	7	9	11	14	17	20	25	30	20
21	1	2	2	3	3	4	5	7	8	10	21
22			1	1	1	1	2	2	3	3	22
23								1	1	1	23
	7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0	
0	825	747	676	611	553	500	453	410	371	335	0
1	5858	5375	4931	4523	4148	3803	3487	3196	2929	2684	1
2	20797	19352	18000	16736	15555	14453	13424	12464	11569	10735	2
3	49219	46444	43799	41282	38889	36614	34455	32407	30465	28626	3
4	87364	83598	79934	76372	72916	69567	66326	63193	60169	57252	4
5	124057	120382	116703	113031	109375	105742	102142	98581	95067	91604	5
6	146800	144458	141989	139405	136718	133940	131082	128156	125171	122138	6
7	148897	148586	148047	147371	146484	145421	144191	142802	141264	139587	7
8	132146	133727	135118	136318	137329	138150	138783	139232	139499	139587	8
9	104249	106982	109596	112084	114440	116660	118737	120668	122449	124077	9
10	74017	77027	80005	82942	85830	88661	91427	94121	96735	99262	10
11	47774	50418	53094	55797	58521	61257	63999	66740	69473	72190	11
12	28267	30251	32299	34408	36575	38796	41066	43381	45736	48127	12
13	15438	16754	18137	19586	21101	22681	24324	26029	27794	29616	13
14	7829	8616	9457	10353	11304	12312	13378	14502	15684	16924	14
15	3706	4136	4603	5107	5652	6238	6867	7541	9260	9026	15
16	1644	1861	2100	2362	2649	2963	3305	3676	4078	4513	16
17	687	788	902	1028	1169	1325	1497	1687	1895	2124	17

تابع الجدول 2

m	λ 7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0	m
18	271	315	366	423	487	559	460	731	832	944	18
19	101	119	141	165	192	224	259	300	346	397	19
20	36	43	51	61	72	85	100	117	137	159	20
21	12	15	18	21	26	31	37	43	51	61	21
22	4	5	6	7	9	11	13	15	18	22	22
23	1	2	2	2	3	4	4	5	6	8	23
24			1	1	1	1	1	2	2	2	24
25								1	1	1	25

m	8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0	m
0	304	275	249	225	203	184	167	151	136	123	0
1	2459	2252	2063	1889	1729	1583	1449	1326	1214	1111	1
2	9958	9234	8560	7933	7350	6808	6304	5836	5402	4998	2
3	26885	25539	23683	22213	20826	19517	18283	17120	16025	14994	3
4	54443	51740	49142	46648	44255	41961	39765	37664	35656	33737	4
5	88198	84854	81576	78368	75233	72174	69192	66289	63467	60727	5
6	119067	115967	112847	109716	106581	103449	100328	97224	94143	91090	6
7	137778	135848	133805	131659	129419	127094	124693	122224	119696	117116	7
8	139500	139244	138823	138242	137508	136626	135604	134446	133161	131756	8
9	125550	126866	128025	129026	129869	130554	131084	131682	131756	131756	9
10	101696	104031	106261	108382	110388	112277	114043	115684	117197	118580	10
11	74885	77550	80179	82764	85300	87780	90197	92547	94823	97020	11
12	50547	52993	55457	57935	60421	62909	65393	67868	70327	72765	12
13	31495	33426	35407	37435	39506	41617	43763	45941	48147	50376	13
14	18222	19578	20991	22461	23986	25565	27196	28877	30608	32384	14
15	9840	10703	11615	12578	13592	14657	15773	16941	18161	19431	15
16	4981	5485	6025	6604	7221	7878	8577	9318	10102	10930	16
17	2373	2646	2942	3263	3610	3985	4389	4823	5289	5786	17
18	1068	1205	1356	1523	1705	1904	2121	2358	2615	2893	18
19	455	520	593	673	763	862	971	1092	1225	1370	19
20	184	213	246	283	324	371	423	481	545	617	20
21	71	83	97	113	131	152	175	201	231	264	21

تابع الجدول 2

m	λ 8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0	m
22	26	31	37	43	51	59	69	81	93	108	22
23	9	11	13	16	19	22	26	31	36	42	23
24	3	4	5	6	7	8	9	11	13	16	24
25	1	1	2	2	2	3	3	4	5	6	25
26				1	1	1	1	1	2	2	26
27									1	1	27

	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	
0	112	101	91	83	75	68	61	55	50	45	0
1	1016	930	850	778	711	560	594	543	497	454	1
2	4624	4276	3954	3655	3378	3121	2883	2663	2459	2270	2
3	14025	13113	12256	11452	10696	9987	9322	8698	8114	7567	3
4	31906	30160	28496	26911	25403	23969	22606	21311	20082	18917	4
5	58069	55494	53002	50593	48266	46020	43855	41770	39763	37833	5
6	88072	85091	82154	79262	76421	73632	70899	68224	65609	63055	6
7	114493	111834	109147	106438	103714	100981	98246	95514	92790	90079	7
8	130236	128609	126883	125065	123160	121178	119123	117004	114827	112599	8
9	131683	131467	131113	130623	130003	129256	128388	127405	126310	125110	9
10	119832	120950	121935	122786	123502	124086	124537	124857	125047	125110	10
11	99133	101158	103090	104926	106661	108293	109819	111236	112542	113736	11
12	75176	77555	79895	82192	84440	86634	88770	90843	92847	94780	12
13	52623	54885	57156	59431	61706	63976	66236	68481	70707	72908	13
14	34205	36067	37968	39904	41872	43869	45892	47937	50000	52077	14
15	20751	22121	23540	25006	26519	28076	29677	31319	33000	34718	15
16	11802	12720	13683	14691	15746	16846	17992	19183	20419	21699	16
17	6318	6884	7485	8123	8799	9513	10266	11058	11891	12764	17
18	3194	3518	3867	4242	4644	5074	5532	6021	6540	7091	18
19	1530	1704	1893	2099	2322	2563	2824	3105	3408	3732	19
20	696	784	880	986	1103	1230	1370	1522	1687	1866	20
21	302	343	390	442	499	563	633	710	795	889	21
22	125	144	165	189	215	245	279	316	358	404	22
23	49	57	67	77	89	102	118	135	154	176	23

تابع الجدول 2

m	λ 9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	m
24	19	22	26	30	35	41	48	55	64	73	24
25	7	8	10	11	13	16	18	22	25	29	25
26	2	3	3	4	5	6	7	8	10	11	26
27	1	1	1	1	2	2	2	3	4	4	27
28					1	1	1	1	1	1	28
29										1	29
	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	
0	41	37	34	30	28	25	23	20	18	17	0
1	415	379	346	317	289	264	241	220	201	184	1
2	2095	1934	1784	1646	1518	1400	1291	1190	1097	1010	2
3	7054	6574	6125	5705	5313	4946	4603	4283	3984	3705	3
4	17811	16764	15773	14834	13946	13107	12313	11564	10856	10189	4
5	35979	34199	32492	30855	29287	27786	26350	24978	23667	22415	5
6	60565	58139	55777	53482	51252	49089	46991	44960	42995	41095	6
7	87387	84716	82072	79458	76878	74334	71830	69367	66949	64577	7
8	110326	108013	105668	103296	100902	98493	96072	93646	91218	88794	8
9	123810	122415	120931	119364	117720	116003	114219	112375	110475	108526	9
10	125048	124863	124559	124139	123606	122963	122215	121365	120418	119378	10
11	114817	115782	116633	117368	117987	118492	118882	119159	119323	119378	11
12	96637	98415	100110	101719	103239	104667	106003	107243	108386	109430	12
13	75080	77218	79318	81375	83385	85344	87248	89094	90877	92595	13
14	54165	56259	58355	60450	62539	64618	66683	68730	70754	72753	14
15	36471	38256	40071	41912	43777	45663	47567	49485	51415	53352	15
16	23022	24388	25795	27243	28729	30252	31810	33403	35026	36680	16
17	13678	14633	15629	16666	17744	18863	20022	21220	22458	23734	17
18	7675	8292	8943	9629	10351	11108	11902	12732	13600	14504	18
19	4080	4451	4848	5271	5720	6197	6703	7237	7802	8397	19
20	2060	2270	2497	2741	3003	3285	3586	3908	4252	4618	20
21	991	1103	1225	1357	1502	1658	1827	2010	2207	2419	21
22	455	511	573	642	717	799	889	987	1093	1210	22
23	200	227	257	290	327	368	413	463	518	578	23

تابع الجدول 2

m	λ										m
	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	
24	84	96	110	126	143	163	184	208	235	265	24
25	34	39	45	52	60	69	79	90	103	117	25
26	13	15	18	21	24	28	32	37	43	49	26
27	5	6	7	8	9	11	13	15	17	20	27
28	2	2	3	3	4	4	5	6	7	8	28
29	1	1	1	1	1	2	2	2	3	3	29
30						1	1	1	1	1	30

	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	
0	15	14	12	11	10	9	8	8	7	6	0
1	168	153	140	128	116	106	97	89	81	74	1
2	931	858	790	727	670	617	568	522	481	442	2
3	3445	3202	2976	2764	2568	2385	2214	2055	1907	1770	3
4	9559	8965	8406	7879	7382	6915	6476	6062	5674	5309	4
5	21221	20082	18997	17936	16979	16043	15153	14307	13504	12741	5
6	39259	37487	35778	34130	32544	31017	29549	28137	26782	254814	6
7	62253	59979	57755	55584	53465	51400	49388	47432	45530	43682	7
8	86376	83970	81579	79206	76856	74529	72231	69962	67725	65523	8
9	106531	104496	102427	100328	98204	96060	93900	91728	89548	87364	9
10	118249	117036	115743	114374	112935	111430	109863	108239	106562	104837	10
11	119324	119164	118899	118533	118068	117508	116854	116110	115281	114368	11
12	110375	111220	111964	112607	113149	113591	113933	114175	114320	114363	12
13	94243	95820	97322	98747	100093	101358	102538	103636	104647	105570	13
14	74721	76656	78553	80409	82219	83982	85694	87350	88950	90489	14
15	55294	57236	59177	61110	63035	64946	66841	68716	70567	72391	15
16	38360	40065	41793	43541	45306	47086	48877	50678	52484	54293	16
17	25047	26396	27780	29198	30648	32129	33639	35176	36739	38325	17
18	15446	16424	17440	18492	19581	20706	21865	23060	24288	25550	18
19	9023	9682	10372	11095	11852	12641	13465	14322	15212	16137	19
20	5008	5422	5860	6324	6815	7332	7877	8450	9051	9682	20
21	2647	2892	3153	3433	3732	4050	4388	4748	5129	5533	21
22	1336	1472	1620	1779	1951	2136	2334	2547	2774	3018	22

تابع الجدول 2

m	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	m
23	645	717	796	882	975	1077	1187	1307	1435	1575	23
24	298	335	375	419	467	521	579	642	712	787	24
25	132	150	169	191	215	242	271	303	339	378	25
26	57	65	74	84	95	108	122	138	155	174	26
27	23	27	31	35	41	46	53	60	68	78	27
28	9	11	12	14	17	19	22	25	29	33	28
29	4	4	5	6	7	8	9	10	12	14	29
30	1	2	2	2	3	3	3	4	5	5	30
31		1	1	1	1	1	1	2	2	2	31
32								1	1	1	32

m	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	m
0	6	5	5	4	4	3	3	3	2	2	0
1	67	61	56	51	47	42	39	35	32	29	1
2	407	374	344	317	291	268	246	226	208	191	2
3	1641	1522	1412	1309	1213	1124	1042	965	894	828	3
4	4966	4643	4341	4057	3791	3541	3307	3088	2882	2690	4
5	12017	11330	10679	10062	9477	8924	8400	7905	7436	6994	5
6	24233	23037	21892	20749	19744	18740	17781	16864	15988	15153	6
7	41889	40151	38467	36836	35258	33733	32259	30837	29464	28141	7
8	63358	61230	59142	57095	55091	53129	51212	49339	47511	45730	8
9	85181	83000	80828	78665	76515	74381	72266	70171	68100	66054	9
10	103069	101261	99418	97544	95644	93720	91777	89819	87849	85870	10
11	113376	112308	111168	109959	108686	107352	105961	104516	103023	101483	11
12	114321	114180	113947	113624	113215	112720	112142	11484	110749	109940	12
13	106406	107153	107811	108380	108860	109251	109554	109769	109897	109940	13
14	91965	93376	94720	95994	97197	98326	99381	100360	101263	102087	14
15	74185	75946	77670	79355	80997	82594	84143	85641	87086	88475	15
16	56103	57909	59709	61500	63279	65043	66788	68513	70213	71886	16
17	39932	41558	43201	44859	46529	48208	49895	51586	53279	54972	17
18	26843	28167	29521	30903	32312	33746	35204	36683	38183	39702	18
19	17095	18086	19111	20168	21258	22379	23531	24713	25925	27164	19

تابع الجدول 2

m	λ 12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	m
20	10342	11033	11753	12504	13286	14099	14942	15816	16721	17657	20
21	5959	6409	6884	7383	7908	8459	9036	9640	10272	10930	21
22	3278	3554	3849	4162	4493	4845	5216	5609	6023	6459	22
23	1724	1885	2058	2244	2442	2654	2880	3122	3378	3651	23
24	869	958	1055	1159	1272	1393	1524	1665	1816	1977	24
25	421	468	519	575	636	702	774	852	973	1028	25
26	196	219	246	274	306	340	378	420	465	514	26
27	88	99	112	126	142	159	178	199	222	248	27
28	38	43	49	56	63	71	81	91	102	115	28
29	16	18	21	24	27	31	35	40	46	52	29
30	6	7	9	10	11	13	15	17	20	22	30
31	2	3	3	4	5	5	6	7	8	9	31
32	1	1	1	2	2	2	2	3	3	4	32
33				1	1	1	1	1	1	2	33
34										1	34

	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	
0	2	2	2	2	1	1	1	1	1	1	0
1	27	24	22	20	19	17	15	14	13	12	1
2	175	161	148	136	125	115	105	97	89	81	2
3	766	709	657	608	562	520	481	445	411	380	3
4	2510	2341	2183	2035	1897	1768	1648	1535	1429	1331	4
5	6575	6180	5807	5455	5123	4810	4514	4236	3974	3727	5
6	14356	13596	12872	12183	11526	10902	10308	9743	9206	8696	6
7	26867	25639	24458	23322	22230	21181	20173	19207	18280	17392	7
8	43994	42304	40661	39064	37512	36007	34547	33132	31762	30435	8
9	64036	62046	60088	58161	56269	54410	52588	50802	49054	47344	9
10	83887	81901	79916	77936	75963	73998	72046	70107	68185	66282	10
11	99901	98281	96626	94940	93227	91489	89730	87953	86162	84359	11
12	109059	108109	107094	106017	104880	103687	102441	101146	99804	98418	12
13	109898	109773	109566	109279	108914	108473	107957	107370	106713	105989	13
14	102833	103500	104087	104595	105024	105373	105644	105836	105951	105989	14

تابع الجدول 2

m	λ 13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	m
15	89807	91080	92291	93439	94522	95539	96488	97369	98181	98923	15
16	73530	75141	76717	78255	79753	81208	82618	83981	85295	86558	16
17	56661	58345	60019	61683	63333	64966	66580	68173	69741	71283	17
18	41237	42786	44348	45920	47500	49386	50675	52266	53856	55442	18
19	28432	29725	31043	32385	33750	35135	36539	37962	39400	40852	19
20	18623	19619	20644	21698	22781	23892	25030	26193	27383	28597	20
21	11617	12332	13074	13846	14645	15473	16329	17213	18125	19064	21
22	6917	7399	7904	8433	8987	9565	10168	10797	11452	12132	22
23	3940	4246	4571	4913	5275	5656	6057	6478	6921	7385	23
24	2151	2336	2533	2743	2967	3205	3457	3725	4008	4308	24
25	1127	1233	1348	1470	1602	1744	1895	2056	2229	2412	25
26	568	626	689	758	832	912	998	1091	1191	1299	26
27	275	306	340	376	416	459	507	558	613	674	27
28	129	144	161	180	201	223	248	275	305	337	28
29	58	66	74	83	93	105	117	131	146	163	29
30	25	29	33	37	42	47	53	60	68	76	30
31	11	12	14	16	18	21	24	27	30	34	31
32	4	5	6	7	8	9	10	12	13	15	32
33	2	2	2	3	3	4	4	5	6	6	33
34	1	1	1	1	1	1	2	2	2	3	34
35						1	1	1	1	1	35

	14.1	14.2	14.3	14.4	14.5	14.6	14.7	14.8	14.9	15.0	
0	1	1	1	1	1						0
1	11	10	9	8	7	7	6	6	5	5	1
2	75	69	63	58	53	49	45	41	38	34	2
3	352	325	300	277	256	237	219	202	186	172	3
4	1239	1153	1073	999	929	864	803	747	694	645	4
5	3494	3275	3070	2876	2694	2523	2362	2211	2069	1936	5
6	8212	7752	7316	6902	6510	6139	5787	5454	5138	4839	6
7	16541	15726	14946	14199	13486	12804	12152	11530	10937	10370	7
8	29153	27913	26715	25559	24443	23367	22330	21331	20370	19444	8

تابع الجدول 2

m	λ										m
	14.1	14.2	14.3	14.4	14.5	14.6	14.7	14.8	14.9	15.0	
9	45673	44040	42447	40894	39380	37907	36472	35078	33723	32407	9
10	64399	62537	60700	58887	57101	55343	53614	51915	50247	48611	10
11	82547	80730	78910	77089	75270	73456	71648	69850	68062	66287	11
12	96993	95530	94034	92507	90951	89371	87769	86148	84510	82859	12
13	105200	104349	103437	102469	101446	100371	99247	98076	96862	95607	13
14	105951	105839	105654	105396	105069	104672	104209	103681	103089	102436	14
15	99594	100195	100723	101181	101567	101881	102125	102298	102402	102436	15
16	87768	88923	90021	91063	92045	92967	93827	94626	95361	96034	16
17	72795	74277	75724	77135	78509	79842	81133	82380	83581	84736	17
18	57023	58596	60158	61708	63243	64761	66259	67735	69187	70613	18
19	42317	43793	45277	46768	48264	49763	51263	52762	54257	55747	19
20	29834	31093	32373	33673	34992	36327	37678	39044	40422	41810	20
21	20231	21025	22045	23090	24161	25256	26375	27517	28680	29865	21
22	122838	13570	14329	15114	15924	16761	17623	18511	19424	20362	22
23	7870	8378	8909	9462	10039	10640	11264	11911	12584	13280	32
24	4624	4957	5308	5677	6065	6472	6899	7345	7812	8300	24
25	2608	2816	3036	3270	3518	3780	4057	4348	4656	4980	25
26	1414	1538	1670	1811	1962	2123	2294	2475	2668	2873	26
27	739	809	884	966	1054	1148	1249	1357	1473	1596	27
28	372	410	452	497	546	598	656	717	784	855	28
29	181	201	223	247	273	301	332	366	403	442	29
30	85	95	106	118	132	147	163	181	200	221	30
31	39	44	49	55	62	69	77	86	96	107	31
32	17	19	22	25	28	32	35	40	45	50	32
33	7	8	9	11	12	14	16	18	20	23	33
34	3	3	4	5	5	6	7	8	9	10	34
35	1	1	2	2	2	2	3	3	4	4	35
36		1	1	1	1	1	1	1	2	2	36
37								1	1	1	37

The Poisson Distribution جدول 3: توزيع بواسون

جدول قيم الدالة: $$g(k,\lambda) = \sum_{m=0}^{k} \frac{\lambda^m e^{-\lambda}}{m!}$$

k \ λ	0.1	0.2	0.3	0.4	0.5	0.6
0	0.904837	0.818731	0.740818	0.670320	0.606531	0.548812
1	0.995321	0.982477	0.963063	0.938448	0.909796	0.878099
2	0.999845	0.998852	0.996390	0.992074	0.985612	0.977885
3	0.999996	0.999943	0.999724	0.999224	0.998248	0.997642
4	1.000000	0.999998	0.999974	0.999939	0.999828	0.999606
5	1.000000	1.000000	0.999999	0.999996	0.999986	0.999962
6	1.000000	1.000000	1.000000	1.000000	0.999999	0.999997
7	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

k \ λ	0.7	0.8	0.9	1.0	2.0	3.0
0	0.496585	0.449329	0.406570	0.367879	0.135335	0.049787
1	0.844195	0.808792	0.772483	0.735759	0.406006	0.199148
2	0.965858	0.952577	0.937144	0.919699	0.676677	0.423190
3	0.994246	0.990920	0.988542	0.981012	0.857124	0.647232
4	0.999214	0.998589	0.997657	0.996340	0.947348	0.815263
5	0.999909	0.999816	0.999658	0.999406	0.983437	0.916082
6	0.999990	0.999980	0.999958	0.999917	0.995467	0.966491
7	0.999998	0.999999	0.999997	0.999990	0.998904	0.988095
8	1.000000	1.000000	1.000000	0.999999	0.999763	0.996196
9				1.000000	0.999954	0.998897
10					0.999992	0.999707
11					0.999999	0.999928
12					1.000000	0.999983
13						0.999996
14						0.999999
15						1.000000

The Exponential Distribution جدول 4: التوزيع الأسي

جدول قيم الدالة: $f(x) = e^{-x}$

x	e^x	e^{-x}
0.00	1.0000	1.0000
0.01	1.0101	0.9900
0.02	1.0202	0.9802
0.03	1.0305	0.9704
0.04	1.0408	0.9608
0.05	1.0513	0.9512
0.06	1.0618	0.9418
0.07	1.0725	0.9324
0.08	1.0833	0.9231
0.09	1.0942	0.9139
0.10	1.1052	0.9048
0.11	1.1163	0.8958
0.12	1.1275	0.8869
0.13	1.1388	0.8781
0.14	1.1503	0.9694
0.15	1.1618	0.8607
0.16	1.1735	0.8521
0.17	1.853	0.8437
0.18	1.1972	0.8353
0.19	1.2092	0.8270
0.20	1.2214	0.8187
0.21	1.2337	0.8106
0.22	1.2461	0.8025
0.23	1.2586	0.7945
0.24	1.2712	0.7866

x	e^x	e^{-x}
0.25	1.2840	0.7788
0.26	1.2969	0.7711
0.27	1.3100	0.7634
0.28	1.3231	0.7558
0.29	1.3364	0.7483
0.30	1.3499	0.7408
0.31	1.13634	0.7334
0.32	1.3771	0.7261
0.33	1.3910	0.7189
0.34	1.4049	0.7118
0.35	1.4191	0.7047
0.36	1.4333	0.6977
0.37	1.4477	.06907
0.38	1.4623	0.6839
0.39	1.4770	0.6771
0.40	1.4918	0.6703
0.41	1.5068	0.6637
0.42	1.5220	0.6570
0.43	1.5373	0.6505
0.44	1.5527	0.6440
0.45	1.5683	0.6376
0.46	1.5683	0.6376
0.47	1.5841	0.6313
0.48	1.6000	0.6250
0.49	1.6161	0.6188

تابع جدول 4.

x	e^x	e^{-x}
0.50	1.6323	0.6126
0.51	1.6487	0.6065
0.52	1.6653	0.6005
0.53	1.6820	0.5945
0.54	1.6989	0.5886
0.55	1.7160	0.5827
0.56	1.7333	0.5769
0.57	1.7507	0.5712
0.58	1.7683	0.5655
0.59	1.7860	0.5599
0.60	1.8040	0.5543
0.61	1.8221	0.5488
0.62	1.8044	0.5434
0.63	1.8589	0.5379
0.64	1.8776	0.5326
0.65	1.8965	0.5273
0.66	1.9155	0.5220
0.67	1.9348	0.5169
0.68	1.9542	0.5117
0.69	1.9739	0.5066
0.70	2.0138	0.4966
0.71	2.0340	0.4916
0.72	2.0544	0.4868
0.73	2.0751	0.4819
0.74	2.0959	0.4771
0.75	2.1170	0.4724
0.76	2.1383	.04677
0.77	2.1598	0.4630
0.78	2.1815	0.4584
0.79	2.2034	0.4538
0.80	2.2255	0.4493
0.81	2.2479	0.4449
0.82	2.2705	0.4404
0.83	2.2933	0.4360
0.84	2.3164	0.4317
0.85	2.3396	0.4274
0.86	2.3632	0.4232
0.87	2.369	0.4190
0.88	2.4109	0.4148
0.89	2.4351	0.4107
0.90	2.4596	0.4066
0.91	2.4843	0.4025
0.92	2.5093	0.3985
0.93	2.5345	0.3946
0.94	2.5600	0.3906
0.95	2.5857	0.3867
0.96	2.6117	0.3829
0.97	2.6379	0.3791
0.98	2.6646	0.3753
0.99	2.6912	0.3716
1.00	2.7183	0.3679
1.05	2.8577	0.3499
1.10	3.0042	0.3329
1.15	3.1582	0.3166

تابع جدول 4

x	e^x	e^{-x}
1.20	3.3201	0.3012
1.25	3.4903	0.2865
1.30	3.6693	0.2725
1.35	3.8574	0.2592
1.40	4.0552	0.2466
1.45	2.631	0.2346
1.50	4.4817	0.2231
1.55	4.7115	0.2122
1.60	4.9530	0.2019
1.65	5.2070	0.1920
1.70	5.4739	0.1827
1.75	5.7546	0.1738
1.80	6.0496	0.1653
1.85	6.3598	0.1572
1.90	6.6859	0.1496
1.95	7.0287	0.1423
2.00	7.3891	0.1353
2.05	7.7679	0.1287
2.10	8.1662	0.1225
2.15	8.5849	0.1165
2.20	9.0250	0.1108
2.25	9.4877	0.1054
2.30	9.7942	0.1003
2.35	10.486	0.0954
2.40	11.023	0.0907
2.45	11.588	0.0863
2.50	12.182	0.0821

x	e^x	e^{-x}
2.55	12.807	0.0781
2.60	13.464	0.0743
2.65	14.154	0.0707
2.70	14.880	0.672
2.75	16.643	0.0639
2.80	16.445	0.0608
2.85	17.288	0.0578
2.90	18.174	0.0550
2.95	19.106	0.0523
3.00	10.086	0.0498
3.05	21.115	0.0474
3.10	22.198	0.0450
3.15	23.336	0.0429
3.20	24.533	0.0408
3.25	25.790	0.0388
3.30	27.113	0.0369
3.35	28.503	0.0351
3.40	29.964	0.0334
3.45	31.500	0.0317
3.50	33.115	0.0302
3.55	34.813	0.0287
3.60	36.598	0.0273
3.65	38.475	0.0260
3.70	40.447	0.0247
3.75	42.521	0.0235
3.80	44.701	0.0224
3.85	46.993	0.0213

تابع جدول 4

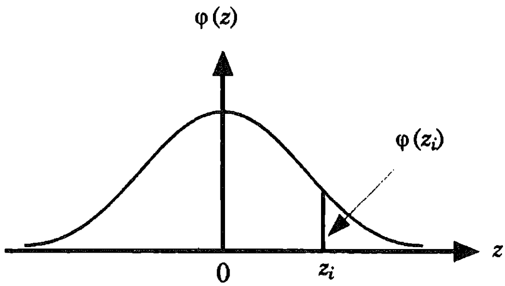
x	e^x	e^{-x}
4.60	99.484	0.0101
4.70	109.95	0.0091
4.80	121.15	0.0082
3.90	49.402	0.0202
3.95	51.935	0.0193
4.00	54.598	0.0183
4.10	60.340	0.0166
4.20	66.686	0.0150
4.30	73.700	0.0136
4.40	81.451	0.0123
4.50	90.017	0.0111

x	e^x	e^{-x}
4.90	134.29	0.0074
5.00	148.41	0.0067
5.20	181.27	0.0055
5.40	221.41	0.0045
5.60	270.43	0.0037
5.80	330.30	0.0030
6.00	403.43	0.0025
7.00	1096.6	0.0009
8.00	2981.0	0.0003
9.00	8103.1	0.0001
10.00	22026.0	0.00005

جدول 5: التوزيع الطبيعي المعياري

The Standard Normal Distribution

جدول قيم الدالة: $\varphi(z) = \dfrac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}}$



z	0	1	2	3	4	5	6	7	8	9
0.0	0.3989	0.3989	0.3989	0.3988	0.3986	0.3984	0.3982	0.3980	0.3977	0.3973
0.1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0.2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0.3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0.4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0.5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0.6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0.7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0.8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0.9	2661	2637	2613	2589	2565	2541	2516	2492	2568	2444
1.0	0.2420	0.2396	0.2371	0.2347	0.2323	0.2299	0.2275	0.2251	0.2227	0.2203
1.1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1.2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1.3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1.4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1.5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1.6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1.7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1.8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1.9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2.0	0.0540	0.0529	0.0519	0.0508	0.0498	0.0488	0.0478	0.0468	0.0459	0.0449
2.1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2.2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2.3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229

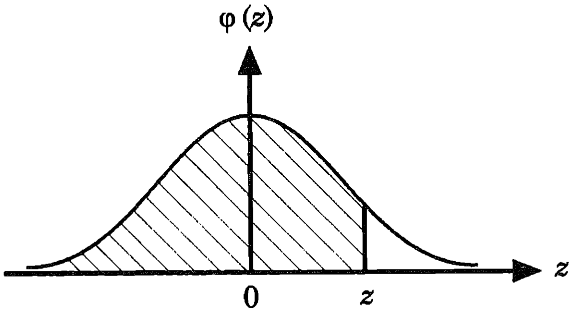
تابع جدول 5

z	0	1	2	3	4	5	6	7	8	9
2.4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2.5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2.6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2.7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2.8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2.9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3.0	0.0044	0.0043	0.0042	0.0041	0.0039	0.0038	0.0037	0.0036	0.0035	0.0034
3.1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3.2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3.3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3.4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3.5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3.6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3.7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3.8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0003
3.9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001
4.0	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
4.1	0.0001338									
4.5	0.0000160									
5.0	0.0000015									

جدول 6: التوزيع الطبيعي المعياري

The Standard Normal Distribution

جدول قيم الدالة: $$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$$



z	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6551	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7221
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7480	.7517	.7549
0.7	.7580	.7611	.7642	.7672	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8238	.8238	.8204	.8289	.8315	.8340	.8366	.8389
1.0	.8413	.8438	.8461	.8185	.8508	.8531	.8551	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890

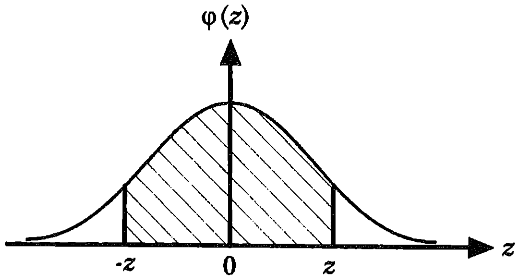
تابع جدول 6

z	0	1	2	3	4	5	6	7	8	9
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9961
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9971
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998									
3.6	.9998									
3.7	.99989									
3.8	.99993									
3.9	.99995									
4.0	.999968									
4.5	.999997									
5.0	.99999997									

جدول 7: التوزيع الطبيعي المعياري

The Standard Normal Distribution

جدول قيم الدالة: $U(z) = \frac{2}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt = P(|Z| < z)$



z	0	1	2	3	4	5	6	7	8	9
0.0	0.0000	0.0080	0.0160	0.0239	0.0319	0.0399	0.0478	0.0558	0.0638	0.0717
0.1	0.0797	0.0876	0.0955	0.1034	0.1113	0.1192	0.1271	0.1350	0.1428	0.1507
0.2	0.1585	0.1663	0.1741	0.1819	0.1897	0.1974	0.2051	0.2128	0.2205	0.2282
0.3	0.2358	0.2434	0.2510	0.2586	0.2661	0.2737	0.2812	0.2886	0.2960	0.3035
0.4	0.3108	0.3182	0.3255	0.3328	0.3401	0.3473	0.3545	0.3616	0.3688	0.3759
0.5	0.3829	0.3899	0.3969	0.4039	0.4108	0.4177	0.4245	0.4313	0.4381	0.4448
0.6	0.4515	0.4581	0.4647	0.4713	0.4778	0.4843	0.4907	0.4971	0.5035	0.5098
0.7	0.5161	0.5223	0.5285	0.5346	0.5407	0.5467	0.5527	0.5587	0.5646	0.5705
0.8	0.5763	0.5821	0.5878	0.5935	0.5991	0.6047	0.6102	0.6157	0.6211	0.6265
0.9	0.6319	0.6372	0.6424	0.6476	0.6528	0.6579	0.6629	0.6679	0.6729	0.6778
1.0	0.6827	0.6875	0.6923	0.6970	0.7017	0.7063	0.7109	0.7154	0.7199	0.7243
1.1	0.7287	0.7330	0.7373	0.7415	0.7457	0.7499	0.7540	0.7580	0.7620	0.7660
1.2	0.7699	0.7737	0.7775	0.7813	0.7850	0.7887	0.7923	0.7959	0.7994	0.8029
1.3	0.8064	0.8098	0.8132	0.8165	0.8198	0.8230	0.8262	0.8293	0.8324	0.8355
1.4	0.8385	0.8415	0.8444	0.8473	0.8501	0.8529	0.8557	0.8584	0.8611	0.8638
1.5	0.8664	0.8690	0.8715	0.8740	0.8764	0.8789	0.8812	0.8836	0.8859	0.8882
1.6	0.8904	0.8926	0.8948	0.8969	0.8990	0.9011	0.9031	0.9051	0.9070	0.9090
1.7	0.9109	0.9127	0.9146	0.9164	0.9181	0.9199	0.9216	0.9233	0.9249	0.9265
1.8	0.9281	0.9297	0.9312	0.9327	0.9342	0.9357	0.9371	0.9385	0.9399	0.9412
1.9	0.9426	0.9439	0.9451	0.9464	0.9476	0.9488	0.9500	0.9512	0.9523	0.9534
2.0	0.9545	0.9556	0.9566	0.9576	0.9586	0.9596	0.9606	0.9616	0.9625	0.9634
2.1	0.9643	0.9651	0.9660	0.9668	0.9676	0.9684	0.9692	0.9700	0.9707	0.9715
2.2	0.9722	0.9729	0.9736	0.9743	0.9749	0.9756	0.9762	0.9768	0.9774	0.9780
2.3	0.9786	0.9791	0.9797	0.9802	0.9807	0.9812	0.9817	0.9822	0.9827	0.9832

تابع جدول 7

z	0	1	2	3	4	5	6	7	8	9
2.4	0.9836	0.9841	0.9845	0.9849	0.9853	0.9857	0.9861	0.9865	0.9869	0.9872
2.5	0.9876	0.9879	0.9883	0.9886	0.9889	0.9892	0.9895	0.9898	0.9901	0.9904
2.6	0.9907	0.9910	0.9912	0.9915	0.9917	0.9920	0.9922	0.9924	0.9926	0.9928
2.7	0.9931	0.9933	0.9935	0.9937	0.9939	0.9940	0.9942	0.9944	0.9946	0.9947
2.8	0.9949	0.9951	0.9952	0.9553	0.9955	0.9956	0.9958	0.9959	0.9960	0.9961
2.9	0.9963	0.9964	9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972
3.0	0.9973	0.9974	0.9975	0.9976	0.9976	0.9977	0.9978	0.9979	0.9979	0.9980
3.1	0.9981	0.9981	0.9982	0.9983	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986
3.2	0.9986	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.3	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.4	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995	0.9995
3.5	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997	0.9997
3.6	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998	0.9998	0.9998
3.7	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
4.0	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

جدول 8: توزيع ذي الحدين *The Binomial Distribution*

95% فترة ثقة (θ_1, θ_2) من أجل المعلمة θ:

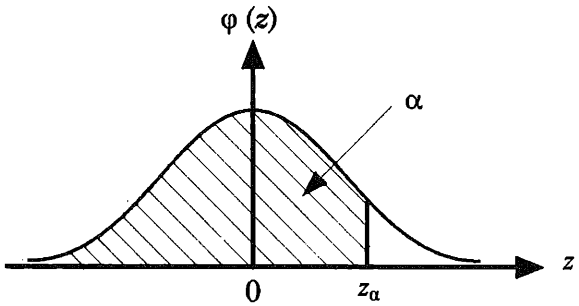
$$P_\theta(\theta_1 < \theta < \theta_2) = 0.95$$

k \ $n-k$	0	1	2	3	4	5	6	7	8	9	10	11	12
0	- -	0.98 0.00	0.84 0.00	0.71 0.00	0.60 0.00	0.52 0.00	0.46 0.00	0.41 0.00	0.37 0.00	0.34 0.00	0.31 0.00	0.29 0.00	0.27 0.00
1	1.00 0.03	0.99 0.01	0.91 0.01	0.81 0.01	0.72 0.01	0.64 0.00	0.58 0.00	0.53 0.00	0.48 0.00	0.45 0.00	0.41 0.00	0.39 0.00	0.36 0.00
2	1.00 0.16	0.99 0.09	0.93 0.07	0.85 0.05	0.78 0.04	0.71 0.04	0.65 0.03	0.60 0.03	0.56 0.03	0.52 0.02	0.48 0.02	0.45 0.02	0.43 0.02
3	1.00 0.29	0.99 0.19	0.95 0.15	0.88 0.12	0.82 0.10	0.76 0.09	0.70 0.08	0.65 0.07	0.61 0.06	0.57 0.06	0.54 0.05	0.51 0.05	0.48 0.04
4	1.00 0.40	1.00 0.29	0.96 0.22	0.90 0.18	0.84 0.16	0.79 0.14	0.74 0.12	0.69 0.11	0.65 0.10	0.61 0.09	0.58 0.08	0.55 0.08	0.52 0.07
5	1.00 0.48	1.00 0.36	0.96 0.29	0.92 0.25	0.86 0.21	0.81 0.19	0.77 0.17	0.72 0.15	0.68 0.14	0.65 0.13	0.62 0.12	0.59 0.11	0.56 0.10
6	1.00 0.54	1.00 0.42	0.97 0.35	0.93 0.30	0.88 0.26	0.83 0.23	0.79 0.21	0.75 0.19	0.71 0.18	0.68 0.16	0.65 0.15	0.62 0.14	0.59 0.13
7	1.00 0.59	1.00 0.47	0.97 0.40	0.93 0.35	0.89 0.31	0.85 0.28	0.81 0.25	0.77 0.23	0.73 0.21	0.70 0.20	0.67 0.18	0.64 0.17	0.62 0.16
8	1.00 0.63	1.00 0.52	0.98 0.44	0.94 0.39	0.90 0.35	0.86 0.32	0.82 .029	0.79 0.27	0.75 0.25	0.72 0.23	0.69 0.22	0.67 0.20	0.64 0.19
9	1.00 0.66	1.00 0.56	0.98 0.48	0.95 0.43	0.91 0.39	0.87 0.35	0.84 0.32	0.80 0.30	0.77 0.28	0.74 0.26	0.71 0.24	0.69 0.23	0.66 0.22
10	1.00 0.69	1.00 0.59	0.98 0.52	0.95 0.46	0.92 0.42	0.88 0.38	0.85 0.35	0.82 0.33	0.79 0.31	0.76 0.29	0.73 0.27	0.70 0.26	0.68 0.24
11	1.00 0.72	1.00 0.62	0.98 0.57	0.95 0.49	0.92 0.45	0.89 0.41	0.86 0.38	0.83 0.36	0.80 0.34	0.77 0.32	0.74 0.30	0.72 0.28	0.69 0.27
12	1.00 0.74	1.00 0.64	0.98 0.57	0.96 0.52	0.93 0.48	0.90 0.44	0.87 0.41	0.84 0.38	0.81 0.36	0.78 0.34	0.76 0.32	0.73 0.31	0.71 0.29

جدول 9: التوزيع الطبيعي المعياري

The Standard Normal Distribution

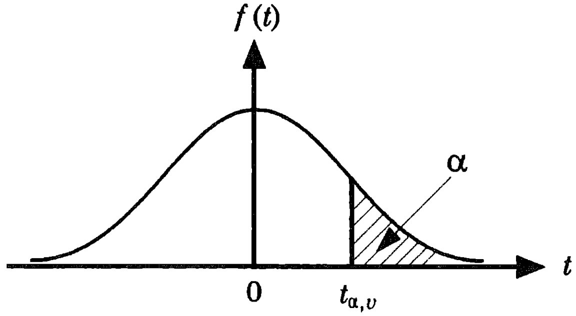
القيم الحرجة z_α للتوزيع:
$$\alpha = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_\alpha} e^{-t^2/2}\, dt$$



α	z_α	α	z_α	α	z_α
0.50	0.000	0.68	0.468	0.86	1.080
0.51	0.025	0.69	0.496	0.87	1.126
0.52	0.050	0.70	0.524	0.88	1.175
0.53	0.075	0.71	0.553	0.89	1.227
0.54	0.100	0.72	0.583	0.90	1.282
0.55	0.126	0.73	0.613	0.91	1.341
0.56	0.151	0.74	0.643	0.92	1.405
0.57	0.176	0.75	0.674	0.93	1.476
0.58	0.202	0.76	0.706	0.94	1.555
0.59	0.228	0.77	0.739	0.95	1.645
0.60	0.253	0.78	0.772	0.96	1.751
0.61	0.279	0.79	0.806	0.97	1.881
0.62	0.305	0.80	0.842	0.98	2.054
0.63	0.332	0.81	0.878	0.99	2.326
0.64	0.358	0.82	0.915	0.999	3.090
0.65	0.385	0.83	0.954	0.9999	3.720
0.66	0.412	0.84	0.994	0.99999	4.265
0.67	0.440	0.85	1.036		

The t Distribution جدول 10: توزيع t

$$\alpha = \int_{t_{\alpha,\nu}}^{\infty} \frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\sqrt{\pi\upsilon}\,\Gamma\left(\frac{\upsilon}{2}\right)} \left(1+\frac{t^2}{\upsilon}\right)^{-\frac{\upsilon+1}{2}} dt$$ قيم $t_{\alpha,\nu}$:



VALUES OF $t_{\alpha,\nu}$ *

ν \ α	.10	.05	.025	.01	.005	ν
1	3.078	6.314	12.706	31.821	63.657	**1**
2	1.886	2.920	4.303	6.965	9.925	**2**
3	1.638	2.353	3.182	4.541	5.841	**3**
4	1.533	2.132	2.776	3.747	4.604	**4**
5	1.476	2.015	2.571	3.365	4.032	**5**
6	1.440	1.943	2.447	3.143	3.707	**6**
7	1.415	1.895	2.365	2.998	3.499	**7**
8	1.397	1.860	2.306	2.896	3.355	**8**
9	1.383	1.833	2.262	2.821	3.250	**9**
10	1.372	1.812	2.228	2.764	3.169	**10**
11	1.363	1.796	2.201	2.718	3.106	**11**
12	1.356	1.782	2.179	2.681	3.055	**12**
13	1.350	1.771	2.160	2.650	3.012	**13**

تابع جدول 10

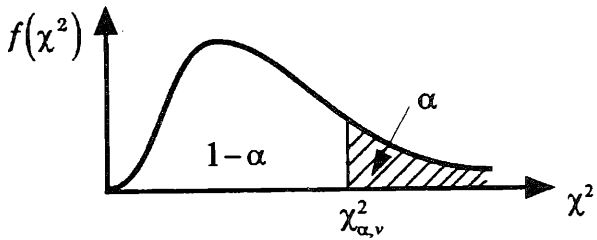
ν \ α	.10	.05	.025	.01	.005	ν
14	1.345	1.761	2.145	2.624	2.977	**14**
15	1.341	1.753	2.131	2.602	2.947	**15**
16	1.337	1.746	2.120	2.583	2.921	**16**
17	1.333	1.740	2.110	2.567	2.898	**17**
18	1.330	1.734	2.101	2.552	2.878	**18**
19	1.328	1.729	2.093	2.539	2.861	**19**
20	1.325	1.725	2.086	2.528	2.845	**20**
21	1.323	1.721	2.080	2.518	2.831	**21**
22	1.321	1.717	2.074	2.508	2.819	**22**
23	1.319	1.714	2.069	2.500	2.807	**23**
24	1.318	1.711	2.064	2.492	2.797	**24**
25	1.316	1.708	2.060	2.485	2.787	**25**
26	1.315	1.706	2.056	2.479	2.779	**26**
27	1.314	1.703	2.052	2.473	2.771	**27**
28	1.313	1.701	2.048	2.467	2.763	**28**
29	1.311	1.699	2.045	2.462	2.756	**29**
inf.	1.282	1.645	1.960	2.326	2.576	**inf.**

* This table is abridged from Table IV of R. A. Fisher, *Statistical Methods for Research Workers,* published by Oliver and Boyd, Ltd., Edinburgh, by permission of the author and publishers.

The Che-Square Distribution χ^2 جدول 11: توزيع

القيم الحرجة $\chi^2_{\alpha,\nu}$ للتوزيع:

$$\alpha = \int_{\chi^2_{\alpha,\upsilon}}^{\infty} f(\chi^2)\,d(\chi^2) = \frac{1}{2^{\upsilon/2}\Gamma(\upsilon/2)} \int_{\chi^2_{\alpha,\upsilon}}^{\infty} f(\chi^2)^{\frac{\upsilon}{2}-1} e^{-\chi^2/2}\,d(\chi^2)$$



VALUES OF $\chi^2_{\alpha,\nu}$ *

ν \ α	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005	ν
1	.0000393	.000157	.000982	.00393	3.841	5.024	6.635	7.879	1
2	.0100	.0201	.0506	.103	5.991	7.378	9.210	10.597	2
3	.0717	.115	.216	.352	7.815	9.348	11.345	12.838	3
4	.207	.297	.484	.711	9.488	11.143	13.277	14.860	4
5	.412	.554	.831	1.145	11.070	12.832	15.086	16.750	5

تابع جدول 11

ν \ α	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005	ν
6	.676	.872	1.237	1.635	12.592	14.449	16.812	18.548	**6**
7	.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278	**7**
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955	**8**
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589	**9**
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188	**10**
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757	**11**
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300	**12**
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819	**13**
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319	**14**
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801	**15**
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267	**16**
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718	**17**
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156	**18**
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582	**19**
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997	**20**
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401	**21**
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796	**22**
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181	**23**

تابع جدول 11

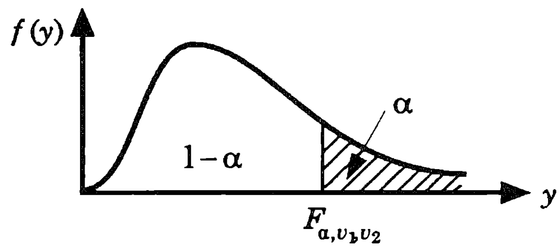
ν \ α	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005	ν
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558	**24**
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928	**25**
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290	**26**
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645	**27**
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993	**28**
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336	**29**
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672	**30**

* This table is based on Table 8 of *Biometrika Tables for Statisticians, Volume I,* by permission of the *Biometrika* trustees.

The F Distribution — جدول 12: توزيع F

القيم الحرجة F_{α, ν_1, ν_2} للتوزيع:

$$\alpha = \int_{F_{\alpha,\upsilon_1,\upsilon_2}}^{\infty} \frac{\Gamma\left(\frac{\upsilon_1+\upsilon_2}{2}\right)}{\Gamma\left(\frac{\upsilon_1}{2}\right)\Gamma\left(\frac{\upsilon_2}{2}\right)} \left(\frac{\upsilon_1}{\upsilon_2}\right)^{\frac{\upsilon_1}{2}} y^{\frac{\upsilon_1}{2}-1}\left(1+\frac{\upsilon_1}{\upsilon_2}y\right)^{-\frac{1}{2}(\upsilon_1+\upsilon_2)} dy$$



VALUES OF $F_{.01,\upsilon_1,\upsilon_2}$ *

ν_1 = Degrees of freedom for numerator

ν_2 = Degrees of freedom for denominator

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4.051	5.000	5.403	5.625	5.764	5.859	5.928	5.982	6.023	6.056	6.106	6.157	6.209	6.235	6.261	6.287	6.313	6.339	6.366
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1

تابع جدول 12

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.9	13.8	13.7	13.7	13.6	13.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	5.86
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	2.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.56	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	2.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26

تابع جدول 12

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.53	2.45	2.36	2.27	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

* This table is reproduced from M. Merrington and C. M. Thompson, "Tables of percentage points of the inverted beta (F) distribution," *Biometrika*, Vol. 33 (1943), by permission of the *Biometrika* trustees.

جدول 12: توزيع F *The F Distribution*

قيم الدالة $F_{0.05,\nu_1,\nu_2}$:

$$\alpha = \int_{F_{0.05,\nu_1,\nu_2}}^{\infty} F_{\nu_1,\nu_2}(y)\,dy =$$

$$= \left(\frac{\upsilon_1}{\upsilon_2}\right)^{\frac{\upsilon_1}{2}} \frac{\Gamma\left(\frac{\upsilon_1+\upsilon_2}{2}\right)}{\Gamma\left(\frac{\upsilon_1}{2}\right)\Gamma\left(\frac{\upsilon_2}{2}\right)} \int_{F_{0.05,\nu_1,\nu_2}}^{\infty} y^{\frac{\nu_1}{2}-1}\left(1+\frac{\upsilon_1}{\upsilon_2}y\right)^{-\frac{(\upsilon_1+\upsilon_2)}{2}} dy$$

VALUES OF $F_{0.05,\upsilon_1,\upsilon_2}$ *

υ_1 = Degrees of freedom for numerator

υ_2 = Degrees of freedom for denominator

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63

تابع جدول 12

$\nu_2 \backslash \nu_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.21	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76

تابع جدول 12

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

* This table is reproduced M. Merrington and C. M. Thompson, "Tables of percentage points of the inverted bats (F) distribution," Biometrika, Vol. 33 (1943), by permission of the Biometrika trustees.