تحليل البيانات
باستعمال البرنامج ألاحصائي
SAS

Version 6.12

أمثلة تطبيقية مع الحلول في العلوم
الطبية
الحياتية
البيطرية
الزراعية

الدكتور فراس رشاد السامرائي
فرع الصحة العامة البيطرية / كلية الطب البيطري
جامعة بغداد

2009

بسم الله الرحمن الرحيم

(وَيَسْـئَلونَكَ عَـنِ الـرُّوحِ قُـلِ الـرُّوحُ مِـنْ أَمْرِ رَبِّي وَمَا أُوتِيتُمْ مِنَ الْعِلْمِ إلا قَلِيلاً)

صدق الله العظيم
سورة الاسراء
الاية 85

الأهداء

الى .....

من يعد شهادة الدكتوراه أول خطوة في سلم العلم....

من يؤمن بان اشاعة العلم هو اسمى رسالة.....

من يدرك بأن العلم والتواضع من أعلى المراتب.....

فراس

	الموضوع	الصفحة
	المقدمة	
الفصل الاول	تنصيب برنامج SAS (6.12) على الحاسوب	1
1 – 1	خطوات التنصيب	1
2 – 1	معلومات اولية عن الصفحة الرئيسية لبرنامج SAS	1
3 – 1	ترتيب البيانات	2
4 – 1	القيم المفقودة	3
5 – 1	تحويل متغير مستمر الى فئات	4
6 – 1	حذف قيم من البيانات	5
7 – 1	دمج الملفات	6
الفصل الثاني	بعض الأوامر المهمة في برنامج SAS	10
1 – 2	الأمر Proc univariate	10
2 – 2	الأمر Proc means	11
3 – 2	الأمر Proc print	12
4 – 2	الأمر Proc sort	14
5 – 2	الأمر Proc transpose	16
6 – 2	الأمر Proc ttest	17
7 – 2	الأمر proc corr	18
8 – 2	الأمر proc chart	19
9 – 2	الآمر proc tabulate	20
10 – 2	الأمر proc npar1way	21
11 – 2	الأمر proc freq	21
12 – 2	الأمر proc anova	21
13 – 2	الأمر proc glm	21

2 – 14	الأمر proc vercomp	21

المحتويات

	الموضوع	**الصفحة**
2 – 15	الأمر proc mixed	21
2 – 16	الأمر proc logestic	22
2 – 17	الأمر proc plot	22
2 – 18	الأمر proc reg	22
2 – 19	الأمر proc lifetest	22
2 – 20	الأمر proc genmod	22
2 – 21	الأمر proc catmod	22
2 – 22	الأمر proc probit	22
2 – 23	الأمر proc phreg	22
الفصل الثالث	اختبار الفرضيات	23
3 – 1	الاختبارات المعلمية	23
3 – 1 – 1	اختبار التوزيع الطبيعي للبيانات	23
3 – 1 – 2	اختبار تجانس التباين	25
3 – 1 – 3	الاختبارات المحددة قبل اجراء التجربة	28
3 – 1 – 3 – 1	المقارنات المستقلة	28
3 – 1 – 3 – 2	تحليل الاتجاه بأستعمال العوامل متعددة الحدود	30
3 – 1 – 4	الاختبارات المقترحة بعد اجراء التجربة	36
3 – 1 – 4 – 1	مقارنة جميع متوسطات المعاملات بمتوسط معاملة المقارنة	36
أ	اختبار Dunnett	36
3 – 1 – 4 – 2	المقارنات المتعددة	36
أ	أختبار أقل فرق معنوي LSD او اختبار t	36

ب	اختبار Tukey	41
ج	اختبار Duncan	41
	الموضوع	**الصفحة**
د	اختبار Scheffe	41
و	اختبار Newman-Keul	41
3 – 2	الاختبارات اللامعلمية	41
أ	اختبار مربع كاي لحسن المطابقة	42
ب	اختبار فيشر المضبوط	47
ج	أختبار الاستقلال اوالاقتران	48
د	اختبار مربع كاي في حالة اكثر من عينتين مستقلتين	52
هـ	اختبار كولمجروف - سيمينوف	53
و	اختبارالعلامة واختبارولكاكسن للمرتبة وعلامتها	58
ز	اختبار ولكاكسن ومان ووتني لعينتين	59
حـ	احصائية كابا	61
ط	اختبار كرسكال ووالس لعدة عينات	62
ي	اختبار P المضبوط	67
ك	اختبار ماكنيمار	70
ل	اختبار الوسيط	72
م	اختبار فان دير فيردن	72
ن	اختبار سافيج	73
س	اختبار كوجران	74
3 – 3	مقاييس الاقتران	75
3 – 3 – 1	معامل سبيرمان لأرتباط الرتب	75
الفصل الرابع	تحليل التباين في التصاميم المختلفة	77
4 – 1	التصميم العشوائي الكامل	77

2 – 4	تصميم القطاعات العشوائية الكاملة	85
3 – 4	التصميم العشوائي الكامل في التجارب العاملية	87
	الموضوع	**الصفحة**
1 – 3 – 4	تأثير عاملين بمشاهدة واحدة لكل عامل	87
2 – 3 – 4	تأثير عاملين ولأكثر من مشاهدة لكل عامل	89
3 – 3 – 4	دراسة تأثير عاملين مع التداخل	95
4 – 4	التصميم المحتضن	105
5 – 4	تصميم المربع اللاتيني	110
6 – 4	تحليل التباين المشترك	115
الفصل الخامس	الانحدار	118
1 – 5	الانحدار البسيط	118
2 – 5	الانحدار المتعدد	120
1 – 2 – 5	طرق اختيار احسن المعادلات	122
1 – 1 – 2 – 5	كل الانحدارات الممكنة	122
2 – 1 – 2 – 5	الانحدار التدريجي	125
3 – 1 – 2 – 5	الاختيارالامامي او المباشر	129
4 – 1 – 2 – 5	الحذف العكسي	132
5 – 1 – 2 – 5	اختيار افضل معادلة اعتمادا على قيمة مالو C_p	132
6 – 1 – 2 – 5	اختيارافضل معادلة اعتمادا على معامل التحديد R^2	132
3 – 5	الانحدار اللاخطي	132
4 – 5	الانحدار اللوحستيكي	135
1 – 4 – 5	العامل المستقل والتابع ثنائي	136
2 – 4 – 5	العامل المستقل ثنائي والعامل التابع نسبة	137
3 – 4 – 5	العوامل المستقلة والمعتمدة الاسمية ذات الفئات	138
4 – 4 – 5	العامل المعتمد يكون بهيئة نسبة	162

5 – 5	الانحدار البواسوني	183
5 – 6	الانحدار التشخيصي	185

	الموضوع	الصفحة

الفصل السادس	الارتباط	196
6 – 1	الارتباط البسيط	196
6 – 2	الارتباط المتعدد	196
الفصل السابع	نسبة الارجحية والخطورة النسبية	198
7 – 1	تقدير نسبة الارجحية والخطورة النسبية	199
الفصل الثامن	تقدير مكونات التباين	268
8 – 1	تقدير المكافىء الوراثي	268
8 – 1 – 1	طرق تقدير المكافىء الوراثي	268
8 – 1 – 1 – 1	تقدير المكافىء الوراثي للأخوة انصاف الاشقاء من مكونات تباين الأب في النماذج الخطية	268
8 – 2	الارتباط الوراثي	279
8 – 3	الارتباط البيئي	279
8 – 4	الارتباط المظهلري	280
8 – 5	تقدير قيم BLUP	280
8 – 1 – 1 – 2	تقدير المكافىء الوراثي عن طريق الأب او الام او كليهما في حالة النموذج المحتضن	286
أ	في حالة تساوي عدد النسل داخل كل مجموعة	286
ب	في حالة عدم تساوي عدد افراد النسل داخل المجاميع	289
8 – 1 – 1 – 3	تقدير المكافىء الوراثي بطريقة انحدار الأبناء على الآباء	294
أ	طريقة انحدار متوسط الابناء على احد الاباء	294
ب	طريقة انحدار الابناء على الام ضمن الأب	298
8 – 6	النماذج المختلطة	300

الفصل التاسع	تحويل البيانات	308
9 – 1	التحويل اللوغاريتمي	308
9 – 2	التحويل الجذري	308
9 – 3	التحويل الزاوي	320
الفصل العاشر	اختبارات التشخيص الطبية	321
10 – 1	تحليل اختبارات التشخيص	321
10 – 2	وصف اختبارات التشخيص	322
10 – 3	تقدير احتمالية الاصابة	322
10 – 4	مقارنة اختبارين للتشخيص	324
الفصل الحادي عشر	تقدير العمر	337
11 – 1	تقدير العمر بأستعمال طريقة تحليل البقاء	337
11 – 2	تقدير الزمن في البيانات المراقبة	345
الفصل الثاني عشر	تقدير قيمة LD50	349
الفصل الثالث عشر	امثلة متنوعة	359

	الموضوع	**الصفحة**

المقدمة

ان التقدم الكبير الذي شهده العالم وفي شتى العلوم يتطلب العمل بهمة اكبرومثابرة أعلى لغرض مواكبته والتواصل معه ، وتمثل البحوث العلمية حجر الزاوية في اسباب ذلك التقدم ، اذ ان القيمة العلمية للبحث تستند الى:

1- تصميم التجربة

2- اختيار التحليل الاحصائي المناسب

3- اختيار الاختبار المناسب

4- تطبيق البرامج الاحصائية الحديثة

هناك اتفاق على وجود علاقات مشتركة بين العلوم المختلفة ، الا ان العلاقة بين تلك العلوم وعلم الاحصاء وتطبيقاته بأستعمال الحاسوب لها خصوصية اكبر واهمية اعظم ، اذ ان الاحصاء وتطبيقاته يمثل قاسما مشتركا لتلك العلوم.

ان زيادة المعرفة العلمية للباحث تعد امرا ضروريا للنهوض بالمستوى العلمي واحد اهم هذه الوسائل هي معرفة طرق التحليل الاحصائي ، اذ ان ذلك سيؤدي الى تفعيل دور الباحث من خلال زيادة قدرته على دراسة تأثير العوامل المختلفة في صفة ما بصورة افضل وادق مقارنة بغيره لأنه ادرى بأختصاصه كما يجعله اكثر قدرة على مواكبة التطورات العلمية والابتعاد عن النمطية عند اجراء البحوث العلمية.

ان صعوبة التحليل الاحصائي تكمن في كونه علما توليفيا اذ يجمع بين الاحصاء وتصميم التجارب والحاسوب ، وهذه الصعوبة تنسحب على الكاتب والقارىء ، وقد حاولت في كتابي هذا ان اسير بتوجس وان اقف على بعض النقاط فأوضحها دون اسهاب حتى لا اوهن رغبة القارىء في الاستمرار بقراءة الكتاب وان انتقي امثلة متنوعة لكي اغطي عدة اختصاصات وبسيطة لكي يسهل فهمها وبذلك اضفي على محاولتي هذه صفة العموم اي زيادة احتمالات الاستفادة من الكتاب من قبل اكبرعدد ممكن من الباحثين ومن اختصاصات مختلفة.

هنالك العديد من البرامج الاحصائية التي يجري استعمالها مثل برنامج Stata او SPSS او Minitab او SAS وغيرها ، واذا كانت هناك اصدارات عن بعض هذه البرامج لاسيما برنامج SPSS مثل كتاب (دليلك الى البرنامج الاحصائي SPSS) اعداد السيد سعد زغلول بشير، فأن ماصدر عن برنامج SAS لايتعدى سوى اصدارا واحدا فقط من اعداد الدكتور نصر نوري الانباري / كلية الزراعة / جامعة بغداد ، ورغم اهميته الا انه لم يتوفر الا بأعداد محدودة

ولم يتضمن الا بضعة امثلة انحصرت في الجانب الزراعي فقط . لذا فأن الحاجة الى كتاب شامل بقيت قائمة نظرا للأهمية الكبيرة لهذا البرنامج والتي تفوق البرامج المذكورة والمتمثلة في كثرة اعتماده في التحليلات الاحصائية وفي مختلف الاختصاصات الطبية والهندسية والبيطرية والزراعية والعلوم الانسانية ، وبناءا على ذلك حاولنا اصدار هذا الكتاب عسى ان يكون دليلا مهما للباحثين وخطوة اولى في زيادة معارفهم ومهاراتهم في التحليل الاحصائي.

ان برنامج SAS هو مختصر لـ (Statistical Analysis System) اذ نفذ وطور من قبل مجموعة من الباحثين في معهد SAS في الولايات المتحدة الامريكية ، وبيعت اول نسخ له عام 1976 ويجرى عليه تطويرمستمر حتى وصلت آخرنسخة معدلة منه الى version 9.2 .

يستعمل هذا البرنامج لاجراء معظم التحليلات الاحصائية ولجميع التصاميم (التصميم العشوائي الكامل ، تصميم القطاعات العشوائية الكاملة ، المربع اللاتيني فضلا عن تصاميم التجارب العاملية للانواع الثلاث المذكورة). كما يمكن استعماله لاجراء مختلف الاختبارات الاحصائية المعلمية واللامعلمية مثل T-Test (Least significance Differences) واختبار Duncan واختبار Tukey واختبار مربع كاي واختبار ولكاكسن وماكنيمار وتقدير حدود الثقة ، كما يستعمل ايضا لتقديرمعامل الارتباط والانحدار البسيط والمتعدد والانحدار اللوجستيكي وتقدير نسبة الخطورة (Risk Ratio) ونسبة الارجحية (Odds Ratio) واختبارات التشخيص فضلا عن امكانية الحصول على رسوم بيانية.كما يمكن اعتماده لغرض تقدير المكافيء الوراثي والارتباطات الوراثية والبيئية والمظهرية علاوة على تقدير قيم BLUP الخاصة بدراسات التقييم الوراثي للحيوانات.

ان هذا الكتاب يمثل خطوة متواضعة لتوفير مصدر ذواسلوب مبسط للباحثين يتيح لهم امكانية اجراء التحليلات الاحصائية ، اذ ضمن الكتاب مجموعة امثلة وفي اختصاصات مختلفة مع النتائج بما يجعل من الباحث اكثر قدرة على اكتساب المهارة في التحليل الاحصائي ، وانني اذ انهي كتابي هذا فأن الامل يحدوني بأن اكون موفقا في مسعاي ولكني وفي الوقت نفسه أأمل ان يساعدني الباحثون من خلال ملاحظاتهم واقتراحاتهم والتي ستساهم بلاشك في زيادة القيمة العلمية للكتاب وذلك بارسالها الى بريدي الالكتروني firas_rashad@yahoo.com ولهم مني خالص شكري وتقديري.

اخيرا يطيب لي ان اتقدم بجزيل شكري وامتناني لكل من مد لي يد العون والمساعدة من زملائي في كلية الطب البيطري وأخص بالذكر منهم الدكتور غالب علوان القيسي استاذ تغذية الدواجن لتشجيعه لي ولأشرافه المباشر على طبع الكتاب وكذلك الدكتور علي حسين الهلالي لتزويده لي ببعض المصادروالدكتور يحيى خالد التميمي والدكتور غسان يوسف بطرس لتشجيعها لي على

انجاز الكتاب ولايفوتني ان اقدم شكري وتقديري الى زملائي في كلية الزراعة/ جامعة بغداد كل من الدكتور نصر نوري الانباري استاذ تربية وتحسين الحيوان والدكتور احمد محمود النداوي استاذ تربية وتحسين الدواجن.

ومن الله التوفيق

الدكتورفراس رشاد السامرائي
فرع الصحة العامة البيطرية
كلية الطب البيطري / جامعة بغداد
آب / 2009

الفصل الأول

تنصيب برنامج SAS (6.12) على الحاسوب

1 – 1 خطوات التنصيب

لغرض تنصيب البرنامج على الحاسوب يجب اتباع الخطوات التالية:

1- يعدل تاريخ الحاسوب الى كانون الثاني 2000 بالنسبة لنسخة البرنامج المرفقة مع هذا الكتاب اذ ان الشركة تعطي لكل نسخة تاريخ لضمان بيع نسخ جديدة من البرنامج.
2- ادخال قرص البرنامج وبعد اكتمال عملية التحميل ستظهر على الشاشة نافذة فيها عدة اختيارات هنا يجب التأشيرعلى الاختيار الاول (set up).
3- يؤشر على Next للنوافذ المتتابعة.
4- تظهر نافذة فيها عدة اختيارات فنختار costum
5- تظهر بعد ذلك نافذة تحوي عدة مربعات تؤشر جميعها بعلامة صح (اي كلك)☑.
6- وبعد ذلك تظهر نافذة توضح ان البرنامج بدأ بالتحميل
7- تظهر بعد ذلك نافذة تشير الى ان التنصيب تم في C ونجد في نفس الصفحة Browse فنبحث عن فايل disk لغرض اختياره.
8- ثم بعد ذلك الضغط على كلمة yes لحين ظهور الصفحة الرئيسية للبرنامج.

1 – 2 معلومات اولية عن الصفحة الرئيسية لبرنامج SAS

يشتمل البرنامج على ثلاثة صفحات رئيسية هي:

Dr. Firas

أ- صفحة محرر البرنامج Program Editor

ب- صفحة المخرجات Output

ج- صفحة Log

في صفحة محرر البرنامج تكتب البيانات المراد تحليلها وينفذ التحليل بالنقر على ايقونة التنفيذ في شريط المهام والتي تمثل شخص يهرول فاذا كان ترتيب البيانات صحيحا تظهر النتائج في صفحة المخرجات اما اذا كان هناك خطأ في الايعازات فلاتظهر نتائج ولغرض معرفة سبب عدم التنفيذ ننقر على الاختيار Globals في شريط المهام الخاص بصفحة البرنامج ونختار Log فستظهز عبارات باللون الاحمر تشير الى الخطأ في الايعازات ، وفي بعض الحالات يجري تنفيذ لجزء من الايعازات مما يعني وجود خطأ لدينا في بعض الايعازات ، ولغرض تصليح قطعة البرنامج ننقر على ايقونة تمثل صفحة ورقة فارغة ثم نضغط على ايقونة التنفيذ (شخص يهرول) ثم F4 فيتم استدعاء البيانات الاصلية لغرض اجراء التعديلات على الايعازات ومعاودة التنفيذ. ويمكن استدعاء البيانات من ايقونة Locals ثم Recall text .

عند ظهور النتائج يمكن خزنها وذلك بالضغط على ايقونة File ثم Save واعطاء اسم للملف.

1 – 3 ترتيب البيانات

يمكن ترتيب البيانات بصورة مباشرة من خلال طبعها في النصف الاسفل من الصفحة الرئيسية للبرنامج ويكون ترتيب العوامل عموديا كما يمكن ترتيبها افقيا ونضع العلامة @@ لكي يقوم البرنامج بقراءة البيانات بصورة افقية، كما يمكن طباعتها في برنامج الاكسل ومن ثم حفظها على save as فيظهر مربع حوار فنقوم باختيار Text (MS-DOS)ومن ثم نعطي اسم للملف لغرض قرائتها باستعمال برنامج SAS. وفي حالة عدم امكانية فتحها باستعمال برنامج SAS نختار من شريط المهام Edit ثم Copy ثم نفتح البرنامج ونعمل Paste في شاشة SAS.

ان تنفيذ اي برنامج يتطلب توفر شروط يجب الالتزام بها وبدونها لايجري التنفيذ تتمثل بما يلي:-

في بداية البرنامج نكتب كلمة Data ونعطي حرف او اسم على ان لايزيد عن 8 حروف ثم فارزة منقوطة ;. مثل Data D; او Data Ali ; وغيرها.

نكتب كلمة Input ونعطي اسماء للمتغيرات لاتتجاوز 8 حروف علما بأن المتغير عندما يكون وصفي مثل كبير ، متوسط ، صغير اومجموعة اسماء ، فيجب وضع علامة $ بعد الرمز الخاص بالمتغير الوصفي حتى يقرأ من قبل البرنامج وفي النهاية نضع فارزة منقوطة ; وفي الاسطر التي تلي كلمة Input يمكن كتابة اي معادلات خاصة بالمتغيرات لغرض اضافتها للبيانات ويعبر عن العمليات الرياضية في البرنامج كالاتي:

لو كان لدينا متغيران هما X, Y:

B =x*x; للدلالة على الضرب

C=(3*Y)/2; للدلالة على الضرب والقسمة

D=log(y); للدلالة على التحويل اللوغاريتمي

C= X**2; للدلالة على الاس التربيعي

G= Y+Y; للدلالة على الجمع

H= Y-X; للدلالة على الطرح

نكتب Cards ثم فارزة منقوطة ; .

نطبع البيانات بحيث ان كل متغير تكون قيمه عموديا وفي النهاية نضع فارزة منقوطة ; .

نختار الامر الخاص بالتحليل الذي يبدأ عادة بـ Proc وتظاف اليه كلمة تختلف حسب التحليل المراد الحصول عليه ونضع فارزة منقوطة ; ثم نختم البرنامج بـ ;run او quit.

ملاحظة مهمة : ان اهمال وضع اي من هذه الشروط سيؤدي الى عدم تنفيذ البرنامج.

1 – 4 القيم المفقودة

في كثير من البيانات قد تفقد قيم لاي سبب من الاسباب، ولكي يستطيع البرنامج التعرف على القيمة المفقودة فيجب التعبير عنها بـ (.) dot.

مثال (1) : البيانات ادناه تمثل قيم ثلاثة صفات وبعض القيم مفقودة ، جد متوسط كل صفة؟

```
Data f ;
 Input id trial1 trial2 trial3 ;
Cards ;
1 1.5  1.4  1.6
2 1.5  .  1.9
3 .  2.0  1.6
4 .  .  2.2
5 2.1  2.3  2.2
6 1.8  2.0  1.9
Option nodate nonumber;
 Proc means ;
 Run ;
```

The SAS System

Variable	N	Mean	Std Dev	Minimum	Maximum
ID	6	3.5000000	1.8708287	1.0000000	6.0000000
TRIAL1	4	1.7250000	0.2872281	1.5000000	2.1000000
TRIAL2	4	1.9250000	0.3774917	1.4000000	2.3000000
TRIAL3	6	1.9000000	0.2683282	1.6000000	2.2000000

ويمكن للقارىء ان يهمل متعمدا كتابة احدى النقاط (.) وينفذ البرنامج سيجد ان النتائج ستكون مختلفة وخاطئة.

ملاحظة: اذا كانت البيانات على شكل تاريخ يوم وشهر وسنة ونريد ازالة الخط المائل:

2/3/99

Dr. Firas

13/11/00

27/9/98

نقوم بتظليل البيانات ثم ننقر على Edit في شريط المهام ونختار Replace فيظهر صندوق حوار نضع الخط المائل / امام الحقل المؤشر امامه Find what ونضع المؤشر على الاختيار الثاني Replace with وننقر على space من لوحة مفاتيح الحاسبة ، ثم ننقر على كلمة Replace all فيختفي الخط المائل او اجراء نفس الخطوات من على صفحة البرنامج.

1 – 5 تحويل متغير مستمر الى فئات

يمكن ايضا استبعاد بعض القيم لمتغير ما من التحليل فمثلا اذا اردنا استبعاد قيم المتغير B التي تزيد عن 44 واستبعاد القيم التي تقل عن 5 ، نكتب الايعازين التاليين بعد العبارة التي تتضمن المتغيرات في السطر الذي يلي كلمة input.

```
If B gt  44 then delete;
If B lt  5 then delete;
```

اذ ان gt تعني اكبر من و lt تعني أقل من.

كثيرا مانحتاج في التحليل الاحصائي الى تحويل متغير ما الى فئات لتقليل عدد مستوياته او للحصول على جدول تحليل التباين ، هنا نحاول ترتيب الفئات من خلال الايعازات بحيث يكون الفرق بين اعلى وادنى قيمة لكل فئة متساوي في جميع الفئات ولكن نحن مخيرون بالفرق الذي نريده كأن يكون 5 او 7 او 3 وذلك حسب طبيعة توزيع قيم المتغير اي نحاول ترتيب البيانات بحيت تكون اعداد المشاهدات لكل فئة متقاربة ضمن كل صف.

مثال (2): لوكان لدينا متغير y له قيم مختلفة واردنا تحويله الى ثلاث فئات ذات عدد متقارب:

```
data s;
input  y;
If y lt 141 then y1=1;
If y gt 140 and y lt 161 then y2=2;
If y gt 160 then y3=3;
cards;
 122
 141
 170
 144
 145
 161
 159
 155
 157
 177
 167
 161
 162
option nodate nonumber;
proc means n;var y1 y2 y3; run;
```

```
The SAS System
  Variable  N
  ----------
  Y1        1
  Y2        6
  Y3        6
  ----------
```

يمكن الحصول على نفس المجموعات بأستعمال الايعاز التالي :

```
If y <=140 then y=1;
else if y <=160 then y=2;
else y=3;
```

يلاحظ من النتائج ان المجموعات الثلاث غير متقاربة في اعدادها لذا يجب اجراء بعض التغيير على الايعازات الخاصة بتحديد حجم كل مجموعة ومعاودة التنفيذ:

```
If y lt 151 then y1=1;
If y gt 150 and y lt 162 then y2=2;
If y gt 161 then y3=3;
```

مثال (3) : لو اردنا تحويل قيم المثال التالي الى لوغاريتيم واضافة قيمه كمتغير جديد:

```
Data f;
Input male female;
M=log(male);
F=log(female);
Cards;
22  32
43  56
76  33
78  54
Option nodate nonumber;
Proc print;
Run;
```

```
                                The SAS System
          OBS   MALE   FEMALE      M        F

           1     22      32     3.09104  3.46574
           2     43      56     3.76120  4.02535
           3     76      33     4.33073  3.49651
           4     78      54     4.35671  3.98898
```

1 – 6 حذف قيم من البيانات

في بعض الحالات قد يرغب الباحث في حذف القيم المتطرفة من البيانات.

مثال (4): احذف القيم التي تساوي او تقل عن 140؟

```
data s;
input  y;
If y <=140 then delete;
cards;
 122
 140
 170
 144
 125
 161
```

Dr. Firas

```
 159
option nodate nonumber;
proc print;
run;
```

OBS	Y
1	170
2	144
3	161
4	159

1– 7 دمج الملفات Merge of files

قد يحتاج الباحث الى دمج ملفين او اكثر لأغراض التحليل الاحصائي ويمكن اجراء ذلك بأستعمال الايعازات الموضحة في المثال الآتي:

مثال (5): ادمج الملف dd مع الملف ss ؟ (يوجد بينهما متغير مشترك).

```
data dd;
input a b c  ;
cards;
11 3 7
2  5 8
3  1 4
9  2 6
;
data ss;
input c f g n;
cards;
1 9 7 33
4 7 2 43
5 8 3 21
6 6 1 11
data growth; set dd ss; run;
proc print data=growth; run;
```

OBS	A	B	C	F	G	N
1	11	3	7	.	.	.
2	2	5	8	.	.	.
3	3	1	4	.	.	.
4	9	2	6	.	.	.
5	.	.	1	9	7	33
6	.	.	4	7	2	43
7	.	.	5	8	3	21
8	.	.	6	6	1	1

مثال (6) يمكن دمج ملفين وكل ملف يحوي متغيرات مختلفة عن الآخر الا ان كلا الملفين يعودان لصفة مشتركة:

```
DATA baseline;
INPUT id sex $ age inc;
CARDS;
 1  F  35 17
```

Dr. Firas

```
17  M  50 14
33  F  45  6
49  M  24 14
65  F  52  9
81  M  44 11
2   F  34 17
18  M  40 14
34  F  47  6
50  M  35 17
;
DATA rating;
INPUT r1 r2 r3;
CARDS;
7 2 2
5 5 3
7 2 7
7 5 7
4 7 7
7 7 7
6 5 3
7 5 2
6 5 6
5 7 5
data survey; merge baseline rating; run;
proc print data=survey; run;
```

The SAS System

OBS	ID	SEX	AGE	INC	R1	R2	R3
1	1	F	35	17	7	2	2
2	17	M	50	14	5	5	3
3	33	F	45	6	7	2	7
4	49	M	24	14	7	5	7
5	65	F	52	9	4	7	7
6	81	M	44	11	7	7	7
7	2	F	34	17	6	5	3
8	18	M	40	14	7	5	2
9	34	F	47	6	6	5	6
10	50	M	35	17	5	7	5

ملاحظة :في بعض الحالات وعندما يراد دمج ملفين غير متساويين في تكرار بياناتهما فأن الملف الناتج من الدمج يكون غير صحيح.

مثال (7): اجري عملية الدمج للملفين التاليين ؟

```
data spacing1;
input P space rep yield96;
 cards;
0  40  1 57
0  40  2 58
0  80  1 57
0  80  2 58
0  80  3 56
0 120  1 49
0 120  2 54
0 120  3 53
25 40  1 53
```

Dr. Firas

```
25  40  2 45
25  40  3 46
25  80  1 54
25  80  2 50
25  80  3 48
25 120  1 63
25 120  2 57
25 120  3 53
;
data spacing2;
input P space rep yield97;
 cards;
0  40  1 35
0  40  2 28
0  40  3 29
0  80  1 38
0  80  2 29
0  80  3 27
0 120  1 10
0 120  2 25
0 120  3 34
25  40  1 24
25  40  2 24
25  40  3 17
25  80  1 25
25  80  2 31
25  80  3 29
25 120  1 44
25 120  3 28
data spacing;
merge spacing1 spacing2;
run;
proc print data=spacing;
run;
```

OBS	P	SPACE	REP	YIELD96	YIELD97
1	0	40	1	57	35
2	0	40	2	58	28
3	0	40	3	57	29
4	0	80	1	58	38
5	0	80	2	56	29
6	0	80	3	49	27
7	0	120	1	54	10
8	0	120	2	53	25
9	0	120	3	53	34
10	25	40	1	45	24
11	25	40	2	46	24
12	25	40	3	54	17
13	25	80	1	50	25
14	25	80	2	48	31
15	25	80	3	63	29
16	25	120	1	57	44
17	25	120	3	53	28

نلاحظ ان عدد المكررات (Rep) لقيم Space التي تساوي 40 هي 2 في الملف الاول و3 في الملف الثاني ولكن عند الدمج اصبحت 3 ، كما نجد ان عدد مكررات قيم Space التي تساوي

Dr. Firas
Dr. Firas

120 هي 3 في الملف الاول و2 في الملف الثاني وعند الدمج اصبحت 2 لكليهما. لذا فأن الايعاز الصحيح في مثل هذه الحالة هو كالآتي:

```
proc sort data=spacing1; by p space rep; run;
proc sort data=spacing2; by p space rep; run;
data spacing;
 merge spacing1 spacing2;
 by p space rep;
run;
proc print data=spacing; run;
```

OBS	P	SPACE	REP	YIELD96	YIELD97
1	0	40	1	57	35
2	0	40	2	58	28
3	0	40	3	.	29
4	0	80	1	57	38
5	0	80	2	58	29
6	0	80	3	56	27
7	0	120	1	49	10
8	0	120	2	54	25
9	0	120	3	53	34
10	25	40	1	53	24
11	25	40	2	45	24
12	25	40	3	46	17
13	25	80	1	54	25
14	25	80	2	50	31
15	25	80	3	48	29
16	25	120	1	63	44
17	25	120	2	57	.
18	25	120	3	53	28

الفصل الثاني

بعض الأوامر المهمة في برنامج SAS

الاوامر في البرنامج تبدأ عادة بكلمة Proc وتلحق معها كلمة اخرى تختلف حسب التحليل المراد اجرائه ، ومن اهم هذه الاوامر هي:

2 – 1 الأمر Proc univariate

يستعمل هذا الامر للتعرف على بعض المقاييس الاحصائية الخاصة بمتغيرما ، والتي على ضوئها يمكن تحديد كون المتغيير يتبع التوزبع الطبيعي ام لا، اذ باعتماد هذا الامر يمكن ان نحصل على قيم المتوسط والوسيط ومعامل الالتواء ومعامل التفلطح. كما يستعمل ايضا للحصول على اختبارات t و Sign و Wilcoxon Signed Rank.

مثـال (8) : فـي دراسـة عـن اوزان واطـوال ثلاثـة طـلاب المطلـوب حسـاب بعـض التقـديرات الاحصائية للطول فقط ؟

نطبق الشروط المذكورة سابقا بخصوص كتابة اي برنامج ونستعمل الآمر Proc univariate .

```
Data A;
Input   weight length;
Cards;
50     177
65         165
75         188
Option nodate nonumber;
Proc univariate; var length;  ← lenght  بعد ترك فراغ بعد  يمكن اضافة weight
Run;
```

عند الايعاز للبرنامج بالتنفيذ ستظهر النتائج كما موضحة ادناه. وعند اضافة weight ستظهر نتائج الصفتين . الايعازين nodate و nonumber هي لغرض عدم كتابة تاريخ التحليل ورقمه.

The SAS System

Univariate Procedure

Variable=LENGTH

Moments				Quantiles(Def=5)			
N	3	Sum Wgts	3	100% Max	188	99%	188
Mean	176.6667	Sum	530	75% Q3	188	95%	188
Std Dev	11.50362	Variance	132.3333	50% Med	177	90%	188
Skewness	-0.13028	Kurtosis	.	25% Q1	165	10%	165
USS	93898	CSS	264.6667	0% Min	165	5%	165
CV	6.511485	Std Mean	6.64162			1%	165
T:Mean=0	26.59994	Pr>|T|	0.0014	Range	23		
Num ^= 0	3	Num > 0	3	Q3-Q1	23		
M(Sign)	1.5	Pr>=|M|	0.2500	Mode	165		
Sgn Rank	3	Pr>=|S|	0.2500				

2 – 2 الأمر Proc means

يستعمل هذا الامر للحصول على عدد البيانات المتوسط العام ،الخطأ القياسي ،الانحراف القياسي ،اعلى وادنى قيمة ، المدى ،معامل الاختلاف وتقدير حدود الثقة.

```
Proc means n mean stderr std var max min range cv clm;var length ;
Run;
```

The SAS System

Analysis Variable : LENGTH

N	Mean	Std Error	Std Dev	Variance	Maximum	Minimum
3	176.6666667	6.6416196	11.5036226	132.3333333	188.0000000	165.0000000

Range	CV
23.0000000	6.5114845

Dr. Firas

مثال (9) : البيانات التالية تمثل عدد الطالبات والطلاب في احد الاقسام جد بعض المقاييس الاحصائية لكل منهم مع حدود الثقة ؟

تمثل كلمة clm ايعاز خاص بحدود الثقة Confidence limites

```
Data f;
Input male female;
Cards;
22  32
43  56
76  33
78  54
Option nodate nonumber;        ← هذه الجملة لغرض الحصول على نتائج بلا تاريخ او رقم
Proc means  n mean stderr std range clm; var male female;
Run;
```

Variable	N	Mean	Std Error	Std Dev	Range	Lower 95.0% CLM	Upper 95.0% CLM
MALE	4	54.7500000	13.5485239	27.0970478	56.0000000	11.6325502	97.8674498
FEMALE	4	43.7500000	6.5112083	13.0224166	24.0000000	23.0284292	64.4715708

2 – 3 الأمر Proc print

وهو الأمر الخاص بطبع النتائج.

مثال (10) : البيانات تمثل عدة متغيرات هي الفئة العمرية وعدد الوفيات والعدد الكلي للفئة العمرية المطلوب اضافة متغيران هما معدل الوفيات لكل فئة ولوغاريتم معدل الوفيات وطبع البيانات ؟

```
data f;
 input age $ deaths pop;
 rate = deaths/pop*100000;
 a = log(rate);
cards;
30-34     1      17742
35-39     5      16554
40-44     5      16059
45-49     12     13083
50-54     25     10784
55-59     38     9645
60-64     54     10706
65-69     65     9933
Option nodate nonumber;
proc print;
run;
```

The SAS System

Dr. Firas

OBS	AGE	DEATHS	POP	RATE	A
1	30-34	1	17742	5.636	1.72924
2	35-39	5	16554	30.204	3.40798
3	40-44	5	16059	31.135	3.43834
4	45-49	12	13083	91.722	4.51876
5	50-54	25	10784	231.825	5.44598
6	55-59	38	9645	393.987	5.97632
7	60-64	54	10706	504.390	6.22335
8	65-69	65	9933	654.384	6.48369

مثال (11): في دراسة تضمنت مجموعتين من الطلاب سجلت اوزان واطوال كل مجموعة ، ويمكن الحصول على متغيرات جديدة بوضع معادلة لكل متغير واعطائه رمز خاص ثم تقديربعض المقاييس الاحصائية ولايشترط اعطاء الآمر Proc print اذ يمكن تقدير المقاييس بدونه.

```
Data s;
input trt  w T;
f=(w*T)/(t-w);
h=log(t);
i=h*f;
j=i/(w*t);
 cards;
1  50 144
1  55 166
1  65 150
1  44 166
2  55 166
2  45 170
2  56 177
2  46 178
option nodate nonumber;
proc print;          ← الآمر الخاص بالطباعة
run;
proc means n mean stderr;var w t f h i j;
run;
```

The SAS System

OBS	TRT	W	T	F	H	I	J
1	1	50	144	76.596	4.96981	380.667	0.052870
2	1	55	166	82.252	5.11199	420.473	0.046054
3	1	65	150	114.706	5.01064	574.749	0.058949
4	1	44	166	59.869	5.11199	306.049	0.041902
5	2	55	166	82.252	5.11199	420.473	0.046054
6	2	45	170	61.200	5.13580	314.311	0.041086
7	2	56	177	81.917	5.17615	424.016	0.042778
8	2	46	178	62.030	5.18178	321.428	0.039256

The SAS System

Variable	N	Mean	Std Error
W	8	52.0000000	2.5213375
T	8	164.6250000	4.2381663
F	8	77.6028303	6.3605235
H	8	5.1012680	0.0264311
I	8	395.2705899	31.1936020

Dr. Nas

J	8	0.0461186	0.0023587

واذا اردنا ان نقدر بعض المقاييس تبعا لكل معاملة فيجب ان نظيف by trt وكما يلي:

```
proc means n mean stderr;by trt;var w t f h i j;
run;
```

The SAS System

---------- TRT=1 ----------

Variable	N	Mean	Std Error
W	4	53.5000000	4.4440972
T	4	156.5000000	5.6199051
F	4	83.3556829	11.4796216
H	4	5.0511060	0.0361243
I	4	420.4843112	56.6264298
J	4	0.0499436	0.0037579

---------- TRT=2 ----------

Variable	N	Mean	Std Error
W	4	50.5000000	2.9011492
T	4	172.7500000	2.8686524
F	4	71.8499777	5.9119050
H	4	5.1514299	0.0166644
I	4	370.0568686	30.1742173
J	4	0.0422936	0.0014451

ملاحظة : يمكن ترتيب البيانات السابقة بصورة افقية لكي تأخذ مساحة اصغر من الصفحة وكما يلي:

```
data s;
input trt  w T @@;
f=(w*T)/(t-w);
h=log(t);
i=h*f;
j=i/(w*t);
cards;
1  50 144 1  55 166 1  65 150 1  44 166
2  55 166 2  45 170 2  56 177 2  46 178
; option nodate nonumber ;
proc print;
run;
proc means n mean stderr;var w t f h i j;
run;
```

الرمز @@ في برنامج SAS يعني قراءة البيانات افقيا وحسب متغيرات تلك البيانات.

2 - 4 الأمر Proc sort

ويعني ترتيب البيانات تصاعديا .

Dr Firas

مثال (12): في دراسة تضمنت الجرعة السمية وعدد الفئران الكلي وعدد الفئران الهالكة المطلوب اضافة متغير آخر يمثل نسبة الهلاكات الى العدد الكلي وترتيب البيانات حسب عدد الفئران؟

```
data t;
 input dose number killed;
 prop = killed/number;
cards;
1.69    59    6
1.72    60    13
1.75    62    18
1.78    56    28
1.81    63    52
1.83    59    53
1.86    62    61
1.88    60    60
Option nodate nonumber;
proc print;
run;
PROC SORT DATA=T;BY NUMBER;
PROC PRINT;
RUN;
```

The SAS System

OBS	DOSE	NUMBER	KILLED	PROP
1	1.69	59	6	0.10169
2	1.72	60	13	0.21667
3	1.75	62	18	0.29032
4	1.78	56	28	0.50000
5	1.81	63	52	0.82540
6	1.83	59	53	0.89831
7	1.86	62	61	0.98387
8	1.88	60	60	1.00000

The SAS System

OBS	DOSE	NUMBER	KILLED	PROP
1	1.78	**56**	28	0.50000
2	1.69	59	6	0.10169
3	1.83	59	53	0.89831
4	1.72	60	13	0.21667
5	1.88	60	60	1.00000
6	1.75	62	18	0.29032
7	1.86	62	61	0.98387
8	1.81	63	52	0.82540

مثال (13): لدينا بيانات لثلاثة متغيرات هي الجنس والوزن والطول المطلوب ترتيب البيانات تصاعديا وحسب الوزن؟

```
data d;
 input sex $ wt height;
 cards;
 M 17 110
 M 15 105
 M 12 100
 F 15 104
```

Dr. Firas

```
 F 16 106
 F 14 102
proc sort data=d;by wt ;
run;
proc print;run;
```

OBS	SEX	WT	HEIGHT
1	M	12	100
2	F	14	102
3	M	15	105
4	F	15	104
5	F	16	106
6	M	17	110

2 – 5 الأمر Proc transpose

يستعمل هذا الامر لتحويل البيانات من الوضع الافقي الى العمودي.

مثال (14): حول البيانات التالية من الوضع الافقي الى العامودي.

```
DATA W;
INPUT ID WT1-WT6 ;
CARDS;
01    155   158   162   149   148   147
02    110   112   114   107   108   109
03    112   133   112   142   121   110
PROC SORT;
  BY ID;
RUN;
PROC TRANSPOSE;
  BY ID;
  VAR WT1-WT6;
run;
 proc print;
run;
```

OBS	ID	_NAME_	COL1
1	1	WT1	155
2	1	WT2	158
3	1	WT3	162
4	1	WT4	149
5	1	WT5	148
6	1	WT6	147
7	2	WT1	110
8	2	WT2	112
9	2	WT3	114
10	2	WT4	107
11	2	WT5	108
12	2	WT6	109
13	3	WT1	112
14	3	WT2	133
15	3	WT3	112
16	3	WT4	142
17	3	WT5	121
18	3	WT6	110

Dr. Firas

2 – 6 الأمر Proc ttest

يستعمل لغرض اجراء اختبار t لاختبار الاختلافات بين متوسطي مجموعتين.

مثـال (15): البيانـات ادنـاه تمثـل اوزان مجموعـة مـن الطالبـات فـي صـفين المطلـوب اختبـار الاختلافات بين المتوسطين ؟

```
data d;
input trt  w;
cards;
1  57.8
1  56.2
1  61.9
1  54.4
1  53.6
1  56.4
1  53.2
2  64.2
2  58.7
2  63.1
2  62.5
2  59.8
2  59.2
proc ttest;
class trt;
var w;
run;
```

Dr. Firas

The SAS System
TTEST PROCEDURE

Variable: W

TRT	N	Mean	Std Dev	Std Error
1	7	56.21428571	3.00245931	1.13482295
2	6	61.25000000	2.30195569	0.93976948

Variances	T	DF	Prob>|T|
Unequal	-3.4177	10.9	0.0058
Equal	-3.3442	11.0	**0.0065**

For H0: Variances are equal, F' = 1.70 DF = (6,5) Prob>F' = 0.5764

يتبين من النتائج ان الاختلافات بين المتوسطين معنوية (p = 0.0065)

مثال (16): البيانات ادنـاه تمثل حمولـة عدد من الشـاحنات تعود الـى شركتين المطلوب اختبـار الاختلافات بين المتوسطين ؟

```
DATA R;
INPUT GROUP$ score;
cards;
C 5.9
C 3.8
C 6.5
C 18.3
C 18.2
```

```
C 16.1
C 7.1
T 7.6
T 0.4
T 1.1
T 3.2
T 6.5
T 4.1
T 4.7
PROC TTEST;
CLASS GROUP;
VAR score;
RUN;
```

The SAS System
TTEST PROCEDURE

Variable: SCORE

GROUP	N	Mean	Std Dev	Std Error
C	7	10.84285714	6.38066275	2.41166383
T	7	3.94285714	2.63619495	0.99638803

Variances	T	DF	Prob>|T|
Unequal	2.6443	8.0	0.0296
Equal	2.6443	12.0	**0.0214**

For H0: Variances are equal, F' = 5.86 DF = (6,6) Prob>F' = 0.0492

تبين من النتائج ان الاختلافات في متوسط الحمولة بين الشركتين كان معنويا وبمستوى5% (p=0.0296).

2 – 7 الأمر proc corr

يستعمل هذا الأمر لتقدير معامل الارتباط البسيط او المتعدد.

مثال (17) جد معامل الارتباط بين المتغيرين x و y ؟

علما ان قيم x = 3 ، 8 ، 6 ، 9 ، 2 و قيم y= 22، 65، 55، 76، 43

```
DATA D;
INPUT X Y;
CARDS;
3  22
8  65
6  55
9  76
2  43
PROC CORR;
VAR X Y;
RUN;
```

Correlation Analysis
2 'VAR' Variables: X Y
Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum

```
X        5  5.600000     3.049590    28.000000     2.000000     9.000000
Y        5  52.200000    20.825465   261.000000    22.000000    76.000000

      Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / N = 5

                          X            Y

                X       1.00000      0.88334
                        0.0          0.0470

                Y       0.88334      1.00000
                        0.0470       0.0
```

اذا معامل الارتباط = 0.88 وهو معنوي (p=0.0470)ونتبع نفس الايعازات في حالة اكثر من متغير.

2 – 8 الآمر proc chart

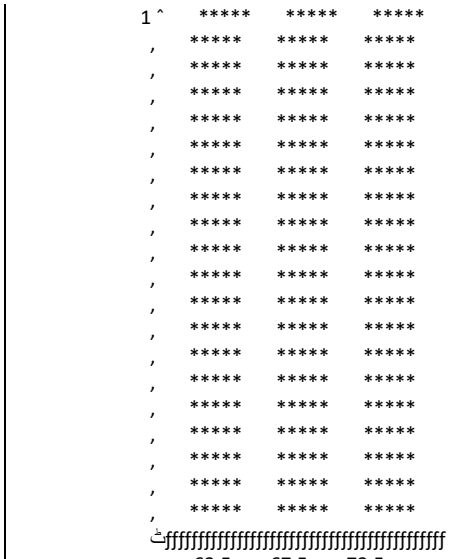
يستعمل للتعرف على التوزيع التكراري للمشاهدات.

مثال (18): ارسم توزيع متغير الوزن؟

```
DATA H;
INPUT NO GENDER   WEIGHT HEIGHT;
DATALINES;
1 1 69.5 155
2 2 61.2  99
3 2 63.0 115
4 1 72.5 205
PROC CHART;
VBAR weight;
RUN;
```

```
                                The SAS System
Frequency

2 ˆ   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
  ,   *****
```

Dr. Firas



2 – 9 الأمر proc tabulate

يستعمل لعمل جدول بالنتائج.

مثال(19): اراد احد اصحاب احد المطاعم بيانات عن معدل اعمار الزبائن ومتوسط دخلهم؟

```
data s;
input age income  gender;
cards;
33  232  1
44  334  2
32  434  1
55  656  2
43  333  1
PROC TABULATE;
CLASS GENDER;
VAR AGE INCOME ;
TABLE (AGE INCOME )*MEAN, GENDER ALL;
RUN;
```

```
                                The SAS System

                    GENDER
                    1     2     ALL
AGE      MEAN       36.00 49.50 41.40
INCOME   MEAN       333.00 495.00 397.80
```

2 – 10 الآمر proc npar1way

يستعمل لغرض اجراء الاختبارات اللامعلمية (سيرد شرحها لاحقا).

2 – 11 الأمر proc freq

يستعمل لغرض اجراء اختبار مربع كاي وتقديرنسبة الارجحية.

2 – 12 الأمر proc anova

يستعمل لغرض عمل جدول تحليل التباين للعوامل المدروسة وعند تساوي عدد المكررات لكل عامل علاوة على اجراء المقارنات بين المتوسطات.

2 – 13 الأمر proc glm

يستعمل لغرض تطبيق طريقة الانموذج الخطي العام عند عمل جدول تحليل التباين للبيانات غير متساوية التكرارات.

2 – 14 الأمر proc vercomp

يستعمل لغرض تقدير مكونات التباين الخاصة بتقدير المعالم الوراثية (المكافىء الوراثي والمعامل التكراري والارتباطات الوراثية والبيئية والمظهرية) للتأثيرات العشوائية او التأثيرات الثابتة والعشوائية في النماذج المختلطة.

2 – 15الأمر proc mixed

يستعمل لغرض تقدير التأثيرات الثابتة والعشوائية في النماذج المختلطة وفي تقدير قيم الجدارة الوراثية BLUP.

2 – 16 الأمر proc logistic

يستعمل لغرض الحصول على نماذج الانحدار اللاخطية اللوغاريتمية للصفات الثنائية (binary) وتقدير نسبة الارجحية (Odds Ratio).

2 – 17 الأمر proc plot

يستعمل لرسم علاقة بيانية بين متغيرين او اكثر.

2 – 18 الأمر proc reg

يستعمل لتقدير معامل الانحدارللمتغيرات المستمرة.

2 – 19 الأمر proc lifetest

ويستعمل لتقدير العمر للافراد اعتمادا بيانات الافراد الميتة والحية.

2 – 20 الأمر proc genmod

يستعمل لمطابقة النموذج الخطي العام وعمل تحليل التباين والانحدارللعوامل التابعة المتقطعة (discrete responses).

2 – 21 الأمر proc catmod

Dr. Firas

يستعمل لمطابقة النماذج الخطية وعمل تحليل التباين وانحدار للعوامل التابعة ذات الفئات (categorical response.).

2 – 22 الأمر proc probit

يستعمل لتقدير معامل الانحدار للصفات الثنائية او ذات الترتيب .

2 – 23 الأمر proc phreg

يستعمل لتقدير نسبة الخطورة .

وهناك اوامر أخرى لايمكن تنفيذ البعض منها في نسخة هذا البرنامج (6.12) وانما في النسخ الاحدث.

proc sql, proc gplot. proc cluster, proc princomp, proc multtest, proc mianalyze, proc mi, proc iml, proc nlin, proc rank, proc factor, proc kde, proc g3d, proc boxplot, proc stdize, proc tree, proc modeclus, proc discrim, proc stepdisc, proc corresp, proc nested, proc allele, proc casecontrol, proc family, proc haplotype, proc psmooth, proc htsnp, proc inbreed, proc transreg, proc qlim, proc gam.

الفصل الثالث

اختبار الفرضيات Hypotheses Test

في اي تحليل احصائي تكون هناك فرضيتان تخضعان للاختبار بأستعمال اختبار F الذي نحصل عليه من جدول تحليل التباين ، وذلك بمقارنة F المحسوبة مع F الجدولية وهذه العملية تنفذ ذاتيا في برنامج SAS اذ تظهر في النتائج كتقدير للاحتمالية.

ان الفرضيتان المراد اختبارهما هما فرضية العدم Null Hypothesis ويرمز لها H_0 والفرضية البديلة Alternative Hypothesis ويرمز لها H_A ، وعادة تفترض فرضية العدم عدم وجود اختلافات بين المعاملات وتوضع هذه الفرضية على أمل رفضها عند مستوى معين (0.05 او 0.01). واذا تم رفض الفرضية مثلا عند مستوى 0.05 فأن ذلك يعني ان قرارنا صحيح بنسبة 95% وان احتمال ان يكون قرارنا خاطىء هو 5% وينطبق الامر نفسه على المستوى 1%.

Dr Firas

ان الاختبار الذي يشير الى وجود اختلافات معنوية يحتاج الى اختبارات اخرى لتحديد اي من المعاملات تختلف عن الاخرى هل ان متوسط المعاملة الاولى يختلف عن الثانية وهل يختلف متوسط المعاملة الثالثة عن الرابعة وبصورة عامة فأن الاختبارات تكون على نوعين:

3 – 1 الاختبارات المعلمية Parametreic Tests

وهي التي تعتمد لاختبار بعض معالم المجتمع مثل المتوسط وتفترض معرفة التوزيع الاحتمالي له اي ان الصفة تتوزع توزيع طبيعي وان التباين متجانس وان العوامل مستقلة.

3 – 1 – 1 اختبار التوزيع الطبيعي للبيانات Test of Normality

يمكن اختباران الصفة ذات توزيع طبيعي اعتمادا على عدة مقاييس:

أ- من قيمتي معامل الالتواء Skewness ومعامل التفلطح Kurtosis اذ تكونان مقاربتان الى الصفر وكذلك من قيمتي المتوسط والوسيط اذ تكون متساوية او متقاربة .

ب- من مخطط الساق والورقة Stem and Leaf اذ يكون التوزيع طبيعي للمتغير عندما تكون التكرارات متماثلة تقريبا عند المتوسط الحسابي. وهذا المخطط يكون مشابه للمدرج التكراري الا انه بدل ان يعبر عن التكرارات باعمدة فانه يعبر عنها افقيا بالقيم الحقيقية.

ج- من مخطط الصندوق Boxplot اذ يستدل على ان التوزيع طبيعي عندما لاتكون هناك قيم متطرفة او شاذة وان موقع الوسيط (الخط المستقيم) في وسط الصندوق. اما اذا كان الوسيط اقرب الى الجزء الاسفل من الصندوق فأن ذلك يعني ان التوزيع ملتو الى اليمين واذا كان الوسيط اقرب الى الجزء العلوي من الصندوق فأن ذلك يعني ان التوزيع ملتو نحو اليسار.

د- من مخطط Normal Probability Plot اذ ترسم كل مشاهدة من البيانات الاصلية على المحور العمودي مقابل قيم التوزيع الطبيعي القياسي المتوقعة لها. وعندما تكون الصفة ذات توزيع طبيعي فأن المشاهدات ستكون بمحاذات الخط المستقيم وبعكس ذلك فأن ذلك سيعني ان الصفة لاتتوزع طبيعيا.

هـ- اختبار كولموجورف- سيميرنوف (سيرد شرحه لاحقا).

عندما تكون قيمتا المتوسط والوسيط متساويتان او متقاربتان فأن ذلك يعني ان التوزيع طبيعي ولكن عندما تكون قيمة المتوسط اكبر من قيمة الوسيط فأن ذلك يعني ان التوزيع منحرف نحو اليمين وعندما حصول العكس فأن ذلك يعني ان التوزيع منحرف نحو اليسار.

مثال (20) : البيانات ادناه تمثل معامل هضم لدى مجموعة من الاغنام المطلوب معرفة توزيع البيانات طبيعي ام لا باستعمال عدة طرق؟

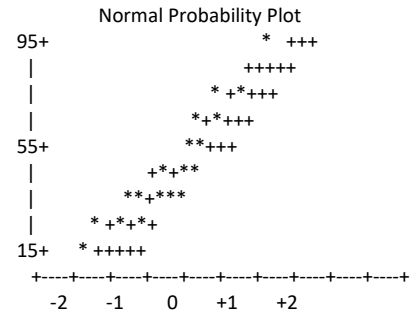
```
data s;
input animal dm;
cards;
```

Dr. Firas

```
2 37
2 25
2 77
2 73
2 41
2 23
2 32
2 91
2 41
2 31
2 44
2 51
2 51
2 50
2 18
2 66
2 66
2 30
2 26
2 39
proc univariate normal plot;
var  dm;
run;
```

```
                                The SAS System
                        Univariate Procedure
Variable=DM

      Moments                           Quantiles(Def=5)

 N           20  Sum Wgts      20      100% Max      91     99%      91
 Mean       45.6 Sum           912      75% Q3       58.5    95%      84
 Std Dev   19.99052 Variance   399.6211  50% Med      41     90%      75
 Skewness  0.754732 Kurtosis   -0.16984  25% Q1       30.5    10%      24
 USS         49180 CSS         7592.8    0% Min      18     5%     20.5
 CV        43.83887 Std Mean   4.470017                     1%      18
 T:Mean=0   10.2013 Pr>|T|     0.0001    Range         73
 Num ^= 0      20  Num > 0       20     Q3-Q1         28
 M(Sign)       10  Pr>=|M|    0.0001    Mode          41
 Sgn Rank     105  Pr>=|S|    0.0001
 W:Normal  0.936501 Pr<W       0.2166

                          Extremes

                     Lowest   Obs    Highest   Obs
                       18(     15)     66(     16)
                       23(      6)     66(     17)
                       25(      2)     73(      4)
                       26(     19)     77(      3)
                       30(     18)     91(      8)

                  Stem Leaf              #          Boxplot
                    9 1                  1            |
                    8                                 |
                    7 37                 2            |
                    6 66                 2            |
                    5 011                3          +-----+
                    4 114                3          *--+--*
                    3 01279              5          +-----+
                    2 356                3            |
                    1 8                  1            |
                      ----+----+----+----+
                  Multiply Stem.Leaf by 10**+1

                                The SAS System
Variable=DM
```

Dr. Firas



نظرا لكون معاملي الالتواء والتفلطح مقاربان للصفر وان قيمة المتوسط 45.6 مقاربة لقيمة الوسيط 41 وان موقع المستقيم في وسط الصندوق في مخطط Boxplot وان البيانات تتركز في الوسط تقريبا في مخطط Stem-Leaf وكذلك ان البيانات بمحاذات الخط المستقيم في مخطط Normal Probability Plot لذا فأن البيانات ذات توزيع طبيعي.

3 – 1 – 2 اختبار تجانس التباين Test of Homogeneity of Variance

يمكن اختبار تجانس التباين بأستعمال اختبارLevene .

مثال (21) : البيانات ادناه تمثل نتائج تجربة تتضمن اربعة معاملات وبواقع 12 مكرر لكل معاملة المطلوب اختبار تجانس التباين بين المعاملات؟

```
data one;
input trt g;
cards;
1 895
1 540
1 1020
1 470
1 428
1 620
1 760
1 537
1 845
1 1050
1 387
1 497
2 1520
2 1610
2 1900
2 1350
2 980
2 1710
2 1930
2 1960
2 1840
2 2410
2 1520
2 1685
3 43300
3 32800
```

Dr. Firas

```
3 28800
3 34600
3 27800
3 32800
3 28100
3 18900
3 31400
3 39500
3 29000
3 22300
4 11000
4 8600
4 8260
4 9830
4 7600
4 9650
4 8900
4 6060
4 10200
4 15500
4 9250
4 7900
proc glm data=one;
class trt;
model g=trt;
means trt/hovtest=levene welch;
run;
```

The SAS System

General Linear Models Procedure

Class Level Information

Class	Levels	Values
TRT	4	1 2 3 4

Number of observations in data set = 48

Dependent Variable: G

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	7035039589.75000000	2345013196.58333000	186.37	0.0001
Error	44	553637612.16666700	12582673.00378780		
Corrected Total	47	7588677201.91666000			

R-Square	C.V.	Root MSE	G Mean
0.927044	33.35186	3547.2063661	10635.70833333

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TRT	3	7035039589.75000000	2345013196.58333000	186.37	0.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	3	7035039589.75000000	2345013196.58333000	186.37	0.0001

Levene's Test for Equality of G Variance

ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
TRT	3	1.409E16	4.697E15	5.5992	**0.0024**

Dr Firas

Error 44 3.691E16 8.389E14

اذا نرفض فرضية العدم حول تجانس تباين المعاملات اذ ان قيمة الاحتمال = 0.0024 ومن ثم نعتمد على جدول Welch لتحليل التباين لاختبار معنوية الاختلافات بين المعاملات.

Welch's ANOVA for G

Source	DF	F Value	Pr > F
TRT	3.0000	145.3	0.0001
Error	21.3988		

The SAS System General Linear Models Procedure

Level of TRT	N	G Mean	G SD
1	12	670.7500	233.92195
2	12	1701.2500	356.53716
3	12	30775.0000	6688.67768
4	12	9395.8333	2326.03626

ويمكن اجراء تحويل للبيانات بأن نحول العامل g الى S=log(g) وننفذ البرنامج وسنجد ان اختبار ليفين سيكون غير معنويا وبذلك يكون اختبار t صحيحا.

```
proc univariate normal plot; var g;run;
proc univariate normal plot; var s;run;
```

يمكن تقسيم الاختبارات حسب الوقت الذي تجرى فيه المقارنات الى نوعين:

3 – 1 – 3 الاختبارات المحددة قبل اجراء التجربة

وتتضمن هذه الاختبارات :

3 – 1 – 3 – 1 المقارنات المستقلة Orthogonal Comparisons

تتضمن معظم التجارب عدة معاملات وان بعض هذه المعاملات تكون ذات اهمية لدى الباحث لذا فأنه قد يحدد قبل اجراء التجربة اي المقارنات التي سيوجه اهتمامه اليها ، وهنا لابد من الاشارة الى نقطة مهمة وهي ان هذه المقارنات التي يريد ان يجريها الباحث ستكون اجراء صحيح حتى وان كان اختبار F غير معنويا.

ان تنفيذ هذه المقارنات في برنامج SAS تستلزم التعرف على كيفية اعطاء قيم المقارنة ، فمثلا لو اردنا اجراء مقارنة بين ثلاث معاملات على ان تجرى المقارنة بين متوسطي المعاملة الثانية والثالثة فأن القيم ستكون كالاتي:

```
contrast ' 2 vs. 3' trt      0  1  -1;
```

وفي حالة مقارنة الاولى مع الثالثة

```
contrast ' 1 vs. 3' trt      1  0  -1;
```

وفي حالة مقارنة الاولى والثانية مع الثالثة

```
contrast ' 1 & 2 vs. 3' trt   1  1  -2;
```

Dr. Firas

يلاحظ من ذلك ان المعاملة المراد اجراء المقارنة لها تأخذ القيمة +1 والمعاملة المقابلة لها – 1 اما المعاملة الغير مشمولة بالمقارنة فتأخذ القيمة 0 بحيث ان مجموع القيم يساوي 0.

مثال (22): اجريت تجربة لدراسة تأثير خمسة سلالات من دجاج اللحم على وزن الافراخ بعمر يوم حلل البيانات واختبر مايلي:

1- المعاملة الاولى ضد باقي المعاملات

2- المعاملة الثانية والثالثة ضد الرابعة والخامسة

3- المعاملة الثانية ضد الثالثة

4- المعاملة الرابعة ضد الخامسة

trt	observations
t1	46 40 42 40
t2	51 48 47 42
t3	36 42 44 46
t4	42 42 45 43
t5	35 36 37 36

```
data d;
input trt rep w;
cards;
1 1 46
1 2 40
1 3 42
1 4 40
2 1 51
2 2 48
2 3 47
2 4 42
3 1 36
3 2 42
3 3 44
3 4 46
4 1 42
4 2 42
4 3 45
4 4 43
5 1 35
5 2 36
5 3 37
5 4 36
proc glm data = d ;
 class trt;
 model w = trt/ ss3 ;
 contrast ' 1 vs. all' trt       4 -1 -1 -1 -1;
 contrast  ' 2 & 3  vs. 4 & 5' trt 0  1  1 -1 -1;
 contrast ' 2 vs. 3' trt         0  1 -1  0  0;
 contrast  ' 4 vs. 5' trt         0  0  0  1 -1;
 run;
```

Dr. FiraS

General Linear Models Procedure
Class Level Information

Class	Levels	Values
TRT	5	1 2 3 4 5

Number of observations in data set = 20

Output 5.13
Three Factors Case with Two Repeated Factors

General Linear Models Procedure

Dependent Variable: W

Source	DF	Sum of Squares	F Value	Pr > F
Model	4	248.00000000	7.15	0.0020
Error	15	130.00000000		
Corrected Total	19	378.00000000		

R-Square	C.V.	W Mean
0.656085	7.009334	42.0000000

Source	DF	Type III SS	F Value	Pr > F
TRT	4	248.00000000	7.15	0.0020

Contrast	DF	Contrast SS	F Value	Pr > F
1 vs. all	1	0.00000000	0.00	1.0000
2 & 3 vs. 4 & 5	1	100.00000000	11.54	0.0040
2 vs. 3	1	50.00000000	5.77	0.0297
4 vs. 5	1	98.00000000	11.31	0.0043

Dr Firas

3 – 1 – 3 – 2 تحليل الاتجاه بأستعمال العوامل متعددة الحدود

Trend analysis by using orthogonal polynomials

يستعمل هذا التحليل عندما يرغب الباحث في دراسة استجابة صفة ما للتغير في مستويات مختلفة لأحد العوامل الكمية ، مثال ذلك التغير في كمية المواد السمية في الغذاء وعلاقتها بوزن الاطفال. هذا التحليل يساعدنا على معرفة نوع العلاقة بين الصفة المدروسة ومستويات المعاملة بمعنى هل ان العلاقة خطية اي تتبع معادلات الدرجة الاولى Linear ام علاقة تربيعية اي تتبع معادلات الدرجة الثانية Quadratic ام علاقة تكعيبية اي تتبع معادلات الدرجة الثالثة Cubic. وهذا التحليل يوضح لنا شكل منحنى الاستجابة فضلا عن تحديد افضل واسوء مستوى للعامل المدروس. وعندما تكون مستويات المعاملات على ابعاد متساوية او منتظمة فأن مكونات العلاقة الخطية او التربيعية وغيرها من العلاقات الاعلى يكون من الممكن تقديرها واختبارها بالاعتماد على جدول (2) .

جدول (2) عوامل معاملات الارتداد الحرة

r	Trend	X=1	2	3	4	5	6	7	sum c(i)**2
2	Linear	-1	1						2
3	Linear	-1	0	1					2
	Quad	1	-2	1					6
4	Linear	-3	-1	1	3				20
	Quad	1	-1	-1	1				4
	Cubic	-1	3	-3	1				20
5	Linear	-2	-1	0	1	2			10
	Quad	2	-1	-2	-1	2			14
	Cubic	-1	2	0	-2	1			10
	Quartic	1	-4	6	-4	1			70
6	Linear	-5	-3	-1	1	3	5		70
	Quad	5	-1	-4	-4	-1	5		84
	Cubic	-5	7	4	-4	-7	5		180
	Quartic	1	-3	2	2	-3	1		28
7	Linear	-3	-2	-1	0	1	2	3	28
	Quad	5	0	-3	-4	-3	0	5	84
	Cubic	-1	1	1	0	-1	-1	1	6
	Quartic	3	-7	1	6	1	-7	3	154
8	Linear	-7	-5	-3	-1	1	3	5	7	168
	Quad	7	1	-3	-5	-5	-3	1	7	168
	Cubic	-7	5	7	3	-3	-7	-5	7	264
	Quartic	7	-13	-3	9	9	-3	-13	7	616
9	Linear	-4	-3	-2	-1	0	1	2	3	4	60
	Quad	28	7	-8	-17	-20	-17	-8	7	28	2772
	Cubic	-14	7	13	9	0	-9	-13	-7	14	990
	Quartic	14	-21	-11	9	18	9	-11	-21	14	2002

Dr Firas

مثال (23): البيانات ادناه تمثل كمية حاصل الذرة الصفراء بالكغم للقطعة التجريبية في احدى التجارب التي استخدم فيها خمسة مستويات من الاسمدة ، حدد منحى استجابة المحصول لمستوى التسميد؟

Level of N	Observations
t1 0	30 28 25 32 24
t2 10kg	30 34 36 35 32
t3 20kg	45 38 34 42 41
t4 30kg	49 47 36 44 40
t5 40kg	48 45 41 53 52

```
data f;
input trt w;
cards;
```

```
1 30
1 28
1 25
1 32
1 24
2 30
2 34
2 36
2 35
2 32
3 45
3 38
3 34
3 42
3 41
4 49
4 47
4 36
4 44
4 40
5 48
5 45
5 41
5 53
5 52
PROC GLM DATA=f;
 CLASS trt;
 MODEL w = trt;
 CONTRAST "a linear" trt -2 -1  0   1 2 ;
 CONTRAST "a quad"   trt  2 -1 -2  -1 2 ;
 CONTRAST "a cubic"  trt -1  2  0  -2 1 ;
RUN;
```

ان اعتماد قيم المقارنة 2 1 0 1- 2- يكون على اساس عدد المعاملات (5) من جدول 2.

The SAS System
General Linear Models Procedure
Class Level Information

Class	Levels	Values
TRT	5	1 2 3 4 5

Number of observations in data set = 25

The SAS System
General Linear Models Procedure

Dependent Variable: W

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	1256.56000000	314.14000000	18.07	0.0001
Error	20	347.60000000	17.38000000		
Corrected Total	24	1604.16000000			

R-Square	C.V.	Root MSE	W Mean
0.783313	10.84530	4.1689327	38.44000000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TRT	4	1256.56000000	314.14000000	18.07	**0.0001**

Source	DF	Type III SS	Mean Square	F Value	Pr > F

Dr. Firas

TRT	4	1256.56000000	314.14000000	18.07	**0.0001**

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
a linear	1	1240.02000000	1240.02000000	71.35	**0.0001**
a quad	1	10.41428571	10.41428571	0.60	0.4479
a cubic	1	0.08000000	0.08000000	0.00	0.9466

يتبين من النتائج ان تأثير السماد كان معنويا وافضل علاقة هي الخطية اذ كانت عالية المعنوية

مثال (24) : حدد نوع العلاقة بين المتغير a الذي يأخذ اربع مستويات مع المتغيرs باستعمال الانحدار ؟

```
data f;
input a rep s @@;
cards;
1 1  7   1 2  9   1 3  9   1 4  9   1 5  9   1 6 11   1 7  12
1 8  12  1 9  14  1 10 14  1 11 13  1 12 18  1 13 18  1 14 20
1 15 20  1 16 20  1 17 21  1 18 22  1 19 22  1 20 24  2 1  13
2 2  18  2 3  12  2 4  17  2 5  16  2 6  20  2 7  21  2 8  22
2 9  22  2 10 22  2 11 25  2 12 25  2 13 21  2 14 23  2 15 25
2 16 24  2 17 22  2 18 26  2 19 27  2 20 26  3 1  19  3 2  20
3 3  19  3 4  19  3 5  19  3 6  20  3 7  21  3 8  22  3 9  23
3 10 23  3 11 25  3 12 22  3 13 24  3 14 24  3 15 25  3 16 25
3 17 26  3 18 26  3 19 26  3 20 28  4 1  13  4 2  12  4 3  14
4 4  11  4 5  16  4 6  17  4 7  17  4 8  18  4 9  22  4 10 21
4 11 19  4 12 23  4 13 23  4 14 23  4 15 24  4 16 24  4 17 25
4 18 25  4 19 29  4 20 25
PROC GLM DATA=f;
 CLASS a;
 MODEL s = a;
 CONTRAST "a linear" a -3 -1  1  3 ;
 CONTRAST "a quad"   a  1 -1 -1  1 ;
 CONTRAST "a cubic"  a -1  3 -3  1 ;
RUN;
```

```
                The SAS System
          General Linear Models Procedure
            Class Level Information

           Class   Levels   Values

           A          4   1 2 3 4
       Number of observations in data set = 80
                The SAS System
         General Linear Models Procedure

Dependent Variable: S

Source        DF     Sum of Squares   Mean Square  F Value    Pr > F
Model          3     652.30000000   217.43333333   10.72    0.0001
Error         76    1541.90000000    20.28815789
Corrected Total 79  2194.20000000
```

```
R-Square       C.V.            Root MSE          S Mean
0.297284     22.69137          4.5042377       19.85000000

Source       DF    Type I SS     Mean Square  F Value   Pr > F
A             3    652.30000000  217.43333333  10.72    0.0001

Source       DF    Type III SS   Mean Square  F Value   Pr > F
A             3    652.30000000  217.43333333  10.72    0.0001

Contrast     DF    Contrast SS   Mean Square  F Value   Pr > F
a linear      1   256.00000000  256.00000000   12.62    0.0007
a quad        1   396.05000000  396.05000000   19.52    0.0001
a cubic       1     0.25000000    0.25000000    0.01    0.9119
```

يتبين من النتائج ان العلاقتين الخطية والتربيعية معنوية ويمكن المفاضلة بين الطريقتين بالاعتماد على قيمة R^2

مثال (25): حدد نوع العلاقة بين مقدار الخصم والمال المدفوع نقدا في البيانات ادناه؟

```
data coupon;
  input discount purchase @@;
cards;
0 32.39   10 98.47   15 71.62   20 60.85
0 38.32   10 74.80   15 59.92   20 46.45
0 35.66   10 52.97   15 75.37   20 68.49
0 74.24   10 46.72   15 77.04   20 63.83
0 63.05   10 76.81   15 72.84   20 75.38
0 66.53   10 69.01   15 52.53   20 70.60
0 46.36   10 53.77   15 80.47   20 52.23
0 41.90   10 54.21   15 72.55   20 57.14
0 44.94   10 83.14   15 78.94   20 60.17
0 41.09   10 49.00   15 64.00   20 60.46
proc glm data=coupon;
  class discount;
  model purchase = discount;
  estimate "linear" discount   -3 -1  1  3;
  estimate "quad"   discount   -2  2  2 -2;
  estimate "cubic"  discount   -1  3 -3  1;
run;
```

```
                                The SAS System
                        Less than Full Rank Model

                      General Linear Models Procedure
                        Class Level Information

                      Class    Levels    Values

                      DISCOUNT     4    0 10 15 20


                    Number of observations in data set = 40

                          The SAS System
                        Less than Full Rank Model
                      General Linear Models Procedure

Dependent Variable: PURCHASE

Source             DF     Sum of Squares     Mean Square  F Value    Pr > F
Model               3     2710.89819000     903.63273000     5.51    0.0032
Error              36     5903.03772000     163.97327000
Corrected Total    39     8613.93591000

      R-Square         C.V.            Root MSE     PURCHASE Mean
      0.314711        20.78548          12.8052048       61.60650000
```

Dr. Firas

```
Source          DF      Type I SS      Mean Square  F Value    Pr > F
DISCOUNT         3    2710.89819000   903.63273000     5.51    0.0032

Source          DF     Type III SS     Mean Square  F Value    Pr > F
DISCOUNT         3    2710.89819000   903.63273000     5.51    0.0032

                      T for H0:       Pr > |T|      Std Error of
Parameter     Estimate        Parameter=0                Estimate

linear      43.9740000            2.43        0.0203       18.1092943
quad        52.8200000            3.26        0.0024       16.1974452
cubic       -0.8020000           -0.04        0.9649       18.1092943
```

يتبين من النتائج ان العلاقة التربيعية هي الافضل.

مثال (26): اختبر نوع العلاقة بين كثافة النبات وعدد الحبوب في السنبلة؟

```
data grain;
input density grain ;
cards;
10 12.2
10 11.4
10 12.4
20 16.0
20 15.5
20 16.5
30 18.6
30 20.2
30 18.2
40 17.6
40 19.3
40 17.1
50 18.0
50 16.4
50 16.6
proc means mean stderr std;
class density;
var grain;
run;
proc glm;
class density;
model grain=density;
contrast 'Linear    ' density -2 -1  0  1 2;
contrast 'Quadratic ' density  2 -1 -2 -1 2;
contrast 'Cubic     ' density -1  2  0 -2 1;
contrast 'Quartic   ' density  1 -4  6 -4 1;
run;
```

عند تنفيذ البرنامج سنلاحظ بأن العلاقة الخطية والعلاقة التربيعية كلاهما معنويتان ولكن عند الاحتكام الى معامل التحديد R^2 سنجد ان قيمته في العلاقة الخطية اكبر من قيمته في العلاقة التربيعية . ويمكن تنفيذ ذلك من خلال اضافة العبارة التالية الى قطعة البرنامج:

```
density1=density*density;
```

واضافة الايعازات التالية الى قطعة البرنامج:

```
proc reg data= grain;
model grain= density;
run;
proc reg data=grain;
```

Dr Firas

```
model grain= density1;
run;
```

3 – 1 – 4 الاختبارات المقترحة بعد اجراء التجربة

هنالك العديد من الاختبارات التي تجرى لمقارنة متوسطات المعاملات ببعضها ، ويعتمد عدد المقارنات على عدد المعاملات وهو يساوي عدد المعاملات x(عدد العاملات – 1)مقسوما على 2. وتختلف طرق الاختبار في دقتها فبعضها يعتمد على قيمة واحدة للمقارنة والبعض الاخر يعتمد عل اكثر من قيمة.

3 – 1 – 4 – 1 مقارنة جميع متوسطات المعاملات بمتوسط معاملة المقارنة.

أ- اختبار Dunnett

3 – 1 – 4 – 2 المقارنات المتعددة

أ- أختبار أقل فرق معنوي LSD او اختبار t

يعتمد هذا الاختبارعلى قيمة احصائية واحدة لاختبار الفروق بين المتوسطات تسمى أقل فرق معنوي LSD ، ويعد من اكثر الاختبارات استعمالا وانتشارا بين الباحثين وذلك لسهولة اجرائه الا انه لايعد اختبارا صحيحا الا في حالة مقارنة معاملتين فقط ، اذ ان استعمال هذا الاختبار في تجارب تضم ثلاث معاملات فأكثر قد يشير الى وجود فروق معنوية في الوقت الذي تكون فيه هذه الفروق غير معنوية ، ويرى بعض الباحثين عكس ذلك اذ اوصى احد الباحثين بأستعمال هذا الاختبار بغض النظر عن عدد المعاملات بعد ان قارنه بأختبارات اخرى.

مثال (27) : هل هناك فروق معنوية في نسبة المادة G في دم الارانب قبل وبعد حقنها بالمضاد الحيوي ؟

```
DATA R;
INPUT GA GB;
D=GA-GB;
cards;
 5.9  7.6
 3.8  0.4
 6.5  1.1
 18.3 3.2
 18.2 6.5
 16.1 4.1
 7.1  4.7
Option nodate  nonumber;
 PROC MEANS N MEAN STDERR t prt ;
 VAR D;
RUN;
```

```
The SAS System
Analysis Variable : D
```

Dr. Firas

N	Mean	Std Error	T	Prob>|T|
7	6.9000000	2.3145502	2.9811407	0.0246

الاختلافات معنوية ($p = 0.0246$)

مثال (28) : اخذت كمية من حاصل البنجر السكري ، وقسم كل رأس من رؤوس البنجر الى نصفين متساويين بهدف تخصيص احد الانصاف وبشكل عشوائي لطريقة v1 لاستخلاص السكر وتخصيص النصف الآخر الى الطريقة v2 وقدر الوقت المستغرق بالدقيقة لمقارنة كفاءة الطريقتين فهل هناك اختلافات معنوية احصائيا بين الطريقتين؟ احسب مجال الثقة للفرق بين الطريقتين؟

ملاحظة: عند دراسة ظاهرة او صفة لمجتمع ما مثل معدل اوزان طلاب المرحلة الابتدائية او معدل اطوالهم او نسبة المتعلمين في بغداد الى عدد سكان بغداد فأننا نأخذ عينة عشوائية من المجتمع ونقدر هذه المقاييس في العينة لنعممها على المجتمع ويكون التقدير بنقطة مثل المعدل اوفترة مثل حدود الثقة اونسبة والتقديرات في حالة النقطة او النسبة لاتساوي المتوسط الحقيقي او النسبة الحقيقية للمجتمع وانما قد تقل اوتزيد عنه بمقدار يعتمد على حجم العينة والصدفة وهذا التقدير لايتضمن اي مجال للثقة اي اننا لانستطيع القول ان ذلك المعدل اوالنسبة هي 95 او 99% تساوي المعدل الحقيقي او النسبة الحقيقية للمجتمع ولكن عندما تكون التقديرات ضمن حد ادنى وحد اعلى فأننا في هذه الحالة يمكننا ان نكون واثقين 95 او 99% بأن المتوسط الحقيقي او النسبة الحقيقية للمجتمع تقع ضمن مدى الحدين.

```
DATA P;
  D = v1 - v2;
DATALINES;
30 18
28 19
33 15
28 17
31 19
23 14
27 11
32 13
25 15
27 10
option nodate nonumber;
 proc means n mean stderr clm t prt;
var d;
run;
```

Dr. Firas

The SAS System
Analysis Variable : D

N	Mean	Std Error	Lower 95.0% CLM	Upper 95.0% CLM	T	Prob>|T|

```
10   13.3000000   1.2115188  10.5593541    16.0406459   10.9779560   0.0001
-------------------------------------------------------------------------------
```

على ضوء نتائج الدراسة يمكن القول بأننا واثقون 95% بـأن الفرق الحقيقي بين الطريقتين يقع بين 10.97 و 16.04.

مثال (29) : اخذت عينتان من البيض تعودان الى سلالتين من الدجاج جد حدود الثقة للفرق بين متوسط وزن البيضة في السلالالتين؟

```
DATA R;
INPUT s1 s2;
d=s2 - s1;
cards;
 58  62
 55  61
 52  59
 54  62
 53  60
PROC MEANS N MEAN STDERR clm t prt ;
VAR D;
RUN;
```

```
                                The SAS System

   Analysis Variable : D

N    Mean    Std Error  Lower 95.0% CLM  Upper 95.0% CLM   T     Prob>|T|
-------------------------------------------------------------------------
5  6.4000000  0.6782330   4.5169233          8.2830767    9.4362852   0.0007
-------------------------------------------------------------------------------
```

الفرق معنوي بين معدلي وزن البيض ويتراوح من 4.51 الى 8.28 غرام

مثال (30): في دراسة لمعرفة هل ان معدل التقييم بين مجموعتين يختلفان بصورة معنوية ؟

في هذا المثال يمكن استعمال الامر الآتي:

```
PROC MEANS N MEAN STDERR clm t prt ;
```

ولكن اذا اردنا الحصول على المتوسطات واختبار الاختلافات بينها فنكتب قطعة البرنامج كالاتي:

```
DATA R;
INPUT GROUP$ score;
cards;
C 5.9
C 3.8
C 6.5
C 18.3
C 18.2
C 16.1
C 7.1
T 7.6
T 0.4
```

Dr. Firas

T 1.1
T 3.2
T 6.5
T 4.1
T 4.7

```
option nodate nonumber;
proc anova;classes group;
model score=group;
means group/lsd;
proc means mean stderr;by group;var score;
run;
```

The SAS System
Analysis of Variance Procedure
Class Level Information

Class	Levels	Values
GROUP	2	C T

Number of observations in data set = 14
The SAS System

Analysis of Variance Procedure
Dependent Variable: SCORE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	166.63500000	166.63500000	6.99	0.0214
Error	12	285.97428571	23.83119048		
Corrected Total	13	452.60928571			

R-Square	C.V.	Root MSE	SCORE Mean
0.368165	66.03293	4.8817200	7.39285714

Source	DF	Anova SS	Mean Square	F Value	Pr > F
GROUP	1	166.63500000	166.63500000	6.99	0.0214

The SAS System
Analysis of Variance Procedure

T tests (LSD) for variable: SCORE
NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 12 MSE= 23.83119
Critical Value of T= 2.18
Least Significant Difference= 5.6854

Means with the same letter are not significantly different.

T Grouping	Mean	N	GROUP
A	10.843	7	C
B	3.943	7	T

توجد بينهما اختلافات معنوية وبمستوى 5%

The SAS System
Analysis Variable : SCORE

---------- GROUP=C ----------

Mean	Std Error
10.8428571	2.4116638

---------- GROUP=T ----------

Dr. Firas

Mean	Std Error
3.9428571	0.9963880

ب - **اختبار Tukey**

يعتمد هذا الاختبارعلى قيمة واحدة لمقارنة الفروق بين المتوسطات وتسمى قيمة الفرق المعنوي الامين .H.S.D (Honest Significant Difference) وتكون عادة القيمة المستعملة في المقارنة اكبر من قيمة اقل فرق معنوي لذا فأن هذا الاختبار اكثر تحفظا من اختبار t الا انه يحمل نفس عيوبه.

جـ - اختبار Duncan

هذا الاختبار يعتمد في تطبيقه على عدد من القيم الاحصائية وليس على قيمة واحدة اذ يأخذ بنظر الاعتبار عدد المتوسطات التي التي تدخل ضمن مدى المقارنة ، وبذلك فهو من اكفأ وادق الاختبارات.

د – اختبار Scheffe

يعتمد هذا الاختبار على قيمة واحدة لمقارنة الفروق بين ازواج متوسطات المعاملات.

و- اختبار Newman-Keul

يعتمد على عدة قيم لمقارنة الفروق بين المتوسطات، وسنذكر مثال عن هذا الاختباروالاختبارات السابقة في الفصل اللاحق.

3 – 2 الاختبارات اللامعلمية Nonparametreic Tests

هي الاختبارات التي تجرى لبعض الحالات التي لايستطيع فيها الباحث معرفة او حتى افتراض التوزيع الاحتمالي للمجتمع او المجتمعات التي جاءت منها البيانات المراد اختبارها وبذلك فأن اختبار الاختلافات بين المتوسطات بأستعمال اختبار t يكون غير دقيق، والاختبارات اللامعلمية مفيدة في اختبار البيانات التي تكون نتائجها على شكل تدريجات والتي تشتمل على المتغيرات الاسمية (Nominal) و ذات الترتيب (Ordinal) و ذات الفترات المتساوية (Interval) و النسب (Ratio)،والمتغيرات الاسمية قد لايكون لها ترتيب ومن الامثلة عليها الجنس (ذكر او انثى) وكذلك الالوان اذ لايوجد اتفاق على ترتيبها ضمن تسلسل معين ولكن عندما تكون المتغيرات على شكل فئات وتوجد دلائل على امكانية وضعها ضمن ترتيب معين فأنها تسمى متغيرات ذات ترتيب (Ordinal) ومثال عليها تصنيف الحالة الاقتصادية الى منخفضة ومتوسطة ومرتفعة او تقسيم مجموعة من الافراد على اساس مستوى تعليمهم اعدادي ، بكالوريوس ، ماجستير ، دكتوراه ، وبالرغم من امكانية اعطائهم ارقام 1، 2 ، 3 ، 4 الا ان

الفرق بين المستويات هي ليست متساوية بمعنى ان الفرق بين مستوى تعلم خريج الكلية والاعدادية لايساوي الفرق بين الدكتوراه والماجستير، ولكن في حالة تساوي هذه الفترات في متغيرات اخرى فأنها في هذه الحالة تسمى متغيرات ذات فترة ومثال عليها الدخل السنوي لثلاث افراد 10000 و 15000 و 20000$ اذ ان الفرق بين الاول والذي يليه 5000$.

ان صحة اجراء التحليل الاحصائي تعتمد على صحة تحديد المتغير فمثلا عندما يكون المتغير لون الشعر فليس من المفيد او المنطقي تقدير متوسط لون الشعر كذلك من غير المفيد تقدير المتوسط للصفات التي لها ترتيب مثل متوسط الشهادة العلمية لان الاختلافات بين مستويات الدراسة ليست متساوية بمعنى اننا نحتاج الى المتوسط للمتغيرات ذات الفترات المتساوية ولكن هذا الامر لايتوفر بصورة دائمة اذ ان بعض المتغيرات تكون بين متغيرات الفترة ومتغيرات الترتيب مثل التقييم الخماسي (مقبول جدا ، مقبول ، اعتيادي ، مرفوض ، مرفوض جدا).

بصوره عامة عند دراسة العلاقة بين متغيرين فأن نوع المتغير يحدد نوع التحليل وكما موضح ادناه:

Independent Variable	**Dependent Variable**	
	Continuous Data	Categorical Data
Continuous Data	Regression	Logistic Regression
Categorical Data	ANOVA	Loglinear Model
Mixed (Continuous & Categorical) Data	ANCOVA	Logistic Regression

ان الاختبارات اللامعلمية في الاحصاء الاستدلالي لاتحتاج الى فرضية ان تكون المتغيرات ذات توزيع طبيعي وهي مناسبة للمتغيرات التي تكون على شكل فئات مرتبة اوغير مرتبة. وترتكز على اختبار الوسيط وشكل التوزيع للمتغيرات ولاعلاقة لها بالمتوسطات.

ويمكن تقسيم الاختبارات اللامعلمية حسب نوع المتغيروحجم ونوع العينة كما في الجدول (3).

أ - اختبار مربع كاي لحسن المطابقة X^2 Test

يستعمل هذا الاختبار لاثبات صحة او عدم صحة فرضية توزيع مفردات مجتمع معين بنسب معينة على عدد من المجموعات.

مثال (31) : في مجتمع ما وجد بأن عدد الاحياء كان 37 وعدد الاموات 23 فهل ان الاحياء يشكلون 80% والاموات 20%؟

```
DATA CHISQ;
INPUT GROUP $  K;
cards;
 ALIVE 37
```

Dr. Firas

```
 DEAD  23
;option nodate nonumber;
PROC FREQ DATA=CHISQ;
  TABLES GROUP /testp=(.8,.2);
  WEIGHT K;
RUN;
```

The SAS System

GROUP	Frequency	Percent	Test Percent	Cumulative Frequency	Cumulative Percent
ALIVE	37	61.7	80.0	37	61.7
DEAD	23	38.3	20.0	60	100.0

Chi-Square Test for Specified Proportions

Statistic = 12.604 DF = 1 **Prob = 0.001**

بما ان الاختبار اوضح ان الفروق عالية المعنوية 0.001 = Prop فأن ذلك يعني ان نسبة الاحياء والاموات تختلف بصورة معنوية عن النسب 80 و 20%.

مثال (32) : اذا كانت نسبة الذكور في مدينة برلين بعد الحرب العالمية الثانية هي 8 اناث لكل ذكر وبعد عشر سنوات اخذت عينة عشوائية مؤلفة من 450 شخصا فكان عدد الذكور فيها 68 فهل النسبة بقيت ثابتة ام تغيرت؟

الحل: الفرضية ان نسبة الذكور هي 1/9 وهي تساوي 0.11 اما الاناث فهي 9/8 وتساوي 0.89 .

```
DATA CHISQ;
INPUT SEX $  K;
cards;
MALE    68
FEMALE  382
option nodate nonumber;
PROC FREQ DATA=CHISQ;
TABLES SEX /testp=(.11,.89);   ←  هنا نضع النسب المراد اختبارها
WEIGHT K;
RUN;
```

The SAS System

SEX	Frequency	Percent	Test Percent	Cumulative Frequency	Cumulative Percent
FEMALE	382	84.9	11.0	382	84.9
MALE	68	15.1	89.0	450	100.0

Chi-Square Test for Specified Proportions

Statistic = 2509.505 DF = 1 Prob = 0.001

اذا نرفض الفرضية ونقول ان نسبة الذكور لاتساوي 0.11 ونسبة الاناث لاتساوي 0.89.

Dr.Firas

مثال (33): اخذت عينات من الدم بصورة عشوائية من مجموعة من الناس لغرض اختبار مدى تطابق مجاميع الدم مع النسب 0.05% ، 1% ، 40% ، 45% وكانت البيانات كما هي ادناه؟

```
DATA blood;
INPUT bltype $ number;
cards;
O  465
A  394
B  96
AB 45
option nodate nonumber;
PROC FREQ;
TABLES bltype / testp=(.45, .4, .1, .05); ←      مجموع النسب يجب ان يساوي 1
WEIGHT number;
run;
```

The SAS System

BLTYPE	Frequency	Percent	Test Percent	Cumulative Frequency	Cumulative Percent
A	394	39.4	45.0	394	39.4
AB	45	4.5	40.0	439	43.9
B	96	9.6	10.0	535	53.5
O	465	46.5	5.0	1000	100.0

Chi-Square Test for Specified Proportions

Statistic = 3766.691 DF = 3 Prob = 0.001

بسبب ان الاحتمال يساوي 0.001 لذا فأن مجاميع الدم لاتتطابق والنسب المذكورة.

مثال (34): سجلت بيانات عن اربعة مجاميع لغرض اختبار ان نسبها تتبع النسب التالية 0.5625 و 0.1875 و 0.1875 و 0.0625؟

```
DATA V;
INPUT b $ number;
cards;
O 89
A 33
B 25
G 13
option nodate nonumber;
PROC FREQ DATA = V ORDER=DATA;
TABLES b / testp=(.5625, .1875, .1875,.0625);
WEIGHT number;
run;
```

The SAS System

B	Frequency	Percent	Test Percent	Cumulative Frequency	Cumulative Percent
O	89	55.6	56.3	89	55.6
A	33	20.6	18.8	122	76.3
B	25	15.6	18.8	147	91.9
G	13	8.1	6.3	160	100.0

Dr. Firas

Chi-Square Test for Specified Proportions

Statistic = 2.044 DF = 3 Prob = 0.563

بسبب ان قيمة مربع كاي غير معنوية (Prob = 0.563) فأن النسب الخاصة بالعينة مطابقة للنسب المفترضة. ويمكن تقدير التكرارات او الاحتمالات لكل خلية بأستعمال الايعاز testf= بدلا عن testp=. وبذلك فأن النتائج ستكون:

The SAS System

B	Frequency	Percent	Cumulative Frequency	Cumulative Percent
O	89	55.6	89	55.6
A	33	20.6	122	76.3
B	25	15.6	147	91.9
G	13	8.1	160	100.0

مثال (35): في دراسة لمعرفة هل ان نسبة الذكور الى الاناث تساوي 1:1 وكانت بيانات المواليد كما موضح ادناه؟

```
DATA BORN;
INPUT  GENDER$  N;
CARDS;
MALE   450
FEMALE  501
PROC FREQ DATA=BORN;
TABLES GENDER /CHISQ;
WEIGHT N;
RUN;
```

The SAS system

GENDER	Frequency	Percent	Cumulative Frequency	Cumulative Percent
FEMALE	501	52.7	501	52.7
MALE	450	47.3	951	100.0

Chi-Square Test for Equal Proportions

Statistic = 2.735 DF = 1 Prob = 0.098

اذا الاختلافات في النسبة الجنسية غير معنوية احصائيا اي ان النسبة تتبع 1:1.

ب- اختبار فيشر المضبوط Fisher's Exact Test

يعد هذا الاختبار مناسبا لنفس حالات اختبار مربع كاي ولكنه يكون اكثر ملائمة عندما يكون حجم العينة صغير وعندما تكون بعض الخلايا ذات تكرار اقل من 5 وان يكون لدينا عاملين مستقلين.

بما ان كلا الاختبارين يستعملان لنفس الحالة فأن من المفضل استعمال مربع كاي اذا كان حجم العينة 20 فأكثر بحيث لاتكون هناك خلايا ذات تكرار اقل من 5، فيما يفضل اختبار فيشر اذا كان حجم العينة اقل من 20.

Dr. Firas

مثال (36) : في دراسة لمقارنة نسب الناجحين الى الراسبين في اربعة مجموعات المطلوب اختبار ان النسب متساوية ؟

```
DATA TREND;
INPUT RESULT $ GROUP $ COUNT;
cards;
FAIL A 5
FAIL B 15
FAIL C 14
FAIL D 25
PASS A 2
PASS B 2
PASS C 86
PASS D 5
OPTION NODATE NONUMBER;
PROC FREQ DATA=TREND;
TABLES RESULT*GROUP / CHISQ EXACT;
WEIGHT COUNT;
RUN;
```

```
                                The SAS System
          TABLE OF RESULT BY GROUP

RESULT    GROUP

Frequency,
Percent ,
Row Pct ,
Col Pct ,A     ,B     ,C     ,D     , Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
FAIL    ,    5,   15,   14,   25,   59
      , 3.25, 9.74, 9.09, 16.23, 38.31
      , 8.47, 25.42, 23.73, 42.37,
      , 71.43, 88.24, 14.00, 83.33,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
PASS    ,    2,    2,   86,    5,   95
      , 1.30, 1.30, 55.84, 3.25, 61.69
      , 2.11, 2.11, 90.53, 5.26,
      , 28.57, 11.76, 86.00, 16.67,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total       7     17    100     30    154
          4.55   11.04   64.94   19.48  100.00


      STATISTICS FOR TABLE OF RESULT BY GROUP

Statistic                    DF    Value      Prob
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Chi-Square                    3   71.915     0.001
Likelihood Ratio Chi-Square   3   76.278     0.001
Mantel-Haenszel Chi-Square    1    0.008     0.928
Fisher's Exact Test (2-Tail)           2.35E-17
Phi Coefficient                0.683
Contingency Coefficient        0.564
Cramer's V                     0.683

Sample Size = 154
WARNING: 25% of the cells have expected counts less
         than 5. Chi-Square may not be a valid test
```

نلاحظ ان كلا الاختبارين عالي المعنوية ولكن هنا لابد من اعتماد اختبار فيشر لوجود مشاهدات تقل عن 5 ، كما ان البرنامج يعطي عبارة تحذيرتشير الى عدم صلاحية مربع كاي كاختبار.

جـ أختبار الاستقلال اوالاقتران Test of association or for independence

يستعمل اختبار مربع كاي لغرض تقرير وجود او عدم وجود العلاقة بين متغيرين اي ان الجدول سيكون r x 2 الا ان الاختبار لايحدد قوة العلاقة وانما يشير الى وجودها او عدم وجودها ولغرض معرفة قوة العلاقة فأننا سنحتاج الى مقاييس اخرى مثل معاملات phi و Contingency و cramar.

مثال (37) : في الجيل الثاني للتهجين مابين نبات ذو بذور لونها اصفر وشكلها مستدير مع نبات ذو بذور لونها اخضر وشكلها مجعد حصل مندل على النتائج التالية ، المطلوب معرفة هل ان هناك علاقة بين شكل البذور ولونها؟

الحل: الفرضية هي ان اللون مستقل عن شكل البذور.

لون البذور	شكل البذور	
	مستديرة	مجعدة
صفراء	320	100
خضراء	110	30

```
data d;
input color$ shap$ freq;
cards;
yallow  c 320
yallow  m 100
green   c 110
green   m 30
PROC FREQ DATA=d;
WEIGHT freq;
TABLES color*shap / Chisq;
RUN;
```

```
                                The SAS System
        TABLE OF COLOR BY SHAP

COLOR    SHAP

Frequency,
Percent ,
Row Pct ,
Col Pct ,c     ,m     , Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
green   ,  110 ,   30 ,   140
       , 19.64 ,  5.36 , 25.00
       , 78.57 , 21.43 ,
       , 25.58 , 23.08 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
yallow  ,  320 ,  100 ,  420
       , 57.14 , 17.86 , 75.00
       , 76.19 , 23.81 ,
       , 74.42 , 76.92 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total      430    130    560
          76.79  23.21  100.00

STATISTICS FOR TABLE OF COLOR BY SHAP
```

Dr. Firas

```
Statistic                  DF     Value       Prob
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Chi-Square                  1     0.334      0.563
Likelihood Ratio Chi-Square   1    0.338      0.561
Continuity Adj. Chi-Square    1    0.214      0.644
Mantel-Haenszel Chi-Square    1    0.333      0.564
Fisher's Exact Test (Left)                 0.754
              (Right)                0.325
              (2-Tail)               0.644
Phi Coefficient                0.024
Contingency Coefficient           0.024
Cramer's V                   0.024

Sample Size = 560
```

الاختبارات الاربعة بعد مربع كاي تعد ايضا من الاختبارات المهمة للاقتران فيما تمثل الاختبارات phi و Contingency Coefficient و Cramer's V تقديرات لقوة العلاقة بين المتغيرات

بما ان قيمة مربع كاي غير معنوية فأننا نقبل الفرضية اي ان الصفتان مستقلتين.

مثال (38): هل هناك علاقة مشتركة بين نوع الحزب وجنس اعضائه ؟

الحزب / الجنس	الجمهوري	المستقل	الديموقراطي	المجموع
النساء	279	73	225	577
الرجال	165	47	191	403
المجموع	444	120	416	980

من بيانات الجدول نلاحظ بأن النساء يميلون اكثر الى الحزب الديمقراطي فيما نجد الرجال يميلون الى الحزب الجمهوري لذا يبدو ان هناك علاقة بين العاملين ولكي نتأكد لابد من اجراء اختبار لاثبات ذلك ، لذ سنفترض عدم وجود علاقة بينهما.

```
data S;
input Gender Party count @@;
cards;
1 1 279    1 2  73    1 3 225
2 1 165    2 2  47    2 3 191
;
proc freq data=s order=data;
weight count;
tables gender*party / chisq;
run;
```

ويمكن كتابة البيانات بالشكل التالي:

```
data S;
input Gender Party count;
cards;
1 1  279
1 2  73
1 3  225
2 1  165
2 2  47
2 3  191
;
```

```
The SAS System
     TABLE OF GENDER BY PARTY

GENDER     PARTY
```

Dr. Firas

```
Frequency,
Expected ,     1,     2,     3, Total
ƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
     1 ,   279 ,    73 ,   225 ,   577
       , 261.42 , 70.653 , 244.93 ,
ƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
     2 ,   165 ,    47 ,   191 ,   403
       , 182.58 , 49.347 , 171.07 ,
ƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total       444     120     416     980


      STATISTICS FOR TABLE OF GENDER BY PARTY

Statistic                     DF       Value      Prob
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Chi-Square                     2       7.010      0.030
Likelihood Ratio Chi-Square    2       7.003      0.030
Mantel-Haenszel Chi-Square     1       6.758      0.009
Phi Coefficient                       0.085
Contingency Coefficient               0.084
Cramer's V                            0.085

Sample Size = 980
```

النظرية المفترضة (نظرية العدم) تفترض ان لا علاقة بين العامل المستقل والتابع ، وبما ان قيمة مربع كاي تشير الى وجود فرق معنوي اذا نرفض الفرضية ونقبل بالفرضية البديلة حول ان العاملان غير مستقلان اي توجد علاقة بينهما.

مثال (39) : البيانات ادناه تمثل عدد الطلبة الذين لديهم سيارات والذين لايملكون سيارات وللمراحل الأربعة فهل توجد علاقة بين المرحلة وتملك السيارات؟

الحالة	المرحلة الاولى	المرحلة الثانية	المرحلة الثالثة	المرحلة الرابعة
يملكون	29	12	8	21
لايملكون	81	118	92	139

```
DATA TREND;
INPUT CAR $ GROUP $ COUNT;
cards;
CAR  A 29
CAR  B 12
CAR  C 8
CAR  D 21
NOCAR A 81
NOCAR B 118
NOCAR C 92
NOCAR D 139
OPTION NODATE NONUMBER;
PROC FREQ DATA=TREND;
TABLES CAR*GROUP / CHISQ;
WEIGHT COUNT;
RUN;
```

```
                                The SAS System

          TABLE OF GROUP BY CAR

          GROUP    CAR
```

Dr. Firas

```
Frequency,
Percent  ,
Row Pct  ,
Col Pct  ,CAR     ,NOCAR   ,  Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆ
A        ,     29 ,     81 ,    110
         ,   5.80 ,  16.20 ,  22.00
         ,  26.36 ,  73.64 ,
         ,  41.43 ,  18.84 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆ
B        ,     12 ,    118 ,    130
         ,   2.40 ,  23.60 ,  26.00
         ,   9.23 ,  90.77 ,
         ,  17.14 ,  27.44 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆ
C        ,      8 ,     92 ,    100
         ,   1.60 ,  18.40 ,  20.00
         ,   8.00 ,  92.00 ,
         ,  11.43 ,  21.40 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆ
D        ,     21 ,    139 ,    160
         ,   4.20 ,  27.80 ,  32.00
         ,  13.13 ,  86.88 ,
         ,  30.00 ,  32.33 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆ
Total          70      430      500
            14.00    86.00   100.00

          The SAS System

      STATISTICS FOR TABLE OF GROUP BY CAR

Statistic                     DF       Value      Prob
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Chi-Square                     3     19.513     0.001
Likelihood Ratio Chi-Square    3     17.866     0.001
Mantel-Haenszel Chi-Square     1      6.323     0.012
Phi Coefficient                       0.198
Contingency Coefficient               0.194
Cramer's V                            0.198

Sample Size = 500
```

بما ان قيمة مربع كاي معنوية لذا نرفض الفرضية ونقول ان تملك السيارات ومرحلة الدراسة عاملان غير مستقلان.

د- اختبار مربع كاي في حالة اكثر من عينتين مستقلتين

في هذا الاختبار يكون الجدول r x k اي اكثر من صفين واكثر من عمودين.

مثال (40) : اخذت عينة من 500 سائق بأعمار مختلفة لدراسة العلاقة بين العمر وعدد الحوادث؟

عدد الحوادث	عمر السائق			
	31-20	40-31	50-41	60-51
0	34	14	6	36
1	82	20	34	64
2	84	16	60	50

الفرضية ان عمر السائق ليس له علاقة بعدد الحوادث.

```
data d;
input age$ acc freq;
cards;
```

Dr.Firas

```
20-31   0 34
20-31   1 82
20-31   2 84
31-40   0 14
31-40   1 20
31-40   2 16
41-50   0 6
41-50   1 34
41-50   2 60
51-60   0 36
51-60   1 64
51-60   2 50
PROC FREQ DATA=d;
WEIGHT freq;
TABLES acc*age / Chisq;
run;
```

```
                TABLE OF ACC BY AGE
          ACC     AGE
          Frequency,
          Percent ,
          Row Pct ,
          Col Pct ,20-31  ,31-40  ,41-50  ,51-60  , Total
          ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
              0 ,   34 ,   14 ,    6 ,   36 ,   90
               , 6.80 , 2.80 , 1.20 , 7.20 , 18.00
               , 37.78 , 15.56 , 6.67 , 40.00 ,
               , 17.00 , 28.00 , 6.00 , 24.00 ,
          ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
              1 ,   82 ,   20 ,   34 ,   64 ,  200
               , 16.40 , 4.00 , 6.80 , 12.80 , 40.00
               , 41.00 , 10.00 , 17.00 , 32.00 ,
               , 41.00 , 40.00 , 34.00 , 42.67 ,
          ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
              2 ,   84 ,   16 ,   60 ,   50 ,  210
               , 16.80 , 3.20 , 12.00 , 10.00 , 42.00
               , 40.00 , 7.62 , 28.57 , 23.81 ,
               , 42.00 , 32.00 , 60.00 , 33.33 ,
          ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
          Total      200      50     100     150     500
                   40.00   10.00   20.00   30.00  100.00

               STATISTICS FOR TABLE OF ACC BY AGE

          Statistic                    DF     Value      Prob
          ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
          Chi-Square                    6    26.693     0.001
          Likelihood Ratio Chi-Square   6    28.261     0.001
          Mantel-Haenszel Chi-Square    1     0.574     0.449
          Phi Coefficient                    0.231
          Contingency Coefficient            0.225
          Cramer's V                         0.163
          Sample Size = 500
```

نتيجة اختبار مربع كاي تشير الى ان هناك اختلافات في نسبة الحوادث بسبب العمر لذا نرفض الفرضية التي مفادها ان المتغيران مستقلين.

هـ - أختبار كولموجورف- سميرنوف لعينتين Kolmogorov-Sminov Test

يستعمل لغرض اختبار ان العينة المأخوذة بصوررة عشوائية ذات توزيع طبيعي . وفي كثير من الحالات يعد من انسب الاختبارات اللامعلمية لاختبار الاختلافات بين عينتين. ولكن اذا افترضنا ان انتشار وشكل التوزيع متشابه في عينتين فأن اختبار ولكاكسن للمرتبة وعلامتها يكون افضل.

مثال (41) : في تجربة لمقارنة معاملي الهضم بين الاغنام والعجول اختبر هل ان المجموعتان لهما نفس شكل التوزيع؟

```
data digest;
input group dm;
cards;
1 57.8
1 56.2
1 61.9
1 54.4
1 53.6
1 56.4
1 53.2
2 64.2
2 58.7
2 63.1
2 62.5
2 59.8
2 59.2
proc npar1way  edf;
class group;
var dm;
run;
```

Dr. Firas

```
                                                The SAS System
                N P A R 1 W A Y  P R O C E D U R E

                 Kolmogorov-Smirnov Test for Variable DM
                    Classified by Variable GROUP

                                                Deviation
                                    EDF            from Mean
GROUP                N          at Maximum          at Maximum

 1                   7          0.857142857          1.04667085
 2                   6          0.000000000         -1.13053373
                    ----        -----------
                    13          0.461538462

              Maximum Deviation Occurred at Observation    1
              Value of DM at Maximum  57.8000000

                 Kolmogorov-Smirnov 2-Sample Test (Asymptotic)
                    KS = 0.427302        D = 0.857143
                   KSa =  1.54066        Prob > KSa = 0.0174
```

المعاملتان تختلفان في شكلي توزيعهما وهما غير متجانستان لذا فأن اختبار t سيكون غير دقيق والاختبارات اللامعلمية هي الافضل.

```
                        The SAS System

                N P A R 1 W A Y  P R O C E D U R E

                Cramer-von Mises Test for Variable DM
                   Classified by Variable GROUP

                                Summed
```

GROUP	N	Deviation from Mean
1	7	0.358410820
2	6	0.418145957

Cramer-von Mises Statistic (Asymptotic)
CM = 0.059735 CMa = 0.776557

يستعمل هذا الاختبار لاختبار تجانس التباين عندما يكون حجم العينة 2000 فأكثر وبعكسه فأن اختبار سمينوف يكون اكثر دقة.

The SAS System
N P A R 1 W A Y P R O C E D U R E
Kuiper Test for Variable DM
Classified by Variable GROUP

GROUP	N	Deviation from Mean
1	7	0.857142857
2	6	0.000000000

Kuiper 2-Sample Test (Asymptotic)
K = 0.857143 Ka = 1.54066 Prob > Ka = 0.1474

ملاحظة: قد تختلف نتائج الاختبارات المختلفة فيما يخص المعنوية وذلك امر طبيعي بمعنى ان الاختبار الذي يشير الى معنوية الاختلافات مثلا اختبار t قد لايطابق اختبار Duncan وكذلك الحال بالنسبة للاختبارات اللامعلمية.

ملاحظة: درج العديد من الباحثين على اعتماد اختبار t للمقارنة بين المتوسطات في جميع التحليلات الاحصائية دون الاهتمام بتوفر شروط ذلك الاختبار ،ورغم ان هذا الاختبار مهم ودقيق عند توفر شروط اعتماده وهي ان يكون توزيع العينات توزيعا طبيعيا وان يكون التباين متجانس بين المعاملات وان تكون العوامل مستقلة فأنه يكون اختبارا خاطئا عند عدم توفر اي من الشروط وهو امر ممكن ان نجده في بعض العينات كما في المثال التالي.

مثال (42): ادناه بيانات معاملتان ونلاحظ ان متوسطيهما مقارب الى الصفرولكن تباينهما يختلف اختلافا واضحا اذ نجد ان المعاملة الاولى تتراوح قيمها من 2.5 الى – 2.5 فيما تتراوح قيم الثانية بين 6 و – 6 واذا كان اختبار t يكشف الاختلاف في التباين الا انه ليس كذلك فيما يخص التوزيع الطبيعي ! فيما نجد ان اختبار كولموجورف- سيرنوف يكشف ذلك ، اذ انه يختبر فرضية العدم التي تنص بأن مشاهدات متغير ما تتبع التوزيع الطبيعي ضد الفرضية البديلة التي تنص على انها لاتتوزع طبيعيا.

```
data s;
input animal dm;
cards;
1 0.22
1 -0.87
1 -2.39
1 -1.79
1 0.37
```

```
1 -1.54
1 1.28
1 -0.31
1 -0.74
1 0.72
1 0.38
1 -0.17
1 -0.62
1 -1.10
1 0.30
1 0.15
1 1.30
1 0.19
1 -0.50
1 -0.09
2 -5.13
2 -2.19
2 -2.43
2 -3.83
2 0.50
2 -3.25
2 4.32
2 1.63
2 5.18
2 -0.43
2 7.11
2 4.87
2 -3.10
2 -5.81
2 3.76
2 6.31
2 2.58
2 0.07
2 5.76
2 3.50
proc ttest;
class animal;
var dm;
run;

proc npar1way anova edf wilcoxon;
class animal;
var dm;
run;
```

```
                    The SAS System
                N P A R 1 W A Y  P R O C E D U R E

                Analysis of Variance for Variable DM
                  Classified by Variable ANIMAL


ANIMAL        N          Mean          Among MS       Within MS
                                    15.1659225      8.76827566
 1           20     -.260500000
 2           20     0.971000000          F Value       Prob > F
                                          1.730         0.1963
                    The SAS System
                N P A R 1 W A Y  P R O C E D U R E
```

Dr. FiraS

Wilcoxon Scores (Rank Sums) for Variable DM
Classified by Variable ANIMAL

ANIMAL	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	20	373.0	410.0	36.9684550	18.6500000
2	20	447.0	410.0	36.9684550	22.3500000

Wilcoxon 2-Sample Test (Normal Approximation)
(with Continuity Correction of .5)

S = 373.000 Z = -.987328 Prob > |Z| = 0.3235

T-Test Approx. Significance = 0.3296

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 1.0017 DF = 1 Prob > CHISQ = 0.3169

The SAS System
N P A R 1 W A Y P R O C E D U R E

Kolmogorov-Smirnov Test for Variable DM
Classified by Variable ANIMAL

ANIMAL	N	EDF at Maximum	Deviation from Mean at Maximum
1	20	1.00000000	1.11803399
2	20	0.50000000	-1.11803399
	40	0.75000000	

Maximum Deviation Occurred at Observation 17
Value of DM at Maximum 1.30000000

Kolmogorov-Smirnov 2-Sample Test (Asymptotic)
KS = 0.250000 D = 0.500000
KSa = 1.58114 Prob > KSa = 0.0135

The SAS System
N P A R 1 W A Y P R O C E D U R E

Cramer-von Mises Test for Variable DM
Classified by Variable ANIMAL

ANIMAL	N	Summed Deviation from Mean
1	20	0.313750000
2	20	0.313750000

Cramer-von Mises Statistic (Asymptotic)
CM = 0.015688 CMa = 0.627500

The SAS System
N P A R 1 W A Y P R O C E D U R E

Kuiper Test for Variable DM
Classified by Variable ANIMAL

ANIMAL	N	Deviation from Mean
1	20	0.500000000
2	20	0.300000000

Dr Firas

```
                Kuiper 2-Sample Test (Asymptotic)
        K = 0.800000      Ka =  2.52982      Prob > Ka = 0.0001
                        The SAS System
                        TTEST PROCEDURE

Variable: DM

ANIMAL   N   Mean    Std Dev   Std Error   Variances      T      DF    Prob>|T|
-------------------------------------------------------  -------------------------
1     20  -0.26050  0.960747  0.21482977   Unequal     -1.3152    21.1     0.2026
2     20   0.971000 4.075967  0.91141414   Equal       -1.3152    38.0     0.1963

For H0: Variances are equal, F' =  18.00   DF = (19,19)   Prob>F' = 0.0000
```

و- اختبارالعلامة واختبار ولكاكسن للمرتبة وعلامتها Wilcoxon signed rank test

يستعمل اختبار العلامة للبيانات التي تكون على شكل ازواج لغرض اختبار تشابه توزيع المجموعتين ولايأخذ هذا الاختبار بنظر الاعتباراختلاف قيم الفروق الموجبة او السالبة على العكس من اختبار ولكاكسن للمرتبة وعلامتها.

مثال (43): البيانات ادناه تمثل درجات عدد من الطلبة قبل وبعد اجراء دورة تقوية في درس الاحصاء (الاختبار من 20) اختبر معنوية الاختلافات بعدة طرق؟

```
data exam ;
input pre post ;
diff = post - pre ;
cards ;
7  11
6  14
10 16
16 17
8  9
13 15
8  9
14 17
16 20
11 12
12 14
13 15
;
proc univariate ;
var diff ;
run ;
```

```
                        The SAS System
                          Univariate Procedure

Variable=DIFF

          Moments                          Quantiles(Def=5)

N           12  Sum Wgts      12      100% Max      8      99%       8
Mean      2.916667  Sum         35      75% Q3        4      95%       8
Std Dev   2.234373  Variance  4.992424    50% Med       2      90%       6
Skewness  1.298107  Kurtosis  1.131608    25% Q1        1      10%       1
USS          157  CSS        54.91667    0% Min        1      5%        1
CV        76.60709  Std Mean  0.645008                       1%        1
```

T:Mean=0	4.521908	Pr>|T|	**0.0009**
Num ^= 0	12	Num > 0	12
M(Sign)	6	Pr>=|M|	**0.0005**
Sgn Rank	39	Pr>=|S|	**0.0005**

Range	7		
Q3-Q1	3		
Mode	1		

Extremes

Lowest	Obs	Highest	Obs
1(	10)	3(	8)
1(	7)	4(	1)
1(	5)	4(	9)
1(	4)	6(	3)
2(	12)	8(	2)

تبين من النتائج ان اختبارات t و Sign و Wilcoxon هي 0.0009 و 0.0006 و 0.0005 وجميعها عالية المعنوية.

ز - اختبار ولكاكسن ومان ووتني لعينتين Wilcoxon – Mann – Whitney Test

يستعمل هذا الاختبار اذا كان لدينا عينتان مستقلتان عن بعضهما البعض والمتغير يكون ذو ترتيب اذ ان اختبار ولكاكسن للمرتبة وعلامتها لاتنطبقان على هذه الحالة بسبب عدم وجود ازواج من البيانات.

مثال (44): في دراسة لمقارنة متوسطي معاملات هضم لمادة علفية على مجموعة من الاغنام والعجول وكانت معاملات الهضم معبر عنها بنسب مؤية اختبر فرضية تساوي وسيطيهما؟

```
data react;
input trt response;
cards;
1 53.2
1 53.6
1 54.4
1 56.2
1 56.4
1 57.8
1 61.9
2 58.7
2 59.2
2 59.8
2 62.5
2 63.1
2 64.2
;option nodate nonumber;
proc npar1way wilcoxon data=react;
class trt;
var response;
exact;
run;
```

Dr. Firas

The SAS System
N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable RESPONSE
Classified by Variable TRT

TRT	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score

1	7	31.0	49.0	7.0	4.4285714
2	6	60.0	42.0	7.0	10.0000000

Wilcoxon 2-Sample Test (Normal Approximation)
(with Continuity Correction of .5)

S = 60.0000 Z = 2.50000 Prob > |Z| = **0.0124**

T-Test Approx. Significance = 0.0279

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 6.6122 DF = 1 Prob > CHISQ = 0.0101

اذا نرفض فرضية تساوي الوسيط لمعاملي الهضم اعتمادا على اختبار ولكاكسن فيما لايصلح اعتماد اختبار كرسكال لانه يخص اكثر من عينتين.

حـ - احصائية Cohen's Kappa (k)

هو مقياس للتوافق بين المتغيرات ذات الفئات المستقلة لمقارنة التوافق المشاهد مع التوافق المتوقع بسبب الصدفة ويكون مناسبا للمتغيرات الاسمية التي لايكون لها ترتيب فيما تستعمل k الموزونة للمتغيرات ذات الترتيب ، وتتراوح قيمة k من 1 الى 0 وقد تكون في بعض الحالات القليلة سالبة.

والجدول ادناه يوضح التقديرات التقريبية المناظرة لقيمة k

k	
اقل من 0	لاتوافق
0 — 0.20	توافق قليل جدا
0.21 — 0.40	توافق قليل
0.41 — 0.60	توافق متوسط
0.61 — 0.80	توافق عالي
0.81 — 1.00	توافق عالي جدا

مثال (45) قدر درجة التوافق في التشخيص الطبي لدى طبيبان A و B مستقلان عن بعضهما لدى تشخيصهم لحالة مرضية معينة لدى 29 شخصا اذ اكد الطبيب A وجود 19 حالة اصابة و10 حالات سليمة وتقديرات الطبيب B المناظرة بلغت 12 و 17 وقد تطابق تشخيصهما بعدم وجود المرض في 10 ووجود المرض في 12 حالة ؟

	Doctor A		
	No	Yes	Total

Dr. Firas

Doctor B	No	10 (34.5%)	7 (24.1%)	17 (58.6%)
	Yes	0 (0.0%)	12 (41.4%)	12 (41.4%)
Total		10 (34.5%)	19 (65.5%)	29

Kappa = (Observed agreement - Chance agreement)/(1 - Chance agreement)
Observed agreement = (10 + 12)/29 = 0.76
Chance agreement = 0.586 * 0.345 + 0.655 * 0.414 = 0.474
Kappa = (0.76 - 0.474)/(1 - 0.474) = 0.54

```
data A;
input a b c;
cards;
1 1 12
1 0 7
0 1 0
0 0 10
;
proc freq;
tables a * b/agree;
weight c;
run;
```

```
                TABLE OF A BY B

A       B

Frequency,
Percent ,
Row Pct ,
Col Pct ,       0,      1,  Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
         0 ,    10 ,     0 ,   10
           , 34.48 ,  0.00 ,  34.48
           ,100.00 ,  0.00 ,
           , 58.82 ,  0.00 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
         1 ,     7 ,    12 ,   19
           , 24.14 , 41.38 , 65.52
           , 36.84 , 63.16 ,
           , 41.18 ,100.00 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total          17      12      29
              58.62   41.38   100.00


          STATISTICS FOR TABLE OF A BY B

                McNemar's Test
                --------------
Statistic = 7.000       DF = 1       Prob = 0.008


             Simple Kappa Coefficient
             ------------------------
                       95% Confidence Bounds
```

Dr Firas

Kappa = 0.542 ASE = 0.134 0.279 0.805
Sample Size = 29

يوجد توافق متوسط بين الطبيبين

ط - اختبار كرسكال ووالس لعدة عينات Kruskal – Wallis Test

يستعمل هذا الاختبار في حالة وجود اكثر من عينتين مستقلتين والمتغير ذو ترتيب . وهذا الاختبار لايفترض بأن العينات ذات توزيع طبيعي وهي الميزة المهمه له وانما يفترض ان العينات المأخوذة من عدة مجتمعات او عشائر لها نفس شكل التوزيع لذا فأن المجتمعات عندما تكون مختلفة في شكل توزيعها مثلا واحدة منحرفة الى اليمين والاخرى منحرفة الى اليسار فأن هذا الاختبارقد يعطي نتائج غير صحيحة.

يستعمل كرسكال لاختبار فرضية ان العينات المأخوذة من عدة مجتمعات هي متطابقة في مواقعها او بمعنى اخر ان متوسط رتب (Mean rank) عينات المجتمعات تكون متساوية، فكرسكال لايختبر فرضية ان العينات لها متوسطات متساوية كما لايختبر فرضية تساوي الوسيط للعينات. ولتوضيح ذلك فأن المثال التالي يعود الى ثلاثة عينات متساوية في المعدل (43.5) والوسيط (27.5) الا ان اختبار كرسكال كان معنوي.

مثال (46): اذا كانت البيانات ادناه تمثل نتائج ثلاثة عينات مستقلة اختبر فرضية انتماء هذه العينات لمجتمع واحد؟

```
data d;
input  Group trait;
cards;
1 1
1 2
1 3
1 4
1 5
1 6
1 7
1 8
1 9
1 46
1 47
1 48
1 49
1 50
1 51
1 52
1 53
1 342
2 10
2 11
2 12
```

Dr. Firas

```
2 13
2 14
2 15
2 16
2 17
2 18
2 37
2 58
2 59
2 60
2 61
2 62
2 63
2 64
2 193
3 19
3 20
3 21
3 22
3 23
3 24
3 25
3 26
3 27
3 28
3 65
3 66
3 67
3 68
3 69
3 70
3 71
3 72
;option nodate nonumber;
proc npar1way wilcoxon data=d;
class group;
var TRAIT;
run;
```

Dr. Firas

The SAS System

N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable TRAIT
Classified by Variable GROUP

GROUP	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	18	367.0	495.0	54.4977064	20.3888889
2	18	495.0	495.0	54.4977064	27.5000000
3	18	623.0	495.0	54.4977064	34.6111111

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 7.3553 DF = 2 Prob > CHISQ = 0.0253

ويمثل جدول تحليل التباين واختبار t طريقة افضل واسهل لاختبار البيانات مالم توجد دلائل قوية على ان توزيع الصفة غير طبيعي ويمكن التأكد من توزيع الصفة حسب المجاميع بأستعمال الايعاز التالي :

```
;option nodate nonumber;
proc univariate normal plot;
run;
```

ولو تفحصنا النتائج سيتضح لنا ان الصفة توزيعها غير طبيعي.

مثال (47): اذا كانت البيانات ادناه تمثل نتائج ثلاثة عينات مستقلة اختبر فرضية انتماء هذه العينات لمجتمع واحد؟

```
data K;
input trt$ TRAIT;
cards;
A 84.4
A 116
A 84
A 68.6
B 75.2
B 62.4
B 62.4
B 73.8
B 69.2
C 56.4
C 83.2
C 90.4
C 85.6
;option nodate nonumber;
proc npar1way wilcoxon data=K;
class trt;
var TRAIT;
run;
```

The SAS System

N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable TRAIT
Classified by Variable TRT

TRT	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
A	4	36.0	28.0	6.47183246	9.00000000
B	5	23.0	35.0	6.82191040	4.60000000
C	4	32.0	28.0	6.47183246	8.00000000

Average Scores Were Used for Ties

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 3.2264 DF = 2 Prob > CHISQ = 0.1992

بما ان الاختبار غير معنوي فأن العينات مأخوذة من نفس المجتمع.

مثال (48): اختبر الاختلافات في كمية الكورتيزول في ثلاث مجاميع من الافراد؟

```
data cortisol;
  input group $ subject cortisol;
  cards;
I  1 262
```

Dr. Firas

```
I   2  307
I   3  211
I   4  323
I   5  454
I   6  339
I   7  304
I   8  154
I   9  287
I  10  356
II  1  465
II  2  501
II  3  455
II  4  355
II  5  468
II  6  362
III 1  343
III 2  772
III 3  207
III 4 1048
III 5  838
III 6  687
proc npar1way wilcoxon;
  class group;
  var cortisol;
run;
```

Dr. Firas

The SAS System

N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable CORTISOL
Classified by Variable GROUP

GROUP	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
I	10	69.0	115.0	15.1657509	6.9000000
II	6	90.0	69.0	13.5646600	15.0000000
III	6	94.0	69.0	13.5646600	15.6666667

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 9.2316 DF = 2 Prob > CHISQ = 0.0099

العينات لاتعود الى نفس المجتمع.

ي– اختبار P المضبوط Exact – P Test

مثال (49) : لو اخذنا المثال السابق واضفنا الى قطعة البرنامج الكلمة exact

```
data react;
input stim time @@;
cards;
1 1.94 1 1.94 1 2.92 1 2.92
1 2.92 1 2.92 1 3.27 1 3.27
1 3.27 1 3.27 1 3.70 1 3.70
1 3.74 2 3.27 2 3.27 2 3.27
2 3.70 2 3.70 2 3.74
proc npar1way wilcoxon data=react;
```

```
class stim;
var time;
exact;
run;
```

```
                                The SAS System
                          N P A R 1 W A Y  P R O C E D U R E
                  Wilcoxon Scores (Rank Sums) for Variable TIME
                          Classified by Variable STIM

            Sum of       Expected    Std Dev       Mean
STIM    N   Scores       Under H0    Under H0      Score

1      13   110.500000    130.0      11.0047836    8.5000000
2       6   79.500000     60.0       11.0047836   13.2500000
                    Average Scores Were Used for Ties

Wilcoxon 2-Sample Test     S =  79.5000

Exact P-Values
(One-sided)  Prob >= S        = 0.0527
(Two-sided)  Prob >= |S - Mean| = 0.1054

Normal Approximation (with Continuity Correction of .5)
Z =  1.72652         Prob > |Z| = 0.0843

T-Test Approx. Significance = 0.1014

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ =  3.1398       DF =  1        Prob > CHISQ = 0.0764
```

ويلاحظ ان اختبار Exact p-Values بلغت 0.0527 مما يعني ان الاختلافات غير معنوية فيما نجد ان اختبار z على الطرفين 0.084 والذي يمثل 0.042 لكل طرف يشير الى وجود فرق معنوي الا ان الاختبار الاول يعد اكثر دقة. اذ ان اختبار z يؤخذ به عندما يكون عدد الازواج اكثر من 25.

مثال (50) : في تجربة على 15 فأر لدراسة تأثير ثلاثة عقارات على فترة بقائها بالايام اختبر الاختلافات بين العقاقير؟

```
data mice;
input trt $ days @@;
cards;
1 1 1 1 1 3 1 3 1 4
2 3 2 4 2 4 2 4 2 15
3 4 3 4 3 10 3 10 3 26
proc npar1way savage data=mice;
class trt;
var days;
exact;
run;
```

```
                          N P A R 1 W A Y  P R O C E D U R E
                  Savage Scores (Exponential) for Variable DAYS
                          Classified by Variable TRT

            Sum of       Expected   Std Dev       Mean
TRT     N   Scores       Under H0   Under H0      Score

1       5   -3.36798017    0.0      1.63455487    -.673596034
2       5    0.09561827    0.0      1.63455487    0.019123654
```

```
3       5     3.27236190     0.0      1.63455487     0.654472379
                  Average Scores Were Used for Ties

              Savage 1-Way Analysis    S =  5.5047

              Exact P-Value     Prob >= S  = 0.0445
               Chi-Square Approximation
                 DF = 2       Prob >  S  = 0.0638
```

يلاحظ من النتائج ان قيمة 0.0445 Exact p-Value تشير الى وجود اختلافات معنوية بمستوى 5% في فترة البقاء للفئران بسبب العقار الا ان اختبار مربع كاي لايشير الى وجود اختلافات معنوية .

مثال (51) في دراسة شملت 59 مريض قسموا الى مجموعتين الاولى (27 مريض) تناولوا عقار والمجموعة الثانية (32 مريض) placebo (لازالة التأثير النفسي)وسجلت نتائج الاستجابة الى ممتاز ، جيد ، متوسط ، ضعيف ، ضعيف جدا. المطلوب اختبار فرضية عدم اختلاف الاستجابة لدى المجموعتين؟

```
data one;
input trt $ resp  ;
cards;
Active 1
Active 1
Active 1
Active 1
Active 1
Active 2
Active 2
Active 2
Active 2
Active 2
Active 2
Active 2
Active 2
Active 2
Active 2
Active 2
Active 2
Active 3
Active 3
Active 3
Active 3
Active 3
Active 3
Active 4
Active 5
Active 5
Active 5
Active 5
Active 5
Placebo 1
Placebo 1
Placebo 2
Placebo 2
```

Dr. Firas

Placebo 2
Placebo 2
Placebo 3
Placebo 3
Placebo 3
Placebo 3
Placebo 3
Placebo 3
Placebo 3
Placebo 4
Placebo 4
Placebo 4
Placebo 4
Placebo 4
Placebo 4
Placebo 4
Placebo 5
Placebo 5
Placebo 5
Placebo 5
Placebo 5
Placebo 5
Placebo 5
Placebo 5
Placebo 5
Placebo 5
Placebo 5
Placebo 5
Placebo 5
Placebo 5

```
proc npar1way wilcoxon data=one;
class trt;
var resp;
exact;
run;
```

The SAS System
N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable RESP
Classified by Variable TRT

TRT	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
Active	27	621.0	810.0	63.9727441	23.0000000
Placebo	32	1149.0	960.0	63.9727441	35.9062500

Average Scores Were Used for Ties

Wilcoxon 2-Sample Test S = 621.000

Exact P-Values
(One-sided) Prob <= S = 0.0014
(Two-sided) Prob >= |S - Mean| = **0.0028**

Normal Approximation (with Continuity Correction of .5)
Z = -2.94657 Prob > |Z| = **0.0032**

T-Test Approx. Significance = 0.0046

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 8.7284 DF = 1 Prob > CHISQ = 0.0031

نتائج اختبار نظرية العدم تشير الى ان معنوية الفروق في اختبار p المضبوط (0.0028) واختبار z (0.0032) والاختبار الاخير اكثر دقة بسبب ان حجم العينة كبير نسبيا.

Dr Firas

ك- اختبار ماكنيمار McNemar Test

اختبار لامعلمي للبيانات لغرض تقرير تساوي تكرارات الصفوف والاعمدة (2× 2) على ان تكون المتغيرات الاسمية ثنائية وان تكون العينتين على شكل ازواج.

مثال (52) في دراسة تضمنت سؤال 200 طبيبا مقيما عن رأيه بخصوص الاستعمال الدوري لل TAP بدلا عن streptokinase في علاج الاصابات بـ myocardial وكان عدد الاطباء الذين يعتمدون استعمال TAP هو 51 ومن لايستعمله هو 149 ثم نفذ برنامج تعليمي باستعمالات TAP وتأثيراته على المرضى ثم اعيد السؤال نفسه على نفس الاطباء وكانت اجاباتهم 130 و 70 فهل كان للبرنامج تأثيرا على تغيير رأيهم بخصوص العلاج اذا علمت ان 57 من مجموع الذين لم يستعملوه (149) بقوا رافضين استعماله و38 من مجموع الراغبين (51) بقوا راغبين ؟

بعد البرنامج	قبل البرنامج		
	Pro-TAP	Anti-TAP	المجموع
Pro-TAP	38	92	130
Anti-TAP	13	57	70
المجموع	51	149	200

```
DATA COMPAR;
INPUT TEST1 TEST2  number;
cards;
1 1 38
1 2 92
2 1 13
2 2 57
PROC FREQ DATA = COMPAR ORDER=DATA;
TABLES TEST1*TEST2 / agree;
WEIGHT number;
run;
```

```
                                        The SAS System

          TABLE OF TEST1 BY TEST2

         TEST1    TEST2

         Frequency,
         Percent ,
         Row Pct ,
         Col Pct ,     1,     2, Total
         ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
             1 ,   38 ,   92 ,  130
               , 19.00 , 46.00 , 65.00
               , 29.23 , 70.77 ,
               , 74.51 , 61.74 ,
         ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
             2 ,   13 ,   57 ,   70
               ,  6.50 , 28.50 , 35.00
```

Dr. Firas

```
          , 18.57 , 81.43 ,
          , 25.49 , 38.26 ,
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
       Total       51     149     200
             25.50   74.50   100.00

    STATISTICS FOR TABLE OF TEST1 BY TEST2

          McNemar's Test
          ---------------
Statistic = 59.438      DF = 1      Prob = 0.001

       Simple Kappa Coefficient
       ------------------------
                 95% Confidence Bounds
Kappa = 0.085    ASE = 0.049     -0.012    0.181
Sample Size = 200
```

هناك تغير معنوي في تفضيل استعمال TAP

ل- اختبار الوسيط Median Test

يستعمل هذا الاختبار في حالة عينتين مستقلتين.

مثال(53): مجموعتين من الفئران عرضت الى محفزين وسجل الوقت لحين حصول الاستجابة

هل هناك فرق بين المجموعتين بسبب نوع المحفز؟

```
data react;
input stim time @@;
cards;
1 1.94 1 1.94 1 2.92 1 2.92
1 2.92 1 2.92 1 3.27 1 3.27
1 3.27 1 3.27 1 3.70 1 3.70
1 3.74 2 3.27 2 3.27 2 3.27
2 3.70 2 3.70 2 3.74
proc npar1way median  data=react;
class stim;
var time;run;
```

```
                                The SAS System
          N P A R 1 W A Y  P R O C E D U R E
      Median Scores (Number of Points Above Median)
            for Variable TIME
          Classified by Variable STIM

          Sum of      Expected      Std Dev       Mean
STIM   N   Scores     Under H0      Under H0      Score

1     13  4.71428571  6.15789474  0.830274031  0.362637363
2      6  4.28571429  2.84210526  0.830274031  0.714285714
          Average Scores Were Used for Ties

      Median 2-Sample Test (Normal Approximation)
      S =  4.28571      Z =  1.73871      Prob > |Z| = 0.0821

      Median 1-Way Analysis (Chi-Square Approximation)
      CHISQ =  3.0231      DF =  1      Prob > CHISQ = 0.0821
```

م- اختبار فان دير فيردن Van der Vaerden Test

يستعمل هذا الاختبار في حالة اكثر من عينة مستقلة للمتغيرات التي لايكون لها ترتيب .

Dr. Firas

مثال (54) في تجربة لدراسة شملت 15 فأرا اعطيت ثلاثة انواع من العقاقير لدراسة تأثير تلك العقاقير على فترة بقائها بالايام ، المطلوب اختبار فرضية ان فترة بقاء الفئران متساوية؟

```
data mice;
input trt $ days @@;
cards;
1 1 1 1 1 3 1 3 1 4
2 3 2 4 2 4 2 4 2 15
3 4 3 4 3 10 3 10 3 26
proc npar1way vw data=mice;
class trt;
var days;
run;
```

The SAS System

N P A R 1 W A Y P R O C E D U R E

Van der Waerden Scores (Normal) for Variable DAYS
Classified by Variable TRT

TRT	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	5	-3.96994900	0.0	1.53844289	-.793989801
2	5	0.71126668	0.0	1.53844289	0.142253336
3	5	3.25868232	0.0	1.53844289	0.651736465

Average Scores Were Used for Ties

Van der Waerden 1-Way Analysis (Chi-Square Approximation)
CHISQ = 7.5729 DF = 2 Prob > CHISQ = 0.0227

تبين من نتائج الاختبار (p=0.0227) رفضنا لفرضية العدم وان الفئران في المجاميع الثلاثة تختلف في فترة بقائها.

ن - اختبار سافيج Savage Test

يستعمل هذا الاختبار في حالة اكثر من عينة مستقلة للمتغيرات التي لايكون لها ترتيب .

مثال (55) استعملت ثلاثة انواع من العقارات بصورة عشوائية على ثلاثة معاملات من الفئران المصابة بمرض ما وسجلت عدد ايام بقائها المطلوب اختبار الاختلافات في بقاء الفئران بسبب اختلاف المعاملة؟

```
data mice;
input trt  days @@;
cards;
1 1 1 1 1 3 1 3 1 4
2 3 2 4 2 4 2 4 2 15
3 4 3 4 3 10 3 10 3 26
proc npar1way savage  data=mice;
class trt;
var days;
run;
```

The SAS System

N P A R 1 W A Y P R O C E D U R E

Savage Scores (Exponential) for Variable DAYS
Classified by Variable TRT

TRT	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score

```
1    5    -3.36798017        0.0     1.63455487    -.673596034
2    5     0.09561827        0.0     1.63455487    0.019123654
3    5     3.27236190        0.0     1.63455487    0.654472379
          Average Scores Were Used for Ties


      Savage 1-Way Analysis (Chi-Square Approximation)
      CHISQ =  5.5047      DF =  2      Prob > CHISQ = 0.0638
```

نلاحظ من النتائج ان اختبار سافيج يشير الى عدم وجود اختلافات معنوية في فترة بقاء الفئران باختلاف العقارات المستعملة. لو اختبرنا بيانات المثال السابق بأستعمال اختبار t سنجد ان الاختبار معنوي ولكن لو اختبرنا توزيع البيانات سنجد ان البيانات توزيعها غير طبيعي مما يعني ان اختبار t غير صحيح (invalid test). ويمكن اجراء ذلك بأستعمال الايعاز الآتي:

```
proc anova;
class trt;
model days = trt;
means trt/lsd ;
run;
proc univariate normal plot;var days;
run;
```

س- اختبار كوكران Cochran Test

في الدراسات التي تكون فيها الاستجابة ذات مظهرين (binary) والتي تقاس عدة مرات **او** خلال اوقات مختلفة او في ظروف مختلفة وتشمل اكثر من فئة للاستجابة فأن هذا الاختبار يستعمل لاختبار فرضية ان احتمال الاستجابة الموجبة لم تتغير خلال الاوقات او ظروف الحالة.
مثال (56) في تجربة تضمنت استعمال ثلاثة عقاقير (A , B , C) اعطيت لمجموعة من المرضى (46 مريضا) لمعالجة احد الامراض المزمنة وكانت الاستجابة معبر عنها بـ F(Favorable) او U (Unfavorable) اختبر تساوي الاستجابة؟

```
data drugs;
   input DrugA $ DrugB $ DrugC $ Count @@;
   datalines;
 F F F 6  U F F 2
 F F U 16  U F U 4
 F U F 2  U U F 6
 F U U 4  U U U 6
 proc freq data=Drugs;
   weight Count;
   tables DrugA DrugB DrugC / nocum;
   tables DrugA*DrugB*DrugC / agree noprint;
   tables DrugA DrugB DrugC /nocum noprint;
 run;
```

```
                                The SAS System
                DRUGA   Frequency   Percent
                ffffffffffffffffffffffffff
                F          28     60.9
                U          18     39.1


                DRUGB   Frequency   Percent
                ffffffffffffffffffffffffff
                F          28     60.9
```

```
U        18    39.1


DRUGC  Frequency  Percent
ffffffffffffffffffffffffff
F        16    34.8
U        30    65.2
```

يتبين من النتائج ان الاستجابة المرغوبة في العقارين A و B تمثلان 61% من مجموع كل حالة فيما بلغت النسبة في عقار C حوالي 35%.
النتائج ادناه توضح التوافق بين العقاقير.

```
          The SAS System
   STATISTICS FOR TABLE 1 OF DRUGB BY DRUGC
        CONTROLLING FOR DRUGA=F

          McNemar's Test
          --------------
Statistic = 10.889      DF = 1      Prob = 0.001


        Simple Kappa Coefficient
        ------------------------
                  95% Confidence Bounds
Kappa = -0.033   ASE = 0.117     -0.262   0.196
Sample Size = 28
```

عدم توافق العقارين عندما يفضل العقار A

```
   STATISTICS FOR TABLE 2 OF DRUGB BY DRUGC
        CONTROLLING FOR DRUGA=U

          McNemar's Test
          --------------
Statistic = 0.400      DF = 1      Prob = 0.527

        Simple Kappa Coefficient
        ------------------------
                  95% Confidence Bounds
Kappa = -0.154   ASE = 0.223     -0.591   0.283
Sample Size = 18
```

توافق العقارين عندما لايفضل العقار A

```
          The SAS System
    SUMMARY STATISTICS FOR DRUGB BY DRUGC
         CONTROLLING FOR DRUGA

        Overall Kappa Coefficient
        ------------------------
                  95% Confidence Bounds
Kappa = -0.059   ASE = 0.103     -0.261   0.144


      Test for Equal Kappa Coefficients
      ---------------------------------
Statistic = 0.231      DF = 1      Prob = 0.631
```

النتائج توضح وجود توافق بين عقار B و C بغض النظر عن العقار A ..

```
   Cochran's Q, for DRUGA by DRUGB by DRUGC
   ----------------------------------------
Statistic = 8.471      DF = 2      Prob = 0.014

Total Sample Size = 46
```

قيمة اختبار كوجران (0.014) تعني رفض فرضية تساوي الاستجابة لدى الانواع الثلاثة من العقاقير.

Dr Firas

3 – 3 – 1 معامل سبيرمان لأرتباط الرتب Spearman Coefficient of Rank Correlation

هو مقياس لقوة العلاقة بين متغيرين غير معروفي التوزيع او ان توزيعهما غير معتدل ويمكن تحديد معامل الارتباط بينهما على اساس الارتباط المشاهد بين مراتبهما المتناظرة.

مثال (57): جد معامل ارتباط الرتبة بين قيم الصفتين x1 و x2 ؟

```
data d;
input x1 x2;
cards;
13.5 307
20.1 79
14.8 71
19.6 192
19.5 122
17.4 404
26.1 55
21.1 82
proc corr spearman;
var x1 x2;
run;
```

Simple Statistics

Variable	N	Mean	Std Dev	Median	Minimum	Maximum
X1	8	19.012500	3.911316	19.550000	13.500000	26.100000
X2	8	164.000000	128.216335	102.000000	55.000000	404.000000

Spearman Correlation Coefficients / Prob > |R| under Ho: Rho=0 / N = 8

	X1	X2
X1	1.00000	-0.54762
	0.0	0.1600
X2	-0.54762	1.00000
	0.1600	0.0

الفصل الرابع

تحليل التباين في التصاميم المختلفة

هناك العديد من التصاميم التي يمكن اعتمادها في التجارب وهي تتفاوت في صعوبتها وسهولتها كما تتفاوت ايضا في ميزاتها وعيوبها. ولكن بصورة عامة هناك اعتبارات هامة يجب ان يأخذها الباحث بنظر الاعتبار قبل اختيار التصميم المناسب.

1- هل التجربة تخص عامل واحد او اكثر من عامل.

Dr. Firas

2- هل الوحدات التجريبية متجانسة ام غير متجانسة.

4 – 1التصميم العشوائي الكامل CRD

هو من اكثر التصاميم تطبيقا وتوزع فيه الوحدات المتجانسة بطريقة عشوائية على المعاملات من غير نظام محدد. ويتميز هذا التصميم ببساطته فضلا عن تميزه في ان درجات حرية الخطأ تكون اعلى مايمكن مما يؤدي الى خفض التباين المقدر وهذا التصميم لايشترط عدد محدد للمعاملات او التكرارات فضلا عن ان فقدان بعض الوحدات بل حتى بعض المعاملات لايمنع اجراء التحليل.

ومعادلة النموذج الرياضي لهذا التصميم هي :

$$Y_{ij} = \mu + T_i + e_{ij}$$

اذ ان

Y_{ij} = قيمة المشاهدة j التي تعود الى المعاملة i

μ = المتوسط العام

T_i = تأثير المعاملة i

e_{ij} = الخطأ القياسي الذي يفترض ان يكون موزع توزيع عشوائي بمتوسط يساوي صفر وتباين σ^2e.

مثال (58): لدينا مجموعتين من الطلاب (trt) واوزان كل فرد في المجموعة ، المطلوب هل ان هناك فرق معنوي بين معدل وزن المجموعتين؟

```
data d;
input trt  w;
cards;
1  57.8
1  56.2
1  61.9
1  54.4
1  53.6
1  56.4
1  53.2
2  64.2
2  58.7
2  63.1
2  62.5
2  59.8
2  59.2
;option nodate nonumber;
proc anova;
classes trt;
model w=trt;
means trt/LSD;   ← ايعاز باجراء مقارنة باستعمال اختبار تي او مايسمى اقل فرق معنوي
proc means n mean stderr;by trt;var w;
run;
```

The SAS System

Analysis of Variance Procedure
Class Level Information

Dr. Firas

Class	Levels	Values
TRT	2	1 2

Number of observations in data set = 13

The SAS System
Analysis of Variance Procedure

Dependent Variable: W

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	81.92719780	81.92719780	11.18	0.0065
Error	11	80.58357143	7.32577922		
Corrected Total	12	162.51076923			

R-Square	C.V.	Root MSE	W Mean
0.504134	4.623657	2.7066176	58.53846154

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TRT	1	81.92719780	81.92719780	11.18	0.0065

The SAS System

Analysis of Variance Procedure

T tests (LSD) for variable: W

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 11 MSE= 7.325779
Critical Value of T= 2.20
Least Significant Difference= 3.3143
WARNING: Cell sizes are not equal.
Harmonic Mean of cell sizes= 6.461538

Means with the same letter are not significantly different.

T Grouping	Mean	N	TRT
A	61	6	2
B	56.214	7	1

معدل المجموعة الثانية

معدل المجموعة الاولى

واختلاف الحروف امام كل متوسط يعني وجود اختلافات معنوية بمستوى 5%

The SAS System

Analysis Variable : W

---------- TRT=1 ----------المعاملة الاولى

N	Mean	Std Error الخطأ القياسي
7	56.2142857	1.1348230

---------- TRT=2 ----------

N	Mean	Std Error

Dr. Firas

```
الخطأ القياسي 0.9397695   61.2500000   6
---------------------------------------
```

ويمكن حل المثال بأستعمال الأمر Proc ttest الا انه لايعطينا قيم لمتوسط المعاملات وانما يعطينا المعنوية فقط كما اوضحنا ذلك سابقا.

```
PROC TTEST;
CLASS trt;
VAR w;
RUN;
```

مثال (59): البيانات التالية تمثل اوزان مجموعة من الاطفال تختلف في نوع تغذيتها اختبر الفروق بينها بأستعمال عدة اختبارات؟

```
data d;
input trt  w;
cards;
1 6
1 8
1 7
1 9
1 10
2 9
2 8
2 11
2 11
2 10
3 7
3 5
3 5
3 6
3 4
4 3
4 3
4 4
4 2
4 3
5 8
5 6
5 9
5 9
5 11
option nodate nonumber;
proc anova;
classes trt;
model w=trt;
means trt/LSD TUKEY DUNCAN DUNNETT;
run;
```

Dr. Firas

```
                                The SAS System

                    Analysis of Variance Procedure
                         Class Level Information

                    Class    Levels    Values
```

TRT 5 1 2 3 4 5

Number of observations in data set = 25

The SAS System
Analysis of Variance Procedure

Dependent Variable: W

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	149.76000000	37.44000000	20.13	0.0001
Error	20	37.20000000	1.86000000		
Corrected Total	24	186.96000000			

R-Square	C.V.	Root MSE	W Mean
0.801027	19.59509	1.3638181	6.96000000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TRT	4	149.76000000	37.44000000	20.13	0.0001

The SAS System
Analysis of Variance Procedure

T tests (LSD) for variable: W
NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 Confidence= 0.95 df= 20 MSE= 1.86
Critical Value of T= 2.08596
Least Significant Difference= 1.7993

Comparisons significant at the 0.05 level are indicated by '***'.

TRT Comparison	Lower Confidence Limit	Difference Between Means	Upper Confidence Limit	
2 - 5	-0.5993	1.2000	2.9993	
2 - 1	0.0007	1.8000	3.5993	***
2 - 3	2.6007	4.4000	6.1993	***
2 - 4	5.0007	6.8000	8.5993	***
5 - 2	-2.9993	-1.2000	0.5993	
5 - 1	-1.1993	0.6000	2.3993	
5 - 3	1.4007	3.2000	4.9993	***
5 - 4	3.8007	5.6000	7.3993	***
1 - 2	-3.5993	-1.8000	-0.0007	***
1 - 5	-2.3993	-0.6000	1.1993	
1 - 3	0.8007	2.6000	4.3993	***
1 - 4	3.2007	5.0000	6.7993	***
3 - 2	-6.1993	-4.4000	-2.6007	***
3 - 5	-4.9993	-3.2000	-1.4007	***
3 - 1	-4.3993	-2.6000	-0.8007	***
3 - 4	0.6007	2.4000	4.1993	***
4 - 2	-8.5993	-6.8000	-5.0007	***
4 - 5	-7.3993	-5.6000	-3.8007	***
4 - 1	-6.7993	-5.0000	-3.2007	***
4 - 3	-4.1993	-2.4000	-0.6007	***

The SAS System
Analysis of Variance Procedure

Tukey's Studentized Range (HSD) Test for variable: W
NOTE: This test controls the type I experimentwise error rate.

Alpha= 0.05 Confidence= 0.95 df= 20 MSE= 1.86

Dr. Firas

Critical Value of Studentized Range= 4.232
Minimum Significant Difference= 2.5811

Comparisons significant at the 0.05 level are indicated by '***'.

TRT Comparison	Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit	
2 - 5	-1.3811	1.2000	3.7811	
2 - 1	-0.7811	1.8000	4.3811	
2 - 3	1.8189	4.4000	6.9811	***
2 - 4	4.2189	6.8000	9.3811	***
5 - 2	-3.7811	-1.2000	1.3811	
5 - 1	-1.9811	0.6000	3.1811	
5 - 3	0.6189	3.2000	5.7811	***
5 - 4	3.0189	5.6000	8.1811	***
1 - 2	-4.3811	-1.8000	0.7811	
1 - 5	-3.1811	-0.6000	1.9811	
1 - 3	0.0189	2.6000	5.1811	***
1 - 4	2.4189	5.0000	7.5811	***
3 - 2	-6.9811	-4.4000	-1.8189	***
3 - 5	-5.7811	-3.2000	-0.6189	***
3 - 1	-5.1811	-2.6000	-0.0189	***
3 - 4	-0.1811	2.4000	4.9811	
4 - 2	-9.3811	-6.8000	-4.2189	***
4 - 5	-8.1811	-5.6000	-3.0189	***
4 - 1	-7.5811	-5.0000	-2.4189	***
4 - 3	-4.9811	-2.4000	0.1811	

The SAS System
Analysis of Variance Procedure

Dunnett's T tests for variable: W
NOTE: This tests controls the type I experimentwise error for comparisons of all treatments against a control.

Alpha= 0.05 Confidence= 0.95 df= 20 MSE= 1.86
Critical Value of Dunnett's T= 2.651
Minimum Significant Difference= 2.2867

Comparisons significant at the 0.05 level are indicated by '***'.

TRT Comparison	Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit	
2 - 1	-0.4867	1.8000	4.0867	
5 - 1	-1.6867	0.6000	2.8867	
3 - 1	-4.8867	-2.6000	-0.3133	***
4 - 1	-7.2867	-5.0000	-2.7133	***

The SAS System
Analysis of Variance Procedure

T tests (LSD) for variable: W
NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 20 MSE= 1.86
Critical Value of T= 2.09
Least Significant Difference= 1.7993

Means with the same letter are not significantly different.

Dr.Firas

```
T Grouping          Mean     N  TRT

      A         9.8000     5  2
      A
B     A         8.6000     5  5
B
B               8.0000     5  1

      C         5.4000     5  3

      D         3.0000     5  4

                The SAS System
          Analysis of Variance Procedure

        Duncan's Multiple Range Test for variable: W
NOTE: This test controls the type I comparisonwise error rate, not the
        experimentwise error rate

          Alpha= 0.05  df= 20  MSE= 1.86

        Number of Means    2    3    4    5
        Critical Range  1.799 1.889 1.945 1.985

    Means with the same letter are not significantly different.

    Duncan Grouping          Mean     N  TRT

          A         9.8000     5  2
          A
          A         8.6000     5  5
          A
          A         8.0000     5  1

          B         5.4000     5  3

          C         3.0000     5  4

                The SAS System
          Analysis of Variance Procedure

        Tukey's Studentized Range (HSD) Test for variable: W
NOTE: This test controls the type I experimentwise error rate, but generally
        has a higher type II error rate than REGWQ.

          Alpha= 0.05  df= 20  MSE= 1.86
        Critical Value of Studentized Range= 4.232
        Minimum Significant Difference= 2.5811

    Means with the same letter are not significantly different.

    Tukey Grouping          Mean     N  TRT

          A         9.8000     5  2
          A
          A         8.6000     5  5
          A
          A         8.0000     5  1

          B         5.4000     5  3
          B
          B         3.0000     5  4
```

ملاحظة في مثالنا السابق نلاحظ ان طريقة عرض المقارنات بين المتوسطات في اختبار LSD و Tukey و Dunnett تكون بأسلوب غيرواضح هنا يمكن وضع اختبار LSD و DUNCAN و SNK و BON وSCHEFFE بذلك ستظهر النتائج باسلوب ابسط.

Dr Firas

The SAS System
Analysis of Variance Procedure
T tests (LSD) for variable: W

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 20 MSE= 1.86
Critical Value of T= 2.09
Least Significant Difference= 1.7993
Means with the same letter are not significantly different.

T Grouping	Mean	N	TRT
A	9.8000	5	2
A			
B A	8.6000	5	5
B			
B	8.0000	5	1
C	5.4000	5	3
D	3.0000	5	4

Analysis of Variance Procedure

Student-Newman-Keuls test for variable: W

NOTE: This test controls the type I experimentwise error rate under the complete null hypothesis but not under partial null hypotheses.

Alpha= 0.05 df= 20 MSE= 1.86

Number of Means	2	3	4	5
Critical Range	1.799258	2.1822472	2.4142354	2.5810868

Means with the same letter are not significantly different.

SNK Grouping	Mean	N	TRT
A	9.8000	5	2
A			
A	8.6000	5	5
A			
A	8.0000	5	1
B	5.4000	5	3
C	3.0000	5	4

The SAS System
Analysis of Variance Procedure

Bonferroni (Dunn) T tests for variable: W

NOTE: This test controls the type I experimentwise error rate, but generally has a higher type II error rate than REGWQ.

Alpha= 0.05 df= 20 MSE= 1.86
Critical Value of T= 3.15
Minimum Significant Difference= 2.72

Means with the same letter are not significantly different.

Bon Grouping	Mean	N	TRT

Dr Firas

```
      A          9.8000     5  2
      A
      A          8.6000     5  5
      A
B     A          8.0000     5  1
B
B     C          5.4000     5  3
      C
      C          3.0000     5  4

                    The SAS System
                Analysis of Variance Procedure

                Scheffe's test for variable: W
NOTE: This test controls the type I experimentwise error rate but generally
      has a higher type II error rate than REGWF for all pairwise comparisons

                Alpha= 0.05  df= 20  MSE= 1.86
                 Critical Value of F= 2.86608
               Minimum Significant Difference= 2.9205

          Means with the same letter are not significantly different.

        Scheffe Grouping          Mean     N  TRT

                  A          9.8000     5  2
                  A
                  A          8.6000     5  5
                  A
            B     A          8.0000     5  1
            B
            B     C          5.4000     5  3
                  C
                  C          3.0000     5  4
```

Dr.Firas

4 – 2 تصميم القطاعات العشوائية الكاملة RCBD

هذا التصميم لايختلف عن التصميم العشوائي الكامل الا بأضافة تأثير القطاع في خطوتين فقط هما Model و Anova لغرض ازالة تأثير القطاع ويستعمل عندما يلاحظ الباحث عدم تجانس البيانات ويعبرعنه بالنموذج الرياضي الآتي:

$$Y_{ijk} = \mu + T_i + B_j + e_{ijk}$$

مثال (60) : في تجربة اعتمدت تصميم القطاعات العشوائية الكاملة لمجموعتين من الافراخ وبواقع 4 قطاعات. المطلوب اختبار الاختلافات في المتوسطات بين المعاملات؟

```
Data s;
Input trt rep w;
Cards;
1 1 40
1 2 37
1 3 39
1 4 38
2 1 41
2 2 42
2 3 42
2 4 44
;proc anova; classes trt rep;
Model w = trt rep;
```

Means trt/lsd; ← lsd **بعد** alpha=0.01 **هنا يمكن وضع**
Proc means n mean stderr; by trt;var w;
Run;

ملاحظة: البرنامج يعطي دائما مقارنات بمستوى 5% ولكن اضافة العبارة السابقة ستؤدي الى المقارنة بمستوى 1%.

The SAS System
Analysis of Variance Procedure
Class Level Information

Class	Levels	Values
TRT	2	1 2
REP	4	1 2 3 4

Number of observations in data set = 8
The SAS System

Analysis of Variance Procedure
Dependent Variable: W

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	30.50000000	7.62500000	3.10	0.1897
Error	3	7.37500000	2.45833333		
Corrected Total	7	37.87500000			

R-Square	C.V.	Root MSE	W Mean
0.805281	3.883362	1.5679073	40.37500000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TRT	1	28.12500000	28.12500000	11.44	0.0430
REP	3	2.37500000	0.79166667	0.32	0.8115

The SAS System

Analysis of Variance Procedure

T tests (LSD) for variable: W
NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 3 MSE= 2.458333
Critical Value of T= 3.18
Least Significant Difference= 3.5283

Means with the same letter are not significantly different.

T Grouping	Mean	N	TRT
A	42.250	4	2
B	38.500	4	1

الحرفان المختلفان A , B يشيران الى ان الفرق بين المتوسطين معنوي من الناحية الاحصائية وبمستوى 5%. وهذا السياق هو عام لجميع التحاليل الاحصائية بمعنى ان الحروف المختلفة تعني وجود اختلافات احصائية وتشابه الحروف تعني عدم وجود اختلافات احصائية.
وعادة يفضل استعمال اختبار Duncan عند مقارنة اكثر من متوسطين.

Dr. Firas

```
The SAS System
                    Analysis Variable : W

------------------------ TRT=1 ---------------
        N        Mean    Std Error
        -------------------------------
        4   38.5000000    0.6454972
        -------------------------------

------------------------ TRT=2 ---------------
        N        Mean    Std Error
        -------------------------------
        4   42.2500000    0.6291529
        -------------------------------
```

4 – 3 التصميم العشوائي الكامل في التجارب العاملية

عندما يكون هناك عاملين او اكثر في تجربة ما فأن التجربة يطلق عليها عاملية وفيها يمكن الحصول على جميع توافيق مستويات العوامل المراد دراسة تأثيرها في صفة ما. اذ ان التجارب العاملية توفر فرصة لمعرفة تأثير التداخل بين مستويات العوامل المدروسة وهو ما قد يختلف عن التأثير الناتج عن نفس المتغيرات اذا مادرست بصورة منفردة.

والتجربة العاملية ليست تصميما تجريبيا خاصا وانما يمكن ان يشمل جميع التصاميم.

4 – 3 – 1 تأثير عاملين بمشاهدة واحدة لكل عامل

مثال (61) : اخذت ثلاثة عينات من الافراد واحصيت عدد الكتب في مكتبة كل منهم من ثلاثة مدن وبواقع عينتين تعود الى منطقتين لكل مدينة المطلوب معرفة تأثير المنطقة والمدينة في حيازة الفرد للكتب؟

هنا لايمكن دراسة التداخل وذلك لوجود قيمة واحدة ضمن مستوى كل عامل.

```
data h;
input  city region book;
cards;
1  1 100
1  2 155
2  1 210
2  2 180
3  1 220
3  2 200
proc glm data=h;
  class city region;
  model book = city region;
  means city region;
run;
```

```
                The SAS System
          General Linear Models Procedure
            Class Level Information

          Class   Levels   Values
```

```
                    CITY       3   1 2 3

                    REGION     2   1 2

                  Number of observations in data set = 6

                              The SAS System
                        General Linear Models Procedure

Dependent Variable: BOOK

Source             DF     Sum of Squares     Mean Square  F Value    Pr > F

Model               3     7729.16666667     2576.38888889   2.39    0.3089
Error               2     2158.33333333     1079.16666667

Corrected Total     5          9887.50000000

      R-Square             C.V.             Root MSE         BOOK Mean
      0.781711             18.50742          32.8506722       177.50000000

Source             DF       Type I SS       Mean Square  F Value    Pr > F

CITY                2     7725.00000000     3862.50000000    3.58    0.2184
REGION              1        4.16666667        4.16666667    0.00    0.9561

Source             DF      Type III SS      Mean Square  F Value    Pr > F

CITY                2     7725.00000000     3862.50000000    3.58    0.2184
REGION              1        4.16666667        4.16666667    0.00    0.9561

                              The SAS System
                        General Linear Models Procedure

                Level of        ----------BOOK----------
                CITY       N       Mean          SD

                1          2    127.500000     38.8908730
                2          2    195.000000     21.2132034
                3          2    210.000000     14.1421356

                Level of        ----------BOOK----------
                REGION     N       Mean          SD

                1          3    176.666667     66.5832812
                2          3    178.333333     22.5462488
```

الاختلافات بحيازة الكتب معنوية بسبب المنطقة والمدينة.

4 – 3 – 2 تأثير عاملين ولأكثر من مشاهدة لكل عامل

مثال (62): في داسة شملت ثلاثة مجموعات (x, y, z) تتضمن ذكورا واناث لمقارنة مجموع الدرجات في 10 دروس.

```
DATA TWOWAY;
INPUT GROUP $ GENDER $ sum ;
CARDS;;
X M 700
X M 850
X M 820
X M 640
X M 920
X F 900
```

Dr. Firas

```
X F 880
X F 899
X F 780
X F 899
Y M 480
Y M 460
Y M 500
Y M 570
Y M 580
Y F 590
Y F 540
Y F 560
Y F 570
Y F 555
Z M 500
Z M 550
Z M 480
Z M 600
Z M 610
Z F 520
Z F 660
Z F 525
Z F 610
Z F 645
OPTION NODATE NONUMBER;
PROC ANOVA; CLASS GROUP GENDER;
  MODEL sum = GROUP|GENDER;
  MEANS GROUP|GENDER / LSD;
RUN;
```

```
                                The SAS System

                          Analysis of Variance Procedure
                               Class Level Information

                          Class    Levels    Values

                          GROUP         3    X Y Z

                          GENDER        2    F M

                          Number of observations in data set = 30

                                The SAS System
                          Analysis of Variance Procedure

                          Dependent Variable: WORDS

                          Source       DF   Sum of Squares   Mean Square  F Value  Pr > F

                          Model         5    531436.16667  106287.2333    23.92    0.0001
Error         24    106659.20000    4444.1333
                          Corrected Total 29  638095.3666667

                             R-Square     C.V.          Root MSE       sum Mean
                           0.832848    10.31264         66.6643332      646.43333333


                          Source       DF  Anova SS     Mean Square    F Value    Pr > F

                        GROUP         2  503215.26666667  251607.63333333  56.62   0.0001
                        GENDER        1  25404.30000000   25404.30000000   5.72    0.0250
                      GROUP*GENDER    2  2816.60000000    1408.30000000    0.32    0.7314

                                The SAS System
```

Dr. Firas

Analysis of Variance Procedure
T tests (LSD) for variable: WORDS

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 24 MSE= 4444.133
Critical Value of T= 2.06
Least Significant Difference= 61.531

Means with the same letter are not significantly different.

T Grouping	Mean	N	GROUP
A	828.80	10	X
B	570.00	10	Z
B	540.50	10	Y

يتبين من نتائج الاختبار وجود فرق معنوي بين مجموعة x والمجموعتان z و y

The SAS System

Analysis of Variance Procedure
T tests (LSD) for variable: WORDS

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.
Alpha= 0.05 df= 24 MSE= 4444.133
Critical Value of T= 2.06
Least Significant Difference= 50.24

Means with the same letter are not significantly different.

T Grouping	Mean	N	GENDER
A	675.53	15	F
B	617.33	15	M

الاختلافات معنوية بين الجنسين.

Level of GROUP	Level of GENDER	SUM N	SUM Mean	SUM SD
X	F	5	871.600000	51.887378
X	M	5	786.000000	113.929803
Y	F	5	563.000000	18.574176
Y	M	5	518.000000	54.037024
Z	F	5	592.000000	66.011363
Z	M	5	548.000000	58.051701

يمكن الحصول على الخطأ القياسي للتقديرات اعلاه بقسمة الانحراف القياسي (SD) على الجذر التربيعي لـ N
الاختلافات معنوية بسبب المجموعة والجنس الا ان التداخل بين العاملين غير معنوي.

Dr. Firas

ملاحظة: في بعض الحالات قد يرغب الباحث في المقارنة بين الجنس ضمن كل معاملة في هذه الحالة يعطى ايعاز لحذف المعاملات وابقاء المعاملة الاولى ثم يجرى الاختبار بعد ذلك تحذف المعاملة الاولى والثالثة لتبقى الثانية فقط ثم نحذف الاولى والثانية لتبقى الثالثة فقط .

مثال (63): في دراسة لمعرفة الاختلافات في عدد اربع انواع من المكائن العاطلة تم تسجيل تلك المكائن لدى 4 مجهزين المطلوب تقدير المتوسط العام للمكائن العاطلة واختبار معنوية الأختلافات من الناحية الاحصائية بين عدد المكائن العاطلة بسبب اختلاف المجهزين وبسبب اختلاف نوع الماكنة؟

```
data Exp;
do Supplier = 'A','B','C','D';
do Machine = 1 to 4;
do rep = 1 to 5;
input Defects @@;
output;
end;
end;
end;
Cards;
2 6 3 3 6 8 6 6 4 4 4 2 4 0 4 5 5 7 8 5
13 12 12 11 12 16 15 14 14 13 11 10 12 12 10 13 13 14 15 12
2 6 3 6 6 6 4 4 6 6 0 3 2 0 2 4 6 7 6 4
20 19 18 21 22 22 24 23 20 20 17 19 18 16 17 23 20 20 22 21
option nodate nonumber center;
proc anova data=Exp;
class Supplier Machine;
model Defects = Supplier Machine;
means Supplier Machine / LSD;
quit;
```

ويمكن ترتيب البيانات بشكل اخر لنحصل على نفس النتائج

نلاحظ ان لدينا اربع مجهزين واربع مكائن وقد اخذنا عدد المكائن العاطلة من مجموع خمسة صفقات بيع. لذا فأن ترتيب البيانات سيكون المجهز الاول للمكينة الاولى للصفقة الاولى وهكذا لغاية الصفقة الخامسة ، ثم نعتمد الاسلوب نفسه لبقية المجهزين وكما موضح في بيانات المثال.

علما بأن Rep تمثل تسلسل الصفقة.

```
data Exp;
input Supplier$ Machine rep Defects;
cards;
A 1 1 2
A 1 2 6
A 1 3 3
A 1 4 3
A 1 5 6
A 2 1 8
A 2 2 6
A 2 3 6
A 2 4 4
A 2 5 4
A 3 1 4
A 3 2 2
A 3 3 4
A 3 4 0
A 3 5 4
A 4 1 5
A 4 2 5
A 4 3 7
A 4 4 8
A 4 5 5
B 1 1 13
B 1 2 12
B 1 3 12
B 1 4 11
```

Dr Firas

B 1 5 12
B 2 1 16
B 2 2 15
B 2 3 14
B 2 4 14
B 2 5 13
B 3 1 11
B 3 2 10
B 3 3 12
B 3 4 12
B 3 5 10
B 4 1 13
B 4 2 13
B 4 3 14
B 4 4 15
B 4 5 12
C 1 1 2
C 1 2 6
C 1 3 3
C 1 4 6
C 1 5 6
C 2 1 6
C 2 2 4
C 2 3 4
C 2 4 6
C 2 5 6
C 3 1 0
C 3 2 3
C 3 3 2
C 3 4 0
C 3 5 2
C 4 1 4
C 4 2 6
C 4 3 7
C 4 4 6
C 4 5 4
D 1 1 20
D 1 2 19
D 1 3 18
D 1 4 21
D 1 5 22
D 2 1 22
D 2 2 24
D 2 3 23
D 2 4 20
D 2 5 20
D 3 1 17
D 3 2 19
D 3 3 18
D 3 4 16
D 3 5 17
D 4 1 23
D 4 2 20
D 4 3 20
D 4 4 22
D 4 5 21
option nodate nonumber center;

Dr. Firas

```
proc anova data=Exp;
class Supplier Machine;
model Defects = Supplier Machine;
means Supplier Machine / lsd;
quit;
```

```
                    The SAS System

                Analysis of Variance Procedure
                  Class Level Information

                  Class    Levels    Values

                  SUPPLIER     4    A B C D

                  MACHINE      4    1 2 3 4

                Number of observations in data set = 80

                    The SAS System
              Analysis of Variance Procedure

Dependent Variable: DEFECTS

Source        DF  Sum of Squares    Mean Square   F Value  Pr > F

Model          6  3604.77500000    600.79583333   304.12  0.0001
Error         73  144.21250000      1.97551370
Corrected Total 79 3748.98750000

 R-Square      C.V.             Root MSE         DEFECTS Mean
 0.961533     13.53097           1.4055296        10.38750000


Source         DF   Anova SS     Mean Square  F Value    Pr > F

SUPPLIER       3   3441.63750000   1147.21250000   580.72  0.0001
MACHINE        3    163.13750000     54.37916667    27.53  0.0001



                    The SAS System
              Analysis of Variance Procedure

           T tests (LSD) for variable: DEFECTS
NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

            Alpha= 0.05  df= 73  MSE= 1.975514
                Critical Value of T= 1.99
            Least Significant Difference= 0.8858

  Means with the same letter are not significantly different.

          T Grouping          Mean     N  SUPPLIER

                A         20.1000    20  D

                B         12.7000    20  B

                C          4.6000    20  A
                C
                C          4.1500    20  C
```

الاختلافات معنوية في عدد المكائن العاطلة بين المجهزين.

```
                    The SAS System
              Analysis of Variance Procedure
```

Dr. Firas

T tests (LSD) for variable: DEFECTS

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 73 MSE= 1.975514
Critical Value of T= 1.99
Least Significant Difference= 0.8858

Means with the same letter are not significantly different.

T Grouping	Mean	N	MACHINE
A	11.7500	20	2
A			
A	11.5000	20	4
B	10.1500	20	1
C	8.1500	20	3

الاختلافات بسبب نوع الماكنة العاطلة كانت معنوية.

4 – 3 – 3 دراسة تأثير عاملين مع التداخل

مثال (64) : في تجربة لمقارنة استعمال مادتين كيمياويتين هما النوع العادي والمحسن تم تصنيع كل منهما في دولتين (العراق ومصر) لمقارنة الوقت اللازم لاستخلاص الالمنيوم كانت النتائج كما موضحة ادناه، هل هناك اختلاف بين المادتين العادي والمحسن وهل هناك اختلافات بين المادتين بسبب البلد المصنع؟ وهل التداخل معنوي بين نوع المادة والبلد المصنع؟

الدولة المصنعة	نوع المادة	الوقت المستغرق
العراق	1	16 17 18
	2	15 13 13
مصر	1	20 21 20
	2	19 18 18

```
Data m;
Input A S REP WT;
CARDS;
1 1 1 16
1 1 2 17
1 1 3 18
1 2 1 15
1 2 2 13
1 2 3 13
2 1 1 20
2 1 2 21
2 1 3 20
2 2 1 19
2 2 2 18
2 2 3 18
PROC ANOVA; CLASSES A S;
MODEL WT= A S A*S;
MEANS A S A*S/LSD;
run;
```

Dr. Firas

The SAS System
Analysis of Variance Procedure
Class Level Information

Class	Levels	Values
A	2	1 2
S	2	1 2

Number of observations in data set = 12

The SAS System
Analysis of Variance Procedure

Dependent Variable: WT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	70.66666667	23.55555556	31.41	0.0001
Error	8	6.00000000	0.75000000		
Corrected Total	11	76.66666667			

R-Square	C.V.	Root MSE	WT Mean
0.921739	4.996300	0.8660254	17.33333333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
A	1	48.00000000	48.00000000	64.00	0.0001
S	1	21.33333333	21.33333333	28.44	0.0007
A*S	1	1.33333333	1.33333333	1.78	0.2191

الاختلافات معنوية بسبب البلد المصنع وبسبب نوع المنتوج بينما التداخل غير معنوي.

The SAS System
Analysis of Variance Procedure

T tests (LSD) for variable: WT

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 8 MSE= 0.75
Critical Value of T= 2.31
Least Significant Difference= 1.153

Means with the same letter are not significantly different.

T Grouping	Mean	N	A
A	19.3333	6	2
B	15.3333	6	1

الاختلافات معنوية بين مصر (2) والعراق (1).اذان المادة المصنعة في العراق استغرقت وقتا اقل.

The SAS System
Analysis of Variance Procedure

T tests (LSD) for variable: WT

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 8 MSE= 0.75
Critical Value of T= 2.31
Least Significant Difference= 1.153

Means with the same letter are not significantly different.

T Grouping	Mean	N	S

Dr.Firas

A 18.6667 6 1

B 16.0000 6 2

الاختلافات في جودة المنتوج المحسن والعادي معنوي. اذ ان المحسن استغرق وقتا اقل.

Level of A	Level of S	N	WT Mean	WT SD
1	1	3	17.0000000	1.00000000
1	2	3	13.6666667	1.15470054
2	1	3	20.3333333	0.57735027
2	2	3	18.3333333	0.57735027

يلاحظ بأن النتائج لاتتضمن اختبارات لقيم التداخل واذا اردنا اختبار قيم التداخل فيمكن اجراء بعض التعديلات على الآوامر ولنفس البيانات وكما يلي:

مثال (65):

```
Data m;
Input A S REP WT;
IF A=1 AND S=1 THEN AS=11;
IF A=1 AND S=2 THEN AS=12;
IF A=2 AND S=1 THEN AS=21;
IF A=2 AND S=2 THEN AS=22;
CARDS;
1 1 1 16
1 1 2 17
1 1 3 18
1 2 1 15
1 2 2 13
1 2 3 13
2 1 1 20
2 1 2 21
2 1 3 20
2 2 1 19
2 2 2 18
2 2 3 18
PROC ANOVA; CLASSES A S;
MODEL WT= A S A*S;
MEANS A S/LSD;
PROC ANOVA; CLASSES AS;
MODEL WT=AS;
MEANS AS/LSD;
RUN;
proc means n mean  stderr;by A; var WT ;
quit;
proc sort; by s;
proc print;
proc means n mean  stderr;by s; var WT ;
run;
```

Dr. Firas

The SAS System

Analysis of Variance Procedure
Class Level Information

Class	Levels	Values
A	2	1 2

S 2 1 2

Number of observations in data set = 12

The SAS System

Analysis of Variance Procedure

Dependent Variable: WT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	70.66666667	23.55555556	31.41	0.0001
Error	8	6.00000000	0.75000000		
Corrected Total	11	76.66666667			

R-Square	C.V.	Root MSE	WT Mean
0.921739	4.996300	0.8660254	17.33333333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
A	1	48.00000000	48.00000000	64.00	0.0001
S	1	21.33333333	21.33333333	28.44	0.0007
A*S	1	1.33333333	1.33333333	1.78	0.2191

The SAS System

Analysis of Variance Procedure

T tests (LSD) for variable: WT

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 8 MSE= 0.75

Critical Value of T= 2.31

Least Significant Difference= 1.153

Means with the same letter are not significantly different.

T Grouping	Mean	N	A
A	19.3333	6	2
B	15.3333	6	1

The SAS System

Analysis of Variance Procedure

T tests (LSD) for variable: WT

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 8 MSE= 0.75

Critical Value of T= 2.31

Least Significant Difference= 1.153

Means with the same letter are not significantly different.

T Grouping	Mean	N	S
A	18.6667	6	1
B	16.0000	6	2

The SAS System

Analysis of Variance Procedure

Class Level Information

Class	Levels	Values
AS	4	11 12 21 22

Number of observations in data set = 12

The SAS System

Dr. Firas

Analysis of Variance Procedure

Dependent Variable: WT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	70.66666667	23.55555556	31.41	0.0001
Error	8	6.00000000	0.75000000		
Corrected Total	11	76.66666667			

R-Square	C.V.	Root MSE	WT Mean
0.921739	4.996300	0.8660254	17.33333333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
AS	3	70.66666667	23.55555556	31.41	0.0001

Analysis of Variance Procedure

T tests (LSD) for variable: WT
Alpha= 0.05 df= 8 MSE= 0.75
Critical Value of T= 2.31
Least Significant Difference= 1.6306

Means with the same letter are not significantly different.

T Grouping	Mean	N	AS
A	20.3333	3	21
B	18.3333	3	22
B	17.0000	3	11
C	13.6667	3	12

B

الاختلافات معنوية بين قيم التداخل

The SAS System
Analysis Variable : WT

---------- A=1 ----------

N	Mean	Std Error
6	15.3333333	0.8432740

---------- A=2 ----------

N	Mean	Std Error
6	19.3333333	0.4944132

The SAS System

OBS	A	S	REP	WT	AS
1	1	1	1	16	11
2	1	1	2	17	11
3	1	1	3	18	11
4	2	1	1	20	21
5	2	1	2	21	21
6	2	1	3	20	21
7	1	2	1	15	12
8	1	2	2	13	12

Dr.Firas

```
 9   1   2   3   13   12
10   2   2   1   19   22
11   2   2   2   18   22
12   2   2   3   18   22
                The SAS System
          Analysis Variable : WT


---------------------------------- S=1 ------------------------------

          N        Mean     Std Error
          ---------------------------------
          6   18.6666667     0.8027730
          ---------------------------------

---------------------------------- S=2 ------------------------------

          N        Mean     Std Error
          ---------------------------------
          6   16.0000000     1.0954451
          ---------------------------------
```

مثال (66) : اخذت مجموعة من الفئران وقسمت عشوائيا الى مجموعتين ، المجموعة الاولى عرضت لمسببات مرضية لمدة ساعة والمجموعة الثانية عرضت لمسببات مرضية لمدة 24 ساعة ، وكل مجموعة قسمت الى ثلاثة مجموعات هي السيطرة ومجموعة اعطيت العقار x ومجموعة اعطيت العقار y اختر التصميم المناسب لدراسة تأثير فترة التعرض وتأثير العقار والتداخل بينهما على فترة بقائها بالايام؟

	Control a=1	Drug x a=2	Drug y a=3
1 houre b=1	1 4 0 7	13 5 7 15	9 16 18 13
24 houres b=2	15 6 10 13	6 18 9 15	14 7 6 13

```
data g;
input b$ a$ c;
cards;
b1   a1  1
b1   a1  4
b1   a1  0
b1   a1  7
b1   a2  13
b1   a2  5
b1   a2  7
b1   a2  15
b1   a3  9
b1   a3  16
```

Dr. Firas

```
b1    a3  18
b1    a3  13
b2    a1  15
b2    a1  6
b2    a1  10
b2    a1  13
b2    a2  6
b2    a2  18
b2    a2  9
b2    a2  15
b2    a3  14
b2    a3  7
b2    a3  6
b2    a3  13
option nodate nonumber center;
proc anova data=g;
class b a;
model c = b a;
means b a / LSD;
quit;
PROC GLM DATA=g;
 CLASS b a ;
 MODEL c = b a b*a;
 MEANS b a b*a ;
RUN;
```

يمكن ترتيب البيانات بطريقة اخرى وكما موضح في المثال التالي:

```
data g;
do b = 'b1','b2';
do a= 1 to 3;
do rep = 1 to 4;
input c @@;
output;
end;
end;
end;
cards;
1  4 0  7 13 5 7 15 9 16 18 13
15 6 10 13 6 18 9 15 14 7 6 13
option nodate nonumber center;
proc anova data=g;
class b a;
model c = b a;
means b a / LSD;
quit;

 PROC GLM DATA=g;
 CLASS b a ;
 MODEL c = b a b*a;
 MEANS b a b*a ;
RUN;
```

The SAS System

Analysis of Variance Procedure
Class Level Information

Class	Levels	Values
B	2	b1 b2
A	3	a1 a2 a3

Number of observations in data set = 24

The SAS System

Analysis of Variance Procedure

Dependent Variable: C

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	136.00000000	45.33333333	1.91	0.1600
Error	20	474.00000000	23.70000000		
Corrected Total	23	610.00000000			

R-Square	C.V.	Root MSE	C Mean
0.222951	48.68265	4.8682645	10.00000000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
B	1	24.00000000	24.00000000	1.01	0.3263
A	2	112.00000000	56.00000000	2.36	0.1199

The SAS System

Analysis of Variance Procedure

T tests (LSD) for variable: C

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 20 MSE= 23.7
Critical Value of T= 2.09
Least Significant Difference= 4.1458

Means with the same letter are not significantly different.

T Grouping	Mean	N	B
A	11.000	12	b2
A			
A	9.000	12	b1

الاختلافات في مدة البقاء بسبب فترة التعرض كانت غير معنوية

The SAS System
Analysis of Variance Procedure

T tests (LSD) for variable: C NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 20 MSE= 23.7
Critical Value of T= 2.09
Least Significant Difference= 5.0775

Means with the same letter are not significantly different.

T Grouping	Mean	N	A
A	12.000	8	a3
A			
A	11.000	8	a2
A			

Dr. Firas
Dr. Firas

A 7.000 8 a1

لم يكن للعقار تأثير على فترة البقاء

The SAS System

General Linear Models Procedure
Class Level Information

Class	Levels	Values
B	2	b1 b2
A	3	a1 a2 a3

Number of observations in data set = 24

The SAS System
General Linear Models Procedure

Dependent Variable: C

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	280.00000000	56.00000000	3.05	0.0361
Error	18	330.00000000	18.33333333		
Corrected Total	23	610.00000000			

R-Square	C.V.	Root MSE	C Mean
0.459016	42.81744	4.2817441	10.00000000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
B	1	24.00000000	24.00000000	1.31	0.2675
A	2	112.00000000	56.00000000	3.05	0.0721
B*A	2	144.00000000	72.00000000	3.93	0.0384

Source	DF	Type III SS	Mean Square	F Value	Pr > F
B	1	24.00000000	24.00000000	1.31	0.2675
A	2	112.00000000	56.00000000	3.05	0.0721
B*A	2	144.00000000	72.00000000	3.93	0.0384

التداخل بين فترة التعرض والعقار كان معنويا.

The SAS System

General Linear Models Procedure

Level of B	N	C Mean	C SD
b1	12	9.0000000	5.96962006
b2	12	11.0000000	4.19956708

Level of A	N	C Mean	C SD
a1	8	7.0000000	5.39841247
a2	8	11.0000000	4.86973159
a3	8	12.0000000	4.27617987

Level of B	Level of A	N	C Mean	C SD
b1	a1	4	3.0000000	3.16227766
b1	a2	4	10.0000000	4.76095229
b1	a3	4	14.0000000	3.91578004
b2	a1	4	11.0000000	3.91578004
b2	a2	4	12.0000000	5.47722558
b2	a3	4	10.0000000	4.08248290

Dr.Firas

4 – 4 التصميم المحتضن Nested Model

في بعض التجارب قد تسجل عدة مشاهدات من كل وحدة تجريبية اي ان المعاملة تحتوي عدد من الوحدات التجريبية وكل وحدة تجريبية تحتوي عدد من العينات العشوائية ومن الممكن ايضا ان تؤخذ من كل عينة عدد من التقديرات ومثل هذه التجارب تسمى المتشعبة او يطلق عليها التصميم المتشعب او المتدرج Nested او Hierarchial ، والنموذج الرياضي لهذا التصميم هو:

$$Y_{ijk} = \mu + T_i + N_{ij} + e_{ijk}$$

مثال (67): في دراسة لمعرفة تأثير المدرسة وتأثير المعلم ضمن المدرسة على مستوى الطلاب اخذت عينة شملت ثلاثة مدارس وكل مدرسة اخذ منها مدرسان وقيم اداء الطلاب في صفين لكل مدرس ، اختبر تأثير العوامل المدروسة؟

```
data training;
 input  A B classes score;
 label A = 'School'
     B = 'Instructor';
cards;
1 1 1  25
1 1 2  29
1 2 1  14
1 2 2  11
2 1 1  11
2 1 2  6
2 2 1  22
2 2 2  18
3 1 1  17
3 1 2  20
3 2 1  5
3 2 2  2
proc glm data=training;
 class A B;
 model score = A B(A);
 MEANS A B(A)/LSD ALPHA=0.01;
 run;
```

```
                    The SAS System
                General Linear Models Procedure
                  Class Level Information

                  Class   Levels   Values

                  A          3   1 2 3

                  B          2   1 2


               Number of observations in data set = 12

                    The SAS System
                General Linear Models Procedure

Dependent Variable: SCORE
```

Dr. Firas

```
Source        DF       Sum of Squares       Mean Square  F Value    Pr > F
Model          5       724.00000000        144.80000000    20.69    0.0010
Error          6        42.00000000          7.00000000
Corrected Total 11        766.00000000

      R-Square              C.V.            Root MSE       SCORE Mean
      0.945170            17.63834           2.6457513       15.00000000


Source       DF       Type I SS          Mean Square  F Value    Pr > F

A             2      156.50000000        78.25000000    11.18     0.0095
B(A)          3       567.50000000       189.16666667    27.02     0.0007

Source       DF       Type III SS        Mean Square  F Value    Pr > F

A             2      156.50000000          78.25000000    11.18    0.0095
B(A)          3       567.50000000         189.16666667    27.02    0.0007


                          The SAS System
                      General Linear Models Procedure

                     T tests (LSD) for variable: SCORE

        NOTE: This test controls the type I comparisonwise error rate not the
              experimentwise error rate.

                        Alpha= 0.01  df= 6  MSE= 7
                         Critical Value of T= 3.71
                     Least Significant Difference= 6.936

            Means with the same letter are not significantly different.

                  T Grouping          Mean     N  A

                       A         19.750     4  1
                       A
                    B  A          14.250     4  2
                    B
                    B              11.000     4  3

              Level of  Level of      ----------SCORE----------
              B         A        N     Mean           SD

              1         1        2    27.0000000     2.82842712
              2         1        2    12.5000000     2.12132034
              1         2        2     8.5000000     3.53553391
              2         2        2    20.0000000     2.82842712
              1         3        2    18.5000000     2.12132034
              2         3        2     3.5000000     2.12132034
```

يمكن ترتيب البيانات بطريقة اخرى لنحصل على نفس النتائج وكما موضح بالمثال التالي:

```
data training;
  DO A=1 TO 3; عدد المدارس
    DO B=1 TO 2; عدد المدرسين
      DO CLASSES=1 TO 2;  عدد الصفوف
        INPUT SCORE @@;
        OUTPUT;
        END;
      END;
    END;
  input score A B classes;
```

Dr.Firas

```
 label A = 'School'
     B = 'Instructor';
cards;
 25  29 14 11      ←     (عدد المدرسين x عدد الصفوف = 4)
 11  6  22 18
 17  20 5  2
PROC GLM;
 CLASSES A B;
  MODEL SCORE= A B(A);
  MEANS A B(A)/ LSD;
   RUN;
```

مثال (68): نفذت دراسة لمعرفة الاختلافات في تركيز الكالسيوم في احد النباتات ، اذ اخذت اربعة نباتات بصورة عشوائية واخذت ثلاث اوراق من كل نبات بصورة عشوائية واخذت عينتين من كل ورقة لتقدير الكالسيوم في كل منها المطلوب اختبار الاختلافات في الكالسيوم؟

```
DATA FG;
INPUT PLANT LEAF SAMPLE CALCIUM;
CARDS;
1 1 1  3.28
1 1 2  3.09
1 2 1  3.52
1 2 2  3.48
1 3 1  2.88
1 3 2  2.80
2 1 1  2.46
2 1 2  2.44
2 2 1  1.87
2 2 2  1.92
2 3 1  2.19
2 3 2  2.19
3 1 1  2.77
3 1 2  2.66
3 2 1  3.74
3 2 2  3.44
3 3 1  2.55
3 3 2  2.55
4 1 1  3.78
4 1 2  3.87
4 2 1  4.07
4 2 2  4.12
4 3 1  3.31
4 3 2  3.31
PROC GLM;
  CLASSES PLANT LEAF;
  MODEL CALCIUM= PLANT  LEAF(PLANT);
  MEANS PLANT LEAF(PLANT)/ LSD;
   RUN;
```

```
The SAS System

General Linear Models Procedure
  Class Level Information

 Class   Levels   Values

 PLANT      4   1 2 3 4
```

Dr. Firas

LEAF 3 1 2 3

Number of observations in data set = 24

The SAS System

General Linear Models Procedure

Dependent Variable: CALCIUM

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	10.19054583	0.92641326	139.22	0.0001
Error	12	0.07985000	0.00665417		
Corrected Total	23	10.27039583			

R-Square	C.V.	Root MSE	CALCIUM Mean
0.992225	2.708195	0.0815730	3.01208333

Source	DF	Type I SS	Mean Square	F Value	Pr > F
PLANT	3	7.56034583	2.52011528	378.73	0.0001
LEAF(PLANT)	8	2.63020000	0.32877500	49.41	0.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
PLANT	3	7.56034583	2.52011528	378.73	0.0001
LEAF(PLANT)	8	2.63020000	0.32877500	49.41	0.0001

The SAS System

General Linear Models Procedure

T tests (LSD) for variable: CALCIUM

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 12 MSE= 0.006654
Critical Value of T= 2.18
Least Significant Difference= 0.1026

Means with the same letter are not significantly different.

T Grouping	Mean	N	PLANT
A	3.74333	6	4
B	3.17500	6	1
C	2.95167	6	3
D	2.17833	6	2

Level of LEAF	Level of PLANT	N	CALCIUM Mean	CALCIUM SD
1	1	2	3.18500000	0.13435029
2	1	2	3.50000000	0.02828427
3	1	2	2.84000000	0.05656854
1	2	2	2.45000000	0.01414214
2	2	2	1.89500000	0.03535534
3	2	2	2.19000000	0.00000000
1	3	2	2.71500000	0.07778175
2	3	2	3.59000000	0.21213203
3	3	2	2.55000000	0.00000000
1	4	2	3.82500000	0.06363961
2	4	2	4.09500000	0.03535534
3	4	2	3.31000000	0.00000000

Dr Firas

يمكن كتابة قطعة البرنامج كما يلي:

```
DATA TURNIP;
DO PLANT=1 TO 4;
DO LEAF=1 TO 3;
DO SAMPLE=1 TO 2;
INPUT CALCIUM @@;
OUTPUT;
END;
END;
END;
 CARDS;
 3.28 3.09 3.52 3.48 2.88 2.80
 2.46 2.44 1.87 1.92 2.19 2.19
 2.77 2.66 3.74 3.44 2.55 2.55
 3.78 3.87 4.07 4.12 3.31 3.31
```

Dr. Firas

4 – 5 تصميم المربع اللاتيني

هو التصميم الذي يتم فيه تجميع الوحدات التجريبية غير المتجانسة الى مجموعات تضم كل منها وحدات تجريبية متجانسة بعدد المعاملات الداخلة في التجربة على ان يتم في اتجاهين يسمى احدهما صفوف والآخر اعمدة ، ومعنى ذلك ان كل صف وكل عمود هو قطاع كامل بحيث ان كل معاملة لاتظهر الا مرة واحدة في كل صف وعمود. ويمكن التعبير عن النموذج الرياضي لهذا التصميم كالاتي:

$$Y_{ijkl} = \mu + T_i + N_j + M_k + e_{ijkl}$$

مثال (69): لمعرفة التغيرات في مستوى سكر الدم في الفئران اجريت تجربة ولمدة اربعة ايام وباستعمال اربعة معاملات والاستجابة هي معدل التغير في سكر الدم ؟

```
data sugar;
  input day group  insulin $ response;
  cards;
 1 1   B  -4.5
 1 2   D  92.33
 1 3   C  59.83
 1 4   A -45.
 2 1   C  91.83
 2 2   A -48.33
 2 3   D 168.99
 2 4   B  89.0
 3 1   D  86.16
 3 2   B -78.16
 3 3   A -24.17
 3 4   C 101.0
 4 1   A   -.17
 4 2   C  68.83
 4 3   B  25.17
 4 4   D 177.17
```

```
proc anova;
 classes day group insulin;
 model response= day group insulin;
 means day group insulin/lsd;
proc means n mean stderr;by day;var response;
 run;
 proc sort data=sugar; by group;
 proc means n mean stderr;by group;var response;
 run;
proc sort data=sugar; by insulin;
 proc means n mean stderr;by insulin;var response;
 run;
```

The SAS System

Analysis of Variance Procedure
Class Level Information

Class	Levels	Values
DAY	4	1 2 3 4
GROUP	4	1 2 3 4
INSULIN	4	A B C D

Number of observations in data set = 16

The SAS System
Analysis of Variance Procedure

Dependent Variable: RESPONSE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	82550.08712500	9172.23190278	14.38	0.0021
Error	6	3826.73425000	637.78904167		
Corrected Total	15	86376.82137500			

R-Square	C.V.	Root MSE	RESPONSE Mean
0.955697	53.16874	25.2544855	47.49875000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
DAY	3	9420.00612500	3140.00204167	4.92	0.0466
GROUP	3	10865.04312500	3621.68104167	5.68	0.0347
INSULIN	3	62265.03787500	20755.01262500	32.54	0.0004

The SAS System
Analysis of Variance Procedure

T tests (LSD) for variable: RESPONSE

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 6 MSE= 637.789
Critical Value of T= 2.45
Least Significant Difference= 43.696

Means with the same letter are not significantly different.

T Grouping	Mean	N	DAY

Dr. Firas
Dr. Firas

	A	75.37	4	2
	A			
B	A	67.75	4	4
B				
B	C	25.67	4	1
	C			
	C	21.21	4	3

The SAS System
Analysis of Variance Procedure

T tests (LSD) for variable: RESPONSE

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 6 MSE= 637.789
Critical Value of T= 2.45
Least Significant Difference= 43.696

Means with the same letter are not significantly different.

T Grouping	Mean	N	GROUP
A	80.54	4	4
A			
A	57.46	4	3
A			
B A	43.33	4	1
B			
B	8.67	4	2

The SAS System
Analysis of Variance Procedure

T tests (LSD) for variable: RESPONSE

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 6 MSE= 637.789
Critical Value of T= 2.45
Least Significant Difference= 43.696

Means with the same letter are not significantly different.

T Grouping	Mean	N	INSULIN
A	131.16	4	D
B	80.37	4	C
C	7.88	4	B
C			
C	-29.42	4	A

The SAS System
Analysis Variable : RESPONSE

---------- DAY=1 ----------

N	Mean	Std Error
4	25.6650000	30.9770599

Dr. Firas

---------- DAY=2 ----------

N	Mean	Std Error
4	75.3725000	45.2061047

---------- DAY=3 ----------

N	Mean	Std Error
4	21.2075000	43.3192489

---------- DAY=4 ----------

N	Mean	Std Error
4	67.7500000	39.1578915

The SAS System

Analysis Variable : RESPONSE

---------- GROUP=1 ----------

N	Mean	Std Error
4	43.3300000	26.4048888

---------- GROUP=2 ----------

N	Mean	Std Error
4	8.6675000	42.2361122

---------- GROUP=3 ----------

N	Mean	Std Error
4	57.4550000	40.9783059

---------- GROUP=4 ----------

N	Mean	Std Error
4	80.5425000	46.1770431

Dr Firas

مثال (70): اختبر فرضية تشابه خمسة انواع من الحنطة وخمسة مستويات من الري ؟

```
data wheat;
do row=1 to 5;
do col=1 to 5;
input trt $ yield @@;
output;
end; end;
cards;
E 59.45 A 47.28 C 54.44 B 50.14 D 59.45
```

```
C 55.16 D 60.89 B 56.59 E 60.17 A 48.71
B 44.41 C 53.72 D 55.87 A 47.99 E 59.45
A 42.26 B 50.14 E 55.87 D 58.74 C 55.87
D 60.89 E 59.45 A 49.43 C 59.45 B 57.31
proc glm;
class trt row col;
model yield=trt row col;
run;
```

ويمكن كتابة قطعة البرنامج بشكل عمودي وكالآتي:

```
data wheat;
input ROW COL trt $ yield ;
CARDS;
1 1 E 59.45
1 2 A 47.28
1 3 C 54.44
1 4 B 50.14
1 5 D 59.45
2 1 C 55.16
2 2 D 60.89
2 3 B 56.59
2 4 E 60.17
2 5 A 48.71
3 1 B 44.41
3 2 C 53.72
3 3 D 55.87
3 4 A 47.99
3 5 E 59.45
4 1 A 42.26
4 2 B 50.14
4 3 E 55.87
4 4 D 58.74
4 5 C 55.87
5 1 D 60.89
5 2 E 59.45
5 3 A 49.43
5 4 C 59.45
5 5 B 57.31
proc glm;
class trt row col;
model yield=trt row col;
run;
```

Dr. Firas

4 – 6 تحليل التباين المشترك Analysis of Covariance (ANCOVA)

هي طريقة احصائية تسمح لنا بتضمين المتغيرات المستمرة وذات الفئات في انموذج واحد عند التحليل. وتفترض هذه الطريقة ان معامل الانحدار متجانس (homogeneous) خلال جميع فئات العامل الاخر.

مثال (71) : في دراسة شملت 30 حالة استعمل فيها 3 انواع من الغذاء (سيطرة ومعاملتين) وقبل بدء التجربة سجلت اطوال الوحدات التجريبية وبعد انتهاء فترة التجربة سجلت الاوزان اختبر الفروق بين المجموعات بسبب اختلاف التغذية؟

ان اجراء التحليل بأستعمال تحليل التباين لدراسة تأثير الغذاء على الوزن قد يؤدي الى حصولنا على استنتاجات غير صحيحة اذا اهملنا تأثير الطول لذا فأن طريقة تحليل التباين المشترك تسمح لنا بأدخال الطول كانحدار ومن ثم يمكننا ان نتخلص من تأثيره والحصول على نتائج اكثر دقة.

```
DATA htwt;
 INPUT id diet height weight ;
CARDS;
 1 1  56 140
 2 1  60 155
 3 1  64 143
 4 1  68 161
 5 1  72 139
 6 1  54 159
 7 1  62 138
 8 1  65 121
 9 1  65 161
10 1  70 145
11 2  56 117
12 2  60 125
13 2  64 133
14 2  68 141
15 2  72 149
16 2  54 109
17 2  62 128
18 2  65 131
19 2  65 131
20 2  70 145
21 3  54 211
22 3  58 223
23 3  62 235
24 3  66 247
25 3  70 259
26 3  52 201
27 3  59 228
28 3  64 245
29 3  65 241
30 3  72 269
PROC GLM DATA=htwt;
 CLASS diet ;
 MODEL weight = height diet ;
 MEANS diet / LSD ;
 run;
```

Dr. Firas

```
                                The SAS System

          General Linear Models Procedure
            Class Level Information

           Class    Levels    Values

           DIET          3    1 2 3
```

Number of observations in data set = 30

The SAS System
General Linear Models Procedure

Dependent Variable: WEIGHT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	67409.81107492	22469.93702497	157.80	0.0001
Error	26	3702.18892508	142.39188173		
Corrected Total	29	71112.00000000			

R-Square	C.V.	Root MSE	WEIGHT Mean
0.947939	6.978250	11.9328069	171.00000000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
HEIGHT	1	683.72455130	683.72455130	4.80	0.0376
DIET	2	66726.08652362	33363.04326181	234.30	0.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
HEIGHT	1	3059.21107492	3059.21107492	21.48	0.0001
DIET	2	66726.08652362	33363.04326181	234.30	0.0001

The SAS System
General Linear Models Procedure

T tests (LSD) for variable: WEIGHT

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 26 MSE= 142.3919
Critical Value of T= 2.06
Least Significant Difference= 10.969

Means with the same letter are not significantly different.

T Grouping	Mean	N	DIET
A	235.900	10	3
B	146.200	10	1
C	130.900	10	2

Dr. Firas

الفصل الخامس
الأنحدار Regression

هو وسيلة احصائية تستخدم لتحليل العلاقة بين متغير مستقل او اكثر ومتغير تابع ، ويعد الانحدار من اكثر الطرق الاحصائية استعمالا ولشتى العلوم لانه يصف العلاقة بين المتغيرات بهيئة معادلة . والانحدار يقسم الى انواع :

5 – 1 الانحدار البسيط Simple regression

عندما يكون هناك متغيران مستمران احدهما مستقل والآخر تابع فيسمى الانحدار البسيط وتكون معادلته كالآتي:

$$\hat{Y} = b0 + b1X1$$

مثال (72): اذا علمت ان y متغيرتابع والمتغيرات x1 x2 x3 متغيرات مستقلة جد معامل الانحدار البسيط والمتعدد؟

Dr. Firas

```
data c;
 input X1 X2 X3 Y;

 cards;
 19 43 29 11
 24 49 28 22
 30 51 37 18
 29 54 31 20
 19 42 30 12
 25 53 23 21
 31 58 27 27
proc reg data=c ;
MODEL Y= X1;
MODEL Y=X2;
MODEL Y=X3;
Run;
```

```
                                The SAS System
                                Model: MODEL1
                            Dependent Variable: Y

                                Analysis of Variance

                Sum of      Mean
Source          DF      Squares     Square      F Value     Prob>F

Model           1    121.19148   121.19148       8.627     0.0324
Error           5     70.23709    14.04742
C Total         6    191.42857

    Root MSE       3.74799    R-square     0.6331
    Dep Mean      18.71429    Adj R-sq     0.5597
    C.V.          20.02742

                            Parameter Estimates

                Parameter     Standard    T for H0:
```

```
                                  Variable  DF     Estimate       Error   Parameter=0    Prob > |T|

                                  INTERCEP   1    -4.057361    7.88112636      -0.515       0.6286
                                    X1       1     0.900574    0.30660656       2.937       0.0324

                                                    The SAS System
                                                   Model: MODEL2
                                                Dependent Variable: Y

                                                  Analysis of Variance

                                 Sum of      Mean
                                              Source      DF     Squares      Square     F Value      Prob>F

                                               Model       1   162.37255   162.37255      27.941      0.0032
Error       5    29.05602     5.81120
C Total     6   191.42857

                                                    Root MSE      2.41064    R-square      0.8482
                                                   Dep Mean      18.71429    Adj R-sq      0.8179
                                                                  C.V.        12.88130

                                                  Parameter Estimates

                                                        Parameter     Standard    T for H0:
                                  Variable  DF     Estimate       Error   Parameter=0    Prob > |T|

                                  INTERCEP   1   -25.893557    8.48798641      -3.051       0.0284
                                    X2       1     0.892157    0.16877884       5.286       0.0032

                                                    The SAS System
                                                   Model: MODEL3
                                                Dependent Variable: Y

                                                  Analysis of Variance

                                 Sum of      Mean
                                              Source      DF     Squares      Square     F Value      Prob>F

                                               Model       1    17.23536    17.23536       0.495      0.5132
Error       5   174.19321    34.83864
C Total     6   191.42857

                                                    Root MSE      5.90243    R-square      0.0900
                                                   Dep Mean      18.71429    Adj R-sq     -0.0920
                                                                  C.V.        31.53969

                                                  Parameter Estimates

                                                        Parameter     Standard    T for H0:
                                  Variable  DF     Estimate       Error   Parameter=0    Prob > |T|

                                  INTERCEP   1    30.336815   16.67413924       1.819       0.1285
                                    X3       1    -0.396867    0.56424179      -0.703       0.5132
```

ويمكن استعمال الايعاز التالي لرسم العلاقة بين الخطأ والقيم المتنبأ بها .

```
plot r.*p.;
run;
```

على ضوء الرسم المبعثر يمكن تحديد مدى مطابقة النموذج للعلاقة بين المتغيرين فأن لم نلاحظ وجود اتجاه محدد للعلاقة فأن ذلك يعني ان العلاقة الخطية غير مناسبة وان من الافضل اختبار معادلة الانحدار من الدرجة الثانية او الثالثة.

Dr. Firas

ملاحظة: هناك عدة ايعازات يمكن استعمالها للحصول على بعض التقديرات اذ تكتب هذه الايعازات بعد الخط المائل الخاص بالنموذج وكما يلي:

```
proc reg data=c ;
MODEL Y= X1/ nointer;
Run;
```

Nointer = ايعاز بكتابة المعالم بدون intercept

P = ايعاز بكتابة القيم المتنبأ بها مع الخطأ

r = ايعاز بكتابة القيم المتنبأ بها اضافة قيم تشخيص الخطأ (Residual diagnostics)

clm = ايعاز بكتابة القيم المتنبأ بها علاوة على حدود الثقة 95% للمتوسطات المحسوبة

cli = ايعاز بكتابة القيم المتنبأ بها علاوة على حدود الثقة 95% للقيم المتنبأ بها

r clm cli = تشمل جميع الايعازات

collin = ايعاز لتشخيص العلاقة بين المتغيرات

collinoint = ايعاز لتشخيص العلاقة بين المتغيرات بدون intercept

slentry = معنوية المستوى لاختيار العامل

slstay = معنوية المستوى لاستبعاد العامل

Include=2 وتعني امر اجباري بشمول اول عاملين

Start=2 وتعني امر بالبدأ بعاملين

5 – 2 الانحدار المتعدد Multiple regression

عندما يكون هناك اكثر من متغير مستقل فيسمى الانحدار المتعدد ومعادلته:

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + \ldots b_nX_n$$

يمكن استعمال الايعازات التالية لنفس بيانات المثال السابق.

```
MODEL Y=X1 X2;
MODEL Y= X1 X3;
MODEL Y= X2 X3;
MODEL Y=X1 X2 X3;
run;
```

يمكن الحصول على جدول تحليل التباين وتقدير معامل الانحدار المتعدد بأستعمال الايعاز الآتي:

```
proc glm;
model y=x1 x2;run;
```

The SAS System
Model: MODEL4
Dependent Variable: Y

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	165.16342	82.58171	12.577	0.0188
Error	4	26.26515	6.56629		
C Total	6	191.42857			

Dr. Firas

Root MSE 2.56248 R-square 0.8628
Dep Mean 18.71429 Adj R-sq 0.7942
C.V. 13.69263

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-30.649895	11.60316757	-2.642	0.0575
X1	1	-0.341446	0.52373625	-0.652	0.5500
X2	1	1.159958	0.44824389	2.588	0.0608

The SAS System
Model: MODEL5
Dependent Variable: Y

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	169.79191	84.89596	15.695	0.0128
Error	4	21.63666	5.40916		
C Total	6	191.42857			

Root MSE 2.32576 R-square 0.8870
Dep Mean 18.71429 Adj R-sq 0.8305
C.V. 12.42773

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	12.503783	7.37856715	1.695	0.1654
X1	1	1.040294	0.19588709	5.311	0.0060
X3	1	-0.686139	0.22890605	-2.997	0.0400

The SAS System
Model: MODEL6
Dependent Variable: Y

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	166.56802	83.28401	13.400	0.0169
Error	4	24.86055	6.21514		
C Total	6	191.42857			

Root MSE 2.49302 R-square 0.8701
Dep Mean 18.71429 Adj R-sq 0.8052
C.V. 13.32147

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-18.860306	12.26104442	-1.538	0.1988
X2	1	0.867818	0.17704212	4.902	0.0080
X3	1	-0.198606	0.24172764	-0.822	0.4574

The SAS System
Model: MODEL7
Dependent Variable: Y

Analysis of Variance

Dr. Firas
Dr. Firas

```
                Sum of      Mean
Source          DF      Squares     Square      F Value     Prob>F

Model           3    174.34350    58.11450      10.204     0.0440
Error           3     17.08507     5.69502
C Total         6    191.42857

                Root MSE       2.38642    R-square     0.9107
                Dep Mean      18.71429    Adj R-sq     0.8215
                C.V.          12.75189

                Parameter Estimates

                Parameter     Standard    T for H0:
Variable  DF    Estimate      Error       Parameter=0    Prob > |T|

INTERCEP   1   115.984697   115.99879797       1.000       0.3911
X1         1     4.408267     3.77269760       1.168       0.3270
X2         1    -2.843782     3.18099148      -0.894       0.4372
X3         1    -2.272353     1.78978257      -1.270       0.2938
```

5 – 2 – 1 أختيار احسن المعادلات Selecting the best regression equation

في كثير من الدراسات قد يعمد الباحث الى ادخال عدد كبير من المتغيرات لدراسة تأثيرها في الصفة المدروسة وبذلك فأن اختيار افضل معادلة تنبوء تكون عملية صعبة ولكن يمكن اعتماد بعض الطرق في التحليل الاحصائي يمكن بموجبها الحصول على افضل معادلة على ان تتضمن اقل عدد ممكن من المتغيرات المستقلة.

5 – 2 – 1 – 1 كل الانحدارات الممكنة All possible regression procedure

يمكن الحصول على جميع الانحدارات الممكنة للمثال السابق بأستعمال ايعاز معين وكما موضح في المثال التالي:

مثال (73):

```
MODEL Y = x1-x3/selection=rsquare adjrsq cp press mse sse;
RUN;
```

```
                                The SAS System
 N = 7    Regression Models for Dependent Variable: Y


Number in    R-square    Adjusted      C(p)       MSE         SSE      Var
 Model                    R-square                          in Model

   1     0.84821481    0.81785777     2.10200    5.811204    29.05602   X2
   1     0.63308981    0.55970777     9.33307   14.047419    70.23709   X1
   1     0.09003546    -.09195745    27.58692   34.838642   174.19321   X3
-------------------------------------------------------------------------------
   2     0.88697268    0.83045902     2.79922    5.409165    21.63666   X1 X3
   2     0.87013146    0.80519719     3.36531    6.215137    24.86055   X2 X3
   2     0.86279398    0.79419096     3.61195    6.566288    26.26515   X1 X2
-------------------------------------------------------------------------------
   3     0.91074964    0.82149928     4.00000    5.695023    17.08507   X1 X2 X3
-------------------------------------------------------------------------------
```

Dr. Firas

مثال (74): في هذا المثال يمكن الحصول على القيم المتنبأ بها مع حدود الثقة بأضافة clm/ الى نموذج الانحدار.

```
data c;
 input x y;
cards;
  80  399
  30  121
  50  221
  90  376
  70  361
  60  224
 120  546
  80  352
 100  353
  50  157
  40  160
  70  252
  90  389
  20  213
 110  435
 100  420
  30  212
  50  268
  90  377
 110  421
  30  273
  90  468
  40  244
  80  342
  70  323
 option nodate nonumber;
proc reg data = c;
 model y = x/ clM;
run;
```

```
                                The SAS System

Model: MODEL1
Dependent Variable: Y

                    Analysis of Variance

                      Sum of       Mean
        Source         DF     Squares     Square     F Value     Prob>F

        Model           1 217938.18687 217938.18687      84.946      0.0001
        Error          23  59008.85313   2565.60231
        C Total        24 276947.04000

         Root MSE     50.65177     R-square      0.7869
         Dep Mean    316.28000     Adj R-sq      0.7777
         C.V.         16.01485

                     Parameter Estimates

                  Parameter     Standard    T for H0:
        Variable  DF     Estimate       Error   Parameter=0    Prob > |T|

        INTERCEP   1    84.042626  27.15779586       3.095       0.0051
        X          1     3.317677   0.35996649       9.217       0.0001
```

Dr.Firas

The SAS System

Obs	Dep Var Y	Predict Value	Std Err Predict	Lower95% Mean	Upper95% Mean	Residual
1	399.0	349.5	10.751	327.2	371.7	49.5432
2	121.0	183.6	17.605	147.2	220.0	-62.5729
3	221.0	249.9	12.428	224.2	275.6	-28.9265
4	376.0	382.6	12.428	356.9	408.3	-6.6335
5	361.0	316.3	10.130	295.3	337.2	44.7200
6	224.0	283.1	10.751	260.9	305.3	-59.1032
7	546.0	482.2	20.653	439.4	524.9	63.8362
8	352.0	349.5	10.751	327.2	371.7	2.5432
9	353.0	415.8	14.807	385.2	446.4	-62.8103
10	157.0	249.9	12.428	224.2	275.6	-92.9265
11	160.0	216.7	14.807	186.1	247.4	-56.7497
12	252.0	316.3	10.130	295.3	337.2	-64.2800
13	389.0	382.6	12.428	356.9	408.3	6.3665
14	213.0	150.4	20.653	107.7	193.1	62.6038
15	435.0	449.0	17.605	412.6	485.4	-13.9871
16	420.0	415.8	14.807	385.2	446.4	4.1897
17	212.0	183.6	17.605	147.2	220.0	28.4271
18	268.0	249.9	12.428	224.2	275.6	18.0735
19	377.0	382.6	12.428	356.9	408.3	-5.6335
20	421.0	449.0	17.605	412.6	485.4	-27.9871
21	273.0	183.6	17.605	147.2	220.0	89.4271
22	468.0	382.6	12.428	356.9	408.3	85.3665
23	244.0	216.7	14.807	186.1	247.4	27.2503
24	342.0	349.5	10.751	327.2	371.7	-7.4568
25	323.0	316.3	10.130	295.3	337.2	6.7200

Sum of Residuals	0
Sum of Squared Residuals	59008.8531
Predicted Resid SS (Press)	71469.4710

5 – 2 – 1 – 2 الانحدار التدريجي Stepwise regression procedure

مثال (75) : حدد معادلة التوقع للقدرة على العيش اعتمادا على تأثير العوامل المؤثرة فيها؟

```
data c;
 input x1 x2 x3 x4 y;
 label x1 = 'blood-clotting'
     x2 = 'prognostic'
       x3 = 'enzyme'
       x4 = 'liver function'
        y = 'survival';
logy = log10(y);
cards;
  6.7  62   81  2.59  200  2.3010
  5.1  59   66  1.70  101  2.0043
  7.4  57   83  2.16  204  2.3096
  6.5  73   41  2.01  101  2.0043
  7.8  65  115  4.30  509  2.7067
  5.8  38   72  1.42   80  1.9031
  5.7  46   63  1.91   80  1.9031
  3.7  68   81  2.57  127  2.1038
  6.0  67   93  2.50  202  2.3054
  3.7  76   94  2.40  203  2.3075
  6.3  84   83  4.13  329  2.5172
  6.7  51   43  1.86   65  1.8129
  5.8  96  114  3.95  830  2.9191
```

M.M. Firas

```
  5.8  83   88  3.95  330  2.5185
  7.7  62   67  3.40  168  2.2253
  7.4  74   68  2.40  217  2.3365
  6.0  85   28  2.98   87  1.9395
  3.7  51   41  1.55   34  1.5315
  7.3  68   74  3.56  215  2.3324
  5.6  57   87  3.02  172  2.2355
  5.2  52   76  2.85  109  2.0374
  3.4  83   53  1.12  136  2.1335
  6.7  26   68  2.10   70  1.8451
  5.8  67   86  3.40  220  2.3424
  6.3  59  100  2.95  276  2.4409
  5.8  61   73  3.50  144  2.1584
  5.2  52   86  2.45  181  2.2577
 11.2  76   90  5.59  574  2.7589
  5.2  54   56  2.71   72  1.8573
  5.8  76   59  2.58  178  2.2504
  3.2  64   65  0.74   71  1.8513
  8.7  45   23  2.52   58  1.7634
  5.0  59   73  3.50  116  2.0645
  5.8  72   93  3.30  295  2.4698
  5.4  58   70  2.64  115  2.0607
  5.3  51   99  2.60  184  2.2648
  2.6  74   86  2.05  118  2.0719
  4.3   8  119  2.85  120  2.0792
  4.8  61   76  2.45  151  2.1790
  5.4  52   88  1.81  148  2.1703
  5.2  49   72  1.84   95  1.9777
  3.6  28   99  1.30   75  1.8751
  8.8  86   88  6.40  483  2.6840
  6.5  56   77  2.85  153  2.1847
  3.4  77   93  1.48  191  2.2810
  6.5  40   84  3.00  123  2.0899
  4.5  73  106  3.05  311  2.4928
  4.8  86  101  4.10  398  2.5999
  5.1  67   77  2.86  158  2.1987
  3.9  82  103  4.55  310  2.4914
  6.6  77   46  1.95  124  2.0934
  6.4  85   40  1.21  125  2.0969
  6.4  59   85  2.33  198  2.2967
  8.8  78   72  3.20  313  2.4955
Option nodate nonumber;
proc reg data = c ;
var x1-x4;
model y = x1 -x4 ;
run;
proc reg data = c;
model logy = x1-x4/ selection = stepwise slentry= .01 slstay= .05;
run;
```

Dr. Firas

The SAS System

Model: MODEL1
Dependent Variable: Y survival

Analysis of Variance

Sum of Mean

Source	DF	Squares	Square	F Value	Prob>F
Model	4	936264.53767	234066.13442	62.788	0.0001
Error	49	182666.96233	3727.89719		
C Total	53	1118931.5			

Root MSE	61.05651	R-square	0.8367
Dep Mean	197.16667	Adj R-sq	0.8234
C.V.	30.96695		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-621.597550	64.80042601	-9.592	0.0001
X1	1	33.163828	7.01727463	4.726	0.0001
X2	1	4.271860	0.56338454	7.582	0.0001
X3	1	4.125738	0.51116093	8.071	0.0001
X4	1	14.091563	12.52532754	1.125	0.2661

Variable	DF	Label
INTERCEP	1	Intercept
X1	1	blood-clotting
X2	1	prognostic
X3	1	enzyme
X4	1	liver function

The SAS System

Stepwise Procedure for Dependent Variable LOGY

Step 1 Variable X4 Entered R-square = 0.52736470 C(p) =787.94706372

	DF	Sum of Squares	Mean Square	F	Prob>F
Regression	1	2.09507677	2.09507677	58.02	0.0001
Error	52	1.87765176	0.03610869		
Total	53	3.97272852			

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	1.69639145	0.07174361	20.18818844	559.10	0.0001
X4	0.18575207	0.02438594	2.09507677	58.02	0.0001

Bounds on condition number: 1, 1

Step 2 Variable X3 Entered R-square = 0.68650294 C(p) =507.80686041

	DF	Sum of Squares	Mean Square	F	Prob>F
Regression	2	2.72728979	1.36364490	55.84	0.0001
Error	51	1.24543873	0.02442037		
Total	53	3.97272852			

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	1.38880640	0.08447159	6.60106867	270.31	0.0001
X3	0.00565213	0.00111085	0.63221303	25.89	0.0001
X4	0.13901554	0.02205792	0.96995100	39.72	0.0001

Bounds on condition number: 1.209789, 4.839155

Step 3 Variable X2 Entered R-square = 0.88287606 C(p) =161.65199200

	DF	Sum of Squares	Mean Square	F	Prob>F

Dr Firas

Regression	3	3.50742690	1.16914230	125.63	0.0001
Error	50	0.46530162	0.00930603		
Total	53	3.97272852			

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	0.94228698	0.07139675	1.62096662	174.18	0.0001
X2	0.00789875	0.00086269	0.78013711	83.83	0.0001
X3	0.00699934	0.00070135	0.92684008	99.60	0.0001
X4	0.08184598	0.01498002	0.27780159	29.85	0.0001

Bounds on condition number: 1.464174, 11.82183

Step 4 Variable X1 Entered R-square = 0.97236206 C(p) = 5.00000000

	DF	Sum of Squares	Mean Square	F	Prob>F
Regression	4	3.86293050	0.96573262	430.98	0.0001
Error	49	0.10979803	0.00224078		
Total	53	3.97272852			

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	0.48873772	0.05023948	0.21206057	94.64	0.0001
X1	0.06852649	0.00544046	0.35550359	158.65	0.0001
X2	0.00925433	0.00043679	1.00587474	448.90	0.0001
X3	0.00947438	0.00039630	1.28070851	571.55	0.0001
X4	0.00191796	0.00971083	0.00008741	0.04	0.8442

Bounds on condition number: 2.555329, 29.28635

Step 5 Variable X4 Removed R-square = 0.97234006 C(p) = 3.03900913

	DF	Sum of Squares	Mean Square	F	Prob>F
Regression	3	3.86284308	1.28761436	585.89	0.0001
Error	50	0.10988544	0.00219771		
Total	53	3.97272852			

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	0.48362257	0.04263391	0.28279576	128.68	0.0001
X1	0.06922866	0.00407844	0.63321777	288.13	0.0001
X2	0.00929460	0.00038255	1.29733876	590.31	0.0001
X3	0.00952328	0.00030644	2.12246912	965.76	0.0001

Bounds on condition number: 1.030813, 9.186356

All variables left in the model are significant at the 0.0500 level.

No other variable met the 0.0100 significance level for entry into the model.

Summary of Stepwise Procedure for Dependent Variable LOGY

Step	Variable Entered	Removed	Number In	Partial R**2	Model R**2	C(p)	F	Prob>F	Label
1	X4		1	0.5274	0.5274	787.9471	58.0214	0.0001	liver function
2	X3		2	0.1591	0.6865	507.8069	25.8888	0.0001	enzyme
3	X2		3	0.1964	0.8829	161.6520	83.8313	0.0001	prognostic

Dr Firas

```
4       X1        4   0.0895  0.9724   5.0000  158.6520  0.0001
    blood-clotting
5           X4    3   0.0000  0.9723   3.0390    0.0390  0.8442
            liver function
```

5 – 2 – 1 - 3 الاختيار المباشراو الامامي Forward selection procedure

مثال (76) : يمكن تحديد افضل العوامل التي يمكن اعتمادها للتنبوء في الصفة التابعة بهذه الطريقة اذ يقوم البرنامج بأختبار عامل ثم عاملين الى ان يدخل تأثير جميع العوامل ويمكن اختيار افضل معادلة اعتمادا على قيم معامل التحديد (R^2).

```
data p054;
 input Y X1 X2 X3 X4 X5 X6 ;
cards;
43 51 30 39 61 92 45
63 64 51 54 63 73 47
71 70 68 69 76 86 48
61 63 45 47 54 84 35
81 78 56 66 71 83 47
43 55 49 44 54 49 34
58 67 42 56 66 68 35
71 75 50 55 70 66 41
72 82 72 67 71 83 31
67 61 45 47 62 80 41
64 53 53 58 58 67 34
67 60 47 39 59 74 41
69 62 57 42 55 63 25
68 83 83 45 59 77 35
77 77 54 72 79 77 46
81 90 50 72 60 54 36
74 85 64 69 79 79 63
65 60 65 75 55 80 60
65 70 46 57 75 85 46
50 58 68 54 64 78 52
50 40 33 34 43 64 33
64 61 52 62 66 80 41
53 66 52 50 63 80 37
40 37 42 58 50 57 49
63 54 42 48 66 75 33
66 77 66 63 88 76 72
78 75 58 74 80 78 49
48 57 44 45 51 83 38
85 85 71 71 77 74 55
82 82 39 59 64 78 39
proc reg data = p054;
model y = x1-x6/ selection = forward slentry = 0.99;
run;
```

```
                                The SAS System
          Forward Selection Procedure for Dependent Variable Y

Step 1   Variable X1 Entered      R-square = 0.68131416   C(p) =  1.41147660

               DF       Sum of Squares      Mean Square        F   Prob>F

     Regression     1       2927.58425270   2927.58425270     59.86   0.0001
```

Dr. Firas

```
      Error         28       1369.38241396     48.90651478
      Total         29       4296.96666667

                Parameter       Standard        Type II
   Variable       Estimate         Error   Sum of Squares        F   Prob>F

   INTERCEP     14.37631941     6.61998601     230.64709952     4.72   0.0385
   X1            0.75460982     0.09753289    2927.58425270    59.86   0.0001

Bounds on condition number:          1,          1
------------------------------------------------------------------------------------

Step 2   Variable X3 Entered       R-square = 0.70801520   C(p) =  1.11481128

          DF       Sum of Squares     Mean Square        F   Prob>F
Regression  2        3042.31769607    1521.15884804     32.74   0.0001
Error      27        1254.64897059     46.46848039
Total      29        4296.96666667

                Parameter       Standard        Type II
   Variable       Estimate         Error   Sum of Squares        F   Prob>F

   INTERCEP      9.87088045     7.06122360     90.80511791     1.95   0.1735
   X1            0.64351764     0.11847743   1370.90743789    29.50   0.0001
   X3            0.21119181     0.13440372    114.73344337     2.47   0.1278
Bounds on condition number:     1.553021,     6.212083
------------------------------------------------------------------------------------

Step 3   Variable X6 Entered       R-square = 0.72559500   C(p) =  1.60270022

              DF       Sum of Squares     Mean Square        F   Prob>F

   Regression     3        3117.85752667    1039.28584222     22.92   0.0001
   Error         26        1179.10913999     45.35035154
   Total         29        4296.96666667

                Parameter       Standard        Type II
   Variable       Estimate         Error   Sum of Squares        F   Prob>F

    INTERCEP     13.57774107     7.54389972     146.90747338     3.24   0.0835
    X1            0.62272975     0.11814642    1259.90769102    27.78   0.0001
    X3            0.31238698     0.15419973     186.12266609     4.10   0.0532
    X6           -0.18695081     0.14485373      75.53983060     1.67   0.2082
                          The SAS System
Bounds on condition number:     2.094593,     15.29215
------------------------------------------------------------------------------------

Step 4   Variable X2 Entered       R-square = 0.72934125   C(p) =  3.28046994

            DF       Sum of Squares     Mean Square        F   Prob>F

   Regression  4        3133.95503514    783.48875878     16.84   0.0001
   Error      25        1163.01163153     46.52046526
   Total      29        4296.96666667

              Parameter       Standard        Type II
   Variable       Estimate         Error   Sum of Squares        F   Prob>F

   INTERCEP     14.30346938     7.73956534     158.88894771     3.42   0.0765
   X1            0.65337761     0.13051133    1165.93982140    25.06   0.0001
   X2           -0.07681738     0.13058767      16.09750847     0.35   0.5616
   X3            0.32394970     0.15740848     197.03481253     4.24   0.0502
   X6           -0.17150977     0.14904033      61.60474913     1.32   0.2607

Bounds on condition number:     2.127773,     28.26995
------------------------------------------------------------------------------------

Step 5   Variable X4 Entered       R-square = 0.73180937   C(p) =  5.06817654

                DF       Sum of Squares     Mean Square        F   Prob>F
```

Dr.Firas

```
      Regression     5       3144.56047808    628.91209562     13.10  0.0001
      Error         24       1152.40618859     48.01692452
      Total         29       4296.96666667

            Parameter       Standard         Type II
  Variable      Estimate         Error   Sum of Squares       F   Prob>F

   INTERCEP    12.79790657     8.49061173    109.09233756     2.27  0.1448
   X1           0.61314610     0.15782694    724.70294635    15.09  0.0007
   X2          -0.07223764     0.13302881     14.15891594     0.29  0.5921
   X3           0.31172229     0.16202277    177.73703177     3.70  0.0663
   X4           0.09795095     0.20842096     10.60544294     0.22  0.6426
   X6          -0.21110682     0.17328146     71.26801701     1.48  0.2350
Bounds on condition number:    2.836099,     56.0349
--------------------------------------------------------------------------------

Step 6   Variable X5 Entered       R-square = 0.73260199   C(p) =  7.00000000

           DF       Sum of Squares      Mean Square        F   Prob>F

 Regression     6       3147.96634184    524.66105697     10.50  0.0001
 Error         23       1149.00032483     49.95653586
 Total         29       4296.96666667

                         The SAS System
            Parameter       Standard         Type II
  Variable      Estimate         Error   Sum of Squares       F   Prob>F
```

```
INTERCEP    10.78707639    11.58925724     43.28013610      0.87  0.3616
X1           0.61318761     0.16098311    724.80036190     14.51  0.0009
X2          -0.07305014     0.13572469     14.47160631      0.29  0.5956
X3           0.32033212     0.16852032    180.50479311      3.61  0.0699
X4           0.08173213     0.22147768      6.80327563      0.14  0.7155
X5           0.03838145     0.14699544      3.40586376      0.07  0.7963
X6          -0.21705668     0.17820947     74.11004403      1.48  0.2356

Bounds on condition number:    3.078226,     76.78153
--------------------------------------------------------------------------------

All variables have been entered into the model.

        Summary of Forward Selection Procedure for Dependent Variable Y

      Variable    Number   Partial    Model
 Step   Entered      In     R**2      R**2      C(p)        F   Prob>F

    1   X1            1   0.6813    0.6813    1.4115    59.8608  0.0001
    2   X3            2   0.0267    0.7080    1.1148     2.4691  0.1278
    3   X6            3   0.0176    0.7256    1.6027     1.6657  0.2082
    4   X2            4   0.0037    0.7293    3.2805     0.3460  0.5616
    5   X4            5   0.0025    0.7318    5.0682     0.2209  0.6426
    6   X5            6   0.0008    0.7326    7.0000     0.0682  0.7963
```

5 – 2 – 1 - 4 الحذف العكسي Backward elimination procedure

يمكن تطبيقها بأستعمال الايعاز التالي:

```
proc reg data = p054;
model y = x1-x6/ selection = backward slstay = 0.01;
run;
```

5 – 2 – 1 - 5 اختيار افضل معادلة اعتمادا على قيمة مالو C_p

يمكن تطبيقها بأستعمال الايعاز التالي:

```
proc reg data = p054 ;
```

Dr. Firas

```
model y = x1-x6/ selection = cp;
run;
```

5 – 2 – 1 - 6 اختيار افضل معادلة اعتمادا على معامل التحديد R^2

يمكن تطبيقها بأستعمال الايعاز التالي:

```
proc reg data = p054;
model y = x1-x6/ selection = rsquare;
run;
```

5 – 3 الانحدار اللاخطي Curvilinear regression

رغم ان الانحدار الخطي او المستقيم يعد من اسهل انواع الانحدار الا ان العلاقة الخطية لاتتواجد بين الكثير من المتغيرات وانما قد تكون العلاقة من الدرجة الثانية او الثالثة ومعادلته:

$$\hat{Y} = b0 + b1Xi + b2Xi^2 + ei$$

مثال (77) :حدد افضل معادلة للتنبوء لقيمة المتغير y بالاعتماد على المتغير x ؟

```
data t;
input   y x;
x2=x*x;
x3=x*x*x;
cards;
150  5
240 10
280 15
330 20
290 24
200 32
proc reg data=t;
model y=x;
model y=x x2;
model y=x x2 x3;
run;
proc plot;
plot y*x;
run;
```

```
                                                The SAS System

Model: MODEL1
Dependent Variable: Y
                  Analysis of Variance

                   Sum of       Mean
  Source       DF    Squares     Square    F Value     Prob>F

  Model         1  2039.47858  2039.47858      0.420     0.5525
  Error         4  19443.85475  4860.96369
  C Total       5  21483.33333

    Root MSE     69.72061    R-square     0.0949
    Dep Mean    248.33333    Adj R-sq    -0.1313
    C.V.        28.07541

                  Parameter Estimates

             Parameter     Standard    T for H0:
  Variable  DF     Estimate       Error   Parameter=0    Prob > |T|

  INTERCEP   1   211.815642   63.15512699       3.354       0.0285
  X          1     2.067039    3.19117269       0.648       0.5525
```

Dr Firas

```
                    The SAS System
Model: MODEL2
Dependent Variable: Y
                  Analysis of Variance

                   Sum of       Mean
 Source       DF    Squares     Square    F Value     Prob>F

 Model         2 20863.04746 10431.52373     50.452     0.0049
 Error         3   620.28588   206.76196
 C Total       5 21483.33333

    Root MSE     14.37922    R-square     0.9711
    Dep Mean    248.33333    Adj R-sq     0.9519
    C.V.         5.79029

                  Parameter Estimates
             Parameter     Standard   T for H0:
 Variable  DF    Estimate       Error  Parameter=0   Prob > |T|

 INTERCEP   1   19.421076  24.00504693       0.809      0.4777
 X          1   29.437477   2.94310690      10.002      0.0021
 X2         1   -0.743361   0.07790840      -9.541      0.0024

                    The SAS System
Model: MODEL3
Dependent Variable: Y
                  Analysis of Variance

                   Sum of       Mean
 Source       DF    Squares     Square    F Value     Prob>F

 Model         3 20889.80266  6963.26755     23.464     0.0412
 Error         2   593.53067   296.76534
 C Total       5 21483.33333

    Root MSE     17.22688    R-square     0.9724
    Dep Mean    248.33333    Adj R-sq     0.9309
    C.V.         6.93700

                  Parameter Estimates

             Parameter     Standard   T for H0:
 Variable  DF    Estimate       Error  Parameter=0   Prob > |T|

 INTERCEP   1   34.836340  58.84595391       0.592      0.6139
 X          1   25.702575  12.92897515       1.988      0.1851
 X2         1   -0.508552   0.78756949      -0.646      0.5847
 X3         1   -0.004204   0.01399954      -0.300      0.7923
```

هنا يمكن الاعتماد على معادلة الدرجة الثانية كأفضل معادلة توقع. ولو اردنا تقدير احدى قيم x المتوقعة اذا كانت قيمة x تساوي 24 باعتماد معادلة الدرجة الاولى فأنها ستساوي :

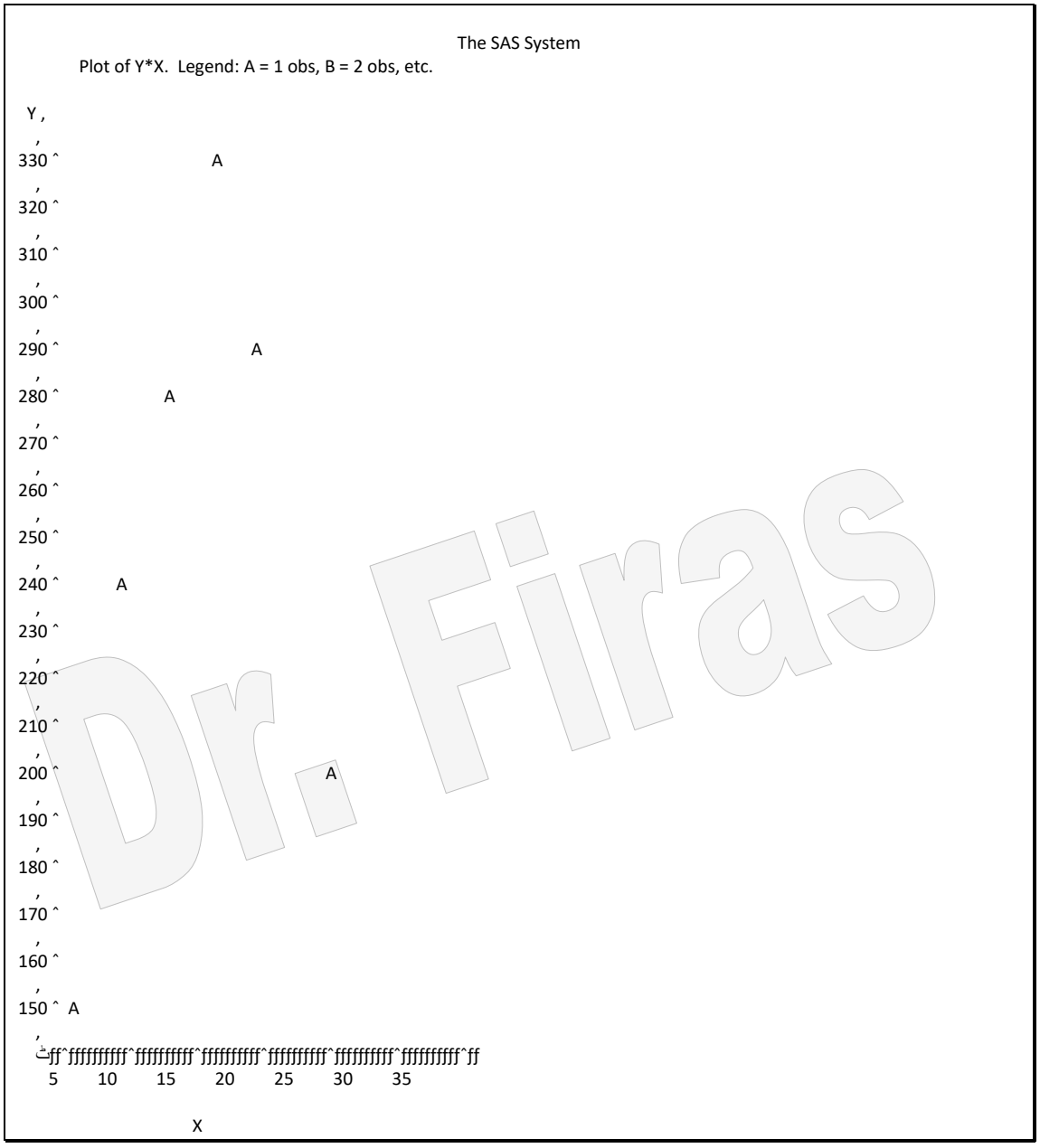
$$Y= 211.81 + 2.06(24) = 261.25$$

اما اذا استعملنا معادلة الدرجة الثانية في التقدير فأن القيمة المتوقعة تساوي:

$$Y = 19.42 + 29.43(24) + 0.74(24)^2 = 299.5$$

ونلاحظ ان القيمة المقدرة في معادلة الدرجة الثانية اكثر قربا الى القيمة الحقيقية.

Dr Firas



المخطط اعلاه يوضح بأن العلاقة لاخطية بين المتغيرين.

5 – 4 الانحدار اللوجستيكي Logistic Regression

ويستعمل عندما يكون المتغير المستقل مستمر والمتغير التابع ثنائي (binary) اي يأخذ قيمتان 1 او صفراو عندما يكون المتغير التابع ذو فئات مرتبة (Ordinal) (واطىء ، متوسط ، عالي) او نسبة. في برنامج SAS يمكن تنفيذ البرنامج بأستعمال عدة طرق (LOGISTIC, PROBIT, GENMOD , CATMOD) ويستعمل الامر Descending مع Logistic

فقط وفيه سيعتبر البرنامج القيمة الاكبر (1) من المتغير هي النجاح و القيمة 0 هي الفشل ، وتعد طريقة genmod اكثر الطرق مرونة اذ يمكن ان تتعامل مع البيانات الاسمية دون الحاجة الى انشاء متغيرات (1، 0) (Dummy variables).

يمكن الحصول على الاحتمالات المتوقعة من نتائج معادلات الانحدار المتحصل عليها من اي من الطرق المذكورة بأستعمال المعادلة الآتية:

$$P = 1/[1 + \exp(- a - bx)]$$

5 – 4 – 1 العامل المستقل والتابع ثنائي

مثال(78): في دراسة لمعرفة تأثير جنس المريض x وتعليمه (متعلم او امي) y على الحالة الصحية (مريض او سليم) z اخذت عينة عشوائية وسجلت اعدادهم كما موضح ادناه فهل ان للعوامل المستقلة تأثير على الحالة الصحية؟

```
DATA zero;
 INPUT x y z f;
 DATALINES;
1  1  1  8
1  0  1  22
1  1  0  5
1  0  0  14
0  1  1  40
0  0  1  11
0  1  0  1
0  0  0  33
;
PROC GENMOD DATA=zero;
FREQ f;
MODEL y = x z / D=B;
RUN;
```

ان D تعني التوزيع فيما يعني B ثنائي

The SAS System

The GENMOD Procedure

Model Information

Description	Value
Data Set	WORK.ZERO
Distribution	BINOMIAL
Link Function	LOGIT
Dependent Variable	Y
Frequency Weight Variable	F
Observations Used	8
Sum Of Frequency Weights	134
Number Of Events	4
Number Of Trials	8

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	131	138.4165	1.0566

Dr. Firas

```
Scaled Deviance     131     138.4165       1.0566
Pearson Chi-Square  131     181.7842       1.3877
Scaled Pearson X2   131     181.7842       1.3877
Log Likelihood        .     -69.2082           .

        Analysis Of Parameter Estimates
Parameter   DF   Estimate    Std Err   ChiSquare  Pr>Chi

INTERCEPT    1    -1.7359     0.4450    15.2153  0.0001
X            1    -1.2589     0.4464     7.9528  0.0048
Z            1     2.6024     0.5098    26.0574  0.0001
SCALE        0     1.0000     0.0000           .       .

NOTE:  The scale parameter was held fixed.
```

يمكن الحصول على نفس النتائج بأستعمال الايعازات التالية:

```
proc logistic;
weight f;
model y= x z;
run;
proc probit;
weight f;
class y;
model y= x z / d=logistic;
run;
```

5 – 4 – 2 العامل المستقل ثنائي والعامل التابع نسبة

تستعمل ثلاثة طرق في تقدير معادلة الانحداروهي : Logistic و Probit و Genmod.

مثال(79): لغرض دراسة تأثير الجنس والحالة الصحية والحالة الزوجية على نسبة المتقاعدين اخذت عينات من عدة دوائر وكانت البيانات كما موضحة ادناه ، فهل لتلك العوامل تأثير على نسبة المتقاعدين؟

```
DATA working;
 INPUT GENDER HEALTH STATUS total working;
 CARDS;
1  1  1 107  85
1  1  0  65  44
0  0  1  66  24
1  0  0 171  17
0  1  1  87  24
1  1  0  65  22
0  0  1  85   1
0  0  0 148   6
PROC GENMOD DATA=working;
MODEL working/total = GENDER HEALTH STATUS / D=B;
RUN;
```

```
                                The SAS System
          The GENMOD Procedure
           Model Information

     Description              Value

     Data Set                 WORK.WORKING
     Distribution             BINOMIAL
```

```
Link Function                LOGIT
Dependent Variable           WORKING
Dependent Variable           TOTAL
Observations Used            8
Number Of Events             223
Number Of Trials             794

        Criteria For Assessing Goodness Of Fit

Criterion          DF       Value      Value/DF

Deviance            4      66.8950      16.7238
Scaled Deviance     4      66.8950      16.7238
Pearson Chi-Square  4      72.9667      18.2417
Scaled Pearson X2   4      72.9667      18.2417
Log Likelihood      .    -350.7287          .

        Analysis Of Parameter Estimates

Parameter   DF   Estimate    Std Err   ChiSquare  Pr>Chi

INTERCEPT    1    -3.4813     0.2662    171.0197  0.0001
GENDER       1     1.7594     0.2684     42.9531  0.0001
HEALTH       1     1.5260     0.2132     51.2190  0.0001
STATUS       1     1.4593     0.2565     32.3599  0.0001
SCALE        0     1.0000     0.0000           .        .

NOTE:  The scale parameter was held fixed.
```

يمكن الحصول على نفس النتائج بأستعمال ايعازات اخرى مثل:

```
proc logistic;
model working/total = GENDER HEALTH STATUS;
run;

proc probit;
model working/total = GENDER HEALTH STATUS / d=logistic;
run;
```

5 – 4 – 3 العوامل المستقلة والمعتمدة الاسمية ذات الفئات Categorical data

عندما يكون المتغير التابع اسمي وله عدة فئات فلغرض مطابقة النموذج للبيانات نستعمل طريقة Catmod.

مثال (80): البيانات في الجدول ادناه تمثل مجموعة من الاطفال يقيمون في مناطق في الريف او في المدينة وقد خضعوا الى فحص طبي مرتان وسجلت عدد مرات الاصابة (هناك ثلاثة احتمالات مصاب مرتان او مرة واحدة او عدم اصابة) وفقا لكل جنس ضمن كل منطقة المطلوب معرفة هل ان معدل مرات الاصابة لها علاقة بالجنس او الموقع اوكليهما؟

بداية لابد من الاشارة الى عدم وجود ادلة تؤكد ان الاصابة ذات توزيع طبيعي لذا فأن تحليل البيانات سيجرى اعتمادا على طريقة proc catmod بدلا من تحليل التباين باستعمال proc anova او proc glm. كما يمكن اجراء الاختبارات اللامعلمية اذا ماارید المقارنة بين المتوسطات.

	عدد مرات الاصابة بالزكام		

العدد الكلي	2	1	0	مكان الاقامة	الجنس
180	71	64	45	قرية	انثى
300	116	104	80	مدينة	انثى
290	82	124	84	قرية	ذكر
310	87	117	106	مدينة	ذكر

```
data colds;
input sex $ residen $ periods count @@;
cards;
female rural 0 45 female rural 1 64 female rural 2 71
female urban 0 80 female urban 1 104 female urban 2 116
male rural 0 84 male rural 1 124 male rural 2 82
male urban 0 106 male urban 1 117 male urban 2 87
;
PROC FREQ DATA=colds;
WEIGHT count;
TABLES sex*residen / Chisq;
run;
```

```
                The SAS System
          Marginal Symmetry, Saturated Model

             TABLE 1 OF SEX BY PERIODS
            CONTROLLING FOR RESIDEN=rural

          SEX      PERIODS

          Frequency,
          Percent ,
          Row Pct ,
          Col Pct ,     0,     1,     2,  Total
          ÍÍÍÍÍÍÍÍÎÍÍÍÍÍÍÎÍÍÍÍÍÍÎÍÍÍÍÍÍÍ
          female  ,    45 ,    64 ,    71 ,   180
                ,  9.57 , 13.62 , 15.11 , 38.30
                ,  25.00 , 35.56 , 39.44 ,
                ,  34.88 , 34.04 , 46.41 ,
          ÍÍÍÍÍÍÍÍÎÍÍÍÍÍÍÎÍÍÍÍÍÍÎÍÍÍÍÍÍÍ
          male    ,    84 ,   124 ,    82 ,   290
                ,  17.87 , 26.38 , 17.45 , 61.70
                ,  28.97 , 42.76 , 28.28 ,
                ,  65.12 , 65.96 , 53.59 ,
          ÍÍÍÍÍÍÍÍÎÍÍÍÍÍÍÎÍÍÍÍÍÍÎÍÍÍÍÍÍÍ
          Total       129     188     153     470
                     27.45   40.00   32.55   100.00


            STATISTICS FOR TABLE 1 OF SEX BY PERIODS
              CONTROLLING FOR RESIDEN=rural

        Statistic                    DF     Value      Prob
        ÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍ
        Chi-Square                    2    6.333     0.042
        Likelihood Ratio Chi-Square   2    6.272     0.043
        Mantel-Haenszel Chi-Square    1    4.249     0.039
        Phi Coefficient                    0.116
        Contingency Coefficient            0.115
        Cramer's V                         0.116

        Sample Size = 470

                The SAS System
          Marginal Symmetry, Saturated Model
```

TABLE 2 OF SEX BY PERIODS
CONTROLLING FOR RESIDEN=urban

SEX PERIODS

Frequency,
Percent ,
Row Pct ,
Col Pct , 0, 1, 2, Total

female	80	104	116	300
	13.11	17.05	19.02	49.18
	26.67	34.67	38.67	
	43.01	47.06	57.14	
male	106	117	87	310
	17.38	19.18	14.26	50.82
	34.19	37.74	28.06	
	56.99	52.94	42.86	
Total	186	221	203	610
	30.49	36.23	33.28	100.00

STATISTICS FOR TABLE 2 OF SEX BY PERIODS
CONTROLLING FOR RESIDEN=urban

Statistic	DF	Value	Prob
Chi-Square	2	8.380	**0.015**
Likelihood Ratio Chi-Square	2	8.405	0.015
Mantel-Haenszel Chi-Square	1	7.854	0.005
Phi Coefficient		0.117	
Contingency Coefficient		0.116	
Cramer's V		0.117	

Sample Size = 610

قيمتا مربع كاي كانتا معنوية في كلا المنطقتين مما يعني ان هناك علاقة بين عدد مرات الاصابة وجنس الطفل.

اذا اردنا وضع نموذج لوصف معدل عدد مرات الاصابة بالزكام فسوف نستعمل proc catmod وفي البداية سنحاول دراسة تأثير العاملين (الجنس والموقع) والتداخل بينهما بأستعمال الايعاز الآتي:

```
;proc catmod data=colds;
weight count;
response means; ← 0 1 2 القيم means يمكن ان نكتب بدلا عن
model periods = sex residen sex*residen;
run;
```

The SAS System

Marginal Symmetry, Saturated Model
CATMOD PROCEDURE

Response: PERIODS	Response Levels (R)= 3
Weight Variable: COUNT	Populations (S)= 4
Data Set: COLDS	Total Frequency (N)= 1080
Frequency Missing: 0	Observations (Obs)= 12

POPULATION PROFILES

Sample	SEX	RESIDEN	Sample Size

```
                1  female  rural      180
                2  female  urban      300
                3  male    rural      290
                4  male    urban      310

                  Response  PERIODS
                  ffffffffffffffff
                     1     0
                     2     1
                     3     2

          Response         DESIGN MATRIX
        Sample  Function     1     2     3     4
        ffffffffffffffffffffffffffffffffffffffffffffffffffffffff
          1    1.14444     1     1     1     1
          2    1.12000     1     1    -1    -1
          3    0.99310     1    -1     1    -1
          4    0.93871     1    -1    -1     1

                The SAS System
            Marginal Symmetry, Saturated Model

                ANALYSIS-OF-VARIANCE TABLE

        Source              DF   Chi-Square     Prob
        ------------------------------------------------------
        INTERCEPT            1     1841.13   0.0000
        SEX                1       11.57   0.0007
        RESIDEN              1        0.65   0.4202
        SEX*RESIDEN            1        0.09   0.7594

        RESIDUAL             0        .      .

          ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

                              Standard   Chi-
      Effect         Parameter  Estimate    Error   Square   Prob
      ---------------------------------------------------------------
      INTERCEPT              1    1.0491   0.0244  1841.13  0.0000
      SEX                  2   0.0832   0.0244    11.57  0.0007
      RESIDEN                3   0.0197   0.0244     0.65  0.4202
      SEX*RESIDEN              4  -0.00749   0.0244     0.09  0.759
```

يلاحظ من النتائج اعلاه من النموذج الذي استعمل لوصف البيانات ان تأثير التداخل غير معنويا لذا يجب اعادة التحليل واستبعاد تأثير التداخل من النموذج ومعاودة التنفيذ.

```
model periods = sex residen;
run;
```

```
                                                The SAS System
                Marginal Symmetry, Saturated Model

                        CATMOD PROCEDURE

      Response: PERIODS              Response Levels (R)=    3
      Weight Variable: COUNT           Populations    (S)=    4
      Data Set: COLDS                Total Frequency (N)=  1080
      Frequency Missing: 0             Observations  (Obs)=   12

                    POPULATION PROFILES
                              Sample
                    Sample   SEX    RESIDEN    Size
                    ffffffffffffffffffffffffffffffffff
                      1  female  rural      180
                      2  female  urban      300
                      3  male    rural      290
                      4  male    urban      310
```

Dr Firas

RESPONSE PROFILES

Response PERIODS

1	0
2	1
3	2

Response DESIGN MATRIX

Sample	Function			
1	1.14444	1	1	1
2	1.12000	1	1	-1
3	0.99310	1	-1	1
4	0.93871	1	-1	-1

مصفوفة القيم وان قيم كل صف فيها يناظر لوصف العشيرة في جدول POPULATION PROFILES

ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
INTERCEPT	1	1882.77	0.0000
SEX	1	12.08	0.0005
RESIDEN	1	0.76	0.3839
RESIDUAL	1	0.09	**0.7594**

The SAS System

Marginal Symmetry, Saturated Model

ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob
INTERCEPT	1	1.0501	0.0242	1882.77	0.0000
SEX	2	0.0842	0.0242	12.08	0.0005
RESIDEN	3	0.0210	0.0241	0.76	0.3839

ان القيمة الغير معنوية للخطأ في اختبار مربع كاي تعني ان النموذج مطابق للبيانات

The SAS System

PREDICTED VALUES FOR RESPONSE FUNCTIONS

Sample	SEX	RESIDEN	Observed Number	Observed Function	Observed Standard Error	Predicted Function	Predicted Standard Error	Residual
1	female	rural	1	1.14444444	0.0588586	1.15529289	0.0470049	-0.0108484
2	female	urban	1	1.12	0.04614952	1.11333067	0.04068788	0.00666933
3	male	rural	1	0.99310345	0.04442608	0.98692295	0.0395775	0.0061805
4	male	urban	1	0.93870968	0.04467893	0.94496073	0.03974364	-0.0062511

يمكن ان نقدر عدد مرات الاصابة بالزكام من الجدول الرئيسي للبيانات:

معدل عدد مرات الاصابة= 0 x 71/180 + 1 x 64/180 + 1 x 45/180 = 1.144 وهو نفس التقدير عندما ادخلنا تأثير كل العوامل(Response function).

اما اذا اردنا الاعتماد على النموذج الثاني فسوف نستعمل قيم متوسط المربعات الموزونة وقيم المصفوفة وكالآتي:

Dr Firas

معدل عدد مرات الاصابة للأطفال في المناطق الريفية

$= 1.050\ 1.155 = 1 \times 0.0210 + 1 \times 0.0842 + 1 \times$

القيم 1 ، 1 ، 1 اخذت من مصفوفة القيم.

لايمكن استعمال logistic او probit او genmod وذلك لان العامل periods ذو ثلاثة فئات .

مثال (81) : في دراسة لمعرفة تأثير بعض العوامل البيئية على اصابة الافراد بالربو تضمنت العوامل المدروسة البيئة التي يعمل بها الفرد (dusty او nodusty) وعدد سنوات العمل ، هل الفرد يدخن ام لا ، حالته (مريض ام سليم) وعدد الافراد ونظمت البيانات لغرض التحليل كما موضح ادناه:

```
data byss;
  input workplac $ em_years $ smoking $ status $ count ;
  cards ;
  dusty      <10  yes  yes 30
  dusty      <10  yes  no  203
  dusty      <10  no   yes 7
  dusty      <10  no   no  119
  dusty      >=10  yes  yes 57
  dusty      >=10  yes  no  161
  dusty      >=10  no   yes 11
  dusty      >=10  no   no  81
  notdusty    <10  yes  yes 14
  notdusty    <10  yes  no  1340
  notdusty    <10  no   yes 12
  notdusty    <10  no   no  1004
  notdusty   >=10  yes  yes 24
  notdusty   >=10  yes  no  1360
  notdusty   >=10  no   yes 10
  notdusty   >=10  no   no  986
 proc catmod order=data;
    weight count;
    response marginals;
    model status = workplac|em_years|smoking;
    run;
 proc catmod order=data;
     weight count;
     response marginals; ← 1 0 يمكن ان نستعيض عن كلمة marginals بالقيم
     model status = workplac|em_years workplac|smoking
                em_years|smoking;
 run;
  proc catmod order=data;
     weight count ;
     response marginals;
     model status = workplac|em_years workplac|smoking;
  run;
 proc catmod order=data;
     weight count;
     response marginals;
     model status = workplac
                em_years(workplace='dusty')
                em_years(workplace='notdusty')
```

Dr Firas

```
            smoking(workplace='dusty')
            smoking(workplace='notdusty') ;
run;
proc catmod order=data;
  weight count ;
  response marginals;
  model status = workplac
            em_years(workplace='dusty')
            smoking(workplace='dusty') / pred;
run;
```

```
                                        The SAS System
              CATMOD PROCEDURE

     Response: STATUS              Response Levels (R)=    2
     Weight Variable: COUNT           Populations     (S)=    8
     Data Set: BYSS                Total Frequency (N)=  5419
     Frequency Missing: 0            Observations  (Obs)=   16

                POPULATION PROFILES
                               Sample
        Sample  WORKPLAC  EM_YEARS  SMOKING    Size
        ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
           1  dusty      <10      yes       233
           2  dusty      <10      no        126
           3  dusty      >=10     yes       218
           4  dusty      >=10     no         92
           5  notdusty   <10      yes      1354
           6  notdusty   <10      no       1016
           7  notdusty   >=10     yes      1384
           8  notdusty   >=10     no        996
```

يمثل الجدول اعلاه توزيع العينة حسب المتغيرات معبرعنها بحجم العينة اي عدد الافراد.

```
                RESPONSE PROFILES
                Response  STATUS
                ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
                    1    yes
                    2    no

      Response                   DESIGN MATRIX
  Sample   Function   1     2     3     4     5     6     7     8
    ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
   1     0.12876     1     1     1     1     1     1     1     1
   2     0.05556     1     1     1     1    -1    -1    -1    -1
   3     0.26147     1     1    -1    -1     1     1    -1    -1
   4     0.11957     1     1    -1    -1    -1    -1     1     1
   5     0.01034     1    -1     1    -1     1    -1     1    -1
   6     0.01181     1    -1     1    -1    -1     1    -1     1
   7     0.01734     1    -1    -1     1     1    -1    -1     1
   8     0.01004     1    -1    -1     1    -1     1     1    -1


                  The SAS System
              ANALYSIS-OF-VARIANCE TABLE
         Source               DF   Chi-Square     Prob
         --------------------------------------------------
         INTERCEPT             1      127.33    0.0000
         WORKPLAC              1       89.61    0.0000
         EM_YEARS              1       13.74    0.0002
         WORKPLAC*EM_YEARS       1       12.35    0.0004
         SMOKING               1       16.44    0.0001
         WORKPLAC*SMOKING        1       14.75    0.0001
         EM_YEARS*SMOKING        1        2.02    0.1551
         WORKPLA*EM_YEARS*SMOKING  1       1.21    0.2714

         RESIDUAL              0        .     .
```

جدول تحليل التباين ويشمل المتغيرات مع التداخلات فضلا عن اختبار معنويتهااعتمادا على اختبار مربع كاي.

ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

Effect	Parameter	Estimate	Standard Error	Chi Square	Prob
INTERCEPT	1	0.0769	0.00681	127.33	0.0000
WORKPLAC	2	0.0645	0.00681	89.61	0.0000
EM_YEARS	3	-0.0252	0.00681	13.74	0.0002
WORKPLAC*EM_YEARS	4	-0.0239	0.00681	12.35	0.0004
SMOKING	5	0.0276	0.00681	16.44	0.0001
WORKPLAC*SMOKING	6	0.0262	0.00681	14.75	0.0001
EM_YEARS*SMOKING	7	-0.00968	0.00681	2.02	0.1551
WORKPLA*EM_YEARS*SMOKING	8	-0.00749	0.00681	1.21	0.2714

تقديرات المربعات الصغرى الموزونة لجميع العوامل والتداخلات وسيتم في الخطوة اللاحقة استبعاد تأثير العامل الأخير لكون تأثيره هو الاقل بالنسبة لبقية العوامل

The SAS System
CATMOD PROCEDURE

Response: STATUS Response Levels (R)= 2
Weight Variable: COUNT Populations (S)= 8
Data Set: BYSS Total Frequency (N)= 5419
Frequency Missing: 0 Observations (Obs)= 16

POPULATION PROFILES

Sample	WORKPLAC	EM_YEARS	SMOKING	Sample Size
1	dusty	<10	yes	233
2	dusty	<10	no	126
3	dusty	>=10	yes	218
4	dusty	>=10	no	92
5	notdusty	<10	yes	1354
6	notdusty	<10	no	1016
7	notdusty	>=10	yes	1384
8	notdusty	>=10	no	996

RESPONSE PROFILES

Response	STATUS
1	yes
2	no

Sample	Response Function	DESIGN MATRIX 1	2	3	4	5	6	7
1	0.12876	1	1	1	1	1	1	1
2	0.05556	1	1	1	1	-1	-1	-1
3	0.26147	1	1	-1	-1	1	1	-1
4	0.11957	1	1	-1	-1	-1	-1	1
5	0.01034	1	-1	1	-1	1	-1	1
6	0.01181	1	-1	1	-1	-1	1	-1
7	0.01734	1	-1	-1	1	1	-1	-1
8	0.01004	1	-1	-1	1	-1	1	1

The SAS System
ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
INTERCEPT	1	131.73	0.0000
WORKPLAC	1	92.90	0.0000
EM_YEARS	1	14.33	0.0002
WORKPLAC*EM_YEARS	1	12.92	0.0003
SMOKING	1	15.46	0.0001
WORKPLAC*SMOKING	1	13.70	0.0002

Dr.Firas

EM_YEARS*SMOKING 1 2.26 0.1324

RESIDUAL 1 1.21 0.2714

ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob
INTERCEPT	1	0.0777	0.00677	131.73	0.0000
WORKPLAC	2	0.0653	0.00677	92.90	0.0000
EM_YEARS	3	-0.0257	0.00680	14.33	0.0002
WORKPLAC*EM_YEARS	4	-0.0244	0.00680	12.92	0.0003
SMOKING	5	0.0248	0.00630	15.46	0.0001
WORKPLAC*SMOKING	6	0.0233	0.00629	13.70	0.0002
EM_YEARS*SMOKING	7	-0.00240	0.00160	2.26	0.1324

التقديرات بعد حذف العامل الأخير وسيجري حذف العامل الاخير هنا بسبب عدم معنويته.

The SAS System
CATMOD PROCEDURE

Response: STATUS Response Levels (R)= 2
Weight Variable: COUNT Populations (S)= 8
Data Set: BYSS Total Frequency (N)= 5419
Frequency Missing: 0 Observations (Obs)= 16

POPULATION PROFILES

Sample	WORKPLAC	EM_YEARS	SMOKING	Sample Size
1	dusty	<10	yes	233
2	dusty	<10	no	126
3	dusty	>=10	yes	218
4	dusty	>=10	no	92
5	notdusty	<10	yes	1354
6	notdusty	<10	no	1016
7	notdusty	>=10	yes	1384
8	notdusty	>=10	no	996

RESPONSE PROFILES

Response	STATUS
1	yes
2	no

Sample	Response Function	DESIGN MATRIX 1	2	3	4	5	6
1	0.12876	1	1	1	1	1	1
2	0.05556	1	1	1	1	-1	-1
3	0.26147	1	1	-1	-1	1	1
4	0.11957	1	1	-1	-1	-1	-1
5	0.01034	1	-1	1	-1	1	-1
6	0.01181	1	-1	1	-1	-1	1
7	0.01734	1	-1	-1	1	1	-1
8	0.01004	1	-1	-1	1	-1	1

The SAS System
ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
INTERCEPT	1	131.56	0.0000
WORKPLAC	1	93.89	0.0000
EM_YEARS	1	14.47	0.0001
WORKPLAC*EM_YEARS	1	12.95	0.0003
SMOKING	1	14.82	0.0001
WORKPLAC*SMOKING	1	13.29	0.0003

Dr. Firas

RESIDUAL 2 3.47 0.1761

ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob
INTERCEPT	1	0.0776	0.00677	131.56	0.0000
WORKPLAC	2	0.0656	0.00677	93.89	0.0000
EM_YEARS	3	-0.0259	0.00680	14.47	0.0001
WORKPLAC*EM_YEARS	4	-0.0245	0.00680	12.95	0.0003
SMOKING	5	0.0242	0.00629	14.82	0.0001
WORKPLAC*SMOKING	6	0.0229	0.00629	13.29	0.0003

التقديرات بعد حذف العامل الأخير.

وبسبب معنوية التداخلات بين عدد السنوات ومكان العمل ومكان العمل والتدخين سنجري تقسيم لمكان العمل حسب السنوات اضافة الى تقسيم مكان العمل حسب حالة الشخص (يدخن ام لا) لتحديد تأثير كل منهم..

The SAS System
CATMOD PROCEDURE

Response: STATUS Response Levels (R)= 2
Weight Variable: COUNT Populations (S)= 8
Data Set: BYSS Total Frequency (N)= 5419
Frequency Missing: 0 Observations (Obs)= 16

POPULATION PROFILES

Sample	WORKPLAC	EM_YEARS	SMOKING	Sample Size
1	dusty	<10	yes	233
2	dusty	<10	no	126
3	dusty	>=10	yes	218
4	dusty	>=10	no	92
5	notdusty	<10	yes	1354
6	notdusty	<10	no	1016
7	notdusty	>=10	yes	1384
8	notdusty	>=10	no	996

RESPONSE PROFILES

Response	STATUS
1	yes
2	no

Sample	Response Function	DESIGN MATRIX 1	2	3	4	5	6
1	0.12876	1	1	1	0	1	0
2	0.05556	1	1	1	0	-1	0
3	0.26147	1	1	-1	0	1	0
4	0.11957	1	1	-1	0	-1	0
5	0.01034	1	-1	0	1	0	1
6	0.01181	1	-1	0	1	0	-1
7	0.01734	1	-1	0	-1	0	1
8	0.01004	1	-1	0	-1	0	-1

The SAS System
ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
INTERCEPT	1	131.56	0.0000
WORKPLAC	1	93.89	0.0000
EM_YEARS(WORKPLAC=dusty)	1	13.90	0.0002
EM_YEA(WORKPLA=notdusty)	1	0.75	0.3860
SMOKING(WORKPLAC=dusty)	1	14.27	0.0002

Dr Firas

SMOKIN(WORKPLA=notdusty) 1 0.64 0.4225

RESIDUAL 2 3.47 0.1761

ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob
INTERCEPT	1	0.0776	0.00677	131.56	0.0000
WORKPLAC	2	0.0656	0.00677	93.89	0.0000
EM_YEARS(WORKPLAC=dusty)	3	-0.0503	0.0135	13.90	0.0002
EM_YEA(WORKPLA=notdusty)	4	-0.00139	0.00161	0.75	0.3860
SMOKING(WORKPLAC=dusty)	5	0.0471	0.0125	14.27	0.0002
SMOKIN(WORKPLA=notdusty)	6	0.00129	0.00160	0.64	0.4225

يتبين من تقديرات الجدول اعلاه ان مكان العمل لكل من السنوات وحالة الفرد من التدخين غير معنوية لذا سيتم استبعادهما من التقديرات.

The SAS System
CATMOD PROCEDURE

Response: STATUS Response Levels (R)= 2
Weight Variable: COUNT Populations (S)= 8
Data Set: BYSS Total Frequency (N)= 5419
Frequency Missing: 0 Observations (Obs)= 16

POPULATION PROFILES

Sample	WORKPLAC	EM_YEARS	SMOKING	Sample Size
1	dusty	<10	yes	233
2	dusty	<10	no	126
3	dusty	>=10	yes	218
4	dusty	>=10	no	92
5	notdusty	<10	yes	1354
6	notdusty	<10	no	1016
7	notdusty	>=10	yes	1384
8	notdusty	>=10	no	996

RESPONSE PROFILES

Response	STATUS
1	yes
2	no

Sample	Response Function	DESIGN MATRIX 1	2	3	4
1	0.12876	1	1	1	1
2	0.05556	1	1	1	-1
3	0.26147	1	1	-1	1
4	0.11957	1	1	-1	-1
5	0.01034	1	-1	0	0
6	0.01181	1	-1	0	0
7	0.01734	1	-1	0	0
8	0.01004	1	-1	0	0

مصفوفة قيم المتغيرات

The SAS System
ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
INTERCEPT	1	131.50	0.0000
WORKPLAC	1	93.98	0.0000
EM_YEARS(WORKPLAC=dusty)	1	13.90	0.0002
SMOKING(WORKPLAC=dusty)	1	14.27	0.0002
RESIDUAL	4	4.68	0.3215

Dr Firas

ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob
INTERCEPT	1	0.0776	0.00677	131.50	0.0000
WORKPLAC	2	0.0656	0.00677	93.98	0.0000
EM_YEARS(WORKPLAC=dusty)	3	-0.0503	0.0135	13.90	0.0002
SMOKING(WORKPLAC=dusty)	4	0.0471	0.0125	14.27	0.0002

التقديرات بعد حذف تأثير العوامل غير المؤثرة وهنا ستكون معادلة التوقع كالآتي:

$$\hat{Y} = 0.0776b1 + 0.0656b2 - 0.0503b3 + 0.0471b4$$

.

The SAS System
PREDICTED VALUES FOR RESPONSE FUNCTIONS

Sample	Function Number	Observed Function	Observed Standard Error	Predicted Function	Predicted Standard Error	Residual
1	1	0.12875536	0.02194193	0.14005252	0.02005695	-0.0112972
2	1	0.05555556	0.02040641	0.04578424	0.01889963	0.00977131
3	1	0.26146789	0.02976224	0.24068285	0.02485574	0.02078504
4	1	0.11956522	0.0338265	0.14641457	0.02640193	-0.0268494
5	1	0.01033973	0.00274909	0.01200316	0.00157957	-0.0016634
6	1	0.01181102	0.00338935	0.01200316	0.00157957	-0.0001921
7	1	0.01734104	0.0035089	0.01200316	0.00157957	0.00533788
8	1	0.01004016	0.003159	0.01200316	0.00157957	-0.001963

التقديرات في العمود الثالث من اليسار (function) تمثل نسبة عدد الافراد المرضى الى عدد الافراد الكلي اي ان 233/30= 0.1287 و 126/7= 0.0556 و 218/57= 0.261 وهكذا لبقية التقديرات. اما القيم المتوقعة فيمكن الحصول عليها من معادلة التوقع المذكورة في الجدول قبل الاخير وكما يلي:

العينة الأولى القيمة المشاهدة 0.1287 اما المتوقعة فهي:

$$\hat{Y} = 0.0776\ (1) + 0.0656(1) - 0.0503(1) + 0.0471(1) = 0.140$$

قيم b1 ، b2 ، b3 ، b4 يمكن الحصول عليها من مصفوفة قيم المتغيرات.

العينة الثالثة القيمة المشاهدة 0.261 اما المتوقعة فهي:

$$\hat{Y} = 0.0776(1) + 0.0656(1) - 0.0503\ (-1) + 0.0471(1) = 0.2406$$

العينة السادسة القيمة المشاهدة 0.011 اما المتوقعة فهي:

$$\hat{Y} = 0.0776\ (1) + 0.0656(1) - 0.0503\ (0) + 0.0471(0) = 0.012$$

مثال (82): في دراسة لتحديد دالة الاستجابة لماركتين تجارية خاصة بالملابس هما M و X من خلال دراسة تأثير ثلاثة متغيرات ذات فئات (categorical variables) المتغير الاول درجة رائقية الماء المستعمل في غسلها (softness) اذ قسم الى soft و med و hard والمتغير الثاني درجة حرارة ماء الغسيل (high او low) والمتغير الثالث هل الشخص المستعمل للماركة سبق وان استعملها (نعم او لا) وسنحاول استخدام نموذجين الاول ويتضمن جميع العوامل والثاني هوالنموذج الذي يضم العوامل المؤثرة فقط؟

Dr. Firas

بما ان المتغيرات اسمية وذات ترتيب وتضم اكثر من فئتين ضمن احد العوامل فلابد من استعمال طريقة catmod .

ملاحظة : يمكن استعمال كلمة marginals بعد كلمة response بدلا من 0 1 المطبقة في حل هذا المثال.

```
data detergen;
    input Softness $ Brand $ Previous $ Temp $ Count @@;
    datalines;
  soft X yes high 19   soft X yes low 57
  soft X no  high 29   soft X no  low 63
  soft M yes high 29   soft M yes low 49
  soft M no  high 27   soft M no  low 53
  med  X yes high 23   med  X yes low 47
  med  X no  high 33   med  X no  low 66
  med  M yes high 47   med  M yes low 55
  med  M no  high 23   med  M no  low 50
  hard X yes high 24   hard X yes low 37
  hard X no  high 42   hard X no  low 68
  hard M yes high 43   hard M yes low 52
  hard M no  high 30   hard M no  low 42
  proc catmod data=detergen;
    response 1 0;
    weight Count;
   model Brand=Softness|Previous|Temp;
  run;
   model Brand=Softness Previous Temp/  noprofile pred;
  run;
```

```
                    The SAS System
                    CATMOD PROCEDURE

        Response: BRAND              Response Levels (R)=    2
        Weight Variable: COUNT         Populations    (S)=   12
        Data Set: DETERGEN           Total Frequency (N)=  1008
        Frequency Missing: 0         Observations  (Obs)=   24
```

يتضح من الجدول ان هناك مستويان للاستجابة و12 عشيرة.

```
                POPULATION PROFILES
                                Sample
            Sample  SOFTNESS  PREVIOUS  TEMP    Size
            ffffffffffffffffffffffffffffffffffffffffffff
               1    hard      no      high      72
               2    hard      no      low      110
               3    hard      yes     high      67
               4    hard      yes     low       89
               5    med       no      high      56
               6    med       no      low      116
               7    med       yes     high      70
               8    med       yes     low      102
               9    soft      no      high      56
              10    soft      no      low      116
              11    soft      yes     high      48
              12    soft      yes     low      106
```

يتضح من الجدول ترتيب مستويات العوامل المستقلة اذ ان حجم العينة 72 يمثل (hard M no high 30 + Hard X no high 42) وحجم العينة 110يمثل (hard X no low + hard M no low 42) وهكذا لبقية العينات.

```
                RESPONSE PROFILES
                Response  BRAND
```

Dr. Firas

ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
1 M
2 X

The SAS System
RESPONSE FREQUENCIES

	Response Number	
Sample	1	2
1	30	42
2	42	68
3	43	24
4	52	37
5	23	33
6	50	66
7	47	23
8	55	47
9	27	29
10	53	63
11	29	19
12	49	57

RESPONSE PROBABILITIES

	Response Number	
Sample	1	2
1	0.41667	0.58333
2	0.38182	0.61818
3	0.64179	0.35821
4	0.58427	0.41573
5	0.41071	0.58929
6	0.43103	0.56897
7	0.67143	0.32857
8	0.53922	0.46078
9	0.48214	0.51786
10	0.4569	0.5431
11	0.60417	0.39583
12	0.46226	0.53774

جدول الاستجابة والتكرارات والاحتمالات. ونلاحظ ان القيمة 41.667 للعينة الاولى هي ناتج قسمة 30 / (30 + 42) والقيمة الثانية 0.58333 هي ناتج قسمة 42/(30+ 42). وهكذا لبقية التقديرات.

The SAS System

	Response	DESIGN MATRIX											
Sample	Function	1	2	3	4	5	6	7	8	9	10	11	12
1	0.41667	1	1	0	1	1	0	1	1	0	1	1	0
2	0.38182	1	1	0	1	1	0	-1	-1	0	-1	-1	0
3	0.64179	1	1	0	-1	-1	0	1	1	0	-1	-1	0
4	0.58427	1	1	0	-1	-1	0	-1	-1	0	1	1	0
5	0.41071	1	0	1	1	0	1	1	0	1	1	0	1
6	0.43103	1	0	1	1	0	1	-1	0	-1	-1	0	-1
7	0.67143	1	0	1	-1	0	-1	1	0	1	-1	0	-1
8	0.53922	1	0	1	-1	0	-1	-1	0	-1	1	0	1
9	0.48214	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	-1
10	0.45690	1	-1	-1	1	-1	-1	-1	1	1	-1	1	1
11	0.60417	1	-1	-1	-1	1	1	1	-1	-1	-1	1	1
12	0.46226	1	-1	-1	-1	1	1	-1	1	1	1	-1	-1

ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
INTERCEPT	1	983.13	0.0000
SOFTNESS	2	0.09	0.9575
PREVIOUS	1	22.68	0.0000
SOFTNESS*PREVIOUS	2	3.85	0.1457
TEMP	1	3.67	0.0555
SOFTNESS*TEMP	2	0.23	0.8914
PREVIOUS*TEMP	1	2.26	0.1324

Dr Firas

SOFTNESS*PREVIOUS*TEMP 2 0.76 0.6850

RESIDUAL 0 . .

ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob
INTERCEPT	1	0.5069	0.0162	983.13	0.0000
SOFTNESS	2	-0.00073	0.0225	0.00	0.9740
	3	0.00623	0.0226	0.08	0.7830
PREVIOUS	4	-0.0770	0.0162	22.68	0.0000
SOFTNESS*PREVIOUS	5	-0.0299	0.0225	1.77	0.1831
	6	-0.0152	0.0226	0.45	0.5007
TEMP	7	0.0310	0.0162	3.67	0.0555
SOFTNESS*TEMP	8	-0.00786	0.0225	0.12	0.7265
	9	-0.00298	0.0226	0.02	0.8953
PREVIOUS*TEMP	10	-0.0243	0.0162	2.26	0.1324
SOFTNESS*PREVIOUS*TEMP	11	0.0187	0.0225	0.69	0.4064
	12	-0.0138	0.0226	0.37	0.5415

جدول تحليل التباين ويتضح منه عدم معنوية جميع التداخلات. لهذا سيتم استبعادها من التحليل وفقا للايعاز الثاني في قطعة البرنامج. ونلاحظ استعمال الأمر noprofile لعدم الحاجة اليه .

The SAS System
CATMOD PROCEDURE

Response: BRAND Response Levels (R)= 2
Weight Variable: COUNT Populations (S)= 12
Data Set: DETERGEN Total Frequency (N)= 1008
Frequency Missing: 0 Observations (Obs)= 24

Sample	Response Function	DESIGN MATRIX 1	2	3	4	5
1	0.41667	1	1	0	1	1
2	0.38182	1	1	0	1	-1
3	0.64179	1	1	0	-1	1
4	0.58427	1	1	0	-1	-1
5	0.41071	1	0	1	1	1
6	0.43103	1	0	1	1	-1
7	0.67143	1	0	1	-1	1
8	0.53922	1	0	1	-1	-1
9	0.48214	1	-1	-1	1	1
10	0.45690	1	-1	-1	1	-1
11	0.60417	1	-1	-1	-1	1
12	0.46226	1	-1	-1	-1	-1

مصفوفة العوامل الرئيسية.

ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob
INTERCEPT	1	0.5080	0.0160	1004.93	0.0000
SOFTNESS	2	-0.00256	0.0218	0.01	0.9066
	3	0.0104	0.0218	0.23	0.6342
PREVIOUS	4	-0.0711	0.0155	20.96	0.0000
TEMP	5	0.0319	0.0161	3.95	0.0468

يتضح من جدول تحليل التباين بأن العاملين previous و Tem لهما تأثير معنوي في تفضيل العلامة m وذلك لأن تسلسل العلامة m هوالاول في جدول الاستجابة . كما يتضح ان النموذج المستعمل ملائم لوصف البيانات بسبب عدم معنوية قيمة مربع كاي الخاصة بالخطأ. اما تأثير رائقية الماء فهي لاتختلف معنويا عن الصفر مما يعني عدم اهميتها. ويمكن ان نستنتج من النتائج بأن الاشخاص الذين سبق وان استعملوا العلامة m هم اكثر تفضيلا لها مقارنة بالاشخاص الذين لم يسبق وان استعملوها. كما نستنتج بأن الذين يفضلون غسل ملابسهم بالماء الساخن

Dr. Firas

(لأن الماء الساخن هو في المرتبة الاولى في جدول مظهر العشيرة population profile) هم اكثر تفضيلا للعلامة m مقارنة بالذين يفضلون غسل ملابسهم بالماء البارد.

The SAS System
PREDICTED VALUES FOR RESPONSE FUNCTIONS

Sample	Function Number	Observed Function	Observed Standard Error	Predicted Function	Predicted Standard Error	Residual
1	1	0.41666667	0.05810139	0.4662642	0.03600714	-0.0495975
2	1	0.38181818	0.0463223	0.402375	0.03265923	-0.0205568
3	1	0.64179104	0.05857707	0.60844093	0.03608142	0.03335012
4	1	0.58426966	0.05224173	0.54455172	0.03440696	0.03971794
5	1	0.41071429	0.06574138	0.47918012	0.03708853	-0.0684658
6	1	0.43103448	0.04598011	0.41529092	0.03282036	0.01574357
7	1	0.67142857	0.05613914	0.62135685	0.03592718	0.05007172
8	1	0.53921569	0.04935487	0.55746764	0.03323007	-0.018252
9	1	0.48214286	0.06677269	0.4610334	0.03856058	0.02110946
10	1	0.45689655	0.04625101	0.39714419	0.03265629	0.05975236
11	1	0.60416667	0.07058525	0.60321012	0.03818853	0.00095655
12	1	0.46226415	0.04842579	0.53932091	0.03390778	-0.0770568

يتضح من الجدول القيم المشاهدة لدالة الاستجابة والقيم المتوقعة بأستعمال معادلة التنبوء المتحصل عليها من النتائج وهي :

$\hat{Y}$ = 0.5080-0.00256(hard)+0.0104(med)-0.0711(no)+0.0319(high)

فمثلا لو اردنا حساب الاحتمالية المتوقعة للعينة الاولى والتي بلغت احتماليتها المحسوبة 0.41667 من تطبيق معادلة التنبوء وبالاعتماد على قيم المصفوفة :

$\hat{Y}$ = 0.5080*1 - 0.00256*1 + 0.0104*0 + 0.0711*1 + 0.0319*1= 0.4662.

وبنفس الطريقة يمكن تقدير القيم الاخرى.

يمكن استعمال طريقة catmod لغرض اجراء اختبار Bhapkar الخاص بتجانس استجابة مجموعتين معبر عنها كمتغيرذو فئات مرتبة.

مثال (83): دراسة عن مجموعة من النساء بأعمار تتراوح من 30 – 39 سنة حول درجة قصرالنظر في العين اليمنى واليسرى المطلوب اختبار تجانس درجات القصر بينهما او بمعنى هل ان توزيع درجات حالات العين اليمنى تساوي التوزيع في العين اليسرى؟

هنا لايستعمل اختبار مربع كاي لاننا لا نريد دراسة العلاقة بين حالات العين اليمنى او اليسرى ولانريد ان نختبر ان توزيع الحالات يتبع نسب معينة لذا نستعمل اختبار Shapkar.

```
data Vision;
input r c count @@;
cards;
1 1 1520 1 2 266 1 3 124 1 4 66 2 1 234 2 2 1512 2 3 432 2 4 78 3 1 117
3 2 362 3 3 1772 3 4 205 4 1 36 4 2 82 4 3 179 4 4 492
;
proc catmod data = Vision;
weight count;
response marginals;
model R * C = _response_ / oneway;
repeated vis 2 / _response_= vis;
run;
```

ملاحظة: لايمكن استعمال قيم بدلا عن marginals وذلك لوجود تكرارات فقط .

The SAS System
CATMOD PROCEDURE

Response: R*C Response Levels (R)= 16
Weight Variable: COUNT Populations (S)= 1
Data Set: VISION Total Frequency (N)= 7477
Frequency Missing: 0 Observations (Obs)= 16

ONE-WAY FREQUENCIES

Variable	Value	Frequency
R	1	1976
	2	2256
	3	2456
	4	789
C	1	1907
	2	2222
	3	2507
	4	841

Sample	Sample Size
1	7477

The SAS System
RESPONSE PROFILES

Response	R	C
1	1	1
2	1	2
3	1	3
4	1	4
5	2	1
6	2	2
7	2	3
8	2	4
9	3	1
10	3	2
11	3	3
12	3	4
13	4	1
14	4	2
15	4	3
16	4	4

Sample	Function Number	Response Function	DESIGN MATRIX 1	2	3	4	5	6
1	1	0.26428	1	0	0	1	0	0
	2	0.30173	0	1	0	0	1	0
	3	0.32847	0	0	1	0	0	1
	4	0.25505	1	0	0	-1	0	0
	5	0.29718	0	1	0	0	-1	0
	6	0.33529	0	0	1	0	0	-1

Dr. Firas
Dr. Firas

ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
INTERCEPT	3	78744.17	0.0000
VIS	3	11.98	0.0075
RESIDUAL	0	.	.

The SAS System

ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob
INTERCEPT	1	0.2597	0.00468	3073.03	0.0000
	2	0.2995	0.00464	4160.17	0.0000
	3	0.3319	0.00483	4725.25	0.0000
VIS	4	0.00461	0.00194	5.65	0.0174
	5	0.00227	0.00255	0.80	0.3726
	6	-0.00341	0.00252	1.83	0.1757

وهذا الاختبار يكون ذو نتائج مطابقة لاختبار ماكنيمار العام الخاص بالمتغيرات ذات الفئات (اكثر من فئتين) علما لاتوجد طريقة حاليا في برنامج ساس للاختبار الاخير لذا يمكن اجراء اختبار Bhakar بدلا عنه لان نتائجهما تكون متقاربة.اما اذا كان المتغير ذو فئتين فيمكن اجراء اختبار ماكنيمار اوباكر اذ تكون نتائجهما متقاربة.

مثال (84): في دراسة شملت 86 مريضا تناولوا عقار لمدة ثلاثة اشهر اختبر تجانس شكلي التوزيع للمجموعة قبل وبعد العلاج بأستعمال اختبار ماكنيمار ؟

	بعد العلاج		
قبل العلاج	كلا	نعم	المجموع الكلي
كلا	60	14	74
نعم	6	6	12
المجموع الكلي	86	20	86

```
data Abnormal;
input r c count @@;
cards;
1 1 60 1 2 14 2 1 6 2 2 6
;
proc freq data = Abnormal;
weight count;
tables r*c/agree;
run;

proc catmod data = Abnormal;
weight count;
response marginals;
model R * C = _response_ / oneway;
repeated abnor 2 / _response_= abnor;
```

Dr. Firas

run;

The SAS System
TABLE OF R BY C

R C

Frequency,
Percent ,
Row Pct ,
Col Pct , 1, 2, Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
1 , 60 , 14 , 74
, 69.77 , 16.28 , 86.05
, 81.08 , 18.92 ,
, 90.91 , 70.00 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
2 , 6 , 6 , 12
, 6.98 , 6.98 , 13.95
, 50.00 , 50.00 ,
, 9.09 , 30.00 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 66 20 86
76.74 23.26 100.00

STATISTICS FOR TABLE OF R BY C

McNemar's Test

Statistic = 3.200 DF = 1 Prob = 0.074

Simple Kappa Coefficient

95% Confidence Bounds
Kappa = 0.243 ASE = 0.121 0.006 0.480

Sample Size = 86

The SAS System
CATMOD PROCEDURE

Response: R*C Response Levels (R)= 4
Weight Variable: COUNT Populations (S)= 1
Data Set: ABNORMAL Total Frequency (N)= 86
Frequency Missing: 0 Observations (Obs)= 4

ONE-WAY FREQUENCIES

Variable Value Frequency

R 1 74
2 12

C 1 66
2 20

Sample
Sample Size
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
1 86

RESPONSE PROFILES

Response R C
ƒƒƒƒƒƒƒƒƒƒƒƒ
1 1 1
2 1 2

Dr Firas

```
   3   2 1
   4   2 2

          The SAS System
                DESIGN
     Function   Response    MATRIX
  Sample  Number   Function    1     2
  ffffffffffffffffffffffffffffffffffffffffffffffffffff
    1       1    0.86047     1     1
            2    0.76744     1    -1


        ANALYSIS-OF-VARIANCE TABLE

  Source              DF  Chi-Square     Prob
  ---------------------------------------------------
  INTERCEPT            1      610.72   0.0000
  ABNOR                1        3.32   0.0683

  RESIDUAL             0           .        .


      ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

                          Standard   Chi-
  Effect        Parameter Estimate   Error   Square  Prob
  ------------------------------------------------------------------
  INTERCEPT             1   0.8140   0.0329   610.72 0.0000
  ABNOR                 2   0.0465   0.0255     3.32 0.0683
```

نلاحظ ان قيمة مربع كاي في الاختبارين هما 3.2 و 3.3 وكلاهما غير معنوي اي ان التجانس قبل وبعد العلاج معنوي او بمعنى اخر ان التوزيع الغير طبيعي لقبل العلاج لايختلف عن التوزيع الغير طبيعي بعد العلاج.

مثال (85): في دراسة عن 88 مريض للتعرف على ان شهيتهم لتناول المواد الغذائية الغنية بالدهون قبل وبعد اسبوعين من استعمال علاج غذائي تجريبي وسجلت درجة شهيتهم الى (ابدا ، احيانا ، دائما) المطلوب معرفة هل ان العلاج الغذائي له تأثير على تكرار درجة الشهية؟

	بعد اسبوعين			
قبل الدراسة	ابدا	احيانا	دائما	المجموع
ابدا	14	6	4	24
احيانا	9	17	2	28
دائما	6	12	8	26
المجموع	21	33	34	88

```
data Symptom;
input r c count @@;
cards;
1 1 14 1 2 6 1 3 4 2 1 9 2 2 17
2 3 2 3 1 6 3 2 12 3 3 8
;

proc catmod data=Symptom;
weight count;
response marginals;
model R * C = _response_ / oneway;
```

```
repeated symp 2 / _response_= symp;
run;
```

```
                        The SAS System
                         CATMOD PROCEDURE

 Response: R*C                    Response Levels (R)=    9
 Weight Variable: COUNT                Populations     (S)=    1
 Data Set: SYMPTOM                  Total Frequency (N)=   78
 Frequency Missing: 0              Observations  (Obs)=    9

                        ONE-WAY FREQUENCIES

                    Variable  Value  Frequency
                    --------------------------
                    R       1       24
                            2       28
                            3       26

                    C       1       29
                            2       35
                            3       14


                             Sample
                       Sample    Size
                       ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
                          1       78


                        RESPONSE PROFILES

                         Response  R  C
                         ƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
                            1   1 1
                            2   1 2
                            3   1 3
                            4   2 1
                            5   2 2
                            6   2 3
                            7   3 1
                            8   3 2
                            9   3 3


                         The SAS System
           Function    Response        DESIGN MATRIX
   Sample   Number     Function     1      2      3      4
   ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
     1        1     0.30769     1      0      1      0
              2     0.35897     0      1      0      1
              3     0.37179     1      0     -1      0
              4     0.44872     0      1      0     -1


                    ANALYSIS-OF-VARIANCE TABLE

      Source              DF   Chi-Square     Prob
      --------------------------------------------------
      INTERCEPT             2      379.45    0.0000
      SYMP                  2        6.50    0.0388

      RESIDUAL              0         .       .

          ANALYSIS OF WEIGHTED-LEAST-SQUARES ESTIMATES

                               Standard    Chi-
 Effect          Parameter  Estimate    Error    Square   Prob
 ----------------------------------------------------------------
 INTERCEPT              1    0.3397    0.0430    62.44  0.0000
```

```
                2   0.4038   0.0435   86.07  0.0000
SYMP            3  -0.0321   0.0318    1.01  0.3142
                4  -0.0449   0.0341    1.73  0.1888
```

5 – 4 – 4 العامل المعتمد يكون بهيئة نسبة

في بعض الحالات يريد الباحث وضع معادلة انحدار لمتغير ما يكون بهيئة نسبة (مثل عدد الذين تحسنت حالتهم الى مجموع المرضى الكلي او عدد الناجحين الى العدد الكلي).

مثال(86): البيانات ادناه تمثل متوسط العمر medage وعدد المرضى الذين تحسنت حالتهم الصحية improve والعدد الكلي للمرضى total.

```
data maoi;
input medage improve total;
cards;
30 3  27
34 3  21
38 6  32
42 6  19
46 11 26
50 12 25
54 10 20
58 12 23
62 12 17
proc logistic;
model improve/total=medage;
run;

proc genmod;
model improve/total=medage /D=B type3;
run;
```

```
                The SAS System
                The LOGISTIC Procedure

Data Set: WORK.MAOI
Response Variable (Events): IMPROVE
Response Variable (Trials): TOTAL
Number of Observations: 9
Link Function: Logit
                  Response Profile

              Ordered  Binary
               Value  Outcome     Count

                  1  EVENT         75
                  2  NO EVENT      135

  Model Fitting Information and Testing Global Null Hypothesis BETA=0

                    Intercept
          Intercept      and
Criterion      Only      Covariates    Chi-Square for Covariates

AIC           275.738      246.693       .
SC            279.085      253.388       .
-2 LOG L      273.738      242.693      31.044 with 1 DF (p=0.0001)
Score            .            .         29.708 with 1 DF (p=0.0001)
```

Dr. Firas

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-4.5303	0.7955	32.4342	0.0001	.	.
MEDAGE	1	0.0854	0.0165	26.6993	0.0001	0.476713	1.089

Association of Predicted Probabilities and Observed Responses

Concordant = 67.7% Somers' D = 0.450
Discordant = 22.7% Gamma = 0.498
Tied = 9.6% Tau-a = 0.208
(10125 pairs) c = 0.725

The SAS System
The GENMOD Procedure

Model Information

Description	Value
Data Set	WORK.MAOI
Distribution	BINOMIAL
Link Function	LOGIT
Dependent Variable	IMPROVE
Dependent Variable	TOTAL
Observations Used	9
Number Of Events	75
Number Of Trials	210

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	7	1.8404	0.2629
Scaled Deviance	7	1.8404	0.2629
Pearson Chi-Square	7	1.8606	0.2658
Scaled Pearson X2	7	1.8606	0.2658
Log Likelihood	.	-121.3467	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	1	-4.5303	0.7955	32.4339	0.0001
MEDAGE	1	0.0854	0.0165	26.6991	0.0001
SCALE	0	1.0000	0.0000	.	.

NOTE: The scale parameter was held fixed.

LR Statistics For Type 3 Analysis

Source	DF	ChiSquare	Pr>Chi
MEDAGE	1	31.0443	0.0001

يلاحظ ان تأثير العمر معنويا اذ ان زيادة العمر سنة واحدة يؤدي الى زيادة نسبة الشفاء بمقدار 9%.

يمكن الحصول على نفس النتائج باستعمال الايعاز التالي:

```
proc probit;
 model improve/total=medage / d=logistic;  run;
```

Dr Firas

مثال (87): البيانات ادناه تمثل مجموعة من الاشخاص باعمار مختلفة ومن كلا الجنسين اختلفوا في رغبتهم في الاشتراك بأحد الصحف وقد عبر عن ذلك 1= يرغب او 0= لايرغب المطلوب وضع معادلة خطية لوصف المتغير subs باستعمال عدة طرق؟

```
data news;
    input sex $ age subs;
    datalines;
  Female    35   0
  Male      44   0
  Male      45   1
  Female    47   1
  Female    51   0
  Female    47   0
  Male      54   1
  Male      47   1
  Female    35   0
  Female    34   0
  Female    48   0
  Female    56   1
  Male      46   1
  Female    59   1
  Female    46   1
  Male      59   1
  Male      38   1
  Female    39   0
  Male      49   1
  Male      42   1
  Male      50   1
  Female    45   0
  Female    47   0
  Female    30   1
  Female    39   0
  Female    51   0
  Female    45   0
  Female    43   1
  Male      39   1
  Male      31   0
  Female    39   0
  Male      34   0
  Female    52   1
  Female    46   0
  Male      58   1
  Female    50   1
  Female    32   0
  Female    52   1
  Female    35   0
  Female    51   0
  ;
  proc format;
    value subscrib 1 = 'accept' 0 = 'reject';
  run;

  proc probit;
    class subs sex;
    model subs=sex age / d=logistic itprint;
```

Dr. Firas

```
   format subs subscrib.;
run;
 proc genmod;
 class sex;
model subs=sex age  / dist=binomial link=logit;run;
```

يمكن تغيير الجنس الى متغير ثنائي وبذلك يمكن استعمال طريقة logistic في تحليل البيانات والحصول على نفس النتائج.

```
data news;
    input sex  age subs;
    datalines;
 1    35   0
 0    44   0
 0    45   1
 1    47   1
 1    51   0
 1    47   0
 0    54   1
 0    47   1
 1    35   0
 1    34   0
 1    48   0
 1    56   1
 0    46   1
 1    59   1
 1    46   1
 0    59   1
 0    38   1
 1    39   0
 0    49   1
 0    42   1
 0    50   1
 1    45   0
 1    47   0
 1    30   1
 1    39   0
 1    51   0
 1    45   0
 1    43   1
 0    39   1
 0    31   0
 1    39   0
 0    34   0
 1    52   1
 1    46   0
 0    58   1
 1    50   1
 1    32   0
 1    52   1
 1    35   0
 1    51   0
 ;
 proc probit;
   class subs;
   model subs=sex age / d=logistic itprint;
 run;
```

Dr. Firas

Dr. Firas

```
proc logistic descending;
 model subs= sex age;
 run;
proc genmod;
model subs=sex age  / dist=binomial link=logit;
run;
```

```
                        Probit Procedure
                      Class Level Information

                   Class    Levels    Values

                   SUBS        2    accept reject

                   Number of observations used = 40

                Logistic Regression of Subscription Status
                        Probit Procedure

Iter Ridge   LogLikelihood     INTERCPT        SEX        AGE
  0    0  -27.7258872224          0          0          0
  1    0 -20.14265929083 -3.634567629 -1.648455751 0.1051634384
  2    0 -19.52245047938 -5.254865196 -2.234724956 0.1506493473
  3    0  -19.4904387863 -5.728485385 -2.409827238 0.1639621828
  4    0 -19.49030280973  -5.76187293 -2.422349862 0.1649007124
  5    0 -19.49030280687   -5.7620267 -2.422407743 0.1649050312
                Logistic Regression of Subscription Status

                        Probit Procedure

        Data Set        =WORK.NEWS
        Dependent Variable=SUBS

           Weighted Frequency Counts for the Ordered Response Categories

                         Level     Count
                         accept       20
                         reject       20

        Log Likelihood for LOGISTIC -19.49030281

                    Last Evaluation of the Gradient

                  INTERCPT            SEX            AGE

                -5.95501E-12     8.768323E-10     -1.636698E-8

                    Last Evaluation of the Hessian

                       INTERCPT            SEX            AGE

            INTERCPT        6.459740        4.604222       292.040518
               SEX          4.604222        4.604222       216.208295
               AGE        292.040518       216.208295          13487
                 Logistic Regression of Subscription Status
                        Probit Procedure

           Variable  DF   Estimate  Std Err ChiSquare  Pr>Chi Label/Value

           INTERCPT   1  5.7620267  2.76345  4.347576  0.0371 Intercept
           SEX        1  2.4224077 0.955899     6.422  0.0113
           AGE        1 -0.16490503 0.065188  6.399204  0.0114
                 Logistic Regression of Subscription Status
```

التقديرات بهذه الطريقة لها عكس الاشارات ولو اردنا الحصول على نفس النتائج فيمكن ذلك بعكس الارقام في حقل subs.

```
                        The LOGISTIC Procedure
```

Dr. Firas

Data Set: WORK.NEWS
Response Variable: SUBS
Response Levels: 2
Number of Observations: 40
Link Function: Logit

Response Profile

Ordered Value	SUBS	Count
1	1	20
2	0	20

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	57.452	44.981	.
SC	59.141	50.047	.
-2 LOG L	55.452	38.981	16.471 with 2 DF (p=0.0003)
Score	.	.	13.640 with 2 DF (p=0.0011)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standard Estimate	Odds Ratio
INTERCPT	1	-5.7620	2.7634	4.3477	0.0371	.	.
SEX	1	-2.4224	0.9559	6.4222	0.0113	-0.645129	0.089
AGE	1	0.1649	0.0652	6.3994	0.0114	0.713137	1.179

Association of Predicted Probabilities and Observed Responses

Concordant = 85.3%		Somers' D = 0.713	
Discordant = 14.0%		Gamma = 0.718	
Tied = 0.8%		Tau-a = 0.365	
(400 pairs)		c = 0.856	

Logistic Regression of Subscription Status

The GENMOD Procedure

Model Information

Description	Value
Data Set	WORK.NEWS
Distribution	BINOMIAL
Link Function	LOGIT
Dependent Variable	SUBS
Observations Used	40
Number Of Events	20
Number Of Trials	40

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	37	38.9806	1.0535
Scaled Deviance	37	38.9806	1.0535
Pearson Chi-Square	37	50.9054	1.3758
Scaled Pearson X2	37	50.9054	1.3758
Log Likelihood	.	-19.4903	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	1	-5.7620	2.7635	4.3476	0.0371

Dr Firas

SEX	1	-2.4224	0.9559	6.4220	0.0113
AGE	1	0.1649	0.0652	6.3992	0.0114
SCALE	0	1.0000	0.0000	.	.

NOTE: The scale parameter was held fixed.

ان القيمة السالبة للجنس تشير الى ان الاناث اقل رغبة في الاشتراك مقارنة بالرجال فيما تمثل القيمة الموجبة بان الرغبة تزداد مع العمر.

يمكن الحصول على نتائج متشابهة عندما نعكس قيم subs ونرفع كلمة descending من طريقة logistic ونظيف العبارة order=data الى طريقة genmod فتكون البيانات والايعازات كالآتي:

```
data news;
   input sex  age subs f;
   datalines;
 1   35   0 1
 0   44   0 1
 0   45   1 0
 1   47   1 0
 1   51   0 1
 1   47   0 1
 0   54   1 0
 0   47   1 0
 1   35   0 1
 1   34   0 1
 1   48   0 1
 1   56   1 0
 0   46   1 0
 1   59   1 0
 1   46   1 0
 0   59   1 0
 0   38   1 0
 1   39   0 1
 0   49   1 0
 0   42   1 0
 0   50   1 0
 1   45   0 1
 1   47   0 1
 1   30   1 0
 1   39   0 1
 1   51   0 1
 1   45   0 1
 1   43   1 0
 0   39   1 0
 0   31   0 1
 1   39   0 1
 0   34   0 1
 1   52   1 0
 1   46   0 1
 0   58   1 0
 1   50   1 0
 1   32   0 1
 1   52   1 0
```

Dr. Firas

```
 1    35   0  1
 1    51   0  1
 ;
 proc probit ;
 class f;
 model f=sex age / d=logistic itprint;
 run;
 proc logistic ;
 model f= sex age;
 run;
 proc genmod rder=data;
 model f=sex age  / dist=binomial link=logit;run;
```

يمكن تقدير الاحتمالية للاشتراك subs في طريقة Logistic للبيانات الاولى بأستعمال الايعاز الآتي:

```
proc logistic descending;
 model subs= sex age;
 output out=prob predicted=phat l=lower u=upper xbeta=logit;
 run;
 proc print data=prob;
 run;
```

وفي طريقة probit نستعمل الايعاز الآتي:

```
 output out=prob p=phat;
 run;
 proc print data=prob ;
 run;
```

مثال (88):اجريت تجربة لمقارنة فعالية 5 عقارات على مجموعة من المرضى وسجلت عدد حالات النجاح اعتمادا على وجود او غياب الاستجابة علما بأن الاستجابة قد قيست ضمن تراكيز مختلفة ولذا اعتبر التركيز متغير مستمر مرافق للعقار (Covariate).

```
data drug;
 input drug$ x r n;
 cards;
 A  .1   1  10
 A  .23  2  12
 A  .67  1   9
 B  .2   3  13
 B  .3   4  15
 B  .45  5  16
 B  .78  5  13
 C  .04  0  10
 C  .15  0  11
 C  .56  1  12
 C  .7   2  12
 D  .34  5  10
 D  .6   5   9
 D  .7   8  10
 E  .2  12  20
 E  .34 15  20
 E  .56 13  15
 E  .8  17  20
 proc genmod data=drug;
```

Dr. Firas

```
class drug;
model r/n=x drug / dist=binomial link=logit;
run;

 proc probit data=drug;
class drug;
model r/n=x drug / d=logistic;
run;
```

فيما لايمكن استعمال طريقة logistic الا بأستعمال الايعاز الآتي:

```
if drug='A' then d1=1; else d1=0;
if drug='B' then d2=1; else d2=0;
if drug='C' then d3=1; else d3=0;
if drug='D' then d4=1; else d4=0;
if drug='E' then d5=1; else d5=0;
```

فيما سيكون ايعاز التنفيذ:

```
proc logistic data=drug;
 model r/n=x d1 d2 d3 d4 d5;
 run;
```

يمكن تقدير الاحتمالية لكل نسبة من العقار بأستعمال الايعاز الآتي:

```
proc logistic data=drug;
 model r/n=x d1 d2 d3 d4 d5;
 output out=prob predicted=phat l=lower u=upper xbeta=logit;
 run;
 proc print data=prob;
 run;
```

ونفس الايعاز بأستعمال طريقة probit:

```
proc probit data=drug;
 class drug;
 model r/n=x drug / d=logistic;
 output out=prob p=phat;
 run;
 proc print data=prob;
 run;
```

```
                    The SAS System
                   Probit Procedure
                  Class Level Information

                Class   Levels   Values

                DRUG       5   A B C D E

               Number of observations used = 18

                    The SAS System
                   Probit Procedure

      Data Set        =WORK.DRUG
      Dependent Variable=R
      Dependent Variable=N
      Number of Observations=  18
      Number of Events    =    99   Number of Trials =    237


      Log Likelihood for LOGISTIC  -114.773151

                    The SAS system
                   Probit Procedure
```

Dr. Firas

```
Variable  DF   Estimate  Std Err ChiSquare  Pr>Chi Label/Value

INTERCPT   1 0.27924116 0.419563  0.442961  0.5057 Intercept
X          1  1.9793946 0.766021  6.677029  0.0098

DRUG       4                    55.16059  0.0001
           1 -2.8955403 0.609224  22.58942  0.0001 A
           1 -2.0161704 0.405161  24.76279  0.0001 B
           1 -3.7952136  0.66546  32.52584  0.0001 C
           1 -0.8548389 0.483816  3.121816  0.0773 D
           0        0        0         .  .  E

                The SAS System
      OBS   DRUG     X     R     N     PHAT

       1     A     0.10    1    10    0.08178
       2     A     0.23    2    12    0.10330
       3     A     0.67    1     9    0.21584
       4     B     0.20    3    13    0.20734
       5     B     0.30    4    15    0.24175
       6     B     0.45    5    16    0.30023
       7     B     0.78    5    13    0.45190
       8     C     0.04    0    10    0.03117
       9     C     0.15    0    11    0.03845
      10     C     0.56    1    12    0.08260
      11     C     0.70    2    12    0.10618
      12     D     0.34    5    10    0.52433
      13     D     0.60    5     9    0.64841
      14     D     0.70    8    10    0.69210
      15     E     0.20   12    20    0.66265
      16     E     0.34   15    20    0.72156
      17     E     0.56   13    15    0.80023
      18     E     0.80   17    20    0.86562
```

كيف نقدر الاحتمالية اعتمادا على معادلة الانحدار:

على ضوء النتائج فأن معادلة التوقع لكل عقار هي:

$$P= 0.2794 + 1.9793x – 2.8955D1$$

$$P= 0.2794 + 1.9793x – 2.0161D2$$

$$P= 0.2794 + 1.9793x – 3.7952D3$$

$$P=0.2794 + 1.9793x – 0.8548\ D4$$

$$P= 0.2794 + 1.9793x – 0\ D5$$

لو اردنا تطبيق المعادلة الخاصة بالعقار 5 واستعمال القانون الآتي:

$$P= 1/ [\ 1 + \exp(– a – bx)]$$

$$– 0.2794 – 1.9793 \times 0.80 = – 1.86284$$

$$\text{Exp} – 1.86284= 0.1557$$

وهذه القيمة يمكن الحصول عليها بأستعمال البرنامج SAS

```
data dose;
input a;
b=exp(a);
```

```
cards;
-1.8626
 proc print;
  run;
```

$p = 1/(1 + 0.1557) = 0.865$

وهي نفس النتيجة المتحصل عليها من التحليل لنتائج المثال وبنفس الطريقة يمكن التنبوء بأي قيمة للعقارات الباقية بأعتماد المعادلة الخاصة بكل عقار.

مثال (89):اجريت تجربة على مجموعة سبائك f عرضت الى وقت محدد من التسخين (T) تم فحصت لمعرفة صلاحيتها على التكور(rolling) (S) اذ اعطيت قيمتان 1 = صالحة و 0 = غير صالحة ، المطلوب معرفة تأثير وقت التسخين على صلاحيتها؟

```
DATA R;
INPUT   F T S;
CARDS;

   1    7    1
   2    7    1
   1   14    0
   2   14    0
   3   14    1
   4   14    1
   1   27    0
   2   27    0
   3   27    0
   4   27    1
   5   27    1
   6   27    1
   1   51    0
   2   51    0
   3   51    0
   4   51    1
   5   51    1
  proc logistic data=R;
 model s=t;
 run;

   proc catmod data=r;
 direct t;
 response logits;
 model s=t;
 run;
  proc probit;
 class s;
 model s=t / d=logistic;
 run;

 proc genmod order=data;
 model s=t  / dist=binomial link=logit;
 run;
```

The SAS System
The LOGISTIC Procedure

Dr. Firas

Data Set: WORK.R
Response Variable: S
Response Levels: 2
Number of Observations: 17
Link Function: Logit

Response Profile

Ordered Value	S	Count
1	0	8
2	1	9

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	25.508	26.491	.
SC	26.341	28.157	.
-2 LOG L	23.508	22.491	1.018 with 1 DF (p=0.3131)
Score	.	.	1.003 with 1 DF (p=0.3165)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standard Estimate	Odds Ratio
INTERCPT	1	-1.0241	1.0486	0.9538	0.3288	.	.
T	1	0.0315	0.0320	0.9685	0.3250	0.285059	1.032

Association of Predicted Probabilities and Observed Responses

Concordant = 51.4%		Somers' D = 0.292	
Discordant = 22.2%		Gamma = 0.396	
Tied = 26.4%		Tau-a = 0.154	
(72 pairs)		c = 0.646	

The SAS System
CATMOD PROCEDURE

Response: S		Response Levels (R)=	2
Weight Variable: None		Populations (S)=	4
Data Set: R		Total Frequency (N)=	17
Frequency Missing: 0		Observations (Obs)=	17

POPULATION PROFILES

Sample	T	Sample Size
1	7	2
2	14	4
3	27	6
4	51	5

RESPONSE PROFILES

Response	S
1	0
2	1

MAXIMUM-LIKELIHOOD ANALYSIS

Iteration	Sub Iteration	-2 Log Likelihood	Convergence Criterion	Parameter Estimates 1	2
0	0	23.567004	1.0000	0	0
1	0	22.491647	0.0456	-0.9903	0.0305

Dr. Firas
Dr. Firas

```
2        0      22.49053   0.0000496     -1.0240     0.0315
3        0      22.49053   4.989E-10     -1.0241     0.0315


                    The SAS System
          MAXIMUM-LIKELIHOOD ANALYSIS-OF-VARIANCE TABLE

       Source               DF  Chi-Square     Prob
       ----------------------------------------------------
       INTERCEPT             1       0.95   0.3288
       T                  1       0.97   0.3250

       LIKELIHOOD RATIO       2       1.90   0.3872

            ANALYSIS OF MAXIMUM-LIKELIHOOD ESTIMATES

                            Standard   Chi-
   Effect         Parameter  Estimate   Error   Square  Prob
   ---------------------------------------------------------------
   INTERCEPT              1  -1.0241   1.0486    0.95  0.3288
   T                  2   0.0315   0.0320    0.97  0.3250

                  Probit Procedure
                 Class Level Information

                 Class    Levels    Values

                 S           2    0 1

              Number of observations used = 17

                    The SAS System
                   Probit Procedure

Data Set        =WORK.R
Dependent Variable=S

     Weighted Frequency Counts for the Ordered Response Categories

                       Level    Count
                         0        8
                         1        9

Log Likelihood for LOGISTIC -11.24526519

                    The SAS System
                   Probit Procedure
    Variable  DF   Estimate  Std Err ChiSquare  Pr>Chi Label/Value

    INTERCPT   1  -1.024119 1.048645  0.953771  0.3288 Intercept
    T          1 0.03150293 0.032011  0.968513  0.3251

          Probit Model in Terms of Tolerance Distribution

                       MU        SIGMA
                    32.50869     31.74308

        Estimated Covariance Matrix for Tolerance Parameters

                             MU          SIGMA

                MU       267.193714       122.216215
              SIGMA      122.216215       1040.381256
                    The SAS System

                  The GENMOD Procedure
                  Model Information

             Description               Value

             Data Set                  WORK.R
```

Dr Firas

Distribution	BINOMIAL
Link Function	LOGIT
Dependent Variable	S
Observations Used	17
Number Of Events	9
Number Of Trials	17

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	15	22.4905	1.4994
Scaled Deviance	15	22.4905	1.4994
Pearson Chi-Square	15	16.9420	1.1295
Scaled Pearson X2	15	16.9420	1.1295
Log Likelihood	.	-11.2453	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	1	-1.0241	1.0486	0.9538	0.3288
T	1	0.0315	0.0320	0.9685	0.3251
SCALE	0	1.0000	0.0000	.	.

NOTE: The scale parameter was held fixed.

يلاحظ من النتائج بأن جميع التقديرات متطابقة .

لو رتبنا بيانات هذا المثال على اساس تكرارعدد حالات النجاح والفشل فان من الممكن تحليل البيانات والوصول الى نتائج مماثلة.

```
DATA R;
INPUT f t  s;
CARDS;
 2    7    1
 2   14    0
 2   14    1
 3   27    0
 3   27    1
 3   51    0
 2   51    1
proc logistic;
 freq f;
 model s=t;
 run;

proc genmod order=data;
 weight f ;
 model s=t / dist=binomial link=logit;
 run;

proc catmod;
 weight f;
direct t;
 response logits;
 model s=t;
 run;

proc probit;
 weight f;
 class s;
```

Dr. Firas

```
model s=t / d=logistic;
run;
```

The SAS System

The LOGISTIC Procedure

Data Set: WORK.R
Response Variable: S
Response Levels: 2
Number of Observations: 7
Frequency Variable: F
Link Function: Logit

Response Profile

Ordered Value	S	Count
1	0	8
2	1	9

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	25.508	26.491	.
SC	26.341	28.157	.
-2 LOG L	23.508	22.491	1.018 with 1 DF (p=0.3131)
Score	.	.	1.003 with 1 DF (p=0.3165)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standard Estimate	Odds Ratio
INTERCPT	1	**-1.0241**	1.0486	0.9538	0.3288	.	.
T	1	**0.0315**	0.0320	0.9685	0.3250	0.285059	1.032

Association of Predicted Probabilities and Observed Responses

Concordant = 51.4%		Somers' D = 0.292	
Discordant = 22.2%		Gamma = 0.396	
Tied = 26.4%		Tau-a = 0.154	
(72 pairs)		c = 0.646	

The SAS System

The GENMOD Procedure

Model Information

Description	Value
Data Set	WORK.R
Distribution	BINOMIAL
Link Function	LOGIT
Dependent Variable	S
Scale Weight Variable	F
Observations Used	7
Number Of Events	4
Number Of Trials	7

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	5	22.4905	4.4981
Scaled Deviance	5	22.4905	4.4981
Pearson Chi-Square	5	16.9420	3.3884
Scaled Pearson X2	5	16.9420	3.3884

Dr Firas

Log Likelihood . -11.2453 .

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	1	**-1.0241**	1.0486	0.9538	0.3288
T	1	**0.0315**	0.0320	0.9685	0.3251
SCALE	0	1.0000	0.0000	.	.

NOTE: The scale parameter was held fixed.

The SAS System
CATMOD PROCEDURE

Response: S Response Levels (R)= 2
Weight Variable: F Populations (S)= 4
Data Set: R Total Frequency (N)= 17
Frequency Missing: 0 Observations (Obs)= 7

POPULATION PROFILES

Sample	T	Sample Size
1	7	2
2	14	4
3	27	6
4	51	5

RESPONSE PROFILES

Response	S
1	0
2	1

MAXIMUM-LIKELIHOOD ANALYSIS

Iteration	Sub Iteration	-2 Log Likelihood	Convergence Criterion	Parameter Estimates 1	2
0	0	23.567004	1.0000	0	0
1	0	22.491647	0.0456	-0.9903	0.0305
2	0	22.49053	0.0000496	-1.0240	0.0315
3	0	22.49053	4.989E-10	-1.0241	0.0315

The SAS System
MAXIMUM-LIKELIHOOD ANALYSIS-OF-VARIANCE TABLE

Source	DF	Chi-Square	Prob
INTERCEPT	1	0.95	0.3288
T	1	0.97	0.3250
LIKELIHOOD RATIO	2	1.90	0.3872

ANALYSIS OF MAXIMUM-LIKELIHOOD ESTIMATES

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob
INTERCEPT	1	**-1.0241**	1.0486	0.95	0.3288
T	2	**0.0315**	0.0320	0.97	0.3250

The SAS System
Probit Procedure
Class Level Information

Dr Firas

```
                    Class    Levels    Values

                    S          2    0 1

                 Number of observations used = 7

                      The SAS System
                     Probit Procedure

Data Set         =WORK.R
Dependent Variable=S
Weight Variable   =F

    Weighted Frequency Counts for the Ordered Response Categories

                        Level    Count
                          0        8
                          1        9

Log Likelihood for LOGISTIC -11.24526519

                      The SAS System
                     Probit Procedure

    Variable  DF   Estimate  Std Err ChiSquare  Pr>Chi Label/Value

    INTERCPT   1  -1.024119 1.048645  0.953771  0.3288 Intercept
    T          1 0.03150293 0.032011  0.968513  0.3251

          Probit Model in Terms of Tolerance Distribution

                        MU       SIGMA
                     32.50869    31.74308

          Estimated Covariance Matrix for Tolerance Parameters

                              MU           SIGMA

                MU       267.193714       122.216215
              SIGMA      122.216215       1040.381256
```

يمكن بطريقة اخرى ترتيب البيانات على اساس عدد حالات الفشل فقط واجراء التحليلات

```
data a;
input T R N;
CARDS;
7  0  2
14 2  4
27 3  6
51 3  5
proc logistic data=A;
 model r/n=t;
 run;

proc probit;
 model r/n=t / d=logistic;
 run;

proc genmod;
 model r/n=t / dist=binomial link=logit;
 run;
```

```
                      The SAS System
                    The LOGISTIC Procedure
```

Dr Firas

Data Set: WORK.A
Response Variable (Events): R
Response Variable (Trials): N
Number of Observations: 4
Link Function: Logit

Response Profile

Ordered Value	Binary Outcome	Count
1	EVENT	8
2	NO EVENT	9

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	25.508	26.491	.
SC	26.341	28.157	.
-2 LOG L	23.508	22.491	1.018 with 1 DF (p=0.3131)
Score	.	.	1.003 with 1 DF (p=0.3165)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standard Estimate	Odds Ratio
INTERCPT	1	-1.0241	1.0486	0.9538	0.3288	.	.
T	1	0.0315	0.0320	0.9685	0.3250	0.285059	1.032

Association of Predicted Probabilities and Observed Responses

Concordant = 51.4%		Somers' D = 0.292	
Discordant = 22.2%		Gamma = 0.396	
Tied = 26.4%		Tau-a = 0.154	
(72 pairs)		c = 0.646	

The SAS System
Probit Procedure

Data Set =WORK.A
Dependent Variable=R
Dependent Variable=N
Number of Observations= 4
Number of Events = 8 Number of Trials = 17
Log Likelihood for LOGISTIC -11.24526519

The SAS System
Probit Procedure

Variable	DF	Estimate	Std Err	ChiSquare	Pr>Chi	Label/Value
INTERCPT	1	-1.024119	1.048645	0.953771	0.3288	Intercept
T	1	0.03150293	0.032011	0.968513	0.3251	

Probit Model in Terms of Tolerance Distribution

MU	SIGMA
32.50869	31.74308

Estimated Covariance Matrix for Tolerance Parameters

	MU	SIGMA
MU	267.193714	122.216215
SIGMA	122.216215	1040.381256

The SAS System
The GENMOD Procedure

Dr Firas

Model Information

Description	Value
Data Set	WORK.A
Distribution	BINOMIAL
Link Function	LOGIT
Dependent Variable	R
Dependent Variable	N
Observations Used	4
Number Of Events	8
Number Of Trials	17

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	2	1.8975	0.9487
Scaled Deviance	2	1.8975	0.9487
Pearson Chi-Square	2	1.3282	0.6641
Scaled Pearson X2	2	1.3282	0.6641
Log Likelihood	.	-11.2453	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	1	-1.0241	1.0486	0.9538	0.3288
T	1	0.0315	0.0320	0.9685	0.3251
SCALE	0	1.0000	0.0000	.	.

NOTE: The scale parameter was held fixed.

مثال (90): البيانات ادناه تعود الى مجموعتين من المرضى الاولى اعطيت placebo والثانية اعطيت عقار ب المطلوب وضع معادلة انحدارلنسبة الذين استجابوا للعلاج بعد التعديل لتأثير العلاج والعمر؟

```
DATA DIARY;
   INPUT PAT TRT AGE SUCC ATTPT @@;
   IF ATTPT = 0 THEN DELETE;
   DATALINES;
 1 0 41  3  6   3 0 44  5 15   5 0 62  0  4   6 0 44  1  2
 8 0 70  3  8  11 0 35  4  8  13 0 72  1  6  15 0 34  5 15
18 0 61  1  7  22 0 35  5  5  24 0 52  6  8  25 0 66  1  7
27 0 35  4 10  30 0 61  4  8  31 0 55  2  5  34 0 41  7  9
37 0 53  2  4  39 0 72  4  6  40 0 68  0  0  41 0 56 12 17
44 0 53  8 15  45 0 45  3  4  48 0 40 14 20   2 1 57  3  8
 4 1 54 10 12   7 1 65  0  0   9 1 51  5  8  10 1 53  8 10
12 1 44 17 22  14 1 66  2  3  16 1 55  9 11  17 1 37  6  8
19 1 40  2  4  20 1 44  9 16  21 1 64  5  9  23 1 78  1  3
26 1 51  6 12  28 1 67  5 11  29 1 44  3  3  32 1 65  7 18
33 1 69  0  2  35 1 53  4 14  36 1 49  5  8  38 1 74 10 15
42 1 39  4  9  43 1 35  8 10  46 1 47  4  5  47 1 46  6  7
;
PROC LOGISTIC DATA = DIARY;
 MODEL SUCC/ATTPT = TRT AGE;
RUN;
proc probit;
 model SUCC/ATTPT = TRT AGE / d=logistic;
 run;
proc genmod;
```

Dr. Firas
Dr. Firas

```
model SUCC/ATTPT = TRT AGE / dist=binomial;
run;
```

هناك بعض الحالات لايمكن فيها استعمال طرق logistic و genmod و probit

5 – 5 الانحدار البواسوني poisson regression

ان هذا التوزيع يعد من اهم التوزيعات الاحتمالية تطبيقا وهو مناسب جدا للبيانات العددية وقد يحاول البعض تطبيق الانحدارالخطي البسيط او المتعدد الا ان من اهم المشاكل التي تنجم عن ذلك عدم تساوي تباين الخطأ (heteroscedasticity) وهو احد الشروط المهمة لتنفيذ الانحدار.

تتمثل تطبيقات هذا الانحدار في التجارب التي تقاس فيها المشاهدات على ضوء وحدة الزمان او المكان مثل عدد المكالمات التليفونية المستلمة في دائرة ما بالدقيقة او عدد الخلايا الجرثومية في كل جزء من الالف من اللتراو عدد حوادث السيارات في شهر ما او عدد الحرائق في منطقة معينة.

مثال (91): سجلت بيانات احد الامراض في مركز السيطرة على الامراض (CDC) في الولايات المتحدة الامريكية من عام 1982 فهل هناك تأثير للسنة في عدد الاصابات؟ بما ان البيانات عددية فأن الطريقة المثلى لوصف التغير في ميل الاعداد خلال السنوات هو انحدار بوسوان.

```
data d;
input Year Cases;
cards;
1982 491
1983 595
1984 1518
1985 2748
1986 1387
1987 2392
1988 4882
1989 8803
1990 7943
1991 9470
1992 9908
1993 8257
1994 13043
1995 11700
1996 16455
1997 12801
proc genmod;
model cases=year/dist=poisson link=log;
run;
```

The SAS System

The GENMOD Procedure

Model Information

Description	Value
Data Set	WORK.D
Distribution	POISSON

Dr. Firas

Link Function LOG
Dependent Variable CASES
Observations Used 16

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	14	11009.9046	786.4218
Scaled Deviance	14	11009.9046	786.4218
Pearson Chi-Square	14	10624.5111	758.8937
Scaled Pearson X2	14	10624.5111	758.8937
Log Likelihood	.	909565.7562	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	1	-311.0496	1.5001	42994.5116	0.0001
YEAR	1	0.1607	0.0008	45545.0973	0.0001
SCALE	0	1.0000	0.0000	.	.

NOTE: The scale parameter was held fixed.

لذا فأن معادلة الانحدار

$Y=e^{-311.04 + 0.1607x}$

ولغرض حساب الزيادة او النقصان في عدد الحالات فأن من الافضل تحويل ذلك الى نسبة اي العدد المتوقع للسنة اللاحقة على العدد المتوقع للسنة السابقة

$Y=e^{-311.04 + 0.1607x+1} / Y=e^{-311.04 + 0.1607x} = 1.17$

ان عدد الحالات يزداد سنويا بمقدار العامل 1.17

5 – 6 الانحدار التشخيصي Diagnostic Regression

يعد الباحث Welch هو اول من اوضح اهمية العمل بالانحدار التشخيصي أذ طلب منه بعض الاصدقاء وضع معادلة توقع لبيانات مصرفية وعند انجازه للمعادلة تبين بأن نتائج التوقع لاحد القيم لم تكن مطابقة لتوقعاتهم ، وعند مراجعته للبيانات وجد ان هناك رقما شاذا ، لذا فقد عمل من اجل التوصل الى طريقة تسمح له بتحديد القيم الشاذة في العينات المدروسة.

هناك عدة فرضيات يجب توفرها لكي تكون معادلة الانحدار صحيحة وهي:

1- ان تكون العلاقة خطية بين المتغيرات المستقلة والتابعة (Linearity).

2- الخطأ موزع توزيعا طبيعيا (Normality).

3- تباين الخطأ يجب ان يكون ثابتا (Homogeneity of Variance)

4- الاخطاء المرافقة لمشاهدة ما لاترتبط باخطاء بقية المشاهدات (Independent)

5- العوامل المستقلة يجب ان تقاس بلا اخطاء

6- يجب ان يكون الموديل المستعمل مناسب بمعنى ان يتضمن العوامل المستقلة المؤثرة ويستبعد العوامل غير المؤثرة.

هناك بعض الحالات التي تواجهنا خلال التحليل الاحصائي الخاص بتحليل الانحداروهي ذات تأثير كبير في نتائج تحليل الانحداروهي :

أ- مشاهدات فردية تؤثر بشكل كبير على تقديرات معامل الانحدار

ب- العوامل المستقلة تكون باستقامة واحدة ، اي مرتبطة خطيا (Collinearity)

ان تحليل الانحدار يمكن ان يتأثر بدرجة كبيرة بمشاهدة واحدة وهناك ثلاثة طرق يمكن ان تؤثر فيها تلك المشاهدة على تحليل الانحدار:

1- المشاهدات الشاذة (Outliers) وتمثل المشاهدات التي تعود الى عامل مستقل والتي تؤثر على قيمة العامل التابع المتنبأ به من معادلة التوقع.

2- المشاهدات المتطرفة (Leverage) للعامل المستقل والتي تكون اكثر بعدا من المشاهدات الشاذة والتي تؤثر على قيم معامل الانحدار.

3- المشاهدة المؤثرة (Influence) وتمثل المشاهدة التي تؤثر في معامل الانحدار عند حذفها وقد تكون احدى قيم المشاهدات الشاذة او المتطرفة.

يمكن تحديد القيم الشاذة او المتطرفة عن طريق قيم Cook's Distance او قيم Dfits وعندما تكون قيم التقديرين مقاربة للصفر او مساوية له فأن ذلك يعني عدم وجود قيم شاذة او متطرفة اما اذا كانت التقديرات عالية فأنها تدلل على وجودها ،وان ازالتها سيقلل بلاشك من الخطا القياسي لمعامل الانحدار، ويمكن اجراء ذلك في الانحدار البسيط او المتعدد باستعمال الايعاز الآتي:

مثال (92): البيانات ادناه تمثل عامل تابع (total) ومجموعة من العوامل المستقلة المطلوب اجراء انحدار تشخيصي للبيانات؟

```
DATA wisc;
 INPUT iq parent ses total coll;
 cards;
1  1  1  353    4
1  1  2  234    2
1  1  3  174    8
1  1  4   52    4
1  2  1   77   13
1  2  2  111   27
1  2  3  138   47
1  2  4   96   39
2  1  1  216    9
2  1  2  208    7
2  1  3  126    6
2  1  4   52    5
2  2  1  105   33
2  2  2  159   64
2  2  3  184   74
2  2  4  213  123
PROC REG DATA=wisc;
 MODEL total= iq parent ses / clm r influence;
RUN;
```

يمكن تحديد القيم الشاذة باستعمال الأمر proc logistic بأضافة الكلمة influence لغرض تقدير الانحدار التشخيصي كرسم مبعثر او كانحراف وتقدير حدود الثقة العليا والدنيا والاحتمالية كما في الايعاز الآتي:

Dr Firas

```
proc logistic;
 model coll/total=iq ses parent/influence iplot;
output out=m2 p=prob lower=LCL upper=UCL;
run;
 proc print data=m2;
run;
```

The LOGISTIC Procedure

Data Set: WORK.WISC
Response Variable (Events): COLL
Response Variable (Trials): TOTAL
Number of Observations: 16
Link Function: Logit

Response Profile

Ordered Value	Binary Outcome	Count
1	EVENT	465
2	NO EVENT	2033

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	2403.030	1779.465	.
SC	2408.853	1802.758	.
-2 LOG L	2401.030	1771.465	629.565 with 3 DF (p=0.0001)
Score	.	.	574.056 with 3 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Stand Estimate	Odds Ratio
INTERCPT	1	-7.7816	0.3754	429.6732	0.0001	.	.
IQ	1	0.6058	0.1233	24.1537	0.0001	0.167013	1.833
SES	1	0.3792	0.0577	43.2065	0.0001	0.222612	1.461
PARENT	1	2.6475	0.1677	249.2622	0.0001	0.723497	14.119

Association of Predicted Probabilities and Observed Responses

Concordant = 81.5%	Somers' D = 0.677		
Discordant = 13.8%	Gamma = 0.711		
Tied = 4.7%	Tau-a = 0.205		
(945345 pairs)	c = 0.838		

The LOGISTIC Procedure
Regression Diagnostics

Case Number	Covariates IQ	SES	PARENT	Pearson Residual Value	(1 unit = 0.22) -8 -4 0 2 4 6 8
1	1.0000	1.0000	1.0000	-0.6385	| * | |
2	1.0000	2.0000	1.0000	-1.4418	| * | |
3	1.0000	3.0000	1.0000	0.9945	| | * |
4	1.0000	4.0000	1.0000	1.0233	| | * |
5	1.0000	1.0000	2.0000	-0.3037	| *| |
6	1.0000	2.0000	2.0000	-0.0572	| * |
7	1.0000	3.0000	2.0000	0.4587	| | * |
8	1.0000	4.0000	2.0000	-0.0753	| * |
9	2.0000	1.0000	1.0000	1.2060	| | * |
10	2.0000	2.0000	1.0000	-0.5036	| * | |
11	2.0000	3.0000	1.0000	-0.5051	| * | |
12	2.0000	4.0000	1.0000	0.3509	| | * |
13	2.0000	1.0000	2.0000	0.5503	| | * |
14	2.0000	2.0000	2.0000	0.7521	| | * |
15	2.0000	3.0000	2.0000	-1.7282	|* | |

Dr Firas

16 2.0000 4.0000 2.0000 0.5079 | | * |

Regression Diagnostics
Deviance Residual Hat Matrix Diagonal

Case Number	Value	(1 unit = 0.22) -8 -4 0 2 4 6 8	Value	(1 unit = 0.03) 0 2 4 6 8 12 16
1	-0.6706	* |	0.1814	*
2	-1.6478	*		0.1465	*
3	0.9380	| *	0.1640	*
4	0.9405	| *	0.0884	*
5	-0.3067	*|	0.2385	*
6	-0.0573	*	0.2552	*
7	0.4565	| *	0.3222	*
8	-0.0753	*	0.3629	*
9	1.1270	| *	0.1882	*
10	-0.5183	* |	0.2126	*
11	-0.5206	* |	0.1886	*
12	0.3428	| *	0.1382	*
13	0.5460	| *	0.3693	*
14	0.7485	| *	0.3077	*
15	-1.7357	*		0.2797	*
16	0.5087	| *	0.5567	*

The SAS System

يتبين من تقديرات Deviance Residual ان القيم المتطرفة هي للمشاهدات 2، 9 ،15

Regression Diagnostics
INTERCPT Dfbeta IQ Dfbeta

Case Number	Value	(1 unit = 0.08) -8 -4 0 2 4 6 8	Value	(1 unit = 0.09) -8 -4 0 2 4 6 8
1	-0.3223	* |	0.1242	*
2	-0.6232	*		0.2696	*
3	0.4100	| *	-0.2009	* |
4	0.2256	| *	-0.1264	*|
5	-0.0748	*|	0.1035	*
6	-0.0139	*	0.0273	*
7	0.0916	*	-0.2980	* |
8	-0.00530	*	0.0467	*
9	0.4263	| *	0.1975	| *
10	-0.1870	* |	-0.0963	*|
11	-0.1430	* |	-0.0837	*|
12	0.0583	*	0.0400	*
13	-0.0177	*	0.1891	| *
14	-0.1307	* |	0.2972	| *
15	0.5694	| *	-0.6990	*
16	-0.4185	* |	0.3426	| *

Regression Diagnostics
SES Dfbeta PARENT Dfbeta

Case Number	Value	(1 unit = 0.07) -8 -4 0 2 4 6 8	Value	(1 unit = 0.06) -8 -4 0 2 4 6 8
1	0.1336	| *	0.2212	| *
2	0.0614	*	0.5047	*
3	0.1161	| *	-0.3932	* |
4	0.1701	| *	-0.2579	* |
5	0.1370	| *	-0.0617	*|
6	0.0152	*	-0.0126	*
7	0.0518	*	0.0987	| *
8	-0.0408	*|	-0.00969	*
9	-0.2797	* |	-0.4620	* |
10	0.0361	*	0.2503	| *
11	-0.0619	*|	0.2405	| *
12	0.0773	*	-0.1267	*|
13	-0.4270	* |	0.1610	| *
14	-0.3152	* |	0.1997	| *
15	-0.1646	*|	-0.3343	* |
16	0.5630	| *|	0.0917	*

Dr. Firas

Regression Diagnostics

C CBAR

Case Number	Value	(1 unit = 0.1) 0 2 4 6 8 12 16	Value	(1 unit = 0.07) 0 2 4 6 8 12 16
1	0.1103	*	0.0903	*
2	0.4181	*	0.3568	*
3	0.2321	*	0.1941	*
4	0.1115	*	0.1016	*
5	0.0379	*	0.0289	*
6	0.00151	*	0.00112	*
7	0.1475	*	0.1000	*
8	0.00507	*	0.00323	*
9	0.4154	*	0.3372	*
10	0.0870	*	0.0685	*
11	0.0731	*	0.0593	*
12	0.0229	*	0.0198	*
13	0.2812	*	0.1773	*
14	0.3631	*	0.2514	*
15	1.6102	*	1.1598	*
16	0.7308	*	0.3240	*

Regression Diagnostics

DIFDEV DIFCHISQ

Case Number	Value	(1 unit = 0.26) 0 2 4 6 8 12 16	Value	(1 unit = 0.26) 0 2 4 6 8 12 16
1	0.5400	*	0.4980	*
2	3.0722	*	2.4356	*
3	1.0739	*	1.1830	*
4	0.9862	*	1.1488	*
5	0.1230	*	0.1212	*
6	0.00440	*	0.00440	*
7	0.3084	*	0.3104	*
8	0.00890	*	0.00890	*
9	1.6072	*	1.7916	*
10	0.3372	*	0.3221	*
11	0.3303	*	0.3144	*
12	0.1373	*	0.1429	*
13	0.4755	*	0.4802	*
14	0.8116	*	0.8171	*
15	4.1724	*	4.1465	*
16	0.5828	*	0.5820	*

The LOGISTIC Procedure

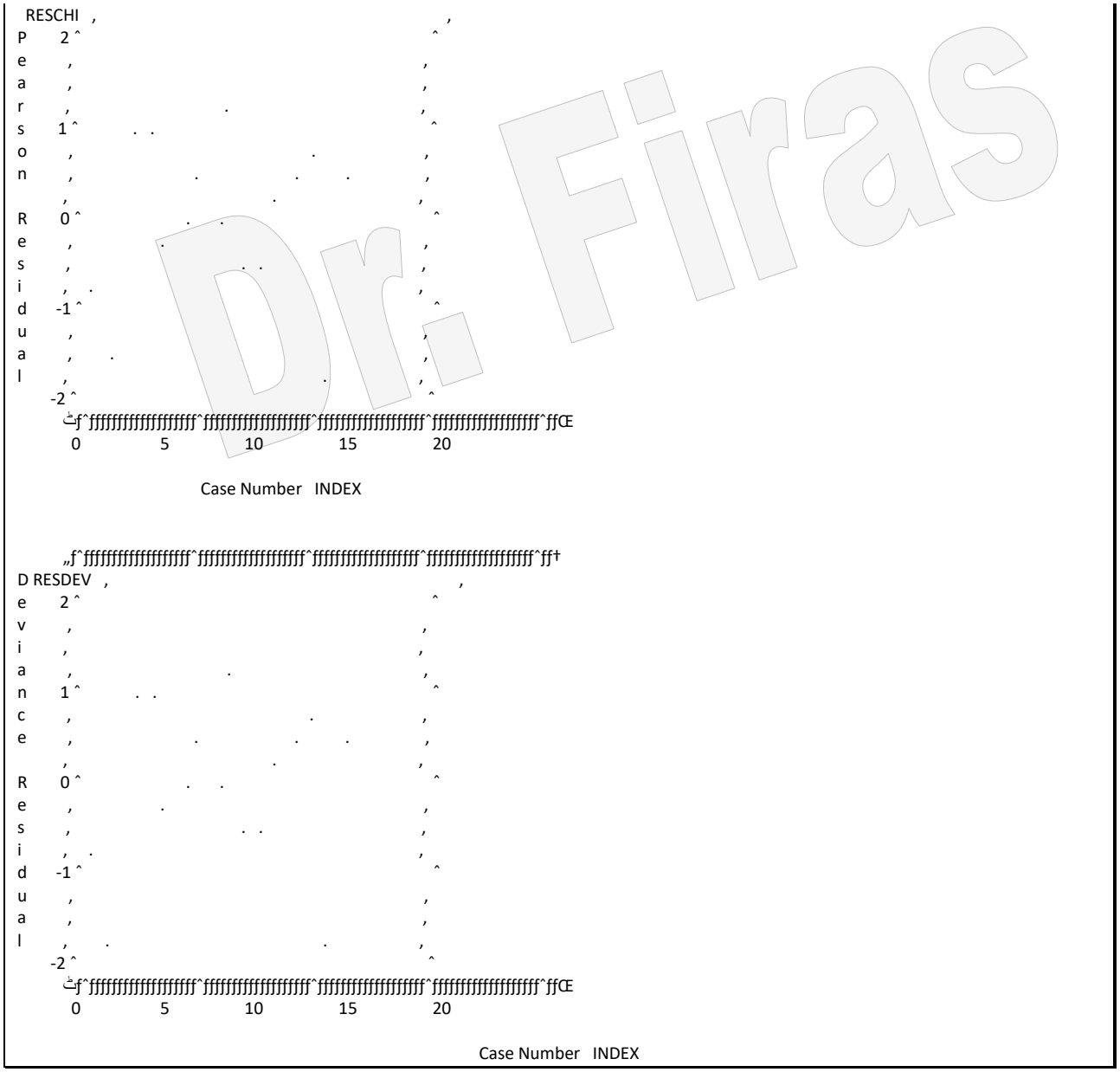
The SAS System

OBS	IQ	PARENT	SES	TOTAL	COLL	PROB	LCL	UCL
1	1	1	1	353	4	0.01553	0.01090	0.02210
2	1	1	2	234	2	0.02254	0.01630	0.03109
3	1	1	3	174	8	0.03259	0.02344	0.04514
4	1	1	4	52	4	0.04692	0.03250	0.06729
5	1	2	1	77	13	0.18219	0.14379	0.22811
6	1	2	2	111	27	0.24558	0.20741	0.28823
7	1	2	3	138	47	0.32233	0.27975	0.36808
8	1	2	4	96	39	0.41003	0.35232	0.47033
9	2	1	1	216	9	0.02810	0.01997	0.03942
10	2	1	2	208	7	0.04054	0.02983	0.05487
11	2	1	3	126	6	0.05815	0.04275	0.07865
12	2	1	4	52	5	0.08275	0.05884	0.11519
13	2	2	1	105	33	0.28992	0.24013	0.34534
14	2	2	2	159	64	0.37366	0.33297	0.41622
15	2	2	3	184	74	0.46573	0.42788	0.50397
16	2	2	4	213	123	0.56019	0.51001	0.60916

The LOGISTIC Procedure

„ƒ`ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ`ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ`ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ`ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ`ƒƒ†

Dr Firas



Dr. FirAS

ويتميز الايعاز proc genmod بامكانيته في تقدير مستويات كل عامل على حدة وذلك بأستعمال الايعاز الآتي:

```
PROC GENMOD DATA=wisc;
classes iq ses parent;
 MODEL coll/total=iq ses parent / D=B TYPE3;run;
```

The SAS System
The GENMOD Procedure

Model Information

Description	Value
Data Set	WORK.WISC
Distribution	BINOMIAL
Link Function	LOGIT
Dependent Variable	COLL
Dependent Variable	TOTAL
Observations Used	16
Number Of Events	465

Number Of Trials 2498

Class Level Information

Class	Levels	Values
IQ	2	1 2
SES	4	1 2 3 4
PARENT	2	1 2

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	10	9.9528	0.9953
Scaled Deviance	10	9.9528	0.9953
Pearson Chi-Square	10	9.6664	0.9666
Scaled Pearson X2	10	9.6664	0.9666
Log Likelihood	.	-885.0597	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	1	0.3074	0.1176	6.8334	0.0089
IQ 1	1	-0.5978	0.1236	23.4053	0.0001
IQ 2	0	0.0000	0.0000	.	.
SES 1	1	-1.1403	0.1865	37.3948	0.0001
SES 2	1	-0.8320	0.1631	26.0079	0.0001
SES 3	1	-0.5416	0.1549	12.2314	0.0005
SES 4	0	0.0000	0.0000	.	.
PARENT 1	1	-2.6486	0.1678	249.0655	0.0001
PARENT 2	0	0.0000	0.0000	.	.
SCALE	0	1.0000	0.0000	.	.

The SAS System

NOTE: The scale parameter was held fixed.

LR Statistics For Type 3 Analysis

Source	DF	ChiSquare	Pr>Chi
IQ	1	23.9032	0.0001
SES	3	45.9558	0.0001
PARENT	1	377.7526	0.0001

كما اوضحنا سابقا ان وجود ارتباط بين العوامل المستقلة (collinearity) هو احد اهم المشاكل التي تواجه تقدير معامل الارتباط المتعدد اذ تؤدي الى اعطاء نموذج غير قادر على وصف البيانات بصورة دقيقة بمعنى ان القيم المتنبأ بها ستكون غير طبيعية . 1ذ ان معادلة التنبوء لاتكون دقيقة مالم يكن التباين المشترك بين العوامل المستقلة يساوي صفر. هناك عدة حلول لهذه الحالة هي :

1- استبعاد تلك العوامل من النموذج فمثلا اذا كان لدينا عاملان مستقلان يرتبطان فيمكن استبعاد احد العوامل والاعتماد على الآخر في وضع معادلة التوقع وهنا يفضل الاحتكام الى قيمة R^2 في اختيار العامل.

2- ان زيادة حجم البيانات لايحل المشكلة اذ ان ذلك لايؤثر على علاقة الارتباط بين المتغيرات المستقلة ، ولكن يمكن استعمال قيم المتغيران بدلالة انحراف تلك القيم عن متوسطها.

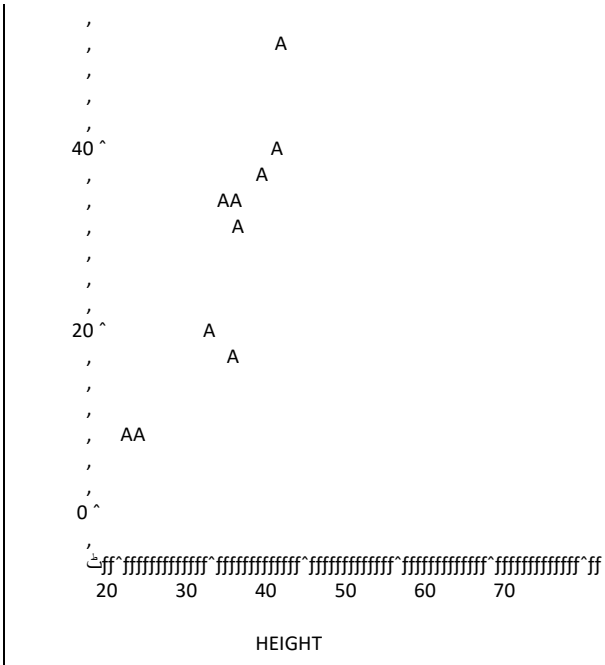
3- استعمال طريقة Orthogonal polynomial regression

مثال (93): نفذت دراسة على مجموعة من الاطفال اذ جمعت البيانات اللازمة لتحديد معادلة التنبوء الخاص بطول القسطرة اللازم لادخالها عن طريق الشريان الرئيسي في منطقة الفخذ وايصالها الى القلب وقد حدد الطول المناسب لها ووصولها الى الموقع الصحيح بأستعمال Fluoroscope . اذ ان الهدف من الدراسة هوتحديد طول القسطرة المناسب اعتمادا على طول ووزن 12 طفل.

```
data f;
input Height Weight Length;
cards;
42.8  40.0  37
63.5  93.5  50
37.5  35.5  34
39.5  30.0  36
45.5  52.0  43
38.5  17.0  28
43.0  38.5  37
22.5  8.5   20
37.0  33.0  34
23.5  9.5   30
33.0  21.0  38
58.0  79.0  47
proc plot;
plot weight*height;
run;
proc reg;
model length= height weight;
run;
proc reg;
model length= height ;
run;
proc reg;
model length= weight;
run;
```

```
                    The SAS System
        Plot of WEIGHT*HEIGHT.  Legend: A = 1 obs, B = 2 obs, etc.

WEIGHT ,
    ,
  100 ˆ
    ,
    ,                                   A
    ,
    ,
    ,
    ,
   80 ˆ                                 A
    ,
    ,
    ,
    ,
    ,
    ,
   60 ˆ
    ,
```

Dr. Firas



The SAS System
Model: MODEL1
Dependent Variable: LENGTH

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	607.18780	303.59390	21.267	0.0004
Error	9	128.47887	14.27543		
C Total	11	735.66667			

Root MSE	3.77828	R-square	0.8254
Dep Mean	36.16667	Adj R-sq	0.7865
C.V.	10.44687		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	20.375764	8.38594862	2.430	0.0380
HEIGHT	1	0.210747	0.34553644	0.610	0.5570
WEIGHT	1	0.191095	0.15827128	1.207	0.2581

The SAS System
Model: MODEL1
Dependent Variable: LENGTH

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	586.37727	586.37727	39.278	0.0001
Error	10	149.28940	14.92894		

Dr. Firas

```
C Total       11   735.66667

  Root MSE      3.86380    R-square     0.7971
  Dep Mean     36.16667    Adj R-sq     0.7768
  C.V.         10.68331

                Parameter Estimates

          Parameter     Standard   T for H0:
Variable  DF    Estimate       Error  Parameter=0   Prob > |T|

INTERCEP  1    11.478976   4.09405157       2.804      0.0187
HEIGHT    1     0.611712   0.09760525       6.267      0.0001

                The SAS System
Model: MODEL1
Dependent Variable: LENGTH

                Analysis of Variance

                  Sum of       Mean
Source       DF   Squares     Square    F Value     Prob>F

Model         1   601.87742   601.87742     44.987     0.0001
Error        10   133.78925   13.37892
C Total      11   735.66667

  Root MSE      3.65772    R-square     0.8181
  Dep Mean     36.16667    Adj R-sq     0.8000
  C.V.         10.11351

                Parameter Estimates

          Parameter     Standard   T for H0:
Variable  DF    Estimate       Error  Parameter=0   Prob > |T|

INTERCEP  1    25.344086   1.92834482      13.143      0.0001
WEIGHT    1     0.283871   0.04232311       6.707      0.0001
```

نلاحظ من النتائج مايلي:

1- ان الرسم البياني بين عاملي الوزن والارتفاع يشير الى وجود علاقة خطية

2- اختبار F يشير الى ان b1 او b2 او كليهما يختلف معنويا عن الصفر كما ان معامل التحديد $R^2 = 0.8254$ مما يعني ان المتغيرين لهما دور كبير في تباين مظهر الصفة وعلى الرغم من ذلك نجد ان اختبار t واحتماليتهما غير معنوية. هذا التضارب في نتائج الاختبارين يعود الى الارتباط بين العاملين وكما اسلفنا فأن ذلك سيؤدي الى حصولنا على قيم متوقعة غير دقيقة وفي هذه الحالة يفضل الاعتماد على متغير واحد وهو الذي يأخذ اعلى قيمة لمعامل التحديد.

3- لو حاولنا تقدير معامل الارتباط بين العاملين المستقلين سنجد ان الارتباط بينهما يساوي 0.96.

4- يمكن اعتماد الانحدار التشخيصي لتحديد القيم التي لها انحراف عن القيم المشاهدة بأستعمال الايعازين p وr :

```
proc reg;
model length= height weight/p r;
run;
```

```
proc reg;
model length= height/ p r ;
run;
proc reg;
model length= weight/p r;
run;
```

الفصل السادس

الارتباط Correlation

هو مقياس لقوة واتجاه العلاقة بين متغيرين مستقلين او اكثر، وتتحدد قيمته بين +1 الى -1.

6 – 1 الارتباط البسيط Simple correlation

هوالعلاقة بين متغيرين مستقلين.

مثال (94) : البيانات ادناه تمثل الكثافة العددية لأشجار الصنوبر في وحدة مساحة قدرها 100 متر مربع(x1) ومساحة قاعدة الاشجار عند ارتفاع متر من الأرض (x2)؟

```
data d;
input x1 x2;
cards;
13.5 307
20.1 79
14.8 71
19.6 192
19.5 122
17.4 404
26.1 55
21.1 82
proc corr ;
var x1 x2;
run;
```

Dr. Firas

```
                    Correlation Analysis
                  2 'VAR' Variables:  X1      X2
                     Simple Statistics

 Variable    N        Mean      Std Dev         Sum      Minimum      Maximum
 X1          8    19.012500     3.911316   152.100000    13.500000    26.100000
 X2          8   164.000000   128.216335  1312.000000    55.000000   404.000000

       Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / N = 8

                            X1           X2
                  X1       1.00000      -0.51950
                           0.0           0.1870

                  X2      -0.51950       1.00000
                                                          0.1870         0.0
```

قيمة الارتباط = - 0.519 وهو غير معنوي اذ ان الاحتمالية = 0.187

6 – 2 الأرتباط المتعدد Multiple correlation

ويجري تقديره عندما يكون هناك عدة متغيرات مستقلة بحيث يمكن تقدير العلاقة بين اي زوج من المتغيرات.

مثال (95) : جد معاملات الارتباط بين المتغيرات X1 , X2 , X3 ؟

```
data c;
 input X1 X2 X3;
 cards;
 19  43  29  11
 24  49  28  22
 30  51  37  18
 29  54  31  20
 19  42  30  12
 25  53  23  21
 31  58  27  27
option nodate nonumber;
 PROC CORR;
VAR X1 X2 X3;
RUN;
```

Dr. Firas

```
                    The SAS System

Correlation Analysis

3 'VAR' Variables:  X1     X2     X3


Simple Statistics

Variable     N      Mean     Std Dev       Sum     Minimum     Maximum

X1           7   25.285714    4.990467   177.000000   19.000000   31.000000
X2           7   50.000000    5.830952   350.000000   42.000000   58.000000
X3           7   29.285714    4.270608   205.000000   23.000000   37.000000


                Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / N = 7

                          X1           X2           X3

                X1     1.00000      0.91641      0.23796
                       0.0          0.0037       0.6074

                X2     0.91641      1.00000     -0.16732
                       0.0037       0.0          0.7199

                X3     0.23796     -0.16732      1.00000
                       0.6074       0.7199       0.0
```

القيم في الصف الاعلى لكل متغير تمثل قيم الارتباط والقيم في الصف الثاني تمثل الاحتمالية (المعنوية).

الفصل السابع

نسبة الأرجحية Odds Ratio والخطورة النسبية Relative Risk

تعد نسبة الارجحية من المقاييس المهمة التي يتم استعمالها في العديد من البحوث الطبية والزراعية الحديثة ،اذ تمثل مقياسا لمقارنة أحتمال حصول حدث ما بين مجموعتين ، وبما انها نسبة فهي ذات تقديرات موجبة دائما ، وعندما يكون تقديرها اقل من 1 فأن ذلك يعني ان احتمال حصول الحدث اقل في المجموعة الثانية مقارنة بالاولى ، اما اذا كان التقدير يساوي 1 فهو يعني تساوي الاحتمال في مجموعتي المقارنة واكبر من 1 وتشير الى ارجحية حصول الحدث في المجموعة الثانية. ويمكن الحصول على نسبة الارجحية من خلال قسمة ارجحية حصول حدث في المجموعة الاولى (عدد مرات حصول الحدث الاول/ عدد مرات عدم حصوله) مقسوما على ارجحية حصول الحدث في المجموعة الثانية (عدد مرات حصول الحدث الثاني/ عدد مرات عدم حصوله) ، وتوفر طريقة الانحدار اللوجستيكي Proc Logistic امكانية تعديل لنسبة الارجحية وفقا للعوامل الداخلة في النموذج الرياضي مما يجعلها مفضلة على مقياس اخر لايقل عنها اهمية وهو نسبة الخطورة RR (Risk Ratio) او قد يسمى الخطورة النسبية Relative (Risk) ، اذ تتصف بصعوبة اجراء تعديل لها ، وتمثل نسبة الخطورة عدد حالات حصول حدث في المجموعة الاولى (عدد الاحداث المرغوبة مقسوما على مجموع الاحداث المرغوبة وغير المرغوبة) مقسوما على عدد حالات حصول نفس الحدث في المجموعة الثانية .

ويمكن الحصول على Odds Ratio بأستعمال الأمر Proc logistic ، فيما يمكن الحصول على التقديرين بأستعمال الأوامر proc phreg و Proc freq.

بصورة عامة عندما تكون البيانات متعلقة باحداث الماضي تسمى (Case-control) فيفضل هنا اعتماد نسبة الارجحية مثال ذلك البحث في العلاقة بين التدخين وسرطان الرئة ، اذ جرى مقارنة مجموعة من المرضى تعاني من سرطان الرئة مع مجموعة لاتعاني من سرطان الرئة ، ثم جرى تصنيف لكل مجموعة الى مجموعتين فرعية مدخن (تعرض للخطر) وغير مدخن. اما الدراسات التي تتعلق بالمستقبل وتسمى (Cohort) فيفضل اعتماد الخطورة النسبية مثال ذلك مقارنة مجموعة من لاعبي كرة القدم كانوا قد تعرضوا في وقت سابق الى تمزق في الغضروف الهلالي للركبة مع مجموعة من اللاعبين لم يصابوا ثم تم مراقبة كلا المجموعتين لمدة عامين لمعرفة فيما اذا كان المصابون بالتمزق اكثر عرضة لحصول تمزق ثاني خلال هذه الفترة. ومن الامثلة الاخرى اجراء مقارنة بين مجموعتين من الاشخاص الاولى ذات كولستيرول مرتفع بينما الثانية ذات مستوى طبيعي من خلال مراقبتهم لمدة خمسة سنوات لتحديد اي مجموعة اكثر عرضة للاصابة بنوبة قلبية.

7 – 1 تقدير نسبة الارجحية والخطورة النسبية

مثال (96) : البيانات التالية تعود الى المسافرين على متن الباخرة تايتانك وكان عدد النساء 462 ، وقد نجون 308 من الغرق و 154 هلكن فيما كانت التقديرات المناظرة لها من الرجال 851 ، 142 و 709 . جد ارجحيةغرق الرجال نسبة الى النساء؟

ويمكن تمثيل البيانات بالجدول الآتي :

	الناجون	الغارقون	العدد الكلي
الرجال	142	709	851
النساء	308	154	462
العدد الكلي	450	863	1313

الحل يدويا:

(154 /308 = 0.5)

(709/142 = 4.993)

Odds Ratio = (4.993 / 0.5) = 9.986

(154 / 462 = 0.3333)

(709 /851 = 0.8331)

Relative Risk = (0.8331/0.3333) = 2.5)

اي ان ارجحية غرق الرجال الى النساء هي اكبر بـ 10 مرا ت.

```
DATA e;
INPUT   group stat freq;
cards;
1 0 308
1 1 154
0 0 142
0 1 709
;proc logistic data=e;
model group=stat;
freq freq;
run;
```

```
                                                The SAS System
                The LOGISTIC Procedure

    Data Set: WORK.E
    Response Variable: stat
    Response Levels: 2
    Number of Observations: 4
    Frequency Variable: FREQ
    Link Function: Logit
                    Response Profile

                  Ordered
                   Value  stat    Count

                     1     0      851
                     2     1      462

        Model Fitting Information and Testing Global Null Hypothesis BETA=0

                          Intercept
                  Intercept      and
```

Dr. Firas

```
Criterion      Only      Covariates    Chi-Square for Covariates

AIC          1705.208     1374.674       .
SC           1710.388     1385.034       .
-2 LOG L     1703.208     1370.674     332.534 with 1 DF (p=0.0001)
Score           .            .         332.057 with 1 DF (p=0.0001)


                Analysis of Maximum Likelihood Estimates

           Parameter  Standard     Wald    Pr >      Standa     Odds
Variable   DF   Estimate     Error     Chi-Sq   Chi-Sq   Estimate   Ratio

INTERCPT   1     -0.7743     0.1014   58.2659  0.0001     .          .
stat       1      2.3012     0.1349  291.0707  0.0001       0.602383      9.986


          Association of Predicted Probabilities and Observed Responses

               Concordant = 55.5%        Somers' D = 0.500
               Discordant =  5.6%        Gamma     = 0.818
               Tied       = 38.9%        Tau-a     = 0.228
               (393162 pairs)            c         = 0.750
```

ويمكن الحصول على نفس النتائج باستعمال الآمر Proc freq.

مثال (97):

```
data s;
input g c count;
cards;
1 2 142
1 1 709
2 2 308
2 1 154
proc freq;
weight count;
tables g*c/cl riskdiff; ← الايعاز الاول لغرض الحصول على حدود الثقة والايعاز الثاني للحصول على الفرق بين نسبتي الخطورة
run;
```

```
                                The SAS System
                TABLE OF G BY C

                G       C

                Frequency,
                Percent ,
                Row Pct ,
                Col Pct ,     1,     2, Total
                ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
                    1 ,   709 ,   142 ,   851
                      , 54.00 , 10.81 , 64.81
                      , 83.31 , 16.69 ,
                      , 82.16 , 31.56 ,
                ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
                    2 ,   154 ,   308 ,   462
                      , 11.73 , 23.46 , 35.19
                      , 33.33 , 66.67 ,
                      , 17.84 , 68.44 ,
                ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
                Total       863     450    1313
                          65.73   34.27  100.00

                     The SAS System
                STATISTICS FOR TABLE OF G BY C
```

Dr. Firas

Statistic	Value	ASE	95% Confidence Bounds	
Gamma	0.818	0.022	0.774	0.862
Kendall's Tau-b	0.503	0.025	0.454	0.552
Stuart's Tau-c	0.456	0.024	0.408	0.504
Somers' D C|R	0.500	0.025	0.450	0.550
Somers' D R|C	0.506	0.025	0.456	0.556
Pearson Correlation	0.503	0.025	0.454	0.552
Spearman Correlation	0.503	0.025	0.454	0.552
Lambda Asymmetric C|R	0.342	0.039	0.266	0.418
Lambda Asymmetric R|C	0.359	0.037	0.287	0.431
Lambda Symmetric	0.351	0.036	0.281	0.421
Uncertainty Coefficient C|R	0.197	0.020	0.157	0.237
Uncertainty Coefficient R|C	0.195	0.020	0.156	0.235
Uncertainty Coefficient Symmetric	0.196	0.020	0.157	0.236

Column 1 Risk Estimates

	Risk	ASE	95% Confidence Bounds (Asymptotic)		95% Confidence Bounds (Exact)	
Row 1	0.833	0.013	0.808	0.858	0.806	0.858
Row 2	0.333	0.022	0.290	0.376	0.290	0.378
Total	0.657	0.013	0.632	0.683	0.631	0.683
Difference (Row 1 - Row 2)	0.500	0.025	0.450	0.550		

The SAS System

Column 2 Risk Estimates

	Risk	ASE	95% Confidence Bounds (Asymptotic)		95% Confidence Bounds (Exact)	
Row 1	0.167	0.013	0.142	0.192	0.142	0.194
Row 2	0.667	0.022	0.624	0.710	0.622	0.710
Total	0.343	0.013	0.317	0.368	0.317	0.369
Difference (Row 1 - Row 2)	-0.500	0.025	-0.550	-0.450		

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Bounds	
Case-Control	**9.986**	7.666	13.008
Cohort (Col1 Risk)	**2.499**	2.189	2.853
Cohort (Col2 Risk)	0.250	0.213	0.295

Sample Size = 1313

مثال (98): اجريت تجربة لمعرفة تأثير جرع مختلفة على هلاك الفئران واستعملت 3 فئران لكل جرعة المطلوب تقدير نسبة الارجحية لزيادة الجرعة ؟

data dose;

Dr. Firas

```
input Dose Deaths Total;
datalines;
0 0 3
1 0 3
2 0 3
3 0 3
4 1 3
5 2 3
run;
option nodate nonumber;
proc logistic data=dose descending;
model Deaths/Total = Dose ;
run;
```

```
                    The SAS System
                  The LOGISTIC Procedure

      Data Set: WORK.DOSE
      Response Variable (Events): DEATHS
      Response Variable (Trials): TOTAL
      Number of Observations: 6
      Link Function: Logit

                      Response Profile

                  Ordered  Binary
                   Value  Outcome     Count

                     1  EVENT         3
                     2  NO EVENT      15

       Model Fitting Information and Testing Global Null Hypothesis BETA=0

                        Intercept
                Intercept      and
      Criterion      Only      Covariates    Chi-Square for Covariates

      AIC           18.220      12.072       .
      SC            19.111      13.853       .
      -2 LOG L      16.220       8.072       8.148 with 1 DF (p=0.0043)
      Score          .           .          5.794 with 1 DF (p=0.0161)


          Analysis of Maximum Likelihood Estimates
                                          Analysis of Maximum Likelihood Estimates

          Parameter  Standard    Wald        Pr >      Standar      Odds
Variable   DF   Estimate  Error    Chi-Square Chi-Square  Estimate       Ratio

INTERCPT   1    -9.4746  5.5677   2.8958     0.0888       .          .
DOSE       1     2.0804  1.2603   2.7249     0.0988       2.015631     8.008

         Association of Predicted Probabilities and Observed Responses
               Concordant = 88.9%        Somers' D = 0.867
               Discordant =  2.2%        Gamma     = 0.951
               Tied       =  8.9%        Tau-a     = 0.255
               (45 pairs)                c         = 0.933
```

ان الزياة في الجرعة تزيد ارجحية الهلاك عن بقائها حية بـ 8 مرات.

مثال (99): قدر ارجحية نجاح الطلاب في المدرسة النموذجية الى نسبة نجاح الطلاب في المدرسة الاعتيادية بأستعمال Odds Ratio اذا كان عدد الطلاب الراسبين والناجحين في المدرسة النموذجية 15 و 60 وفي المدرسة الاعتيادية 65 و 25 على التوالي؟

Dr.Fics

```
DATA EXAMPLE;
INPUT SCHOOL $ LEVEL $ freq;
CARDS;
E P 15
E F 60
N P 65
N F 25
PROC FREQ DATA=example;
WEIGHT freq;
TABLES SCHOOL*LEVEL / CL RISKDIFF;
RUN;
```

```
                                        The SAS System
              TABLE OF SCHOOL BY LEVEL

            SCHOOL    LEVEL

            Frequency,
            Percent ,
            Row Pct ,
            Col Pct ,F     ,P     , Total
            ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
            E      ,   60 ,   15 ,   75
                   , 36.36 , 9.09 , 45.45
                   , 80.00 , 20.00 ,
                   , 70.59 , 18.75 ,
            ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
            N      ,   25 ,   65 ,   90
                   , 15.15 , 39.39 , 54.55
                   , 27.78 , 72.22 ,
                   , 29.41 , 81.25 ,
            ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
            Total       85     80     165
                     51.52   48.48   100.00

                  The SAS System
          STATISTICS FOR TABLE OF SCHOOL BY LEVEL

                                        95%
Statistic                    Value    ASE    Confidence Bounds
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Gamma                        0.825   0.060     0.708   0.941
Kendall's Tau-b              0.520   0.066     0.391   0.649
Stuart's Tau-c               0.518   0.066     0.389   0.647

Somers' D C|R                0.522   0.066     0.393   0.652
Somers' D R|C                0.518   0.066     0.389   0.648

Pearson Correlation          0.520   0.066     0.391   0.649
Spearman Correlation         0.520   0.066     0.391   0.649

Lambda Asymmetric C|R        0.500   0.084     0.336   0.664
Lambda Asymmetric R|C        0.467   0.090     0.291   0.643
Lambda Symmetric             0.484   0.078     0.330   0.637

Uncertainty Coefficient C|R  0.206   0.055     0.098   0.315
Uncertainty Coefficient R|C  0.207   0.055     0.099   0.316
Uncertainty Coefficient Symmetric  0.207   0.055     0.098   0.315

                Column 1 Risk Estimates

                      95% Confidence Bounds  95% Confidence Bounds
              Risk    ASE      (Asymptotic)        (Exact)
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Row 1         0.800   0.046     0.709   0.891      0.692   0.884
Row 2         0.278   0.047     0.185   0.370      0.189   0.382
Total         0.515   0.039     0.439   0.591      0.436   0.594
```

Dr. Firas

```
Difference     0.522    0.066     0.393    0.652
(Row 1 - Row 2)


                     The SAS System
                  Column 2 Risk Estimates

                     95% Confidence Bounds  95% Confidence Bounds
           Risk     ASE       (Asymptotic)          (Exact)
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Row 1         0.200   0.046      0.109   0.291       0.116   0.308
Row 2         0.722   0.047      0.630   0.815       0.618   0.811
Total        0.485    0.039     0.409    0.561       0.406    0.564

Difference    -0.522    0.066     -0.652   -0.393
(Row 1 - Row 2)


           Estimates of the Relative Risk (Row1/Row2)
                               95%
       Type of Study          Value     Confidence Bounds
       ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
       Case-Control          10.400       5.012     21.581
       Cohort (Col1 Risk)     2.880       2.026      4.094
       Cohort (Col2 Risk)     0.277       0.173      0.443

       Sample Size = 165
```

ارجحية النجاح تساوي 10.4 مرة اكبر في النموذجية والخطورة النسبية للمدرسة الاعتيادية = 2.8

ملاحظة: عندما يكون احتمال حصول حدث ما قليل فأن تقديرات نسبة الأرجحية والخطورة النسبية تكون متقاربة.

مثال (100): في دراسةعن مرض ما وجد ان عدد النساء المصابات 6 وعدد النساء غير المصابة 354 فيما كانت التقديرات المناظرة لها في الرجال 70 و 1010 ، جد ارجحية اصابة الرجال عن النساء؟

```
data s;
input g c count;
cards;
1 1 354
1 2 6
2 1 1010
2 2 70
proc freq;
weight count;
tables g*c/cl riskdiff;
run;
```

```
                                  The SAS System

                                 TABLE OF G BY C

            G       C

            Frequency,
            Percent  ,
            Row Pct  ,
            Col Pct  ,      1,      2,  Total
            ƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆ
                 1 ,    354 ,     6 ,    360
                   , 24.58 ,   0.42 ,  25.00
```

, 98.33 , 1.67 ,
, 25.95 , 7.89 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
2 , 1010 , 70 , 1080
, 70.14 , 4.86 , 75.00
, 93.52 , 6.48 ,
, 74.05 , 92.11 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 1364 76 1440

94.72 5.28 100.00

The SAS System
STATISTICS FOR TABLE OF G BY C

Statistic	Value	ASE	95% Confidence Bounds	
Gamma	0.607	0.136	0.341	0.873
Kendall's Tau-b	0.093	0.018	0.059	0.128
Stuart's Tau-c	0.036	0.008	0.021	0.051
Somers' D C|R	0.048	0.010	0.028	0.068
Somers' D R|C	0.181	0.033	0.116	0.246
Pearson Correlation	0.093	0.018	0.059	0.128
Spearman Correlation	0.093	0.018	0.059	0.128
Lambda Asymmetric C|R	0.000	0.000	0.000	0.000
Lambda Asymmetric R|C	0.000	0.000	0.000	0.000
Lambda Symmetric	0.000	0.000	0.000	0.000
Uncertainty Coefficient C|R	0.026	0.011	0.004	0.048
Uncertainty Coefficient R|C	0.010	0.004	0.001	0.018
Uncertainty Coefficient Symmetric	0.014	0.006	0.002	0.026

Column 1 Risk Estimates

	Risk	ASE	95% Confidence Bounds (Asymptotic)		95% Confidence Bounds (Exact)	
Row 1	0.983	0.007	0.970	0.997	0.964	0.994
Row 2	0.935	0.007	0.921	0.950	0.919	0.949
Total	0.947	0.006	0.936	0.959	0.934	0.958
Difference (Row 1 - Row 2)	0.048	0.010	0.028	0.068		

~~The SAS System~~

~~Column 2 Risk Estimates~~

	~~Risk~~	~~ASE~~	~~95% Confidence Bounds (Asymptotic)~~		~~95% Confidence Bounds (Exact)~~	
~~Row 1~~	~~0.017~~	~~0.007~~	~~0.003~~	~~0.030~~	~~0.006~~	~~0.036~~
~~Row 2~~	~~0.065~~	~~0.007~~	~~0.050~~	~~0.079~~	~~0.051~~	~~0.081~~
~~Total~~	~~0.053~~	~~0.006~~	~~0.041~~	~~0.064~~	~~0.042~~	~~0.066~~
~~Difference (Row 1 - Row 2)~~	~~0.048~~	~~0.010~~	~~-0.068~~	~~-0.028~~		

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Bounds	
Case-Control	4.089	1.761	9.495
Cohort (Col1 Risk)	1.051	1.030	1.073

Dr Firas

Cohort (Col2 Risk)	0.257	0.113	0.587

Sample Size = 1440

ارجحية اصابة الرجال مقارنة بالنساء تساوي 4.089 مرة.

مثال (101) : في دراسة شملت عينة من الاغنام في احد المجازر تضمنت اغنام بأعمار مختلفة وكانت بينها ذكور واناث ولمدة اربعة فصول ، لغرض دراسة تأثير العوامل المذكورة على اصابتها بالثايليريا والبابيزيا والانابلازما من خلال تقدير ارجحية الاصابة.

```
data f;
input no sex age TH BAB ANA SE;
if age lt 1.1 then age=1 ;
if age gt 1  then age =2;
if se=3 then se=2;
if se=4 then se=1;
cards;
1    1    2    0    1    0    4
2    1    1    0    1    0    4
3    1    1    0    1    0    4
4    2    1.6  0    1    0    4
5    2    2    0    1    0    4
6    2    1    0    0    0    4
7    2    1    0    0    0    4
8    2    1    0    0    0    4
9    2    1    0    0    0    3
10   2    1    0    0    0    3
11   2    1    0    0    0    3
12   2    1    0    0    0    3
13   2    1    0    0    0    3
14   2    1.6  1    1    1    3
15   2    2    0    0    0    3
16   2    2    0    0    0    2
17   2    1    0    0    0    2
18   2    1    0    0    0    2
19   2    1.8  0    0    0    2
20   2    1.8  0    0    0    2
21   2    1.8  0    0    0    2
22   2    1.8  0    0    0    2
23   2    1    0    0    0    2
24   2    1.4  0    0    0    1
25   2    1    0    0    0    1
26   1    1.4  0    0    0    1
27   1    1    0    1    0    1
28   1    1.4  0    0    0    1
29   2    1    1    0    0    1
30   1    1.4  0    0    0    1
31   1    1.6  0    0    0    2
32   1    1.6  0    1    0    2
33   1    1.6  1    0    0    2
34   2    1.6  1    0    0    2
35   2    1.6  0    0    0    2
36   2    1.6  0    0    0    3
37   2    1.6  0    0    0    3
```

Dr Firas

```
38    2    1.6   0    0    0    3
39    2    1.6   0    0    0    3
40    2    1     0    0    0    3
41    1    1     1    1    1    2
42    1    1     1    1    1    1
43    2    2     1    1    0    3
44    2    2     1    1    0    1
45    2    1     0    0    0    1
46    2    1     0    0    0    1
47    2    1     0    0    0    1
48    2    1     0    0    0    2
49    2    1     0    0    0    2
50    2    1     0    0    0    2
51    2    1     0    0    0    2
52    2    1     0    0    0    4
53    2    1     0    0    0    4
54    1    1     0    0    0    4
55    1    1     0    0    0    4
56    2    1     0    0    0    4
57    2    1     0    0    0    4
58    2    1     0    0    0    4
59    1    1     0    0    0    4
60    1    2     0    0    0    4
;option ps=2000 ls=70 nodate nonumber;
 proc glm;classes sex age se;
 model   th = sex se age;
 means sex se age/LSD ;
 run;
 PROC LOGISTIC DATA=f DESCENDING;
 MODEL th  = sex age se;
 RUN;
 PROC LOGISTIC DATA=f DESCENDING;
 MODEL  ana = sex age se;
 RUN;
 PROC LOGISTIC DATA=f DESCENDING;
 MODEL  bab  = sex age se;
 RUN;
```

Dr Firas

```
                The SAS System

          General Linear Models Procedure
            Class Level Information

            Class   Levels   Values

            SEX       2    1 2

            AGE       2    1 2

            SE        2    1 2


        Number of observations in data set = 60
                The SAS System

          General Linear Models Procedure

Dependent Variable: TH
Source              DF    Sum of Squares     F Value     Pr > F
```

Model	3	0.25889538	0.72	0.5419
Error	56	6.67443796		
Corrected Total	59	6.93333333		

R-Square	C.V.	TH Mean
0.037341	258.9254	0.13333333

Source	DF	Type I SS	F Value	Pr > F
SEX	1	0.06401515	0.54	0.4667
SE	1	0.10398400	0.87	0.3543
AGE	1	0.09089622	0.76	0.3862

Source	DF	Type III SS	F Value	Pr > F
SEX	1	0.08193122	0.69	0.4106
SE	1	0.05032913	0.42	0.5185
AGE	1	0.09089622	0.76	0.3862

The SAS System

General Linear Models Procedure

T tests (LSD) for variable: TH

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 56 MSE= 0.119186
Critical Value of T= 2.00
Least Significant Difference= 0.2019
WARNING: Cell sizes are not equal.
Harmonic Mean of cell sizes= 23.46667

Means with the same letter are not significantly different.

T Grouping	Mean	N	SEX
A	0.1875	16	1
A			
A	0.1136	44	2

The SAS System

General Linear Models Procedure

T tests (LSD) for variable: TH

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha= 0.05 df= 56 MSE= 0.119186
Critical Value of T= 2.00
Least Significant Difference= 0.1787
WARNING: Cell sizes are not equal.
Harmonic Mean of cell sizes= 29.96667

Means with the same letter are not significantly different.

T Grouping	Mean	N	SE
A	0.16129	31	2
A			
A	0.10345	29	1

Dr Firas

The SAS System

General Linear Models Procedure

T tests (LSD) for variable: TH

NOTE: This test controls the type I comparisonwise error rate
not the experimentwise error rate.

Alpha= 0.05 df= 56 MSE= 0.119186
Critical Value of T= 2.00
Least Significant Difference= 0.1802
WARNING: Cell sizes are not equal.
Harmonic Mean of cell sizes= 29.46667

Means with the same letter are not significantly different.

T Grouping	Mean	N	AGE
A	0.19231	26	2
A			
A	0.08824	34	1

The SAS System

The LOGISTIC Procedure

Data Set: WORK.F
Response Variable: TH
Response Levels: 2
Number of Observations: 60
Link Function: Logit

Response Profile

Ordered Value	TH	Count
1	1	8
2	0	52

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	49.121	52.873	.
SC	51.215	61.250	.
-2 LOG L	47.121	44.873	2.248 with 3 DF (p=0.5226)
Score	.	.	2.240 with 3 DF (p=0.5240)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate
INTERCPT	1	-2.6044	1.9919	1.7095	0.1911	.
SEX	1	-0.7679	0.8946	0.7368	0.3907	-0.188802
AGE	1	0.7132	0.8174	0.7612	0.3829	0.196496
SE	1	0.6069	0.8940	0.4608	0.4972	0.168610

Analysis of Maximum Likelihood Estimates

Variable	Odds Ratio

INTERCPT .
SEX 0.464
AGE 2.041
SE 1.835

ارجحية اصابة الذكور الى الاناث 0.464 والاصابة بعمر اكبر من سنة الى عمر سنة فأقل 2.041 والاصابة في بقية الفصول الى الشتاء 1.835.

Association of Predicted Probabilities and Observed Responses

Concordant = 60.1% Somers' D = 0.337
Discordant = 26.4% Gamma = 0.389
Tied = 13.5% Tau-a = 0.079
(416 pairs) c = 0.668

The SAS System

The LOGISTIC Procedure

Data Set: WORK.F
Response Variable: ANA
Response Levels: 2
Number of Observations: 60
Link Function: Logit

Response Profile

Ordered
Value ANA Count

1 1 3
2 0 57

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	25.822	27.373	.
SC	27.916	35.750	.
-2 LOG L	23.822	19.373	4.449 with 3 DF (p=0.2169)
Score	.	.	4.517 with 3 DF (p=0.2108)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate
INTERCPT	1	-0.1051	2.9743	0.0012	0.9718	.
SEX	1	-2.8140	1.5292	3.3863	0.0657	-0.691868
AGE	1	-1.2761	1.4675	0.7562	0.3845	-0.351580
SE	1	2.0609	1.5714	1.7201	0.1897	0.572589

Analysis of Maximum Likelihood Estimates

Variable	Odds Ratio
INTERCPT	.

Dr Firas

SEX 0.060
AGE 0.279
SE 7.853

Association of Predicted Probabilities and Observed Responses

Concordant = 71.3% Somers' D = 0.538
Discordant = 17.5% Gamma = 0.605
Tied = 11.1% Tau-a = 0.052
(171 pairs) c = 0.769

The SAS System

The LOGISTIC Procedure

Data Set: WORK.F
Response Variable: BAB
Response Levels: 2
Number of Observations: 60
Link Function: Logit

Response Profile

Ordered Value	BAB	Count
1	1	12
2	0	48

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	62.048	59.155	.
SC	64.143	67.532	.
-2 LOG L	60.048	51.155	8.894 with 3 DF (p=0.0307)
Score	.	.	9.337 with 3 DF (p=0.0251)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate
INTERCPT	1	0.8235	1.6041	0.2636	0.6077	.
SEX	1	-1.5243	0.7422	4.2176	0.0400	-0.374775
AGE	1	0.9224	0.7490	1.5166	0.2181	0.254140
SE	1	-0.7579	0.8003	0.8967	0.3437	-0.210560

Analysis of Maximum Likelihood Estimates

Variable	Odds Ratio
INTERCPT	.
SEX	0.218
AGE	2.515
SE	0.469

Association of Predicted Probabilities and Observed Responses

Concordant = 77.4% Somers' D = 0.627
Discordant = 14.8% Gamma = 0.680
Tied = 7.8% Tau-a = 0.204

Dr Firas

(576 pairs) c = 0.813

ونلاحظ انه في حالة تقدير نسبة الارجحية لجنس الحيوان فقط فأن نسبة الارجحية المقدرة ستختلف اذ لايجري تعديل لها بسبب العاملين الآخرين.

```
PROC LOGISTIC DATA=f DESCENDING;
 MODEL th  = sex ;
RUN;
```

The SAS System

The LOGISTIC Procedure

Data Set: WORK.F
Response Variable: TH
Response Levels: 2
Number of Observations: 60
Link Function: Logit

Response Profile

Ordered Value	TH	Count
1	1	8
2	0	52

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	49.121	50.599	.
SC	51.215	54.788	.
-2 LOG L	47.121	46.599	0.522 with 1 DF (p=0.4700)
Score	.	.	0.554 with 1 DF (p=0.4567)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate
INTERCPT	1	-0.8786	1.3663	0.4135	0.5202	.
SEX	1	-0.5878	0.7974	0.5433	0.4611	-0.144516

Analysis of Maximum Likelihood Estimates

Variable	**Odds Ratio**
INTERCPT	.
SEX	**0.556**

Association of Predicted Probabilities and Observed Responses

Concordant = 28.1% Somers' D = 0.125
Discordant = 15.6% Gamma = 0.286
Tied = 56.3% Tau-a = 0.029

(416 pairs) c = 0.563

Dr. Firas

يمكن الحصول على نفس التقدير وذلك بحساب عدد الحيوانات المصابة والتي تساوي عدد الذكور في نسبة الاصابة 16 *0.1875 = 3 مصاب وغير المصاب 16 – 3 = 13 وكذلك بالنسبة للأناث 44*0.1136 = 5 مصابة وغير المصابة 44 – 5 = 39.

ثم نرتب البيانات ونطبق الايعاز التالي للحصول على نفس نسبة الارجحية وهي غير معدلة.

مثال (102):

```
DATA e;
INPUT   sex th freq;
cards;
1 0 3
1 1 13
0 0 5
0 1 39
PROC LOGISTIC DATA=e DESCENDING;
WEIGHT freq;
MODEL sex = th;run;
```

```
                                                The SAS System

          The LOGISTIC Procedure

          Data Set: WORK.E
          Response Variable: SEX
          Response Levels: 2
          Number of Observations: 4
          Weight Variable: FREQ
          Sum of Weights: 60
          Link Function: Logit
                              Response Profile

                    Ordered                    Total
                     Value    SEX    Count      Weight

                         1     1        2    16.000000
                         2     0        2    44.000000
Model Fitting Information and Testing Global Null Hypothesis BETA=0

                       Intercept
                Intercept     and
 Criterion         Only     Covariates   Chi-Square for Covariates

 AIC             71.590     73.068       .
 SC              70.976     71.840       .
-2 LOG L         69.590     69.068       0.522 with 1 DF (p=0.4700)
 Score              .          .         0.554 with 1 DF (p=0.4567)

          Analysis of Maximum Likelihood Estimates

          Parameter   Standard   Wald   Pr >       Standardized      Odds
Variable  DF Estimate     Error   Chi-Square Chi-Square     Estimate     Ratio

INTERCPT   1 -0.5108     0.7303    0.4893   0.4843       .            .
TH         1 -0.5878     0.7974    0.5433   0.4611      -0.492653     0.556
Association of Predicted Probabilities and Observed Responses

                    Concordant = 25.0%        Somers' D = 0.000
                    Discordant = 25.0%        Gamma     = 0.000
                    Tied       = 50.0%        Tau-a     = 0.000
                    (4 pairs)                 c         = 0.500
```

Dr. Firas

مثال (103):في دراسة طبية لمعرفة العلاقة بين كل من العمر والاسبيرين والاصابة بنوبة قلبية تضمنت ثلاثة مجموعات وبأعمار مختلفة 40 – 49 ، 50 – 59 ، 60 – 69 ، 70 – 84 . وكانت البيانات كما موضح ادناه:

العمر 40 – 49 سنة

تعاطى الاسبيرين	لم يتعاطى الاسبرين Placebo	
27	24	اصيب بنوبة قلبية
4500	4500	لم يصاب بنوبة قلبية

العمر 50 – 59 سنة

تعاطى الاسبيرين	لم يتعاطى الاسبرين Placebo	
51	87	اصيب بنوبة قلبية
3674	3638	لم يصاب بنوبة قلبية

العمر 60 – 69 سنة

تعاطى الاسبيرين	لم يتعاطى الاسبرينPlacebo	
39	84	اصيب بنوبة قلبية
2006	1961	لم يصاب بنوبة قلبية

العمر 70 – 84 سنة

تعاطى الاسبيرين	لم يتعاطى الاسبرين Placebo	
22	44	اصيب بنوبة قلبية
718	696	لم يصاب بنوبة قلبية

```
DATA e;
 INPUT   group drug freq;
DATALINES;
1 0 27
1 1 4500
0 0 24
0 1 4500
PROC LOGISTIC DATA=e DESCENDING;
WEIGHT freq;
MODEL group = drug;run;
```

```
                                The SAS System
                          The LOGISTIC Procedure

      Data Set: WORK.E
      Response Variable: GROUP
      Response Levels: 2
      Number of Observations: 4
      Weight Variable: FREQ
      Sum of Weights: 9051
      Link Function: Logit


                        Response Profile

                    Ordered                  Total
                      Value  GROUP    Count      Weight
```

Dr. Firas

1 1 2 4527.0000
2 0 2 4524.0000

Model Fitting Information and Testing Global Null Hypothesis BETA=0.

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	12549.349	12551.174	.
SC	12548.736	12549.946	.
-2 LOG L	12547.349	12547.174	0.176 with 1 DF (p=0.6752)
Score	.	.	0.175 with 1 DF (p=0.6753)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Stand Error	Wald Chi-Square	Pr > Chi-Square	Stand Estimate	Odds Ratio
INTERCPT	1	0.1178	0.2805	0.1763	0.6746	.	.
DRUG	1	-0.1178	0.2813	0.1753	0.6755	-0.266988	**0.889**

ان اختبار نسبة الارجحية يشير الى عدم معنوية التقديرفي مجموعة 40 – 49 سنة.

Association of Predicted Probabilities and Observed Responses

Concordant = 25.0% Somers' D = 0.000
Discordant = 25.0% Gamma = 0.000
Tied = 50.0% Tau-a = 0.000
(4 pairs) c = 0.500

```
DATA e;
INPUT   group drug freq;
cards;
1 0 51
1 1 3674
0 0 87
0 1 3638
PROC LOGISTIC DATA=e DESCENDING;
WEIGHT freq;
MODEL group = drug;run;
```

The SAS System

The LOGISTIC Procedure

Data Set: WORK.E
Response Variable: GROUP
Response Levels: 2
Number of Observations: 4
Weight Variable: FREQ
Sum of Weights: 7450
Link Function: Logit

Response Profile

Ordered Value	GROUP	Total Count	Weight
1	1	2	3725.0000
2	0	2	3725.0000

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates

Dr Firas

AIC 10329.893 10322.215 .
SC 10329.279 10320.988 .
-2 LOG L 10327.893 10318.215 9.678 with 1 DF (p=0.0019)
Score . . 9.569 with 1 DF (p=0.0020)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Sq	Pr > Chi-Sq	Stand Estimate	Odds Ratio
INTERCPT	1	-0.5341	0.1764	9.1713	0.0025	.	.
DRUG	1	0.5439	0.1779	9.3481	0.0022	2.014987	**1.723**

تبين بأن ارجحية حصول النوبة القلبية للذين لم يتناولوا الاسبرين كانت معنوية وهي اعلى بحوالي 1.7 مرة مقارنة بالمجموعة التي تناولت الاسبرين للمجموعة التي تتراوح اعمارهم من 50 – 59 سنة

Association of Predicted Probabilities and Observed Responses

Concordant = 25.0% Somers' D = 0.000
Discordant = 25.0% Gamma = 0.000
Tied = 50.0% Tau-a = 0.000
(4 pairs) c = 0.500

```
DATA e;
INPUT   group drug freq;
DATALINES;
1 0 39
1 1 2006
0 0 84
0 1 1961
PROC LOGISTIC DATA=e DESCENDING;
WEIGHT freq;
MODEL group = drug;
RUN;
```

The SAS System

The LOGISTIC Procedure

Data Set: WORK.E
Response Variable: GROUP
Response Levels: 2
Number of Observations: 4
Weight Variable: FREQ
Sum of Weights: 4090
Link Function: Logit

Response Profile

Ordered Value	GROUP	Total Count	Weight
1	1	2	2045.0000
2	0	2	2045.0000

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	5671.944	5656.582	.
SC	5671.330	5655.354	.
-2 LOG L	5669.944	5652.582	17.362 with 1 DF (p=0.0001)
Score	.	.	16.974 with 1 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Dr Firas

Variable	DF	Parameter Estimate	Stand Error	Wald Chi-Sq	Pr > Chi-Sq	Standardized Estimate	Odds Ratio
INTERCPT	1	-0.7673	0.1938	15.6790	0.0001	.	.
DRUG	1	0.7899	0.1964	16.1853	0.0001	2.746427	**2.203**

تبين بأن ارجحية حصول النوبة القلبية للذين لم يتناولوا الاسبرين كانت معنوية وهي اعلى بحوالي 2 مرة مقارنة بالمجموعة التي تناولت الاسبرين للمجموعة التي تتراوح اعمارهم من 60 – 69 سنة

Association of Predicted Probabilities and Observed Responses

Concordant = 25.0%	Somers' D = 0.000
Discordant = 25.0%	Gamma = 0.000
Tied = 50.0%	Tau-a = 0.000
(4 pairs)	c = 0.500

وبنفس الطريقة نستخرج نسبة الارجحية للمجموعة الرابعة والتي تبلغ 2.06 . اذا فأن نسب الارجحية للأعمار 40 – 49 ، 50 -59 ، 60 – 69 ، 70 – 84 سنة هي 0.89 ، 1.72 ، 2.20 ، 2.06 على التوالي ، وبناءا على هذه النتائج فأننا نجد ان استعمال الاسبيرين ذو تأثير كبير في الاعمار الاكبر مع ان اعمار المجموعة الرابعة حصل انخفاض قليل فيها كما يمكن ان نوصي بعدم استعمال الاسبيرين للأعمار اقل من 50 سنة.

مثال (104) : وجد بين مجموعة من مرضى السرطان ان 72 من 129 لديهم صلع في مقدمة الرأس او في قمة الرأس مقارنة بمجموعة السيطرة اذ كان عدد الرجال الذين لديهم صلع هو 82 من اصل 182. كما موضح ادناه:

	عدد المصابين	مجموعة السيطرة	المجموع
اصلع	72	82	154
اعتيادي	55	57	112
المجموع	137	129	266

هنا يمكن تقدير ارجحية حصول الصلع لدى المصابين بالسرطان ولكن لايمكن تقدير ارجحية الاصابة بالسرطان لدى المصابين بالصلع بسبب ان الحالة تخضع لتصميم الحالة القياسي. اي ان الصلع يعد حدث من الزمن الماضي.

```
DATA example;
INPUT Group $ CASE $ freq;
CARDS;
G1 0     82
G1 1     57
G2 0     72
G2 1     55
PROC FREQ DATA=example;
WEIGHT freq;
TABLES Group*case / CL RISKDIFF RELRISK RDIFF;
RUN;
```

Dr. Firas

على العكس من ذلك يمكن دراسة العلاقة بين الصلع والاصابة بامراض القلب وذلك بتقدير الخطورة النسبية.

مثال (105):

	مصاب	سليم	المجموع
اصلع	(%9.4) 127	(%90.6)1224	1351
اعتيادي	(6.7) 548	(%93.3) 7611	8159
المجموع	675	8835	9510

Relative Risk = 1.4=9.4/6.7

```
data s;
input g c count;
cards;
1 1 127
1 2 1224
2 1 548
2 2 7611
proc freq;
weight count;
tables g*c/cl riskdiff;
run;
```

```
                                The SAS System
                        TABLE OF G BY C

                G       C

                Frequency,
                Percent ,
                Row Pct ,
                Col Pct ,     1,     2, Total
                ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
                    1 ,   127 ,  1224 ,  1351
                      ,  1.34 , 12.87 , 14.21
                      ,  9.40 , 90.60 ,
                      , 18.81 , 13.85 ,
                ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
                    2 ,   548 ,  7611 ,  8159
                      ,  5.76 , 80.03 , 85.79
                      ,  6.72 , 93.28 ,
                      , 81.19 , 86.15 ,
                ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
                Total      675    8835    9510
                         7.10   92.90  100.00

                      The SAS System
                STATISTICS FOR TABLE OF G BY C
                                        95%
        Statistic                 Value    ASE   Confidence Bounds
        ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
        Gamma                    0.181   0.050    0.083   0.279
        Kendall's Tau-b          0.036   0.011    0.014   0.059
        Stuart's Tau-c           0.013   0.004    0.005   0.021

        Somers' D C|R            0.027   0.008    0.010   0.043
        Somers' D R|C            0.050   0.015    0.019   0.080

        Pearson Correlation      0.036   0.011    0.014   0.059
        Spearman Correlation     0.036   0.011    0.014   0.059

        Lambda Asymmetric C|R    0.000   0.000    0.000   0.000
```

Dr Firas

Lambda Asymmetric R|C	0.000	0.000	0.000	0.000
Lambda Symmetric	0.000	0.000	0.000	0.000
Uncertainty Coefficient C|R	0.002	0.001	-0.000	0.005
Uncertainty Coefficient R|C	0.002	0.001	-0.000	0.003
Uncertainty Coefficient Symmetric	0.002	0.001	-0.000	0.004

Column 1 Risk Estimates

		95% Confidence Bounds (Asymptotic)		95% Confidence Bounds (Exact)	
Risk	ASE				
Row 1	0.094	0.008	0.078	0.110	0.079	0.111
Row 2	0.067	0.003	0.062	0.073	0.062	0.073
Total	0.071	0.003	0.066	0.076	0.066	0.076
Difference (Row 1 - Row 2)	0.027	0.008	0.010	0.043		

The SAS System

Column 2 Risk Estimates

		95% Confidence Bounds (Asymptotic)		95% Confidence Bounds (Exact)	
Risk	ASE				
Row 1	0.906	0.008	0.890	0.922	0.889	0.921
Row 2	0.933	0.003	0.927	0.938	0.927	0.938
Total	0.929	0.003	0.924	0.934	0.924	0.934
Difference (Row 1 - Row 2)	-0.027	0.008	-0.043	-0.010		

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Bounds	
Case-Control	1.441	1.177	1.764
Cohort (Col1 Risk)	1.400	1.164	1.683
Cohort (Col2 Risk)	0.971	0.954	0.989

ان الخطورة النسبية للاشخاص الذين ليس لديهم صلع اعلى وبصورة معنوية لحصول امراض القلب مقارنة بالاشخاص الذين لديهم صلع.

Sample Size = 9510

مثال (106): في دراسة تضمنت حالتين مرضية 50 حالة حادة و50 حالة شبه حادة وقد استعمل عقارين لدراسة ارجحيتهما في الحالتين وكان عدد المرضى الذين تماثلوا للشفاء في المجموعة الاولى العقار الاول 16 والذين لم يستجيبوا للعلاج 9 وللعقار الثاني 20 و 5 فيما كانت الاعداد المناظرة لها في المجموعة الثانية 14 ، 11 و 17 ، 8 قارن بين افضلية العقارين ضمن كل حالة؟

```
data temp;
input clinic treat response count;
cards;
1 1 1 9
1 1 2 16
1 2 1 5
1 2 2 20
2 1 1 11
2 1 2 14
2 2 1 8
```

2 2 2 17
proc freq data=temp;
weight count;by clinic;
table clinic*treat*response/ cmh relrisk;
run;

The SAS System

---------- CLINIC=1 ----------

TABLE 1 OF TREAT BY RESPONSE
CONTROLLING FOR CLINIC=1

TREAT RESPONSE

Frequency, 1, 2, Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
1 , 9 , 16 , 25
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
2 , 5 , 20 , 25
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 14 36 50

The SAS System

---------- CLINIC=1 ----------

SUMMARY STATISTICS FOR TREAT BY RESPONSE
CONTROLLING FOR CLINIC

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	1.556	0.212
2	Row Mean Scores Differ	1	1.556	0.212
3	General Association	1	1.556	0.212

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Bounds	
Case-Control (Odds Ratio)	Mantel-Haenszel	2.250	0.629	8.047
	Logit	2.250	0.628	8.057
Cohort (Col1 Risk)	Mantel-Haenszel	1.800	0.715	4.533
	Logit	1.800	0.702	4.618
Cohort (Col2 Risk)	Mantel-Haenszel	0.800	0.563	1.136
	Logit	0.800	0.562	1.139

The confidence bounds for the M-H estimates are test-based.

Total Sample Size = 50

يتبن بأن الخطورة النسبية لعدم شفاء المرض المزمن للعقار الاول 1.8 مرة مقارنة بالمرضى الذين تناولوا العلاج الثاني.

---------- CLINIC=2 ----------

TABLE 1 OF TREAT BY RESPONSE
CONTROLLING FOR CLINIC=2

TREAT RESPONSE

Frequency, 1, 2, Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
1 , 11 , 14 , 25
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
2 , 8 , 17 , 25

Dr Firas

```
                    fffffffff^fffffffff^fffffffff^
                    Total       19      31      50

                              The SAS System
---------------------------------------- CLINIC=2 ----------------------------------

                    SUMMARY STATISTICS FOR TREAT BY RESPONSE
                         CONTROLLING FOR CLINIC

                  Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

                 Statistic  Alternative Hypothesis   DF     Value     Prob
                 fffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffff
                   1        Nonzero Correlation      1     0.749    0.387
                   2        Row Mean Scores Differ   1     0.749    0.387
                   3        General Association      1     0.749    0.387


                     Estimates of the Common Relative Risk (Row1/Row2)
                                                  95%
                 Type of Study   Method          Value   Confidence Bounds
                 fffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffff
                 Case-Control    Mantel-Haenszel  1.670     0.523    5.332
                  (Odds Ratio)   Logit            1.670     0.527    5.290

                 Cohort          Mantel-Haenszel  1.375     0.668    2.829
                  (Col1 Risk)    Logit            1.375     0.668    2.832

                 Cohort          Mantel-Haenszel  0.824     0.530    1.278
                  (Col2 Risk)    Logit            0.824     0.531    1.278

                 The confidence bounds for the M-H estimates are test-based.

                 Total Sample Size = 50
```

الخطورة النسبية لعدم شفاء مجموعة المرض المزمن 1.37 مقارنة بالمجموعة الثانية.

ويمكن استعمال الايعاز التالي للحصول على نفس النتائج :

```
proc freq data=temp;
weight count;by clinic;
table clinic*treat*response/  CL RISKDIFF;
run;
```

يمكن ايضا المقارنة بين الشفاء في الحالات الحادة والحالات شبه الحادة بأستعمال الايعاز الآتي:

```
proc freq data=temp;
 weight count;
 table treat*clinic*response/ cmh;
run;
```

مثال (107): في دراسة لمعرفة تأثير بعض العوامل على وفيات المواليد ، تضمنت البيانات عدد المواليد الهالكة وعدد المواليد الكلي وهل الام مدخنة ام لا (1= مدخنة ، 2 = غير مدخنة) وعمر الام (اكبر من 40 سنة = 1، واصغر من 40 سنة = 2) وطول مدة الحمل (اقل من 8 اشهر = 1 ، واكثر من 8 اشهر = 2).

```
data mortal;
```

Dr Firas

```
 input deaths tbirths cigs age gestpd;
 agexcigs=age*cigs;
 agexgest=age*gestpd;
cards;
50 365  1 1 1
9  49   2 1 1
41 188  1 2 1
4  15   2 2 1
24 4036 1 1 2
6  465  2 1 2
14 1508 1 2 2
1  125  2 2 2;
 proc genmod data=mortal;
  model deaths/tbirths=cigs age gestpd / link=logit
                        dist=binomial;
run;
proc genmod data=mortal;
  model deaths/tbirths = cigs age gestpd
              age*cigs age*gestpd / link=logit
                        dist=binomial;
```

لغرض الحصول على نسبة الارجحية والتعديل للتداخلات يجب كتابة الايعاز الثاني بطريقة اخرى وكالآتي:

```
run;
proc logistic data=mortal;
  model deaths/tbirths = cigs age gestpd agexcigs agexgest;
run;
```

The GENMOD Procedure

Model Information

Description	Value
Data Set	WORK.MORTAL
Distribution	BINOMIAL
Link Function	LOGIT
Dependent Variable	DEATHS
Dependent Variable	TBIRTHS
Observations Used	8
Number Of Events	149
Number Of Trials	6751

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	4	1.3694	0.3424
Scaled Deviance	4	1.3694	0.3424
Pearson Chi-Square	4	1.3794	0.3449
Scaled Pearson X2	4	1.3794	0.3449
Log Likelihood	.	-541.4166	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	1	0.5643	0.4830	1.3648	0.2427
CIGS	1	0.4162	0.2621	2.5211	0.1123
AGE	1	0.4866	0.1805	7.2683	0.0070
GESTPD	1	-3.2878	0.1847	316.8595	0.0001
SCALE	0	1.0000	0.0000	.	.

Dr Firas

NOTE: The scale parameter was held fixed.

تبين من النتائج بأن التدخين لم يؤثر على نسبة المواليد الهالكة في هذه الدراسة.

The GENMOD Procedure

Model Information

Description	Value
Data Set	WORK.MORTAL
Distribution	BINOMIAL
Link Function	LOGIT
Dependent Variable	DEATHS
Dependent Variable	TBIRTHS
Observations Used	8
Number Of Events	149
Number Of Trials	6751

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	2	0.6228	0.3114
Scaled Deviance	2	0.6228	0.3114
Pearson Chi-Square	2	0.6299	0.3150
Scaled Pearson X2	2	0.6299	0.3150
Log Likelihood	.	-541.0432	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	1	-0.3816	1.1987	0.1013	0.7502
CIGS	1	0.8996	0.7985	1.2690	0.2599
AGE	1	1.1968	0.8516	1.9750	0.1599
GESTPD	1	-2.9851	0.5564	28.7855	0.0001
CIGS*AGE	1	-0.3765	0.5997	0.3942	0.5301
AGE*GESTPD	1	-0.2203	0.3864	0.3252	0.5685
SCALE	0	1.0000	0.0000	.	.

كان لطول مدة الحمل تأثيرا معنويا على نسبة المواليد الهالكة.

NOTE: The scale parameter was held fixed.

The LOGISTIC Procedure

Data Set: WORK.MORTAL
Response Variable (Events): DEATHS
Response Variable (Trials): TBIRTHS
Number of Observations: 8
Link Function: Logit

Response Profile

Ordered Value	Binary Outcome	Count
1	EVENT	149
2	NO EVENT	6602

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates

Dr. Firas
Dr. Firas

```
AIC          1433.110     1094.086       .
SC           1439.927     1134.991       .
-2 LOG L     1431.110     1082.086     349.023 with 5 DF (p=0.0001)
Score           .            .         717.554 with 5 DF (p=0.0001)

            Analysis of Maximum Likelihood Estimates

          Parameter  Standard    Wald      Pr >      Standardized    Odds
Variable    DF  Estimate     Error  Chi-Square  Chi-Square  Estimate     Ratio

INTERCPT    1   -0.3816     1.1987    0.1013      0.7502     .           .
CIGS        1    0.8996     0.7985    1.2690      0.2599    0.146709     2.459
AGE         1    1.1968     0.8516    1.9750      0.1599    0.293633     3.310
GESTPD      1   -2.9851     0.5564   28.7855      0.0001   -0.474287     0.051
AGEXCIGS    1   -0.3765     0.5997    0.3942      0.5301   -0.124956     0.686
AGEXGEST    1   -0.2203     0.3864    0.3252      0.5685   -0.112031     0.802
```

تقديرات نسبة الارجحية لهلاك المواليد حسب العوامل المدروسة.

```
        Association of Predicted Probabilities and Observed Responses

              Concordant = 76.0%        Somers' D = 0.665
              Discordant =  9.4%        Gamma     = 0.779
              Tied       = 14.6%        Tau-a     = 0.029
              (983698 pairs)            c         = 0.833
```

مثال (108) : في دراسة لمعرفة الاختلافات في الاصابة بالسرطان بسبب التدخين ونوع المركزتم تقديرنسب الارجحية.

```
data cmh;
input center smoke cancer count @@;
cards;
1 1 1 126   1 1 2 100   1 2 1  35   1 2 2  61
2 1 1 908   2 1 2 688   2 2 1 497   2 2 2 807
3 1 1 913   3 1 2 747   3 2 1 336   3 2 2 598
4 1 1 235   4 1 2 172   4 2 1  58   4 2 2 121
5 1 1 402   5 1 2 308   5 2 1 121   5 2 2 215
6 1 1 182   6 1 2 156   6 2 1  72   6 2 2  98
7 1 1  60   7 1 2  99   7 2 1  11   7 2 2  43
8 1 1 104   8 1 2  89   8 2 1  21   8 2 2  36
;
proc freq data=cmh;
  weight count;by center;
  table center*smoke*cancer/ cmh;
run;
```

بعد تنفيذ البرنامج سنحصل على تقيرات لنسبة الارجحية لجميع المراكز ووفقا لها يمكن ترتيب المراكز.

```
                    Odds Ratio by Center

---------------------------------------- CENTER=1 ----------------------------------

                  TABLE 1 OF SMOKE BY CANCER
                  CONTROLLING FOR CENTER=1

                SMOKE    CANCER

                Frequency,     1,     2, Total
                fffffffff^fffffff^fffffff^
                     1 ,   126 ,   100 ,   226
```

Dr. Firas

ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
2 , 35 , 61 , 96
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 161 161 322

Odds Ratio by Center
---------- CENTER=1 ----------

SUMMARY STATISTICS FOR SMOKE BY CANCER
CONTROLLING FOR CENTER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	10.002	0.002
2	Row Mean Scores Differ	1	10.002	0.002
3	General Association	1	10.002	0.002

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Bounds	
Case-Control (Odds Ratio)	Mantel-Haenszel	2.196	1.349	3.576
	Logit	2.196	1.343	3.590
Cohort (Col1 Risk)	Mantel-Haenszel	1.529	1.175	1.990
	Logit	1.529	1.146	2.041
Cohort (Col2 Risk)	Mantel-Haenszel	0.696	0.556	0.871
	Logit	0.696	0.564	0.860

The confidence bounds for the M-H estimates are test-based.

Total Sample Size = 322

Odds Ratio by
---------- CENTER=2 ----------

TABLE 1 OF SMOKE BY CANCER
CONTROLLING FOR CENTER=2

SMOKE CANCER

Frequency, 1, 2, Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
1 , 908 , 688 , 1596
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
2 , 497 , 807 , 1304
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 1405 1495 2900

Odds Ratio by Center
---------- CENTER=2 ----------

SUMMARY STATISTICS FOR SMOKE BY CANCER
CONTROLLING FOR CENTER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	101.292	0.001
2	Row Mean Scores Differ	1	101.292	0.001
3	General Association	1	101.292	0.001

Dr. Firas

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Bounds	
Case-Control (Odds Ratio)	Mantel-Haenszel	2.143	1.847	2.486
	Logit	2.143	1.846	2.488
Cohort (Col1 Risk)	Mantel-Haenszel	1.493	1.381	1.614
	Logit	1.493	1.376	1.619
Cohort (Col2 Risk)	Mantel-Haenszel	0.697	0.649	0.747
	Logit	0.697	0.649	0.748

The confidence bounds for the M-H estimates are test-based.

Total Sample Size = 2900

Odds Ratio by Center

---------- CENTER=3 ----------

TABLE 1 OF SMOKE BY CANCER
CONTROLLING FOR CENTER=3

SMOKE	CANCER		
Frequency,	1,	2,	Total
1	913	747	1660
2	336	598	934
Total	1249	1345	2594

Odds Ratio by Center

---------- CENTER=3 ----------

SUMMARY STATISTICS FOR SMOKE BY CANCER
CONTROLLING FOR CENTER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	86.627	0.001
2	Row Mean Scores Differ	1	86.627	0.001
3	General Association	1	86.627	0.001

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Bounds	
Case-Control (Odds Ratio)	Mantel-Haenszel	2.175	1.847	2.562
	Logit	2.175	1.844	2.565
Cohort (Col1 Risk)	Mantel-Haenszel	1.529	1.398	1.672
	Logit	1.529	1.389	1.683
Cohort (Col2 Risk)	Mantel-Haenszel	0.703	0.653	0.757
	Logit	0.703	0.654	0.755

The confidence bounds for the M-H estimates are test-based.

Total Sample Size = 2594

Odds Ratio by Center

---------- CENTER=4 ----------

TABLE 1 OF SMOKE BY CANCER
CONTROLLING FOR CENTER=4

Dr Firas

SMOKE CANCER

Frequency, 1, 2, Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
1 , 235 , 172 , 407
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
2 , 58 , 121 , 179
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 293 293 586

Odds Ratio by Center
---------- CENTER=4 ----------

SUMMARY STATISTICS FOR SMOKE BY CANCER
CONTROLLING FOR CENTER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic Alternative Hypothesis DF Value Prob
ƒƒ
1 Nonzero Correlation 1 31.871 0.001
2 Row Mean Scores Differ 1 31.871 0.001
3 General Association 1 31.871 0.001

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Bounds	
Case-Control (Odds Ratio)	Mantel-Haenszel	2.850	1.981	4.100
	Logit	2.850	1.969	4.125
Cohort (Col1 Risk)	Mantel-Haenszel	1.782	1.458	2.178
	Logit	1.782	1.420	2.237
Cohort (Col2 Risk)	Mantel-Haenszel	0.625	0.531	0.736
	Logit	0.625	0.537	0.728

The confidence bounds for the M-H estimates are test-based.

Total Sample Size = 586
Odds Ratio by Center
----------CENTER=5----------

TABLE 1 OF SMOKE BY CANCER
CONTROLLING FOR CENTER=5

SMOKE CANCER

Frequency, 1, 2, Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
1 , 402 , 308 , 710
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
2 , 121 , 215 , 336
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 523 523 1046

Odds Ratio by Center
---------- CENTER=5 ----------

SUMMARY STATISTICS FOR SMOKE BY CANCER
CONTROLLING FOR CENTER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Dr. Firas

Statistic Alternative Hypothesis DF Value Prob

1	Nonzero Correlation	1	38.706	0.001
2	Row Mean Scores Differ	1	38.706	0.001
3	General Association	1	38.706	0.001

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Bounds	
Case-Control	Mantel-Haenszel	2.319	1.779	3.023
(Odds Ratio)	Logit	2.319	1.775	3.031
Cohort	Mantel-Haenszel	1.572	1.363	1.813
(Col1 Risk)	Logit	1.572	1.345	1.838
Cohort	Mantel-Haenszel	0.678	0.600	0.766
(Col2 Risk)	Logit	0.678	0.604	0.761

The confidence bounds for the M-H estimates are test-based.

Total Sample Size = 1046

Odds Ratio by Center

---------- CENTER=6 ----------

TABLE 1 OF SMOKE BY CANCER
CONTROLLING FOR CENTER=6

SMOKE CANCER

Frequency,	1,	2,	Total
1 ,	182 ,	156 ,	338
2 ,	72 ,	98 ,	170
Total	254	254	508

Odds Ratio by Center

---------- CENTER=6 ----------

SUMMARY STATISTICS FOR SMOKE BY CANCER
CONTROLLING FOR CENTER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic Alternative Hypothesis DF Value Prob

1	Nonzero Correlation	1	5.965	0.015
2	Row Mean Scores Differ	1	5.965	0.015
3	General Association	1	5.965	0.015

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Bounds	
Case-Control	Mantel-Haenszel	1.588	1.096	2.302
(Odds Ratio)	Logit	1.588	1.095	2.303
Cohort	Mantel-Haenszel	1.271	1.049	1.542
(Col1 Risk)	Logit	1.271	1.040	1.555
Cohort	Mantel-Haenszel	0.801	0.670	0.957
(Col2 Risk)	Logit	0.801	0.674	0.952

The confidence bounds for the M-H estimates are test-based.

Dr Firas

Total Sample Size = 508

Odds Ratio by Center
---------- CENTER=7 ----------

TABLE 1 OF SMOKE BY CANCER
CONTROLLING FOR CENTER=7

SMOKE CANCER

Frequency,	1,	2,	Total
1 ,	60 ,	99 ,	159
2 ,	11 ,	43 ,	54
Total	71	142	213

Odds Ratio by Center
---------- CENTER=7 ----------

SUMMARY STATISTICS FOR SMOKE BY CANCER
CONTROLLING FOR CENTER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	5.444	0.020
2	Row Mean Scores Differ	1	5.444	0.020
3	General Association	1	5.444	0.020

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Bounds	
Case-Control	Mantel-Haenszel	2.369	1.148	4.889
(Odds Ratio)	Logit	2.369	1.135	4.945
Cohort	Mantel-Haenszel	1.852	1.104	3.109
(Col1 Risk)	Logit	1.852	1.054	3.256
Cohort	Mantel-Haenszel	0.782	0.636	0.961
(Col2 Risk)	Logit	0.782	0.652	0.937

The confidence bounds for the M-H estimates are test-based.

Total Sample Size = 213

Odds Ratio by Center
---------- CENTER=8 ----------

TABLE 1 OF SMOKE BY CANCER
CONTROLLING FOR CENTER=8

SMOKE CANCER

Frequency,	1,	2,	Total
1 ,	104 ,	89 ,	193
2 ,	21 ,	36 ,	57
Total	125	125	250

Dr Firas

Odds Ratio by Center

---------- CENTER=8 ----------

SUMMARY STATISTICS FOR SMOKE BY CANCER
CONTROLLING FOR CENTER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	5.093	0.024
2	Row Mean Scores Differ	1	5.093	0.024
3	General Association	1	5.093	0.024

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Bounds	
Case-Control	Mantel-Haenszel	2.003	1.096	3.663
(Odds Ratio)	Logit	2.003	1.091	3.680
Cohort	Mantel-Haenszel	1.463	1.051	2.035
(Col1 Risk)	Logit	1.463	1.016	2.105
Cohort	Mantel-Haenszel	0.730	0.556	0.959
(Col2 Risk)	Logit	0.730	0.569	0.938

The confidence bounds for the M-H estimates are test-based.

Total Sample Size = 250

Center	Odds Ratio
4	2.850
7	2.369
5	2.319
1	2.196
3	2.175
2	2.143
8	2.003
6	1.688

ويمكن الحصول على نفس النتائج بأضافة relrisk الى عبارة التنفيذ وسنحصل على قيم odds ratio و relative risk لكل مستشفى اضافة الى قيمة واحدة لكل من المقياسين بغض النظر عن عدد المستشفيات. وهذه القيمتين يمكن الحصول عليها من البيانات السابقة بحذف عبارة by center وتنفيذ البرنامج.

يمكن استعمال ايعاز اخر لاختبار العلاقة بين التدخين والسرطان لكل مركز ولجميع المراكز.

```
proc freq data=cmh;
by center;
 tables cancer*smoke/chisq;
weight count;
run;
proc freq data=cmh;
 tables cancer*smoke/chisq;
weight count;
run;
```

Dr Firas

مثال (109) : جد ارجحية حصول الفتق لدى سواق الشاحنات عن غيرهم؟

عدد المستشفيات=7 ، سائق سيارة=1 ، ليس لديه سيارة =2 ،ليس لديهم فتق=1 ، لديهم فتق=2 ،

```
data cmh;
input hospital driving hernia count @@;
cards;
1 1 1 8   1 1 2 1   1 2 1 47   1 2 2 26
2 1 1 5   2 1 2 0   2 2 1 17   2 2 2 21
3 1 1 4   3 1 2 4   3 2 1 13   3 2 2 77
4 1 1 2   4 1 2 10  4 2 1 12   4 2 2 65
5 1 1 1   5 1 2 3   5 2 1 5    5 2 2 10
6 1 1 1   6 1 2 2   6 2 1 3    6 2 2 11
7 1 1 2   7 1 2 2   7 2 1 12   7 2 2 37
proc freq data=cmh;
 weight count;
 table hospital*driving*hernia/ cmh  relrisk;
run;
```

```
                                    Odds Ratio by hospital
                                TABLE 1 OF DRIVING BY HERNIA
                                CONTROLLING FOR HOSPITAL=1

     DRIVING    HERNIA

    Frequency,     1,     2, Total
    ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
        1,    8,    1,    9
    ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
        2,   47,   26,   73
    ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
    Total       55     27     82


   STATISTICS FOR TABLE 1 OF DRIVING BY HERNIA
        CONTROLLING FOR HOSPITAL=1

        Estimates of the Relative Risk (Row1/Row2)

                            95%
  Type of Study         Value     Confidence Bounds
  ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
  Case-Control          4.426     0.524    37.365
  Cohort (Col1 Risk)    1.381     1.036     1.840
  Cohort (Col2 Risk)    0.312     0.048     2.031

             Sample Size = 82


TABLE 2 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=2

        DRIVING    HERNIA

        Frequency,     1,     2, Total
            ƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
            1,    5,    0,    5
            ƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
            2,   17,   21,   38
             ˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
        Total       22     21     43

              Odds Ratio by hospital
```

Dr. Firas

STATISTICS FOR TABLE 2 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=2

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Bounds	
Cohort (Col1 Risk)	2.235	1.570	3.183

Sample Size = 43
One or more risk estimates not computed --- zero cell.

TABLE 3 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=3

DRIVING HERNIA

Frequency,	1,	2,	Total
	1 ,	4 ,	4 , 8
	2 ,	13 ,	77 , 90
Total	17	81	98

STATISTICS FOR TABLE 3 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=3

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Bounds	
Case-Control	5.923	1.315	26.688
Cohort (Col1 Risk)	3.462	1.470	8.149
Cohort (Col2 Risk)	0.584	0.291	1.175

Sample Size = 98

Odds Ratio by hospital
TABLE 4 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=4

DRIVING HERNIA

Frequency,	1,	2,	Total
	1 ,	2 ,	10 , 12
	2 ,	12 ,	65 , 77
Total	14	75	89

STATISTICS FOR TABLE 4 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=4

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Bounds	
Case-Control	1.083	0.210	5.575
Cohort (Col1 Risk)	1.069	0.272	4.199
Cohort (Col2 Risk)	0.987	0.753	1.294

Dr Firas

Sample Size = 89

TABLE 5 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=5

DRIVING	HERNIA	
Frequency,	1,	2, Total
1 ,	1 ,	3 , 4
2 ,	5 ,	10 , 15
Total	6	13 19

Odds Ratio by hospital
STATISTICS FOR TABLE 5 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=5

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Bounds	
Case-Control	0.667	0.054	8.161
Cohort (Col1 Risk)	0.750	0.119	4.732
Cohort (Col2 Risk)	1.125	0.576	2.197

Sample Size = 19
TABLE 6 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=6

DRIVING	HERNIA	
Frequency,	1,	2, Total
1 ,	1 ,	2 , 3
2 ,	3 ,	11 , 14
Total	4	13 17

STATISTICS FOR TABLE 6 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=6

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Bounds	
Case-Control	1.833	0.121	27.797
Cohort (Col1 Risk)	1.556	0.235	10.283
Cohort (Col2 Risk)	0.848	0.364	1.976

Sample Size = 17

Odds Ratio by hospital
TABLE 7 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=7

DRIVING	HERNIA	
Frequency,	1,	2, Total
1 ,	2 ,	2 , 4

Dr. Firas

```
       2,  12,  37,  49
    ƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
    Total      14    39    53


STATISTICS FOR TABLE 7 OF DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL=7

Estimates of the Relative Risk (Row1/Row2)

                          95%
Type of Study       Value     Confidence Bounds
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Case-Control        3.083     0.391    24.320
Cohort (Col1 Risk)  2.042     0.682     6.111
Cohort (Col2 Risk)  0.662     0.245     1.787


Sample Size = 53
Odds Ratio by hospital
SUMMARY STATISTICS FOR DRIVING BY HERNIA
CONTROLLING FOR HOSPITAL


Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic  Alternative Hypothesis   DF     Value     Prob
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
1      Nonzero Correlation      1     8.993    0.003
2      Row Mean Scores Differ   1     8.993    0.003
3      General Association      1     8.993    0.003


Estimates of the Common Relative Risk (Row1/Row2)
                                  95%
Type of Study   Method          Value   Confidence Bounds
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Case-Control    Mantel-Haenszel   3.002    1.464   6.159
(Odds Ratio)    Logit *           2.789    1.288   6.037

Cohort          Mantel-Haenszel   1.659    1.192   2.309
(Col1 Risk)     Logit             1.720    1.399   2.115

Cohort          Mantel-Haenszel   0.702    0.557   0.885
(Col2 Risk)     Logit *           0.897    0.720   1.117

The confidence bounds for the M-H estimates are test-based.

* denotes that the logit estimators use a correction
of 0.5 in every cell of those tables that contain a zero.

Breslow-Day Test for Homogeneity of the Odds Ratios

Chi-Square =  6.137       DF =  6       Prob = 0.408


Total Sample Size = 401
```

تبين من النتائج ان ارجحية حصول الفتق لدى سائقي الشاحنات كان عاليا في جميع المستشفيات بأستثناء مستشفى رقم 5 اذ بلغ 0.667 وبمعدل عام = 3 وكذلك كانت تقديرات الخطورةالنسبية بمعدل عام = 1.65

يمكن الحصول على نفس النتائج اذا استعملنا الايعاز by hospital وكما موضح ادناه:

```
proc freq data=cmh;
 weight count; by hospital;
 table driving*hernia/ cmh norow nocol nopercent relrisk;run;
```

واذا حذفنا by hospital فأننا سنحصل على نسبة ارجحية بغض النظر عن المستشفى.

Dr Firas

```
proc freq data=cmh;
 weight count;
 table driving*hernia/ cmh norow nocol nopercent relrisk;
run;
```

ويمكن اختبار علاقة الاستقلال بين السياقة وحصول الفتق باستعمال الايعاز التالي:

```
proc freq data=cmh;
 weight count;
 table driving*hernia/ chisq;
run;
```

يمكن حل المثال السابق بترتيب البيانات بشكل عمودي كما موضح ادناه:

```
data cmh;
input hospital driving hernia count ;
cards;
1 1 1 8
1 1 2 1
1 2 1 47
1 2 2 26
2 1 1 5
2 1 2 0
2 2 1 17
2 2 2 21
3 1 1 4
3 1 2 4
3 2 1 13
3 2 2 77
4 1 1 2
4 1 2 10
4 2 1 12
4 2 2 65
5 1 1 1
5 1 2 3
5 2 1 5
5 2 2 10
6 1 1 1
6 1 2 2
6 2 1 3
6 2 2 11
7 1 1 2
7 1 2 2
7 2 1 12
7 2 2 37
proc freq data=cmh;
 weight count;
 table hospital*driving*hernia/ cmh norow nocol nopercent relrisk;
run;
```

مثال (110) : في دراسة شملت 65 مريض استعمل عقار جديد لعلاجهم ، وقد توفي 48 مريض خلال الدراسة فيما بقي على قيد الحياة 17 مريض ، وشملت الدراسة مجموعة من العوامل هي الوقت Time وتشيرالى المدة الزمنية بالاشهر من بداية التشخيص ، المتغير VSTATUS ويمثل قيمتين 0 و 1 وهما يشيران الى ان المريض حي او ميت لغاية نهاية التجربة ، واذا كانت قيمة VSTATUS تساوي صفر فأن قيمة المتغير Time تسمى مراقب censored (اي لايزال

Dr. Firas

حـي) او 1 وتعنـي غيـر مراقـب uncensored (ميـت) ،امـا المتغيـر LOGBUN فيمثـل لوغـارتيم bun عنـد التشـخيص، HGB ويمثـل الهيموغلـوبين عنـد التشـخيص، PLATELET وتمثل الاقراص الدموية اذ ان القيمة 0 = غير طبيعية والقيمة 1 = طبيعية ، Age وتمثـل العمـر بالسـنوات عنـد التشـخيص ، LOGWBC وتمثـل لوغـارتيم عـدد كريـات الـدم البيضـاء عنـد التشـخيص ، Frac وتمثـل fractures وبقيمتـين 0= لاتوجـد و 1= موجـودة ، LOGPAM وتمثـل لوغـارتيم عـدد خلايـا البلازمـا و Protein وتمثـل بـروتين اليوريـا و SCALC وتمثـل الكالسيوم في السيرم.

المطلوب تحديد عوامل التشخيص المهمة من بين جميع العوامل المدروسة.

```
data myeloma;
  input time vstatus logbun hgb platelet age logwbc frac
      logpbm protein scalc;
  label time='survival time'
      vstatus='0=alive 1=dead';
  cards;
 1.25  1  2.2175   9.4  1  67  3.6628  1  1.9542  12  10
 1.25  1  1.9395  12.0  1  38  3.9868  1  1.9542  20  18
 2.00  1  1.5185   9.8  1  81  3.8751  1  2.0000   2  15
 2.00  1  1.7482  11.3  0  75  3.8062  1  1.2553   0  12
 2.00  1  1.3010   5.1  0  57  3.7243  1  2.0000   3   9
 3.00  1  1.5441   6.7  1  46  4.4757  0  1.9345  12  10
 5.00  1  2.2355  10.1  1  50  4.9542  1  1.6628   4   9
 5.00  1  1.6812   6.5  1  74  3.7324  0  1.7324   5   9
 6.00  1  1.3617   9.0  1  77  3.5441  0  1.4624   1   8
 6.00  1  2.1139  10.2  0  70  3.5441  1  1.3617   1   8
 6.00  1  1.1139   9.7  1  60  3.5185  1  1.3979   0  10
 6.00  1  1.4150  10.4  1  67  3.9294  1  1.6902   0   8
 7.00  1  1.9777   9.5  1  48  3.3617  1  1.5682   5  10
 7.00  1  1.0414   5.1  0  61  3.7324  1  2.0000   1  10
 7.00  1  1.1761  11.4  1  53  3.7243  1  1.5185   1  13
 9.00  1  1.7243   8.2  1  55  3.7993  1  1.7404   0  12
11.00  1  1.1139  14.0  1  61  3.8808  1  1.2788   0  10
11.00  1  1.2304  12.0  1  43  3.7709  1  1.1761   1   9
11.00  1  1.3010  13.2  1  65  3.7993  1  1.8195   1  10
11.00  1  1.5682   7.5  1  70  3.8865  0  1.6721   0  12
11.00  1  1.0792   9.6  1  51  3.5051  1  1.9031   0   9
13.00  1  0.7782   5.5  0  60  3.5798  1  1.3979   2  10
14.00  1  1.3979  14.6  1  66  3.7243  1  1.2553   2  10
15.00  1  1.6021  10.6  1  70  3.6902  1  1.4314   0  11
16.00  1  1.3424   9.0  1  48  3.9345  1  2.0000   0  10
16.00  1  1.3222   8.8  1  62  3.6990  1  0.6990  17  10
17.00  1  1.2304  10.0  1  53  3.8808  1  1.4472   4   9
17.00  1  1.5911  11.2  1  68  3.4314  0  1.6128   1  10
18.00  1  1.4472   7.5  1  65  3.5682  0  0.9031   7   8
19.00  1  1.0792  14.4  1  51  3.9191  1  2.0000   6  15
19.00  1  1.2553   7.5  0  60  3.7924  1  1.9294   5   9
24.00  1  1.3010  14.6  1  56  4.0899  1  0.4771   0   9
25.00  1  1.0000  12.4  1  67  3.8195  1  1.6435   0  10
26.00  1  1.2304  11.2  1  49  3.6021  1  2.0000  27  11
32.00  1  1.3222  10.6  1  46  3.6990  1  1.6335   1   9
```

```
35.00  1  1.1139   7.0  0  48  3.6532  1  1.1761   4  10
37.00  1  1.6021  11.0  1  63  3.9542  0  1.2041   7   9
41.00  1  1.0000  10.2  1  69  3.4771  1  1.4771   6  10
41.00  1  1.1461   5.0  1  70  3.5185  1  1.3424   0   9
51.00  1  1.5682   7.7  0  74  3.4150  1  1.0414   4  13
52.00  1  1.0000  10.1  1  60  3.8573  1  1.6532   4  10
54.00  1  1.2553   9.0  1  49  3.7243  1  1.6990   2  10
58.00  1  1.2041  12.1  1  42  3.6990  1  1.5798  22  10
66.00  1  1.4472   6.6  1  59  3.7853  1  1.8195   0   9
67.00  1  1.3222  12.8  1  52  3.6435  1  1.0414   1  10
88.00  1  1.1761  10.6  1  47  3.5563  0  1.7559  21   9
89.00  1  1.3222  14.0  1  63  3.6532  1  1.6232   1   9
92.00  1  1.4314  11.0  1  58  4.0755  1  1.4150   4  11
 4.00  0  1.9542  10.2  1  59  4.0453  0  0.7782  12  10
 4.00  0  1.9243  10.0  1  49  3.9590  0  1.6232   0  13
 7.00  0  1.1139  12.4  1  48  3.7993  1  1.8573   0  10
 7.00  0  1.5315  10.2  1  81  3.5911  0  1.8808   0  11
 8.00  0  1.0792   9.9  1  57  3.8325  1  1.6532   0   8
12.00  0  1.1461  11.6  1  46  3.6435  0  1.1461   0   7
11.00  0  1.6128  14.0  1  60  3.7324  1  1.8451   3   9
12.00  0  1.3979   8.8  1  66  3.8388  1  1.3617   0   9
13.00  0  1.6628   4.9  0  71  3.6435  0  1.7924   0   9
16.00  0  1.1461  13.0  1  55  3.8573  0  0.9031   0   9
19.00  0  1.3222  13.0  1  59  3.7709  1  2.0000   1  10
19.00  0  1.3222  10.8  1  69  3.8808  1  1.5185   0  10
28.00  0  1.2304   7.3  1  82  3.7482  1  1.6721   0   9
41.00  0  1.7559  12.8  1  72  3.7243  1  1.4472   1   9
53.00  0  1.1139  12.0  1  66  3.6128  1  2.0000   1  11
57.00  0  1.2553  12.5  1  66  3.9685  0  1.9542   0  11
77.00  0  1.0792  14.0  1  60  3.6812  0  0.9542   0  12
title 'Example 1. Stepwise Regression';
proc phreg data=myeloma;
  model time*vstatus(0)=logbun hgb platelet age logwbc frac
              logpbm protein scalc /
              selection=stepwise slentry=0.25
              slstay=0.15 details;
run;
```

تم تحليل البيانات بطريقة الانحدار المتدرج اذ يتم في الخطوة الاولى اختبار جميع العوامل المدروسة ثم يؤخذ العامل الاكثر معنوية والذي كان LOGBUN اعتمادا على قيمة مربع كاي 0.0035 ثم نقدر نسبة الارجحية Odds Ratio لهذا العامل.

ملاحظة : يحدد الامر slentry ان يكون احتمال العامل يساوي 0.25 فأقل قبل شموله بالتحليل فيما يحدد الأمر slstay ان العامل يبقى في النموذج اذا كانت احتماليته بعد ادخاله للنموذج تساوي 0.15 فأقل.

```
                Example 1. Stepwise Regression
                  The PHREG Procedure

    Data Set: WORK.MYELOMA
      Dependent Variable: TIME     survival time
    Censoring Variable: VSTATUS   0=alive 1=dead
    Censoring Value(s): 0
```

Dr. Firas

Ties Handling: BRESLOW

Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
65	48	17	26.15

Analysis of Variables Not in the Model

Variable	Score Chi-Square	Pr > Chi-Square
LOGBUN	8.5164	0.0035
HGB	5.0664	0.0244
PLATELET	3.1816	0.0745
AGE	0.0183	0.8924
LOGWBC	0.5658	0.4519
FRAC	0.9151	0.3388
LOGPBM	0.5846	0.4445
PROTEIN	0.1466	0.7018
SCALC	1.1109	0.2919

Residual Chi-square = 18.4550 with 9 DF (p=0.0302)

Step 1: Variable LOGBUN is entered. The model contains the following explanatory variables.

LOGBUN

Example 1. Stepwise Regression

The PHREG Procedure

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L	309.716	301.959	7.757 with 1 DF (p=0.0053)
Score	.	.	8.516 with 1 DF (p=0.0035)
Wald	.	.	8.339 with 1 DF (p=0.0039)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
LOGBUN	1	1.745947	0.60460	8.33919	0.0039	5.731

يأخذ البرنامج بعد ذلك العامل الثاني الاعلى تأثير HGB ويقدر نسبة الارجحية

Analysis of Variables Not in the Model

Variable	Score Chi-Square	Pr > Chi-Square
HGB	4.3468	0.0371
PLATELET	2.0183	0.1554
AGE	0.7159	0.3975
LOGWBC	0.0704	0.7908
FRAC	1.0354	0.3089
LOGPBM	1.0334	0.3094
PROTEIN	0.5214	0.4703
SCALC	1.4150	0.2342

Residual Chi-square = 9.3164 with 8 DF (p=0.3163)

Dr Firas

Step 2: Variable HGB is entered. The model contains the following explanatory variables.

LOGBUN HGB

Example 1. Stepwise Regression
The PHREG Procedure

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L	309.716	297.767	11.949 with 2 DF (p=0.0025)
Score	.	.	12.725 with 2 DF (p=0.0017)
Wald	.	.	

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
LOGBUN	1	1.674399	0.61209	7.48330	0.0062	5.336
HGB	1	-0.118987	0.05751	4.28112	0.0385	0.888

يأخذ البرنامج العامل الثالث SCALC ويقدر نسبة الارجحية وذلك لان هذا العامل كان ذو اقل قيمة احتمالية وكما موضح ادناه:

Analysis of Variables Not in the Model

Variable	Score Chi-Square	Pr > Chi-Square
PLATELET	0.2266	0.6341
AGE	1.3508	0.2451
LOGWBC	0.3785	0.5384
FRAC	1.0491	0.3057
LOGPBM	0.6741	0.4116
PROTEIN	0.6592	0.4168
SCALC	1.8225	0.1770

Residual Chi-square = 5.3635 with 7 DF (p=0.6157)

Step 3: Variable SCALC is entered. The model contains the following explanatory variables.

LOGBUN HGB SCALC

Example 1. Stepwise Regression

The PHREG Procedure

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L	309.716	296.078	13.638 with 3 DF (p=0.0034)
Score	.	.	15.305 with 3 DF (p=0.0016)
Wald	.	.	14.454 with 3 DF (p=0.0023)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
LOGBUN	1	1.635924	0.62359	6.88214	0.0087	5.134
HGB	1	-0.126428	0.05868	4.64194	0.0312	0.881

Dr Firas

SCALC	1	0.132856	0.09868	1.81254	0.1782	1.142

بسبب عدم معنوية تأثير العامل الثالث يقوم البرنامج بحذفه

Step 4: Variable SCALC is removed. The model contains the following explanatory variables.

LOGBUN HGB

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L	309.716	297.767	11.949 with 2 DF (p=0.0025)
Score	.	.	12.725 with 2 DF (p=0.0017)
Wald	.	.	12.190 with 2 DF (p=0.0023)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
LOGBUN	1	1.674399	0.61209	7.48330	0.0062	5.336
HGB	1	-0.118987	0.05751	4.28112	0.0385	0.888

يتبين من النتائج ان زيادة قيمة لوغاريتيم خلايا البلازما تزيد من نسبة الخطورة فيما يعمل انخفاض الهيموغلوبين على تقليل نسبة الخطورة.

Example 1. Stepwise Regression

The PHREG Procedure

Analysis of Variables Not in the Model

Variable	Score Chi-Square	Pr > Chi-Square
PLATELET	0.2266	0.6341
AGE	1.3508	0.2451
LOGWBC	0.3785	0.5384
FRAC	1.0491	0.3057
LOGPBM	0.6741	0.4116
PROTEIN	0.6592	0.4168
SCALC	1.8225	0.1770

Residual Chi-square = 5.3635 with 7 DF (p=0.6157)

NOTE: Model building terminates because the variable to be entered is the variable that was removed in the last step.

Summary of Stepwise Procedure

Step	Variable Entered	Removed	Number In	Score Chi-Square	Wald Chi-Square	Pr > Chi-Square
1	LOGBUN		1	8.5164	.	0.0035
2	HGB		2	4.3468	.	0.0371
3	SCALC		3	1.8225	.	0.1770
4		SCALC	2	.	1.8125	0.1782

لو حاولنا تحديد بقاء العامل بقيمة اخرى مثل slstay=0.20 سنجد ان البرنامج سيدخل عامل العمر لانه كما موضح في جدول النتائج ذو احتماليته = 0.2451 ويمكن اعادة التنفيذ بتغيير الأمر لنحصل في جدول ملخص طريقة الانحدار المتدرج على عامل آخر هو العمر.

Dr.Firas

مثال (111): في دراسة عن 840 شخص تضمنت 538 امرأة و 302 رجل ، كان عدد النساء المصابة بمرض ما = 222 وعدد الرجال المصابين 105 احسب نسبة ارجحية الاصابة للنساء مقارنة بالرجال؟

الصفة المراد التحري عنها هي الاصابة ونلاحظ في ترتيب البيانات ان عدم الاصابة اخذت رقم 1 فيما اخذت الاصابة 0 لذا نرتب البيانات تنازليا بالايعاز Descending.

```
data raw;
input  w m freq;
cards;
1 1 316
1 0 222
2 1 197
2 0 105
PROC LOGISTIC DATA=raw DESCENDING;
 WEIGHT freq;
 MODEL m = w;
RUN;
```

Dr. Firas

The SAS System

The LOGISTIC Procedure

Data Set: WORK.RAW
Response Variable: M
Response Levels: 2
Number of Observations: 4
Weight Variable: FREQ
Sum of Weights: 840
Link Function: Logit

Response Profile

Ordered Value	M	Total Count	Weight
1	1	2	513.00000
2	0	2	327.00000

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	1124.958	1123.502	.
SC	1124.345	1122.275	.
-2 LOG L	1122.958	1119.502	3.456 with 1 DF (p=0.0630)
Score	.	.	3.433 with 1 DF (p=0.0639)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Stand Estimate	Odds Ratio
INTERCPT	1	0.0769	0.2128	0.1306	0.7178	.	.
W	1	0.2762	0.1492	3.4251	0.0642	1.222630	1.318

Association of Predicted Probabilities and Observed Responses

Concordant = 25.0% Somers' D = 0.000

```
Discordant = 25.0%        Gamma    = 0.000
Tied      = 50.0%         Tau-a    = 0.000
(4 pairs)                 c        = 0.500
```

ارجحية اصابة النساء 1.31 مرة بقدر اصابة الرجال

مثال (112) : تقدير نسبة الخطورة ونسبة الارجحية للمثال السابق.

يلاحظ عدم وجود كلمة descending وذلك لاننا غيرنا الترتيب فلا نحتاج اليها.

```
data c;
input W M COUNT;
CARDS;
1 0 222
1 1 316
2 0 105
2 1 197
;OPTION NODATE NONUMBER;
proc freq data=c;
 weight count;
 table W*M/ cmh norow nocol nopercent ;
run;
```

ملاحظة: يمكن الحصول على قيم الاقتران باضافة الايعاز Measures بعد الخط المائل.

```
                                The SAS System
                                TABLE OF W BY M

        W       M

       Frequency,     0,     1, Total
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
           1 ,   222 ,   316 ,   538
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
           2 ,   105 ,   197 ,   302
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
       Total       327     513     840


SUMMARY STATISTICS FOR W BY M

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic  Alternative Hypothesis   DF     Value     Prob
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
1      Nonzero Correlation      1     3.429    0.064
2      Row Mean Scores Differ   1     3.429    0.064
3      General Association      1     3.429    0.064

Estimates of the Common Relative Risk (Row1/Row2)
                                    95%
Type of Study   Method          Value   Confidence Bounds
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Case-Control    Mantel-Haenszel   1.318    0.984   1.766
(Odds Ratio)    Logit             1.318    0.984   1.766

Cohort          Mantel-Haenszel   1.187    0.990   1.423
(Col1 Risk)     Logit             1.187    0.987   1.427

Cohort          Mantel-Haenszel   0.900    0.806   1.006
(Col2 Risk)     Logit             0.900    0.808   1.004

The confidence bounds for the M-H estimates are test-based.
Total Sample Size = 840
```

Dr. Firas

مثال (113): في دراسة عن مرضى القلب وجد ان 51 مريضا من الرجال احتاج الى تداخل جراحي فيما لم يحتاج الى ذلك 6 مرضى وكانت الاعداد المناظرة لها من النساء 22 و 21 على التوالي ، ماهي ارجحية الحاجة الى تداخل جراحي بين الرجال مقارنة بالنساء؟

```
DATA e;
 INPUT CHD  aged  freq;

cards;
0 0 51
0 1 6
1 0 22
1 1 21
 proc freq data=e;
  tables CHD*aged/chisq;
weight freq;
run;
proc logistic data=e descending;
  model  CHD= aged;
weight freq;
run;
```

Dr. Firas

```
                                        The SAS System
            TABLE OF CHD BY AGED

          CHD     AGED

          Frequency,
          Percent ,
          Row Pct ,
          Col Pct ,     0,     1, Total
          ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
              0 ,   51 ,    6 ,   57
                , 51.00 ,  6.00 , 57.00
                , 89.47 , 10.53 ,
                , 69.86 , 22.22 ,
          ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
              1 ,   22 ,   21 ,   43
                , 22.00 , 21.00 , 43.00
                , 51.16 , 48.84 ,
                , 30.14 , 77.78 ,
          ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
          Total       73     27    100
                 73.00   27.00  100.00

          STATISTICS FOR TABLE OF CHD BY AGED

     Statistic                 DF     Value      Prob
     ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
     Chi-Square                 1    18.252      0.001
     Likelihood Ratio Chi-Square   1    18.704      0.001
     Continuity Adj. Chi-Square    1    16.360      0.001
     Mantel-Haenszel Chi-Square    1    18.069      0.001
     Fisher's Exact Test (Left)                 1.000
                (Right)                 2.24E-05
                (2-Tail)                2.74E-05
     Phi Coefficient                   0.427
     Contingency Coefficient           0.393
     Cramer's V                        0.427

     Sample Size = 100
```

القيمة المعنوية لمربع كاي تشير الى وجود فروق معنوية بين الرجال والنساء في الحاجة الى التداخل الجراحي

The SAS System

The LOGISTIC Procedure

Data Set: WORK.E
Response Variable: CHD
Response Levels: 2
Number of Observations: 4
Weight Variable: FREQ
Sum of Weights: 100
Link Function: Logit

Response Profile

Ordered Value	CHD	Count	Total Weight
1	1	2	43.000000
2	0	2	57.000000

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	138.663	121.959	.
SC	138.049	120.732	.
-2 LOG L	136.663	117.959	18.704 with 1 DF (p=0.0001)
Score	.	.	18.252 with 1 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standard Estimate	Odds Ratio
INTERCPT	1	-0.8408	0.2551	10.8652	0.0010	.	.
AGED	1	2.0935	0.5285	15.6899	0.0001	2.958530	8.114

Association of Predicted Probabilities and Observed Responses

Concordant = 25.0% Somers' D = 0.000
Discordant = 25.0% Gamma = 0.000
Tied = 50.0% Tau-a = 0.000
(4 pairs) c = 0.500

ارجحية حاجة الرجال الى التداخل الجراحي مقارنة بالنساء حوالي 8 مرات.

مثال: (114) في دراسة لمعرفة تأثير التدخين على الاصابة بالسرطان ، جمعت بيانات عن مرضى السرطان في ثمانية مراكز طبية لغرض اختبار تجانس ارجحية الاصابة لدى المدخنين باختلاف المراكز؟

```
data cmh;
input center smoke cancer count @@;
cards;
1 1 1 126   1 1 2 100   1 2 1  35   1 2 2  61
2 1 1 908   2 1 2 688   2 2 1 497   2 2 2 807
3 1 1 913   3 1 2 747   3 2 1 336   3 2 2 598
4 1 1 235   4 1 2 172   4 2 1  58   4 2 2 121
5 1 1 402   5 1 2 308   5 2 1 121   5 2 2 215
6 1 1 182   6 1 2 156   6 2 1  72   6 2 2  98
7 1 1  60   7 1 2  99   7 2 1  11   7 2 2  43
8 1 1 104   8 1 2  89   8 2 1  21   8 2 2  36
```

Dr Firas

```
proc freq data=cmh;
 weight count;
 table center*smoke*cancer/ cmh norow nocol nopercent ;
run;
```

```
                    Odds Ratio by Center
                  TABLE 1 OF SMOKE BY CANCER
                   CONTROLLING FOR CENTER=1

                SMOKE    CANCER

                Frequency,     1,     2, Total
                ffffffff^fffffff^fffffff^
                    1 ,   126 ,   100 ,  226
                ffffffff^fffffff^fffffff^
                    2 ,    35 ,    61 ,   96
                ffffffff^fffffff^fffffff^
                Total      161    161    322


                  TABLE 2 OF SMOKE BY CANCER
                   CONTROLLING FOR CENTER=2

                SMOKE    CANCER

                Frequency,     1,     2, Total
                ffffffff^fffffff^fffffff^
                    1 ,   908 ,   688 ,  1596
                ffffffff^fffffff^fffffff^
                    2 ,   497 ,   807 ,  1304
                ffffffff^fffffff^fffffff^
                Total     1405   1495   2900


                  TABLE 3 OF SMOKE BY CANCER
                   CONTROLLING FOR CENTER=3

                SMOKE    CANCER

                Frequency,     1,     2, Total
                ffffffff^fffffff^fffffff^
                    1 ,   913 ,   747 ,  1660
                ffffffff^fffffff^fffffff^
                    2 ,   336 ,   598 ,   934
                ffffffff^fffffff^fffffff^
                Total     1249   1345   2594

                    Odds Ratio by Center
                  TABLE 4 OF SMOKE BY CANCER
                   CONTROLLING FOR CENTER=4

                SMOKE    CANCER

                Frequency,     1,     2, Total
                ffffffff^fffffff^fffffff^
                    1 ,   235 ,   172 ,   407
                ffffffff^fffffff^fffffff^
                    2 ,    58 ,   121 ,   179
                ffffffff^fffffff^fffffff^
                Total      293    293    586


TABLE 5 OF SMOKE BY CANCER
CONTROLLING FOR CENTER=5


          SMOKE    CANCER

          Frequency,     1,     2, Total
          ffffffff^fffffff^fffffff^
```

Dr. Firas
Dr. Firas

```
          1 ,  402 ,  308 ,  710
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
          2 ,  121 ,  215 ,  336
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
        Total      523    523   1046


TABLE 6 OF SMOKE BY CANCER
CONTROLLING FOR CENTER=6

         SMOKE    CANCER

       Frequency,     1,     2, Total
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
          1 ,  182 ,  156 ,  338
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
          2 ,   72 ,   98 ,  170
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
       Total      254    254    508

Odds Ratio by Center
TABLE 7 OF SMOKE BY CANCER
CONTROLLING FOR CENTER=7

         SMOKE    CANCER

      Frequency,     1,     2, Total
      ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
         1 ,   60 ,   99 ,  159
      ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
         2 ,   11 ,   43 ,   54
      ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
      Total       71    142    213

TABLE 8 OF SMOKE BY CANCER
CONTROLLING FOR CENTER=8

         SMOKE    CANCER

       Frequency,     1,     2, Total
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
          1 ,  104 ,   89 ,  193
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
          2 ,   21 ,   36 ,   57
       ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
       Total      125    125    250
Odds Ratio by Center
SUMMARY STATISTICS FOR SMOKE BY CANCER
CONTROLLING FOR CENTER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic   Alternative Hypothesis    DF      Value      Prob
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
1       Nonzero Correlation      1     280.138     0.001
2       Row Mean Scores Differ   1     280.138     0.001
3       General Association      1     280.138     0.001
Estimates of the Common Relative Risk (Row1/Row2)
                                        95%
Type of Study    Method          Value   Confidence Bounds
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Case-Control    Mantel-Haenszel    2.174     1.985    2.382
(Odds Ratio)   Logit              2.173     1.983    2.382

Cohort          Mantel-Haenszel    1.519     1.447    1.595
(Col1 Risk)    Logit              1.513     1.436    1.594

Cohort          Mantel-Haenszel    0.700     0.671    0.730
(Col2 Risk)    Logit              0.701     0.673    0.730

The confidence bounds for the M-H estimates are test-based.
```

Dr Firas

```
Breslow-Day Test for Homogeneity of the Odds Ratios

Chi-Square =  5.200       DF =  7       Prob = 0.636
          Total Sample Size = 8419
```

تبين من النتائج ان ارجحية الاصابة لدى المدخنين هي 2.17 وبحدود ثقة تتراوح من 1.98 الى 2.38 وان اختبار بريسلو للتجانس يشير الى ان نسبة الارجحية متقاربة في جميع المراكز.

ويمكن تقدير ارجحية الاصابة لدى المدخنين في كل مركز بأضافة كلمتي ;by center الى قطعةالبرنامج السابق وتنفيذ البرنامج فأن النتائج ستكون ضمن كل مركز، اي ان قطعة البرنامج السابق بعد الاضافة ستكون كما يلي:

```
data cmh;
input center smoke cancer count @@;
cards;
1 1 1 126   1 1 2 100   1 2 1  35   1 2 2  61
2 1 1 908   2 1 2 688   2 2 1 497   2 2 2 807
3 1 1 913   3 1 2 747   3 2 1 336   3 2 2 598
4 1 1 235   4 1 2 172   4 2 1  58   4 2 2 121
5 1 1 402   5 1 2 308   5 2 1 121   5 2 2 215
6 1 1 182   6 1 2 156   6 2 1  72   6 2 2  98
7 1 1  60   7 1 2  99   7 2 1  11   7 2 2  43
8 1 1 104   8 1 2  89   8 2 1  21   8 2 2  36
proc freq data=cmh;
  weight count;by center;
  table center*smoke*cancer/ cmh norow nocol nopercent ;
run;
```

مثال (115): في دراسة لمعرفة تأثير التدخين على انخفاض وزن المواليد في ثلاثة اجناس المطلوب معرفة تأثير التدخين على اوزان المواليد لكل جنس وتقدير ارجحية النساء المدخنات ضمن كل جنس ومن ثم تقدير الارجحية بغض النظر عن الجنس؟

العرق= 1، 2 ، 3 والوزن عند الميلاد منخفض (اقل من 2500 غرام)=0 واعتيادي =1 وحالة المرأة مدخنة=0 وغير مدخنة = 1.

```
DATA e;
  INPUT race birthw  smoke  freq;
cards;
1 0 0 40
1 0 1 33
1 1 0 4
1 1 1 19
2 0 0 11
2 0 1 4
2 1 0 5
2 1 1 6
3 0 0 35
3 0 1 7
```

Dr. Firas

```
3 1 0 20
3 1 1 5
proc freq data=e;
 by race;
 tables birthw*smoke/chisq;
 weight freq;
run;
proc logistic data=e descending;
 by race;
 model  birthw= smoke;
 weight freq;
run;
proc freq data=e;
 tables birthw*smoke/chisq;
 weight freq;
run;
proc logistic data=e descending;
 model  birthw= smoke;
 weight freq;
run;
```

TABLE OF BIRTHW BY SMOKE

BIRTHW SMOKE

Frequency, Percent , Row Pct , Col Pct ,	0	1	Total
0	40 41.67 54.79 90.91	33 34.38 45.21 63.46	73 76.04
1	4 4.17 17.39 9.09	19 19.79 82.61 36.54	23 23.96
Total	44 45.83	52 54.17	96 100.00

STATISTICS FOR TABLE OF BIRTHW BY SMOKE

Statistic	DF	Value	Prob
Chi-Square	1	9.856	**0.002**
Likelihood Ratio Chi-Square	1	10.636	0.001
Continuity Adj. Chi-Square	1	8.407	0.004
Mantel-Haenszel Chi-Square	1	9.753	0.002
Fisher's Exact Test (Left)			1.000
(Right)			1.43E-03
(2-Tail)			1.79E-03
Phi Coefficient		0.320	
Contingency Coefficient		0.305	
Cramer's V		0.320	

Sample Size = 96

قيمة مربع كاي المعنوية (0.002) تعني ان التدخين يؤثر في انخفاض وزن المواليد في الجنس الاول.

The SAS System

D Firas

---------- RACE=2 ----------

TABLE OF BIRTHW BY SMOKE

BIRTHW SMOKE

Frequency,
Percent ,
Row Pct ,
Col Pct , 0, 1, Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
0 , 11 , 4 , 15
, 42.31 , 15.38 , 57.69
, 73.33 , 26.67 ,
, 68.75 , 40.00 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
1 , 5 , 6 , 11
, 19.23 , 23.08 , 42.31
, 45.45 , 54.55 ,
, 31.25 , 60.00 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 16 10 26
61.54 38.46 100.00

STATISTICS FOR TABLE OF BIRTHW BY SMOKE

Statistic	DF	Value	Prob
Chi-Square	1	2.084	0.149
Likelihood Ratio Chi-Square	1	2.091	0.148
Continuity Adj. Chi-Square	1	1.072	0.300
Mantel-Haenszel Chi-Square	1	2.004	0.157
Fisher's Exact Test (Left)			0.968
(Right)			0.150
(2-Tail)			0.228
Phi Coefficient			0.283
Contingency Coefficient			0.272
Cramer's V			0.283

Sample Size = 26
WARNING: 25% of the cells have expected counts less
than 5. Chi-Square may not be a valid test.

قيمة مربع كاي غير المعنوية (0.149) تعني ان التدخين لم يؤثر في انخفاض وزن المواليد في الجنس الثاني.

---------- RACE=3 ----------

TABLE OF BIRTHW BY SMOKE

BIRTHW SMOKE

Frequency,
Percent ,
Row Pct ,
Col Pct , 0, 1, Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
0 , 35 , 7 , 42
, 52.24 , 10.45 , 62.69
, 83.33 , 16.67 ,
, 63.64 , 58.33 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
1 , 20 , 5 , 25
, 29.85 , 7.46 , 37.31
, 80.00 , 20.00 ,
, 36.36 , 41.67 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ

Dr Iras

Total	55	12	67
	82.09	17.91	100.00

STATISTICS FOR TABLE OF BIRTHW BY SMOKE

Statistic	DF	Value	Prob
Chi-Square	1	0.118	0.731
Likelihood Ratio Chi-Square	1	0.117	0.732
Continuity Adj. Chi-Square	1	0.000	0.988
Mantel-Haenszel Chi-Square	1	0.117	0.733
Fisher's Exact Test (Left)			0.752
(Right)			0.487
(2-Tail)			0.751
Phi Coefficient			0.042
Contingency Coefficient			0.042
Cramer's V			0.042

Sample Size = 67
WARNING: 25% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

قيمة مربع كاي غير المعنوية (0.731) تعني ان التدخين لم يؤثر في انخفاض وزن المواليد في الجنس الثالث.

The SAS System

---------- RACE=1 ----------

The LOGISTIC Procedure

Data Set: WORK.E
Response Variable: BIRTHW
Response Levels: 2
Number of Observations: 4
Weight Variable: FREQ
Sum of Weights: 96
Link Function: Logit

Response Profile

Ordered Value	BIRTHW	Total Count	Weight
1	1	2	23.000000
2	0	2	73.000000

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	107.715	99.079	.
SC	107.101	97.852	.
-2 LOG L	105.715	95.079	10.636 with 1 DF (p=0.0011)
Score	.	.	9.856 with 1 DF (p=0.0017)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Sq	Pr > Chi-Sq	Standard Estimate	Odds Ratio
INTERCPT	1	-2.3026	0.5244	19.2796	0.0001	.	.
SMOKE	1	1.7505	0.5983	8.5611	0.0034	2.720249	5.758

Association of Predicted Probabilities and Observed Responses

Concordant = 25.0%		Somers' D = 0.000	
Discordant = 25.0%		Gamma = 0.000	
Tied = 50.0%		Tau-a = 0.000	

Dr Firas

(4 pairs) c = 0.500

ارجحية المدخنات في الجنس الاول = 5.75 مرة لولادة مواليد منخفضة الوزن مقارنة بغير المدخنات.

The SAS System

---------- RACE=2 ----------

The LOGISTIC Procedure

Data Set: WORK.E
Response Variable: BIRTHW
Response Levels: 2
Number of Observations: 4
Weight Variable: FREQ
Sum of Weights: 26
Link Function: Logit

Response Profile

Ordered Value	BIRTHW	Count	Total Weight
1	1	2	11.000000
2	0	2	15.000000

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	37.426	37.335	.
SC	36.812	36.108	.
-2 LOG L	35.426	33.335	2.091 with 1 DF (p=0.1482)
Score	.	.	2.084 with 1 DF (p=0.1489)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Stand Estimate	Odds Ratio
INTERCPT	1	-0.7885	0.5394	2.1370	0.1438	.	.
SMOKE	1	1.1939	0.8412	2.0146	0.1558	0.942757	3.300

ارجحية المدخنات في الجنس الثاني = 3.3 مرة على لولادة مواليد منخفضة الوزن.

Association of Predicted Probabilities and Observed Responses

Concordant = 25.0% Somers' D = 0.000
Discordant = 25.0% Gamma = 0.000
Tied = 50.0% Tau-a = 0.000
(4 pairs) c = 0.500

The SAS System

---------- RACE=3 ----------

The LOGISTIC Procedure

Data Set: WORK.E
Response Variable: BIRTHW
Response Levels: 2
Number of Observations: 4
Weight Variable: FREQ
Sum of Weights: 67
Link Function: Logit

Response Profile

Ordered Value	BIRTHW	Count	Total Weight
1	1	2	25.000000
2	0	2	42.000000

Dr. Firas

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	90.521	92.404	.
SC	89.907	91.176	.
-2 LOG L	88.521	88.404	0.117 with 1 DF (p=0.7322)
Score	.	.	0.118 with 1 DF (p=0.7307)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-0.5596	0.2803	3.9858	0.0459	.	.
SMOKE	1	0.2231	0.6492	0.1182	0.7310	0.222930	1.250

ارجحية المدخنات في الجنس الثالث = 1.25 مرة على ولادة مواليد منخفضة الوزن.

Association of Predicted Probabilities and Observed Responses

Concordant = 25.0% Somers' D = 0.000
Discordant = 25.0% Gamma = 0.000
Tied = 50.0% Tau-a = 0.000
(4 pairs) c = 0.500

The SAS System

TABLE OF BIRTHW BY SMOKE

BIRTHW SMOKE

Frequency,
Percent ,
Row Pct ,
Col Pct , 0, 1, Total

BIRTHW	SMOKE 0	SMOKE 1	Total
0	86	44	130
	45.50	23.28	68.78
	66.15	33.85	
	74.78	59.46	
1	29	30	59
	15.34	15.87	31.22
	49.15	50.85	
	25.22	40.54	
Total	115	74	189
	60.85	39.15	100.00

STATISTICS FOR TABLE OF BIRTHW BY SMOKE

Statistic	DF	Value	Prob
Chi-Square	1	4.924	0.026
Likelihood Ratio Chi-Square	1	4.867	0.027
Continuity Adj. Chi-Square	1	4.236	0.040
Mantel-Haenszel Chi-Square	1	4.898	0.027
Fisher's Exact Test (Left)			0.991
(Right)			0.020
(2-Tail)			0.036
Phi Coefficient			0.161
Contingency Coefficient			0.159
Cramer's V			0.161

Sample Size = 189

Dr Firas

اختبار مربع كاي (0.026) يشير الى وجود علاقة بين التدخين والمواليد قليلة الوزن.

The SAS System

The LOGISTIC Procedure

Data Set: WORK.E
Response Variable: BIRTHW
Response Levels: 2
Number of Observations: 12
Weight Variable: FREQ
Sum of Weights: 189
Link Function: Logit

Response Profile

Ordered Value	BIRTHW	Total Count	Weight
1	1	6	59.00000
2	0	6	130.00000

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	236.672	233.805	.
SC	237.157	234.774	.
-2 LOG L	234.672	229.805	4.867 with 1 DF (p=0.0274)
Score	.	.	4.924 with 1 DF (p=0.0265)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-1.0871	0.2147	25.6280	0.0001	.	.
SMOKE	1	0.7041	0.3196	4.8518	0.0276	0.785340	2.022

ارجحية المدخنات في جميع الاجناس = 2.02 مرة على ولادة مواليد منخفضة الوزن.

Association of Predicted Probabilities and Observed Responses

Concordant = 25.0% Somers' D = 0.000
Discordant = 25.0% Gamma = 0.000
Tied = 50.0% Tau-a = 0.000
(36 pairs) c = 0.500

مثال (116): في دراسة شملت مركزين وبواقع 90 مريض في كل مركز، اعطي لنصفهم في المركز الاول عقار A واعطي للنصف الآخر placebo لازالة التأثير النفسي ونفذ نفس العمل على المركز الثاني وسجلت اعداد الذين استجابوا والذين لم يستجيبوا للعلاج في كلا المركزين المطلوب اختبار الاستجابة وتقدير ارجحية شفاء الذين تناولوا العقار مقارنة بمجموعة السيطرة؟

```
data respire;
  input center treatmnt $ response $ count ;
  cards;
  1 test   y 29
  1 test   n 16
```

Dr.Firas

```
  1 placebo y 14
  1 placebo n 31
  2 test    y 37
  2 test    n  8
  2 placebo y 24
  2 placebo n 21
proc freq order=data;
    weight count;
    tables center*treatmnt*response /
     nocol nopct chisq cmh;
                                                          run;
```

```
                                The SAS System

              TABLE 1 OF TREAT BY RESPONSE
               CONTROLLING FOR CENTER=1

             TREAT    RESPONSE

             Frequency,
             Row Pct ,y     ,n     , Total
             ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
             test    ,   29 ,   16 ,   45
                  , 64.44 , 35.56 ,
             ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
             placebo ,   14 ,   31 ,   45
                  , 31.11 , 68.89 ,
             ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
             Total       43     47     90


           STATISTICS FOR TABLE 1 OF TREAT BY RESPONSE
                CONTROLLING FOR CENTER=1

          Statistic                    DF     Value      Prob
          ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
          Chi-Square                    1    10.020     0.002
          Likelihood Ratio Chi-Square   1    10.216     0.001
          Continuity Adj. Chi-Square    1     8.728     0.003
          Mantel-Haenszel Chi-Square    1     9.908     0.002
          Fisher's Exact Test (Left)                    1.000
                    (Right)                     1.46E-03
                    (2-Tail)                    2.92E-03
          Phi Coefficient                     0.334
          Contingency Coefficient             0.317
          Cramer's V                          0.334

          Sample Size = 90
```

بما ان قيمة مربع كاي معنوية فأن العقار اثر في الاستجابة في المركز الاول.

```
                The SAS System
              TABLE 2 OF TREAT BY RESPONSE
               CONTROLLING FOR CENTER=2

             TREAT    RESPONSE

             Frequency,
             Row Pct ,y     ,n     , Total
             ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
             test    ,   37 ,    8 ,   45
                  , 82.22 , 17.78 ,
             ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
             placebo ,   24 ,   21 ,   45
                  , 53.33 , 46.67 ,
             ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
             Total       61     29     90
```

Dr Firas

STATISTICS FOR TABLE 2 OF TREAT BY RESPONSE
CONTROLLING FOR CENTER=2

Statistic	DF	Value	Prob
Chi-Square	1	8.598	0.003
Likelihood Ratio Chi-Square	1	8.832	0.003
Continuity Adj. Chi-Square	1	7.326	0.007
Mantel-Haenszel Chi-Square	1	8.503	0.004
Fisher's Exact Test (Left)			0.999
(Right)			3.14E-03
(2-Tail)			6.28E-03
Phi Coefficient			0.309
Contingency Coefficient			0.295
Cramer's V			0.309

Sample Size = 90

بما ان قيمة مربع كاي معنوية فأن العقار اثر في الاستجابة في المركز الثاني.

The SAS System
SUMMARY STATISTICS FOR TREAT BY RESPONSE
CONTROLLING FOR CENTER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	18.411	0.001
2	Row Mean Scores Differ	1	18.411	0.001
3	General Association	1	18.411	0.001

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Bounds	
Case-Control	Mantel-Haenszel	4.029	2.132	7.614
(Odds Ratio)	Logit	4.029	2.106	7.707
Cohort	Mantel-Haenszel	1.737	1.350	2.235
(Col1 Risk)	Logit	1.676	1.294	2.170
Cohort	Mantel-Haenszel	0.462	0.324	0.657
(Col2 Risk)	Logit	0.474	0.326	0.688

The confidence bounds for the M-H estimates are test-based.

Breslow-Day Test for Homogeneity of the Odds Ratios

Chi-Square = 0.000 DF = 1 Prob = 0.990

Total Sample Size = 180

ارجحية الاستجابة للذين اخذوا العقار في المركزين = 4.02 مرة مقارنة بالسيطرة

ويمكن تقدير ارجحية الاستجابة في كل مركز باضافة كلمتي by center وبذلك تصبح قطعة البرنامج كما يلي:

```
proc freq order=data;
    weight count;by center;
    tables treatmnt*response /nocol nopct chisq cmh;
run;
```

Dr Firas

مثال (117): في دراسة لمعرفة العلاقة بين مرض ما ودرجة الاجهاد لدى مجموعتين (سكان القرى وسكان المدن) وكانت الاعداد كما موضحة ادناه ، فهل تؤثر درجة الاجهاد في المرض في كلا المنطقتين ومادرجية المعرضين للاجهاد العالي في حصول المرض؟

```
data stress;
  input region $ stress $ outcome $ count @@;
  cards;
 urban low  f 48 urban  low  u  12
 urban high f 96 urban  high u  94
 rural low  f 55 rural  low  u 135
 rural high f  7 rural  high u  53
proc freq order=data;
  weight count;
  tables region*stress*outcome / chisq cmh nocol nopct;
run;
```

```
                TABLE 1 OF STRESS BY OUTCOME
                CONTROLLING FOR REGION=urban

              STRESS    OUTCOME

              Frequency,
              Row Pct ,f     ,u     , Total
              ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
              low    ,   48 ,   12 ,   60
                   , 80.00 , 20.00 ,
              ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
              high   ,   96 ,   94 ,  190
                   , 50.53 , 49.47 ,
              ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
              Total      144    106    250


          STATISTICS FOR TABLE 1 OF STRESS BY OUTCOME

              CONTROLLING FOR REGION=urban

        Statistic                 DF    Value      Prob
        ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
        Chi-Square                 1   16.220     0.001
        Likelihood Ratio Chi-Square  1   17.352     0.001
        Continuity Adj. Chi-Square   1   15.035     0.001
        Mantel-Haenszel Chi-Square   1   16.155     0.001
        Fisher's Exact Test (Left)               1.000
                  (Right)             3.25E-05
                  (2-Tail)            4.55E-05
        Phi Coefficient              0.255
        Contingency Coefficient      0.247
        Cramer's V                   0.255

        Sample Size = 250
```

بسبب معنوية قيمة مربع كاي فأن درجة الاجهاد تؤثر في المرض في المنطقة الاولى.

```
              TABLE 2 OF STRESS BY OUTCOME

              CONTROLLING FOR REGION=rural

            STRESS    OUTCOME

            Frequency,
            Row Pct ,f     ,u     , Total
            ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
            low    ,   55 ,  135 ,  190
                 , 28.95 , 71.05 ,
```

Dr Firas

```
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
high   ,    7 ,   53 ,   60
     , 11.67 , 88.33 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total       62    188    250

STATISTICS FOR TABLE 2 OF STRESS BY OUTCOME
   CONTROLLING FOR REGION=rural

Statistic                DF    Value      Prob
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Chi-Square                1    7.302      0.007
Likelihood Ratio Chi-Square  1    8.198      0.004
Continuity Adj. Chi-Square   1    6.404      0.011
Mantel-Haenszel Chi-Square   1    7.272      0.007
Fisher's Exact Test (Left)                   0.999
          (Right)                 4.06E-03
          (2-Tail)                6.08E-03
Phi Coefficient                   0.171
Contingency Coefficient           0.168
Cramer's V                        0.171

Sample Size = 250
```

بسبب معنوية قيمة مربع كاي فأن درجة الاجهاد تؤثر في المرض في المنطقة الثانية.

```
SUMMARY STATISTICS FOR STRESS BY OUTCOME

CONTROLLING FOR REGION

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic  Alternative Hypothesis   DF     Value     Prob
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
  1      Nonzero Correlation       1     23.050    0.001
  2      Row Mean Scores Differ    1     23.050    0.001
  3      General Association       1     23.050    0.001


   Estimates of the Common Relative Risk (Row1/Row2)
                                95%
Type of Study   Method          Value   Confidence Bounds
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Case-Control    Mantel-Haenszel   3.542    2.114   5.935
 (Odds Ratio)   Logit             3.559    2.081   6.089

Cohort          Mantel-Haenszel   1.752    1.393   2.202
 (Col1 Risk)    Logit             1.629    1.356   1.956

Cohort          Mantel-Haenszel   0.661    0.558   0.783
 (Col2 Risk)    Logit             0.773    0.682   0.877

The confidence bounds for the M-H estimates are test-based.

   Breslow-Day Test for Homogeneity of the Odds Ratios

Chi-Square =   0.183        DF =   1        Prob = 0.669

Total Sample Size = 500
```

ارجحية المعرضين للاجهاد العالي للمنطقتين = 3.5 مرة اكثر للاصابة بالمرض مقارنة بالاجهاد القليل.

مثال (118) : في دراسة لمعرفة تأثير عمل المرأة على دخل العائلة ، جد ارجحية ارتفاع الدخل او انخفاضه بسبب عمل المرأة؟

Dr. Firas

```
data s;
input inc wifework;
cards;
10 0
10 1
10 1
11 0
11 1
11 1
11 1
11 1
12 0
12 1
12 1
12 1
12 1
12 1
12 1
12 1
12 1
;
proc logistic data=s
descending;
model wifework = inc;
run;
```

```
                                                The SAS System
                        The LOGISTIC Procedure

          Data Set: WORK.S
          Response Variable: WIFEWORK
          Response Levels: 2
          Number of Observations: 17
          Link Function: Logit

                             Response Profile

                        Ordered
                          Value  WIFEWORK     Count

                            1       1       14
                            2       0        3

           Model Fitting Information and Testing Global Null Hypothesis BETA=0

                             Intercept
                    Intercept       and
          Criterion      Only      Covariates    Chi-Square for Covariates

          AIC           17.844       19.102       .
          SC            18.677       20.769       .
          -2 LOG L      15.844       15.102       0.742 with 1 DF (p=0.3891)
          Score          .            .           0.781 with 1 DF (p=0.3770)


                        Analysis of Maximum Likelihood Estimates

                   Parameter   Standard     Wald       Pr >   Standardized   Odd Variable    DF   Estimate      Error     Chi-Square Chi-Square
Estimate     Ratio

 INTERCPT    1   -6.2383     8.9795     0.4827     0.4872       .       .
 INC         1    0.6931     0.8072     0.7373     0.3905     0.300335  2.000
```

Dr. Firas
Dr. Firas

Association of Predicted Probabilities and Observed Responses

Concordant = 47.6%	Somers' D = 0.286
Discordant = 19.0%	Gamma = 0.429
Tied = 33.3%	Tau-a = 0.088
(42 pairs)	c = 0.643

ارجحية الدخل العالي في العوائل التي تكون فيها النساء ذات عمل = 2 مرة مقارنة بالعوائل التي لاتعمل فيها النساء.

الفصل الثامن

تقدير مكونات التباين

Estimation of Variance Components

لكل عشيرة من الحيوانات خصائصها التي تميزها عن غيرها من العشائر، ومن بين هذه الخصائص متوسط العشيرة وتباينها الوراثي والبيئي ونسبة هذه التباينات الى بعضها.

ان اهمية تقديرمكونات التباين تتمثل في امكانية تقدير المكافىء الوراثي والارتباطات الوراثية والمظهرية بين الصفات الاقتصادية المرغوبة ، وبذلك نكون قد خطونا اولى الخطوات لغرض وضع سياسة صحيحة لتحسين اداء العشيرة في الحيوانات والنباتات.

8 – 1 تقدير المكافىء الوراثي Heritability Estimation

يعد المكافىء الوراثي من اهم المقاييس الوراثية اذ يعد مقياسا لقوة العلاقة بين مظهر الصفة على الحيوان وقابليته الوراثية ، اي يمثل دليلا للمربين يوضح مدى امكانية الاعتماد على مظهر الحيوان كمؤشر عن تركيبه الوراثي . ان تقدير المكافىء الوراثي لصفة ما مثل وزن الجسم عند البلوغ في قطيع من الاغنام لايمكن بدون سجلات النسب ، اي ان الحيوانات المدروسة تكون معروفة الأب او الأم اوكليهما

8 – 1 – 1 طرق تقدير المكافىء الوراثي

هناك عدة طرق لتقدير مكونات التباين عن طريق الأب او الأم للأخوة انصاف الاشقاء او الاشقاء اوكلا الابوين عند حساب المكافىء الوراثي.

8 – 1 – 1 – 1 تقدير المكافىء الوراثي للأخوة انصاف الاشقاء من مكونات تباين الأب في النماذج الخطية

يمكن تقدير ذلك بأربعة طرق.

أ - طريقة MIVQUE

ب - طريقة TYPE1

جـ - طريقة Maximum Likelihood

د - طريقة Restricted Maximum Likelihood

مثال (119) : البيانات التالية تمثل الوزن عند الولادة والوزن عند الفطام (كغم) لمجموعة من الحيوانات تعود الى اربعة آباء خلال أربعة فصول ، المطلوب تقدير المكافيء الوراثي للصفتين بأربعة طرق.

```
Data F;
input s p x y se;
xy=x+y;
cards;
1 1 4.5 27.1 1
1 2 4.3 25.5 2
1 3 4.4 26.4 2
1 4 3.9 24.3 3
1 5 4.5 27.0 1
1 6 4.1 24.3 4
1 7 4.3 25.7 4
1 8 4.4 26.1 2
1 9 4.6 28   3
1 10 4.4 29.14
2 1 4.1 24.8 1
2 2 4.2 25.1 2
2 3 3.6 23.0 4
2 4 3.8 24.1 2
2 5 4.1 25.0 3
2 6 4   25   4
2 7 4.1 25.2 1
2 8 4   24.9 2
2 9 4.2 26   4
2 10 4.3 26.14
3 1 3.9 23.3 3
3 2 3.8 24   2
3 3 3.9 25   3
3 4 4.1 25.4 2
3 5 4   25.1 1
3 6 4.1 25.3 1
3 7 3.9 24.3 3
3 8 4.4 26.2 4
3 9 4.3 26   1
3 10 4  25.1 3
4 1 4.1 25   2
4 2 4.2 26   2
4 3 4.6 27.5 1
4 4 4.8 30.1 4
4 5 4.3 26.3 4
4 6 4.2 25.9 3
4 7 4.3 26.2 2
4 8 3.8 24.1 1
4 9 3.7 23.8 3
4 10 4.1 25.1 4
```

Dr. Firas

```
 ; option nodate nonumber;
;proc GLM; classes se  ;
model x y  = se   /ss3;
means se   /duncan;
lsmeans se  /stderr pdiff; ← يستعمل هذا الايعاز للحصول على متوسط المربعات الصغرى
 quit;
;proc varcomp method= MIVQUE0; classes se  s  ;
model x y xy  = se  s  /fixed=1;
quit;
;proc varcomp method= type1; classes se s  ;
model x y xy = se  s  /fixed=1;
quit;
 ;proc varcomp method= ml; classes se s  ;
model x y xy = se  s  /fixed=1;
quit;
;proc varcomp method= rmel; classes se s  ;
model x y xy = se  s  /fixed=1;
quit;
proc glm ; classes se s ;
 model y = se s ;
 random s /test ;
run;
 ;proc mixed;classes se  s;
model x =  se  ;
random s/solution;
quit;
;proc mixed;classes se  s;
model y =  se  ;
random s/solution;
quit;
```

```
                    The SAS System

                General Linear Models Procedure
                  Class Level Information

                 Class    Levels    Values

                 SE          4    1 2 3 4

                Number of observations in data set = 38


                    The SAS System

                General Linear Models Procedure

Dependent Variable: X

Source          DF    Sum of Squares   Mean Square  F Value  Pr > F

Model            3     0.14010965      0.04670322    0.64   0.5952
Error           34    2.48541667      0.07310049
Corrected Total 37    2.62552632

 R-Square         C.V.          Root MSE          X Mean

 0.053364       6.490271          0.2703710        4.16578947


Source          DF    Type III SS   Mean Square  F Value    Pr > F

SE              3     0.14010965     0.04670322     0.64    0.5952
```

Dr. Firas

The SAS System

General Linear Models Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1.84631082	0.61543694	0.29	0.8341
Error	34	72.78119444	2.14062337		
Corrected Total	37	74.62750526			

R-Square	C.V.	Root MSE	Y Mean
0.024740	5.705916	1.4630869	25.64157895

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SE	3	1.84631082	0.61543694	0.29	0.8341

The SAS System

General Linear Models Procedure
Least Squares Means

SE	X	Std Err LSMEAN	Pr > |T| LSMEAN	Pr > |T| H0: LSMEAN(i)=LSMEAN(j) H0:LSMEAN=0	i/j	1	2	3	4
	1	4.23750000	0.09559059	0.0001	1	.	0.5698	0.2022	0.7770
	2	4.16666667	0.07804939	0.0001	2	0.5698	.	0.4075	0.7815
	3	4.06666667	0.09012367	0.0001	3	0.2022	0.4075	.	0.3029
4		4.20000000	0.09012367	0.0001	4	0.7770	0.7815	0.3029	.

SE	Y	Std Err LSMEAN	Pr > |T| LSMEAN	Pr > |T| H0: LSMEAN(i)=LSMEAN(j) H0:LSMEAN=0	i/j	1	2	3	4
	1	25.9125000	0.5172793	0.0001	1	.	0.7004	0.3815	0.8146
	2	25.6533333	0.4223568	0.0001	2	0.7004	.	0.5689	0.8885
	3	25.2822222	0.4876956	0.0001	3	0.3815	0.5689	.	0.5073
4		25.7444444	0.4876956	0.0001	4	0.8146	0.8885	0.5073	.

NOTE: To ensure overall protection level, only probabilities associated with pre-planned comparisons should be used.

The SAS System

Variance Components Estimation Procedure
Class Level Information

Class	Levels	Values
SE	4	1 2 3 4
S	4	1 2 3 4

Number of observations in data set = 38

The SAS System

MIVQUE(0) Variance Component Estimation Procedure

SSQ Matrix

Dr Firas

Source	S	Error	X	Y	XY
S	256.84259259	27.69444444	5.39447917	118.11034954	173.09937037
Error	27.69444444	34.00000000	2.48541667	72.78119444	99.78311111

Variance Component	Estimate X	Estimate Y	Estimate XY
Var(S)	0.01438424	**0.25109227**	0.39192464
Var(Error)	0.06138391	**1.93609804**	2.61555811

The SAS System

Variance Components Estimation Procedure
Class Level Information

Class	Levels	Values
SE	4	1 2 3 4
S	4	1 2 3 4

Number of observations in data set = 38

The SAS System

Variance Components Estimation Procedure (Type 1)

Dependent Variable: X

Source	DF	Type I SS	Type I MS
SE	3	0.14010965	0.04670322
S	3	0.57109104	0.19036368
Error	31	1.91432563	0.06175244
Corrected Total	37	2.62552632	

Source	Expected Mean Square
SE	Var(Error) + 0.2597 Var(S) + Q(SE)
S	Var(Error) + 9.2315 Var(S)
Error	Var(Error)

Variance Component	Estimate	
Var(S)	**0.01393181**	← تباين الآباء للصفة الاولى
Var(Error)	**0.06175244**	← تباين الخطأ للصفة الاولى

The SAS System

Variance Components Estimation Procedure (Type 1)

Dependent Variable: Y

Source	DF	Type I SS	Type I MS
SE	3	1.84631082	0.61543694
S	3	12.55732326	4.18577442
Error	31	60.22387118	1.94270552
Corrected Total	37	74.62750526	

Source	Expected Mean Square
SE	Var(Error) + 0.2597 Var(S) + Q(SE)
S	Var(Error) + 9.2315 Var(S)

Dr. Firas

Error Var(Error)

Variance Component Estimate

Var(S) **0.24298038** ← تباين الآباء للصفة الثانية

Var(Error) **1.94270552** ← تباين الخطأ للصفة الثانية

The SAS System

Variance Components Estimation Procedure

Dependent Variable: XY

Source	DF	Type I SS	Type I MS
SE	3	3.00202573	1.00067524
S	3	18.39335166	6.13111722
Error	31	81.38975945	2.62547611
Corrected Total	37	102.78513684	

Source	Expected Mean Square
SE	Var(Error) + 0.2597 Var(S) + Q(SE)
S	Var(Error) + 9.2315 Var(S)
Error	Var(Error)

Variance Component Estimate

Var(S) **0.37974849** ← تباين الآباء للصفتين

Var(Error) **2.62547611** ← تباين الخطأ للصفتين

The SAS System

Variance Components Estimation Procedure
Class Level Information

Class	Levels	Values
SE	4	1 2 3 4
S	4	1 2 3 4

Number of observations in data set = 38

The SAS System

Maximum Likelihood Variance Components Estimation Procedure

Dependent Variable: X

Iteration	Objective	Var(S)	Var(Error)
0	-105.49578381	0.01290109	0.05505463
1	-105.59140163	0.00920865	0.05628186
2	-105.59145922	0.00912863	0.05631531
3	-105.59145922	0.00912820	0.05631549

Convergence criteria met.

Asymptotic Covariance Matrix of Estimates

	Var(S)	Var(Error)
Var(S)	0.0001148395	-.0000193642
Var(Error)	-.0000193642	0.0001864263

The SAS System

Dr. Firas

Maximum Likelihood Variance Components Estimation Procedure

Dependent Variable: Y

Iteration	Objective	Var(S)	Var(Error)
0	24.11999812	0.22484448	1.73370913
1	24.03060251	0.14588147	1.77104071
2	24.03051648	0.14370298	1.77232109
3	24.03051647	0.14367316	1.77233873
4	24.03051647	**0.14367316**	**1.77233873**

Convergence criteria met.

Asymptotic Covariance Matrix of Estimates

	Var(S)	Var(Error)
Var(S)	0.0565712085	-.0194243273
Var(Error)	-.0194243273	0.1847464308

The SAS System

Maximum Likelihood Variance Components Estimation Procedure

Dependent Variable: XY

Iteration	Objective	Var(S)	Var(Error)
0	35.88812245	0.35103494	2.34267559
1	35.79798467	0.23511532	2.39335600
2	35.79790636	0.23208038	2.39500863
3	35.79790635	0.23204475	2.39502816
4	35.79790635	0.23204475	2.39502816

Convergence criteria met.

Asymptotic Covariance Matrix of Estimates

	Var(S)	Var(Error)
Var(S)	0.1208169641	-.0354029190
Var(Error)	-.0354029190	0.3373321895

The SAS System

Variance Components Estimation Procedure
Class Level Information

Class	Levels	Values
SE	4	1 2 3 4
S	4	1 2 3 4

Number of observations in data set = 38

The SAS System

Restricted Maximum Likelihood Variance Components Estimation Procedure

Dependent Variable: X

Iteration	Objective	Var(S)	Var(Error)
0	-91.34742227	0.01441886	0.06153165

Dr.Firas

1	-91.34883446	0.01378375	0.06170554
2	-91.34883474	0.01379264	0.06170302
3	-91.34883474	0.01379264	0.06170302

Convergence criteria met.

Asymptotic Covariance Matrix of Estimates

	Var(S)	Var(Error)
Var(S)	0.0002782556	-.0000253496
Var(Error)	-.0000253496	0.0002452505

The SAS System

Restricted Maximum Likelihood Variance Components Estimation Procedure

Dependent Variable: Y

Iteration	Objective	Var(S)	Var(Error)
0	24.85012996	0.25129677	1.93767491
1	24.84963330	0.24290287	1.94094028
2	24.84963315	0.24304805	1.94088277
3	24.84963315	**0.24304530**	**1.94088385**

Convergence criteria met.

Asymptotic Covariance Matrix of Estimates

	Var(S)	Var(Error)
Var(S)	0.1374357303	-.0253105377
Var(Error)	-.0253105377	0.2426054600

The SAS System

Restricted Maximum Likelihood Variance Components Estimation Procedure

Dependent Variable: XY

Iteration	Objective	Var(S)	Var(Error)
0	35.32872696	0.39233317	2.61828448
1	35.32810683	0.37864636	2.62320957
2	35.32810666	0.37887086	2.62312711
3	35.32810666	0.37887086	2.62312711

Convergence criteria met.

Asymptotic Covariance Matrix of Estimates

	Var(S)	Var(Error)
Var(S)	0.2934061190	-.0461708128
Var(Error)	-.0461708128	0.4431706985

The SAS System

The SAS System

General Linear Models Procedure

Source Type III Expected Mean Square

SE Var(Error) + Q(SE)

S Var(Error) + 9.2315 Var(S)

Dr Firas

The SAS System
General Linear Models Procedure
Tests of Hypotheses for Mixed Model Analysis of Variance

Dependent Variable: Y

Source: SE
Error: MS(Error)

DF	Type III MS	Denominator DF	Denominator MS	F Value	Pr > F
3	0.4865155317	31	1.942705522	0.2504	0.8604

Source: S
Error: MS(Error)

DF	Type III MS	Denominator DF	Denominator MS	F Value	Pr > F
3	4.1857744206	31	1.942705522	2.1546	0.1134

تباين الاباء = (4.18 – 1.94)\9.23 = 0.24
تباين الخطأ = 1.94

The MIXED Procedure

Class Level Information

Class	Levels	Values
SE	4	1 2 3 4
S	4	1 2 3 4

REML Estimation Iteration History

Iteration	Evaluations	Objective	Criterion
0	1	-45.98248967	
1	2	-48.39003121	0.00000026
2	1	-48.39003740	0.00000000

Convergence criteria met.

Covariance Parameter Estimates (REML)

Cov Parm	Estimate
S	0.01379232
Residual	0.06170311

Model Fitting Information for X

Description	Value
Observations	38.0000
Res Log Likelihood	-7.0489
Akaike's Information Criterion	-9.0489
Schwarz's Bayesian Criterion	-10.5753
-2 Res Log Likelihood	14.0978

Solution for Random Effects

Effect	S	Estimate	SE Pred	DF	t	Pr > |t|
S	1	0.12303267	0.08193537	31	1.50	0.1433
S	2	-0.08396375	0.08313462	31	-1.01	0.3203
S	3	-0.06898983	0.08365250	31	-0.82	0.4158
S	4	0.02992090	0.08188530	31	0.37	0.7173

Dr. Firas

The SAS System
Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
SE	3	31	0.68	0.5715

The SAS System
The MIXED Procedure

Class Level Information

Class	Levels	Values
SE	4	1 2 3 4
S	4	1 2 3 4

REML Estimation Iteration History

Iteration	Evaluations	Objective	Criterion
0	1	68.83609804	
1	2	67.80843122	0.00000002
2	1	67.80843050	0.00000000

Convergence criteria met.

Covariance Parameter Estimates (REML)

Cov Parm	Estimate
S	0.24304487
Residual	1.94088402

Model Fitting Information for Y

Description	Value
Observations	38.0000
Res Log Likelihood	-65.1481
Akaike's Information Criterion	-67.1481
Schwarz's Bayesian Criterion	-68.6745
-2 Res Log Likelihood	130.2963

Solution for Random Effects

Effect	S	Estimate	SE Pred	DF	t	Pr > |t|
S	1	0.40351100	0.37824617	31	1.07	0.2943
S	2	-0.35793661	0.38341431	31	-0.93	0.3578
S	3	-0.24472406	0.38562926	31	-0.63	0.5303
S	4	0.19914967	0.37802515	31	0.53	0.6021

The SAS System
Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
SE	3	31	0.28	0.8421

Dr. Firas

ومن النتائج المتحصل عليها من مثالنا يمكن تقدير المكافىء الوراثي بخمسة طرق (MIVQUE , TYPE1, RMEL, ML متوسط المربعات المتوقع) ويلاحظ ايضا ان تقديرات المكافىء الوراثي كانت متساوية في كل من TYPE1 و RMEL ومتوسط المربعات المتوقع ، كما

يمكن تقدير الارتباطات الوراثية والبيئية والمظهرية وذلك بتطبيق القوانين الخاصة بكل منها وسنعتمد على نتائج طريقة RMEL وكما موضح ادناه:

المكافىء الوراثي (h^2) للصفة Y المقدر حسب كل طريقة ووفق المعادلة التالية:

$h^2_y = 4 [\sigma^2 s/ (\sigma^2 s + \sigma^2 e)]$

MIVQUE = (4* 0.25)/(1.93 + 0.25)= 0.46

TYPE 1 = (4*0.24)/(1.94 + 0.24) = 0.44

ML = (4* 0.14)/(1.77 + 0.14) = 0.29

REML = (4*0.24)/ (1.94 + 0.24) = 0.44

8 – 2 الارتباط الوراثي Genetic correlation

نطبق المعادلة الخاصة بتقدير معامل الارتباط الوراثي (rG) بعد تقدير التباين المشترك (Covariance) بين الصفتين وحسب المعادلة التالية:

Covariance xy = 1/2 [Variance xy – Variance x – Variance y]

rG = Cov xy/ $\sqrt{}$Vx*Vy

Cov xy = ½ [0.37 – 0.013– 0.24] = 0.058

$\sqrt{}$Vx*Vy = $\sqrt{}$0.013*0.24 = 0.055

rG = 0.058 / 0.055 = 1.05

والتباين المقصود هنا هو تباين الآباء (S)

8 – 3 الارتباط البيئي Environmental Correlation

نستخرج التباين المشترك للأبناء الذي يمثله تباين الخطأ

Cov $_{progeny}$ = 1/2 [Variance xy $_{error}$ – Variance x $_{error}$ – Variance y $_{error}$]

Cov $_{progeny}$ = 1/2[2.62 – 1.94 – 0.061] = 0.31

نطبق المعادلة الخاصة بتقدير الارتباط البيئي بين الصفتين

rE = (Cov $_{progeny}$ – Cov s) / $\sqrt{}$(σ^2Error(x) - σ^2s (x)) (σ^2Error(y) - σ^2s (y))

rG= (0.31 – 0.058) /[$\sqrt{}$ (0.061 – 0.013)* (1.94 – 0.24)] = 0.89

8 – 4 الارتباط المظهري Phenotypic Correlation

Dr Firas

لتقدير الارتباط المظهري نطبق المعادلة الخاصة بتقدير الارتباط المظهري:

$$rP = (Cov_{progeny} + Cov\ s) / \sqrt{(\sigma^2 Error(x) + \sigma^2 s(x))(\sigma^2 Error(y) + \sigma^2 s(y))}$$

$$rP = (0.31 + 0.058)/ [\sqrt{(0.061 + 0.013)*(1.94 + 0.24)}] = 0.92$$

8 – 5 تقدير قيم BLUP

لغرض تقدير افضل تنبوء خطي غير منحاز للآباء للصفات المدروسة (BLUP) فيمكن الحصول عليها باستعمال الايعاز proc mixed :

```
;proc mixed;classes se  s;
model y= se s ;
random s /solution;
run;
```

وبسبب قلة عدد البيانات فأننا لم نحصل على نتائج قيم BLUP. لذا سنفترض ان البيانات كبيرة فأننا سنحصل على النتائج وكما موضح ادناه:

```
                                The SAS System

                        The MIXED Procedure

                        Class Level Information

Class     Levels  Values

SE             4  1 2 3 4
YB            12  88 89 90 91 92 93 94 95 96 97
                  98 99
SIR           93  0675 0761 0773 0800 0801 0802
                  1 1060 14H0689 14H0699 1505
                  1633 1671 1818002 2 2H53 3432
                  404 45 4510 49 525 537 543
                  660 753 754 755 761 7651 7652
                  7664 7665 7667 7669 767 7673
                  7677 7687 7696 7H1530 7H2717
                  7H800 7H935 800 9133 916 9163
                  92233 93323 96232 96283 97865
                  98765 99044 9913 99144 99229
                  99300 99302 99333 N651 N652
                  N653 N654 N655 N656 N657 N658
                  N659 N661 N662 N663 N665 N669
                  N671 N672 N674 N751 N752 N756
                  N757 N759 N760 N762 N765 N767
                  N771 N772 N774 N962 NB1 NB2


                REML Estimation Iteration History

Iteration  Evaluations    Objective    Criterion

    0          1  7012.2588790
    1          2  6981.9059551    0.00092186
    2          1  6977.8174055    0.00031666
    3          1  6976.4959709    0.00004617
    4          1  6976.3195244    0.00000121
    5          1  6976.3152438    0.00000000

                Convergence criteria met.

                Covariance Parameter Estimates (REML)

                Cov Parm      Estimate
```

Dr Firas

SIR 1.20946337
Residual 10.50027799

The SAS System

Model Fitting Information for AFI

Description	Value
Observations	2053.000
Res Log Likelihood	-5360.95
Akaike's Information Criterion	-5362.95
Schwarz's Bayesian Criterion	-5368.57
-2 Res Log Likelihood	10721.91

Solution for Random Effects

Effect	SIR	Estimate	SE Pred	DF	t	Pr > |t|
SIR	0675	-0.09721264	1.04189386	1946	-0.09	0.9257
SIR	0761	0.19370136	0.99284085	1946	0.20	0.8453
SIR	0773	-0.02277102	1.04181110	1946	-0.02	0.9826
SIR	0800	-0.18220623	0.95112705	1946	-0.19	0.8481
SIR	0801	-0.17341019	0.99301717	1946	-0.17	0.8614
SIR	0802	0.04377677	1.04186223	1946	0.04	0.9665
SIR	1	0.88019910	0.38556314	1946	2.28	0.0225
SIR	1060	-1.06667820	0.67514212	1946	-1.58	0.1143
SIR	14H0689	0.84065183	0.74903716	1946	1.12	0.2619
SIR	14H0699	-0.67746451	0.88159147	1946	-0.77	0.4423
SIR	1505	0.03646578	0.36729880	1946	0.10	0.9209
SIR	1633	-1.09138583	0.61074976	1946	-1.79	0.0741
SIR	1671	-0.26511901	0.77238302	1946	-0.34	0.7314
SIR	1818002	-1.10024248	0.74080751	1946	-1.49	0.1377
SIR	2	0.60234141	0.85459692	1946	0.70	0.4810
SIR	2H53	0.03771205	0.99803117	1946	0.04	0.9699
SIR	3432	2.98111236	0.99810840	1946	2.99	0.0029
SIR	404	0.28514792	0.56627424	1946	0.50	0.6146
SIR	45	-0.14220892	0.61246179	1946	-0.23	0.8164
SIR	4510	1.66852634	0.60679932	1946	2.75	0.0060
SIR	49	-0.00577426	1.04212313	1946	-0.01	0.9956
SIR	525	-0.28488890	1.04334165	1946	-0.27	0.7848
SIR	537	1.08424217	0.89442393	1946	1.21	0.2256
SIR	543	0.86114109	0.45866598	1946	1.88	0.0606
SIR	660	0.19703509	0.95123121	1946	0.21	0.8359
SIR	753	0.68195805	0.85237278	1946	0.80	0.4238
SIR	754	0.05371663	0.68034729	1946	0.08	0.9371
SIR	755	0.09416176	0.87962551	1946	0.11	0.9148
SIR	761	-0.03918380	0.82415216	1946	-0.05	0.9621
SIR	7651	-1.23517821	0.55816128	1946	-2.21	0.0270
SIR	7652	0.69951711	1.04303435	1946	0.67	0.5025
SIR	7664	-0.51507067	0.89884905	1946	-0.57	0.5667
SIR	7665	0.58597515	0.86864610	1946	0.67	0.5000
SIR	7667	0.02397316	0.34525670	1946	0.07	0.9446
SIR	7669	-0.22261359	0.29771526	1946	-0.75	0.4547
SIR	767	1.50261060	0.78433988	1946	1.92	0.0555

نأخذ النتائج الخاصة بقيم BLUP ونعمل ملف جديد ونرتب البيانات حسب قيم BLUP تصاعديا ونحدد اختيارنا في قطعة البرنامج برقم الاب وقيمة BLUP.

```
DATA D;
INPUT A$ B$ C D E F G;
CARDS;
```

SIR 0675 -0.09721264 1.04189386 1946 -0.09 0.9257
SIR 0761 0.19370136 0.99284085 1946 0.20 0.8453

SIR	0773	-0.02277102	1.04181110	1946	-0.02	0.9826
SIR	0800	-0.18220623	0.95112705	1946	-0.19	0.8481
SIR	0801	-0.17341019	0.99301717	1946	-0.17	0.8614
SIR	0802	0.04377677	1.04186223	1946	0.04	0.9665
SIR	1	0.88019910	0.38556314	1946	2.28	0.0225
SIR	1060	-1.06667820	0.67514212	1946	-1.58	0.1143
SIR	14H0689	0.84065183	0.74903716	1946	1.12	0.2619
SIR	14H0699	-0.67746451	0.88159147	1946	-0.77	0.4423
SIR	1505	0.03646578	0.36729880	1946	0.10	0.9209
SIR	1633	-1.09138583	0.61074976	1946	-1.79	0.0741
SIR	1671	-0.26511901	0.77238302	1946	-0.34	0.7314
SIR	1818002	-1.10024248	0.74080751	1946	-1.49	0.1377
SIR	2	0.60234141	0.85459692	1946	0.70	0.4810
SIR	2H53	0.03771205	0.99803117	1946	0.04	0.9699
SIR	3432	2.98111236	0.99810840	1946	2.99	0.0029
SIR	404	0.28514792	0.56627424	1946	0.50	0.6146
SIR	45	-0.14220892	0.61246179	1946	-0.23	0.8164
SIR	4510	1.66852634	0.60679932	1946	2.75	0.0060
SIR	49	-0.00577426	1.04212313	1946	-0.01	0.9956
SIR	525	-0.28488890	1.04334165	1946	-0.27	0.7848
SIR	537	1.08424217	0.89442393	1946	1.21	0.2256
SIR	543	0.86114109	0.45866598	1946	1.88	0.0606
SIR	660	0.19703509	0.95123121	1946	0.21	0.8359
SIR	753	0.68195805	0.85237278	1946	0.80	0.4238
SIR	754	0.05371663	0.68034729	1946	0.08	0.9371
SIR	755	0.09416176	0.87962551	1946	0.11	0.9148
SIR	761	-0.03918380	0.82415216	1946	-0.05	0.9621
SIR	7651	-1.23517821	0.55816128	1946	-2.21	0.0270
SIR	7652	0.69951711	1.04303435	1946	0.67	0.5025
SIR	7664	-0.51507067	0.89884905	1946	-0.57	0.5667
SIR	7665	0.58597515	0.86864610	1946	0.67	0.5000
SIR	7667	0.02397316	0.34525670	1946	0.07	0.9446
SIR	7669	-0.22261359	0.29771526	1946	-0.75	0.4547
SIR	767	1.50261060	0.78433988	1946	1.92	0.0555

```
 ;proc sort;by C ;
proc print; id B;var C;
quit;
```

The SAS System

B	C
7651	-1.23518
1818002	-1.10024
1633	-1.09139
1060	-1.06668
14H0699	-0.67746
7664	-0.51507
525	-0.28489
1671	-0.26512
7669	-0.22261
0800	-0.18221
0801	-0.17341
45	-0.14221
0675	-0.09721
761	-0.03918
0773	-0.02277
49	-0.00577
7667	0.02397
1505	0.03647
2H53	0.03771
0802	0.04378
754	0.05372
755	0.09416
0761	0.19370

Dr. Firas

```
660       0.19704
404       0.28515
7665      0.58598
2         0.60234
753       0.68196
7652      0.69952
14H0689   0.84065
543       0.86114
1         0.88020
537       1.08424
767       1.50261
4510      1.66853
3432      2.98111
```

مثال (120) : في قطيع كبير من الاغنام غير مربى تربية داخلية ،اختير 5 كباش بصورة عشوائية وكل كبش لقح ثمانية نعاج وفي كل تلقيحة كان الناتج ذكرا واحدا ، وقد اخذت صفة معدل الزيادة اليومية للحملان من الفطام لغاية عمر 6 أشهر بالغرامات ، المطلوب تقدير المكافىء الوراثي للصفة؟ (غزال ، 1982).

```
data s;
input sire w;
cards;
1 187
1 191
1 293
1 175
1 200
1 253
1 204
1 217
2 118
2 180
2 92
2 183
2 131
2 191
2 194
2 232
3 118
3 187
3 263
3 247
3 178
3 237
3 231
3 103
4 100
4 157
4 169
4 106
4 218
4 193
4 169
4 148
5 217
5 158
```

Dr. Firas

```
5 174
5 111
5 178
5 288
5 150
5 190
 proc glm data=s; classes sire ;
 model w = sire ;
 random sire /test ;
run;
 ;proc varcomp method= MIVQUE0; classes sire  ;
model w  =   sire  /fixed=0;

quit;
;proc varcomp method= type1; classes sire  ;
model w = sire  /fixed=0;
quit;
```

```
                                        The SAS System
                    General Linear Models Procedure
                      Class Level Information

                     Class    Levels    Values

                     SIRE        5    1 2 3 4 5

                   Number of observations in data set = 40

                         The SAS System
                    General Linear Models Procedure

Dependent Variable: W

Source             DF      Sum of Squares     Mean Square    F Value    Pr > F

Model               4     17197.60000000     4299.40000000     1.84    0.1428
Error              35     81686.37500000     2333.89642857

Corrected Total    39     98883.97500000

     R-Square              C.V.             Root MSE          W Mean
     0.173917             26.35952          48.3104173       183.27500000


Source             DF      Type I SS     Mean Square  F Value    Pr > F

SIRE                4     17197.60000000  4299.40000000     1.84    0.1428

Source             DF     Type III SS     Mean Square    F Value    Pr > F
SIRE                4     17197.60000000  4299.40000000     1.84    0.1428


                         The SAS System
                    General Linear Models Procedure

Source     Type III Expected Mean Square
SIRE       Var(Error) + 8 Var(SIRE)

                         The SAS System
                    General Linear Models Procedure
          Tests of Hypotheses for Random Model Analysis of Variance

Dependent Variable: W
```

Dr Firas

Source: SIRE
Error: MS(Error)

DF	Type III MS	Denominator DF	Denominator MS	F Value	Pr > F
4	4299.4	35	2333.8964286	1.8422	0.1428

تباين الآباء يساوي (4299.4 – 2333.9) \ 8 = 245.68
المكافىء الوراثي يساوي 4 (245.68)\(245.68 + 2333.9) = 0.38

The SAS System
Variance Components Estimation Procedure
Class Level Information

Class	Levels	Values
SIRE	5	1 2 3 4 5

Number of observations in data set = 40

The SAS System
MIVQUE(0) Variance Component Estimation Procedure

SSQ Matrix

Source	SIRE	Error	W
SIRE	256.0000000	32.0000000	137580.8000000
Error	32.0000000	39.0000000	98883.9750000

Variance Component	Estimate W
Var(SIRE)	245.68794643
Var(Error)	2333.89642857

المكافىء الوراثي = 4(245.68)\ (245.68 + 2333.9) = 0.38

The SAS System
Variance Components Estimation Procedure
Class Level Information

Class	Levels	Values
SIRE	5	1 2 3 4 5

Number of observations in data set = 40
The SAS System
Variance Components Estimation Procedure

Dependent Variable: W

Source	DF	Type I SS	Type I MS
SIRE	4	17197.60000000	4299.40000000
Error	35	81686.37500000	2333.89642857
Corrected Total	39	98883.97500000	

Source	Expected Mean Square
SIRE	Var(Error) + 8 Var(SIRE)
Error	Var(Error)

Variance Component	Estimate
Var(SIRE)	245.68794643
Var(Error)	2333.89642857

المكافىء الوراثي = 4(245.68)\ (245.68 + 2333.9) = 0.38

Dr Firas

8 – 1 – 1 – 2 تقدير المكافىء الوراثي عن طريق الأب او الام او كليهما في حالة النموذج المحتضن

أ - في حالة تساوي عدد النسل داخل كل مجموعة

النموذج الرياضي:

$Y_{ijk} = \mu + S_i + D_{ij} + e_{ijk}$

Y_{ijk} = سجل النسل k العائد الى الام i التي تزاوجت مع الاب j ، μ = المتوسط العام ، S_i = تأثير الاب i ، D_{ij}= تأثير الام j التي لقحت من الاب i ، e_{ijk} = الخطأ العشوائي.

مثال (121) : في قطيع كبير من الدجاج غير مربى تربية داخلية اختيرت 5 ذكور وكل منها لقح ثلاث امهات وكل منها انتج ثلاثة افراد من السل البنات واخذ وزن الجسم للبنات بعمر 8 أسابيع بالغرامات المطلوب تقدير المكافىء الوراثي للصفة؟ (غزال ، 1982).

```
data s;
input sire dam rep w;
cards;
1   1   1   965
1   1   2   813
1   1   3   765
1   2   1   803
1   2   2   640
1   2   3   714
1   3   1   644
1   3   2   753
1   3   3   705
2   1   1   740
2   1   2   798
2   1   3   941
2   2   1   701
2   2   2   847
2   2   3   909
2   3   1   909
2   3   2   800
2   3   3   853
3   1   1   696
3   1   2   807
3   1   3   800
3   2   1   752
3   2   2   863
3   2   3   739
3   3   1   686
3   3   2   832
3   3   3   796
4   1   1   979
4   1   2   798
4   1   3   788
4   2   1   905
4   2   2   880
4   2   3   770
4   3   1   797
```

Dr. Firas

```
4    3    2    721
4    3    3    765
5    1    1    809
5    1    2    756
5    1    3    775
5    2    1    887
5    2    2    935
5    2    3    937
5    3    1    872
5    3    2    811
5    3    3    925
proc glm data=s;
 class sire dam;
 model w = sire dam(sire);
 random sire dam(sire)/test ;
run;
```

```
                    The SAS System
                General Linear Models Procedure
                  Class Level Information

                    Class    Levels    Values

                    SIRE        5    1 2 3 4 5

                    DAM         3    1 2 3

                 Number of observations in data set = 45

                    The SAS System
                General Linear Models Procedure

Dependent Variable: W

Source             DF    Sum of Squares   Mean Square    F Value    Pr > F

Model              14   151321.20000000  10808.65714286     1.96    0.0603
Error              30   165732.00000000   5524.40000000
Corrected Total    44   317053.20000000

     R-Square             C.V.           Root MSE       W Mean
     0.477274         9.193491          74.3263075      808.46666667


Source             DF    Type I SS     Mean Square    F Value    Pr > F

SIRE                4   63208.31111111  15802.07777778     2.86    0.0404
DAM(SIRE)          10   88112.88888889   8811.28888889     1.59    0.1561

Source             DF    Type III SS    Mean Square    F Value    Pr > F

SIRE                4   63208.31111111  15802.07777778     2.86    0.0404
DAM(SIRE)          10   88112.88888889   8811.28888889     1.59    0.1561



                    The SAS System
                General Linear Models Procedure

Source     Type III Expected Mean Square

SIRE       Var(Error) + 3 Var(DAM(SIRE)) + 9 Var(SIRE)
DAM(SIRE)  Var(Error) + 3 Var(DAM(SIRE))


                    The SAS System
                General Linear Models Procedure
```

Dr Firas

Tests of Hypotheses for Random Model Analysis of Variance

Dependent Variable: W

Source: SIRE
Error: MS(DAM(SIRE))

DF	Type III MS	Denominator DF	Denominator MS	F Value	Pr > F
4	15802.077778	10	8811.2888889	1.7934	0.2068

Source: DAM(SIRE)
Error: MS(Error)

DF	Type III MS	Denominator DF	Denominator MS	F Value	Pr > F
10	8811.2888889	30	5524.4	1.5950	0.1561

من جدول تحليل التباين يمكن الحصول على التباينات (الارقام مؤشر تحتها خط) :

تباين الأبناء يساوي 5524

تباين الامهات يساوي (8811 – 5524)\ 3= 1095

تباين الآباء يساوي (15802 – 8811) \ 9 = 776

التباين الكلي (المظهري) يساوي 776 + 1095 + 5524 = 7395

المكافىء الوراثي عن طريق الآباء يساوي 4(776) \ 7395 = 0.42

المكافىء الوراثي عن طريق الامهات يساوي 4(1095)\ 7395 = 0.59

المكافىء الوراثي عن طريق الآباء والامهات يساوي 2(776 + 1095)\ 7395 = 0.51

ب - في حالة عدم تساوي عدد افراد النسل داخل المجاميع

مثال (122): في احدى التجارب الوراثية اختيرت 5 ذكور (ديكة) وكل منها لقح عدد مختلف من الاناث وانتجت عدد مختلف من الافراخ المطلوب تقدير المكافىء الوراثي لوزن الافراخ بعمر يوم واحد ؟ (غزال ، 1982).

```
data s;
input sire dam rp w;
cards;
1    1 1   40
1    1 2   41
1    1 3   45
1    1 4   39
1    1 5   43
1    2 1   41
1    2 2   43
1    2 3   42
1    2 4   46
1    3 1   40
1    3 2   41
1    3 3   42
1    3 4   45
1    3 5   43
```

Dr Firas

1	3	6	47
2	1	1	39
2	1	2	41
2	1	3	40
2	1	4	42
2	2	1	42
2	2	2	40
2	2	3	41
2	2	4	37
2	2	5	39
2	2	6	42
2	2	7	41
2	2	8	40
2	2	9	42
2	2	10	39
2	3	1	38
2	3	2	39
2	3	3	40
2	3	4	41
2	3	5	40
2	3	6	40
2	4	1	42
2	4	2	41
2	4	3	43
2	4	4	45
2	4	5	38
2	4	6	41
2	4	7	43
2	5	1	40
2	5	2	43
2	5	3	40
2	5	4	41
2	5	5	39
2	5	6	45
2	5	7	42
2	5	8	39
2	5	9	38
3	1	1	40
3	1	2	41
3	1	3	37
3	1	4	41
3	1	5	40
3	1	6	39
3	1	7	40
3	2	1	41
3	2	2	42
3	2	3	40
3	2	4	41
3	2	5	43
3	3	1	43
3	3	2	41
3	3	3	42
3	3	4	40
3	4	1	44
4	1	2	42
4	1	3	43
4	1	4	40

Dr. Firas

```
4    1 5   41
4    1 6   42
4    1 7   43
4    1 8   49
4    1 9   39
4    2 1   41
4    2 2   42
4    2 3   40
4    2 4   43
4    2 5   37
4    2 6   42
5    1 1   43
5    2 2   42
5    2 3   41
5    2 4   46
5    2 5   43
5    2 6   37
5    3 1   40
5    3 2   45
5    3 3   41
5    4 1   39
5    4 2   41
5    4 3   40
5    4 4   41
5    4 5   43
5    4 6   40
5    5 1   40
5    5 2   42
5    5 3   43
5    5 4   42
proc glm data=s;
 class sire dam;
 model w = sire dam(sire);
 random sire dam(sire)/test;run;
```

The SAS System
General Linear Models Procedure
Class Level Information

Class	Levels	Values
SIRE	5	1 2 3 4 5
DAM	5	1 2 3 4 5

Number of observations in data set = 101

The SAS System
General Linear Models Procedure

Dependent Variable: W

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	18	106.97135392	5.94285300	1.38	0.1644
Error	82	353.26626984	4.30812524		
Corrected Total	100	460.23762376			

R-Square	C.V.	Root MSE	W Mean
0.232426	5.028444	2.0756023	41.27722772

Dr. Firas
Dr. Firas

Source	DF	Type I SS	Mean Square	F Value	Pr > F
SIRE	4	44.84004403	11.21001101	2.60	0.0419
DAM(SIRE)	14	62.13130989	4.43795071	1.03	0.4328

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SIRE	4	42.14213598	10.53553399	2.45	0.0529
DAM(SIRE)	14	62.13130989	4.43795071	1.03	0.4328

The SAS System
General Linear Models Procedure

Source Type III Expected Mean Square

SIRE Var(Error) + 4.5249 Var(DAM(SIRE)) + 15.765 Var(SIRE)

DAM(SIRE) Var(Error) + 5.0685 Var(DAM(SIRE))

The SAS System
General Linear Models Procedure
Tests of Hypotheses for Random Model Analysis of Variance

Dependent Variable: W

Source: SIRE
Error: 0.8928*MS(DAM(SIRE)) + 0.1072*MS(Error)

DF	Type III MS	Denominator DF	Denominator MS	F Value	Pr > F
4	10.535533994	17.42	4.424027283	2.3814	0.0913

Source: DAM(SIRE)
Error: MS(Error)

DF	Type III MS	Denominator DF	Denominator MS	F Value	Pr > F
14	4.4379507066	82	4.308125242	1.0301	0.4328

The SAS System
Variance Components Estimation Procedure

Dependent Variable: W

Source	DF	Type I SS	Type I MS
SIRE	4	44.84004403	11.21001101
Error	96	415.39757973	4.32705812
Corrected Total	100	460.23762376	

Source	Expected Mean Square
SIRE	Var(Error) + 19.391 Var(SIRE)
Error	Var(Error)

Variance Component	Estimate
Var(SIRE)	0.35495443
Var(Error)	4.32705812

Dr. Firas

ولغرض الحصول على تباين الآباء من تحليل التباين S1 وليس S3 نقدر مكونات التباين للآباء بأستعمال طريقة Type1.

من جدول تحليل التباين يمكن الحصول على التباينات (الارقام مؤشر تحتها خط) :

تباين الأبناء يساوي 4.31

تباين الامهات يساوي (4.44 – 4.31)\ 5.07= 0.026

تباين الآباء يساوي (11.21 – 4.44) \ 19.39 = 0.349

التباين الكلي (المظهري) يساوي 0.349 + 0.026 + 4.31 = 4.68

المكافىء الوراثي عن طريق الآباء يساوي 4(0.349) \ 4.68 = 0.30

المكافىء الوراثي عن طريق الامهات يساوي 4(0.026)\ 4.68 = 0.02

المكافىء الوراثي عن طريق الآباء والامهات يساوي 2(0.349 + 0.026)\ 4.68 = 0.16

8 – 1 – 1 – 3 تقدير المكافىء الوراثي بطريقة انحدار الأبناء (النسل) على الآباء

أ - طريقة انحدار متوسط الابناء على احد الاباء

يستعمل هذا التصميم في تقدير انحداري الاب – الابن او الاب – البنت وذلك عندما يقوم الذكر بتلقيح عدد من الاناث وكل انثى تعطي مولودا واحدا فقط. ويكون النموذج الرياضي كالاتي:

$Z_i = b\ X_i + e_{ij}$

اذ ان:

Z_i = متوسط الابناء للأب i ، X = المشاهدة الخاصة بالاب i ، b = انحدار z على x ، e_{ij} = الخطأ العشوائي.

مثال (123) : اخذت بيانات من مجتمع كبير من الدجاج غير مربى تربية داخلية وقد سجلت اوزان الافراد بعمر 8 أسابيع وعندما بلغت افراد العشيرة اخذ منها 17 ديك ولقح عدد من الاناث بصورة عشوائية ثم سجلت اوزان ابنائها عند عمر 8 أسابيع جد المكافىء الوراثي للصفة؟ (غزال ، 1982).

```
data s;
input y x;
cards;
601 910
733 983
793 976
795 1050
818 1080
838 1040
854 1040
880 1025
```

Dr. Firas

```
882 994
895 1030
952 1021
953 1078
961 964
979 976
995 1110
997 1041
1040 1035
proc reg;
model x=y;
run;
```

```
                                The SAS System
Model: MODEL1
Dependent Variable: X
                  Analysis of Variance
                    Sum of       Mean
      Source        DF     Squares     Square     F Value     Prob>F

      Model          1  9256.15665  9256.15665       4.716     0.0463
      Error         15  29442.90218  1962.86015
      C Total       16  38699.05882

        Root MSE     44.30418    R-square      0.2392
        Dep Mean   1020.76471    Adj R-sq      0.1885
        C.V.          4.34029

                  Parameter Estimates

             Parameter     Standard    T for H0:
      Variable  DF     Estimate       Error   Parameter=0    Prob > |T|

      INTERCEP   1   831.442739  87.84239499       9.465       0.0001
      Y          1     0.215052  0.09903154       2.172       0.0463
```

المكافىء الوراثي يساوي 2 0.43 = 0.215 x.

مثال(124) في دراسة لتقدير المكافىء الوراثي لحجم القذفة للسائل المنوي لثلاثة ديكة لكهورن اذ قيس حجم القذفة لها خلال ثلاثة اعمار 24 و 28 و32 شهر كما قيس حجم القذفة لنسلها (اي كل ديك أب لديه 9 ابناء من الديكة) وخلال نفس المراحل العمرية قدر المكافىء الوراثي لحجم القذفة؟

نظرا لكون بيانات الآباء منفصلة عن بيانات الأبناء لذا يجب دمج البيانات وكما موضح ادناه:

```
DATA SEMPA;
INPUT S AGE SVPA ;
CARDS;
 1 24 0.35
 2 24 0.35
 3 24 0.45
 1 28 0.30
 2 28 0.40
 3 28 0.55
 1 32 0.25
```

← S= رقم الاب Age= عمر الاب SVPA= حجم القذفة للآباء

Dr.Firas

```
 2 32 0.55
 3 32 0.30
;OPTION PS=5000 LS=78 NODATE NONUMBER NOCENTER;
PROC SORT;BY S AGE;

DATA SEMPR;
INPUT S D PRG HN AGE SVPR;⟵ رقم الاب =S  رقم الام  = D  رقم الابن = PRG  حجم قذفة الابن = SVPR
CARDS;
 1   5 218  3  24 0.30
 1   6  45  2  24 0.40
 1   6  61  1  24 0.20
 2   7  53  1  24 0.30
 2   7  85  1  24 0.25
 2   8 141  1  24 0.30
 3  17 312  3  24 0.25
 3  18  65  1  24 0.30
 3  18 194  2  24 0.55
 1   1  11  1  28 0.35
 1   1 210  1  28 0.15
 1   1 249  1  28 0.35
 2   9 222  1  28 0.45
 2   9 555  3  28 0.45
 2  11  78  3  28 0.45
 3  13  43  2  28 0.40
 3  13 215  2  28 0.25
 3  13 256  1  28 0.30
 1   5 218  3  32 0.25
 1   6  45  2  32 0.40
 1   6  61  1  32 0.45
 2   7  53  1  32 0.60
 2   7  85  1  32 0.50
 2   8 141  1  32 0.50
 3  15 273  3  32 0.25
 3  17 312  3  32 0.50
 3  18  65  1  32 0.50
PROC SORT;BY S AGE;

DATA NEW1;
MERGE SEMPA SEMPR;BY S AGE;
PROC PRINT;
RUN;
PROC REG DATA=NEWI;
MODEL SVPR=SVPR;
RUN;
```

The SAS System

OBS	S	AGE	SVPA	D	PRG	HN	SVPR
1	1	24	0.35	5	218	3	0.30
2	1	24	0.35	6	45	2	0.40
3	1	24	0.35	6	61	1	0.20
4	1	28	0.30	1	11	1	0.35
5	1	28	0.30	1	210	1	0.15
6	1	28	0.30	1	249	1	0.35
7	1	32	0.25	5	218	3	0.25
8	1	32	0.25	6	45	2	0.40
9	1	32	0.25	6	61	1	0.45
10	2	24	0.35	7	53	1	0.30
11	2	24	0.35	7	85	1	0.25

Dr Firas

```
12  2   24  0.35   8  141   1  0.30
13  2   28  0.40   9  222   1  0.45
14  2   28  0.40   9  555   3  0.45
15  2   28  0.40  11   78   3  0.45
16  2   32  0.55   7   53   1  0.60
17  2   32  0.55   7   85   1  0.50
18  2   32  0.55   8  141   1  0.50
19  3   24  0.45  17  312   3  0.25
20  3   24  0.45  18   65   1  0.30
21  3   24  0.45  18  194   2  0.55
22  3   28  0.55  13   43   2  0.40
23  3   28  0.55  13  215   2  0.25
24  3   28  0.55  13  256   1  0.30
25  3   32  0.30  15  273   3  0.25
26  3   32  0.30  17  312   3  0.50
27  3   32  0.30  18   65   1  0.50
```

البيانات بعد الدمج ستخضع الى التحليل الاحصائي وكالاتي:

```
DATA S;
INPUT OBS   S   AGE   SVPA    D   PRG   HN   SVPR;
CARDS;

 1   1   24   0.35   5   218   3   0.30
 2   1   24   0.35   6   45    2   0.40
 3   1   24   0.35   6   61    1   0.20
 4   1   28   0.30   1   11    1   0.35
 5   1   28   0.30   1   210   1   0.15
 6   1   28   0.30   1   249   1   0.35
 7   1   32   0.25   5   218   3   0.25
 8   1   32   0.25   6   45    2   0.40
 9   1   32   0.25   6   61    1   0.45
10   2   24   0.35   7   53    1   0.30
11   2   24   0.35   7   85    1   0.25
12   2   24   0.35   8   141   1   0.30
13   2   28   0.40   9   222   1   0.45
14   2   28   0.40   9   555   3   0.45
15   2   28   0.40   11  78    3   0.45
16   2   32   0.55   7   53    1   0.60
17   2   32   0.55   7   85    1   0.50
18   2   32   0.55   8   141   1   0.50
19   3   24   0.45   17  312   3   0.25
20   3   24   0.45   18  65    1   0.30
21   3   24   0.45   18  194   2   0.55
22   3   28   0.55   13  43    2   0.40
23   3   28   0.55   13  215   2   0.25
24   3   28   0.55   13  256   1   0.30
25   3   32   0.30   15  273   3   0.25
26   3   32   0.30   17  312   3   0.50
27   3   32   0.30   18  65    1   0.50

PROC REG;
MODEL SVPR=SVPA;
RUN;
                          The SAS System
Model: MODEL1
Dependent Variable: SVPR
                       Analysis of Variance

                         Sum of       Mean
          Source    DF   Squares     Square    F Value    Prob>F

          Model      1   0.02599    0.02599      1.970    0.1727
          Error     25   0.32975    0.01319
          C Total   26   0.35574

            Root MSE      0.11485    R-square      0.0731
            Dep Mean      0.36852    Adj R-sq      0.0360
            C.V.         31.16487

                       Parameter Estimates
```

WWW Firas

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	0.250394	0.08700960	2.878	0.0081
SVPA	1	0.303748	0.21639984	1.404	0.1727

المكافىء الوراثي = 2 0.60 = x 0.30

ب - طريقة انحدار الابناء على الام ضمن الأب

في بعض الانواع الحيوانية يكون بمقدور ذكرواحد ان يلقح عدة اناث وكل انثى يمكن ان تنتج عدة ابناء لذا يمكن ان تقدير المكافىء الوراثي عن طريق الام – البنات ضمن الاب اي يجب ازالة تأثير الاب. ويكون النموذج الرياضي لهذا التصميم كالاتي:

$$Z_{ij} = \mu + S_i + b(X_{ij} - X..) + e_{ij}$$

اذ ان :

Z_{ij} = متوسط سجلات النسل الناتج من تزاوج الذكر i والانثى j ، μ = المتوسط العام ، S_i = تأثير الذكر i ، b = معامل انحدار z على x ، X_{ij} = سجل الام j التي لقحها الذكر I ، x.. = المتوسط المظهري ، e_{ij} = الخطأ العشوائي.

مثال (125) قدر المكافىء الوراثي لوزن الجسم عند عمر اربعة اسابيع في قطيع من دجاج اللحم اذا علمت ان التجربة تضمنت 6 ذكور تزاوجت مع اعداد مختلفة من الاناث وسجلت بيانات العمر للامهات ومعدل اوزان نسل كل منها باستعمال النموذج المحتضن ؟ (غزال ، 1982).

```
DATA S;
INPUT SIRE DAMNO DAM PRO;
CARDS;
1 1 754 808
1 2 648 700
1 3 881 720
2 4 740 725
2 5 712 840
2 6 812 800
3 7 765 780
3 8 807 840
4 9 969 850
4 10 849 802
4 11 732 830
5 12 740 806
5 13 741 835
6 14 831 830
6 15 639 800
6 16 733 504
PROC glm;classes sire;
MODEL PRO= sire dam/solution;
RUN;
```

Dr.Firas

The SAS System

General Linear Models Procedure

Class Level Information

Class	Levels	Values
SIRE	6	1 2 3 4 5 6

Number of observations in data set = 16

The SAS System

General Linear Models Procedure

Dependent Variable: PRO

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	31134.05006206	5189.00834368	0.58	0.7409
Error	9	80929.69993794	8992.18888199		
Corrected Total	15	112063.75000000			

R-Square	C.V.	Root MSE	PRO Mean
0.277824	12.16708	94.8271526	779.37500000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
SIRE	5	30330.58333333	6066.11666667	0.67	0.6534
DAM	1	803.46672872	803.46672872	0.09	0.7718

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SIRE	5	24792.81372254	4958.56274451	0.55	0.7347
DAM	1	803.46672872	803.46672872	0.09	0.7718

Parameter	Estimate	T for H0: Parameter=0	Pr > |T|	Std Error of Estimate
INTERCEPT	637.7152160 B	2.53	0.0324	252.2943565
SIRE 1	28.6599564 B	0.37	0.7216	77.9408706
2	74.9615501 B	0.96	0.3600	77.7257832
3	93.4869989 B	1.06	0.3172	88.2822372
4	104.4042276 B	1.21	0.2587	86.6005504
5	108.5484483 B	1.25	0.2416	86.5896538
6	0.0000000 B	.	.	.
DAM	**0.1002516**	0.30	0.7718	0.3353824

NOTE: The X'X matrix has been found to be singular and a generalized inverse was used to solve the normal equations.
Estimates followed by the letter 'B' are biased, and are not unique estimators of the parameters.

بمـا ان معامـل انحدارمعـدل النسـل علـى الام بعـد التعـديل لتـأثير الأب = 0.10 فـأن المكـافىء الوراثي يساوي 2x 0.10 = 0.20.

8 – 6 النماذج المختلطة Mixed Models

ان اسـتعمال النمـاذج المختلطـة لايقتصـر علـى الدراسـات الخاصـة بتقـدير المكـافىء الـوراثي وتقـديرقيم BLUP او القيم التربويـة وانمـا يمكـن اسـتعمالها ايضـا لغـرض التعـرف علـى تـأثير العوامـل الثابتـة والعوامـل العشـوائية ولغـرض ايجـاد نمـاذج خطيـة ذات دقـة اعلـى فـي وصـف

البيانات، فمثلا لو اردنا معرفة تأثير عقارين على الاصابة بالامراض الطفيلية في الحيوانات نستطيع اختبارتأثير العقارين على مجموعة من الحيوانات في قطيع ما من خلال اعطائه لها بصورة عشوائية هنا سيمثل العقار عامل ثابت (Fixed Effect) ذو مستويين ويمكن استعمال نموذج العوامل الثابتة والذي يكون فيه فقط الخطأ عشوائي لوصف التباين بين الحيوان ضمن كل عقار وحتى اذا كان عدد الحيوانات غير كافي في هذا القطيع فأن من الممكن اختيار عدد اخر من الحيوانات من قطيع اخرويعطى لها العقارين بصورة عشوائية ، في هذه الحالة اذا اردنا اعطاء تعميم اوسع للنتائج فأن من الافضل اعتبار تأثيرالقطيع عشوائي وبذلك فان النموذج المختلط (Mixed Model) سيكون افضل عند تحليل البيانات. والواقع ان النماذج المختلطة تتفوق على النماذج الثابتة في كونها لاتعتمد على الخطأ فقط كمصدر للعشوائية وانما تعتمد على عوامل اخرى لذا فأن كل من اختبار الفرضية والخطأ القياسي والانحراف القياسي سيكون اكثر دقة ، وبذلك فأن معادلة الخط المستقيم المتحصل عليها ستكون ادق.

ان النماذج الخطية استعملت لأول مرة في بحوث وراثة الحيوان ثم جرى تطبيقها على البحوث البيطرية والبحوث الانسانية.

وفي برنامج SAS تستعمل طريقة proc mixed التي تعتمد على طريقة REML او ML لتقدير معالم التباين وهما بصورة عامة يفضل استعمالهما مقارنة باستعمال ANOVA كما يمكن استعمال طريقة MIVQUE التي تعطي تقديرات تشابه تقديرات ANOVA.

مثال (126): البيانات ادناه تعود الى اطوال (انج) مجموعة من الافراد في اربعة عوائل ولكلا الجنسين ويمكن اجراء تحليل باستعمال Two-way وبالاعتماد على نموذج مختلط.

```
data heights;
input Family Gender$ Height @@;
datalines;
1 F 67 1 F 66 1 F 64 1 M 71 1 M 72 2 F 63
2 F 63 2 F 67 2 M 69 2 M 68 2 M 70 3 F 63
3 M 64 4 F 67 4 F 66 4 M 67 4 M 67 4 M 69
run;

proc mixed;
class Family Gender;
model Height = Gender Family Family*Gender;
run;
```

The MIXED Procedure

Class Level Information

Class	Levels	Values
FAMILY	4	1 2 3 4
GENDER	2	F M

Covariance Parameter Estimates (REML)

Cov Parm Estimate

Residual 2.10000000

Model Fitting Information for HEIGHT

Description	Value
Observations	18.0000
Res Log Likelihood	-20.7894
Akaike's Information Criterion	-21.7894
Schwarz's Bayesian Criterion	-21.9407
-2 Res Log Likelihood	41.5789

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
GENDER	1	10	17.63	0.0018
FAMILY	3	10	5.90	0.0139
FAMILY*GENDER	3	10	2.89	0.0889

نلاحظ من النتائج انها تشابه نتائج طريقة glm ولكن في كثير من الحالات لاتكون النتائج متطابقة ويتضح ان تأير التداخل كان غير معنويا فيما كان تأثير العاملين الآخرين معنويا.

ان الفرضية التي جرى التحليل وفقا لها تقضي بأن تكون العوامل المستقلة ذات توزيع طبيعي وهي مستقلة عن بعضها وذات تباين ثابت. وشرط التوزيع الطبيعي قد يكون متوفرا طالما ان الصفة تخص الطول ولكن بسبب ان البيانات تقع ضمن مجاميع (عوائل) فأن هناك احتمال كبير بأن تكون بيانات العائلة الواحدة ذات ارتباط مما يعني عدم توفر شرط استقلال العوامل.

ان طريقة proc mixed لاتزال قائمة على اساس التوزيع الطبيعي ولكن يمكن ازالة فرضية الاستقلال من خلال وضع نموذج للارتباط بطرق مختلفة كما يمكن وضع نموذج للتعامل مع البيانات ذات التباين المختلف ، وفيما يخص الطول فأن ابسط طريقة لنموذج الارتباط هو من خلال استعمال العوامل العشوائية اي اعتبار العائلة والتداخل بين العائلة والجنس عوامل عشوائية ويمكن تفسير تأثير التداخل من خلال كون الانثى في عائلة ما تظهر ارتباطا مع الاناث في عائلتها اكثر من ذكور العائلة نفسها وكذلك ذكور العوائل الاخرى. وبذلك فأن النموذج في شكله الاخير اصبح اكثر قدرة على وصف البيانات. وبذلك نستعمل الايعاز التالي:

```
proc mixed;
class Family Gender;
model Height = Gender;
random Family Family*Gender;
run;
```

The SAS System

The MIXED Procedure
Class Level Information

Class	Levels	Values
FAMILY	4	1 2 3 4
GENDER	2	F M

Dr. Firas

REML Estimation Iteration History

Iteration	Evaluations	Objective	Criterion
0	1	44.70471526	
1	2	42.11010697	0.01441208
2	1	41.73242684	0.00412226
3	1	41.63010249	0.00058188
4	1	41.61678451	0.00001689
5	1	41.61642597	0.00000002
6	1	41.61642563	0.00000000

Convergence criteria met.

Covariance Parameter Estimates (REML)

Cov Parm	Estimate
FAMILY	2.40099764
FAMILY*GENDER	1.76567389
Residual	2.16676858

Model Fitting Information for HEIGHT

Description	Value
Observations	18.0000
Res Log Likelihood	-35.5112
Akaike's Information Criterion	-38.5112
Schwarz's Bayesian Criterion	-39.6701
-2 Res Log Likelihood	71.0225

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
GENDER	1	3	7.95	0.0667

يلاحظ من النتائج ان تأثير الجنس اصبح غير معنويا . ان نتائج التحليل باستعمال العوامل الثابتة تكون مقتصرة على االعوائل المشمولة بالدراسة بينما في حالة النموذج المختلط فأن النتائج يمكن تعميمها على كل العشيرة.

مثال(127): في دراسة لمقارنة نوعين من العقار وبمستويين لكل عقار على مجموعة من الحيوانات تعود الى 6 مناطق مختلفة ، حيث اخذ من كل منطقة قطيعين واعطيت لبعض الحيوانات المصابة في كل قطيع وبصورة عشوائية تركيز عالي او تركيز واطىء ثم حسب معدل الاستجابة للحيوانات المعاملة لكل تركيز في القطيعين ضمن كل منطقة والاستجابة هنا تتمثل بمعدل قيم PCV بعد شهر من تناول العقار.

ان التجربة يمكن تحليلها بطريقة نموذج العوامل الثابتة او بالطريقة الافضل وهي طريقة النموذج المختلط.

	Drug			
	Berenil		Samorin	
	L	H	L	H

Dr.Firas

Region	1	21.8	22.6	16.4	19.1
	2	28.8	29.0	18.2	25.3
	3	23.7	24.0	16.0	19.5
	4	26.3	27.7	16.6	21.9
	5	23.3	28.3	19.3	25.2
	6	21.8	21.1	16.3	17.9

```
data ex125;
input region $ drug $ dose $ PCV;
cards;
1  BERENIL  H  22.6
1  BERENIL  L  21.8
1  SAMORIN  H  19.1
1  SAMORIN  L  16.4
2  BERENIL  H  29.0
2  BERENIL  L  28.8
2  SAMORIN  H  25.3
2  SAMORIN  L  18.2
3  BERENIL  H  24.0
3  BERENIL  L  23.7
3  SAMORIN  H  19.5
3  SAMORIN  L  16.0
4  BERENIL  H  27.7
4  BERENIL  L  26.3
4  SAMORIN  H  21.9
4  SAMORIN  L  16.6
5  BERENIL  H  28.3
5  BERENIL  L  23.3
5  SAMORIN  H  25.2
5  SAMORIN  L  19.3
6  BERENIL  H  21.1
6  BERENIL  L  21.8
6  SAMORIN  H  17.9
6  SAMORIN  L  16.3
proc mixed data=ex125;
class drug dose region;
model pcv=drug dose drug*dose;
random region drug*region/solution;
run;
```

```
                                The SAS System
          The MIXED Procedure

          Class Level Information

        Class     Levels  Values

        DRUG          2  BERENIL SAMORIN
        DOSE          2  H L
        REGION        6  1 2 3 4 5 6

        REML Estimation Iteration History

     Iteration  Evaluations    Objective     Criterion

         0          1   67.81993192
         1          1   55.62103313    0.00000000

            Convergence criteria met.
```

Dr. Firas

Covariance Parameter Estimates (REML)

Cov Parm	Estimate
REGION	5.15083333
DRUG*REGION	0.38700000
Residual	2.09641667

Model Fitting Information for PCV

Description	Value
Observations	24.0000
Res Log Likelihood	-46.1893
Akaike's Information Criterion	-49.1893
Schwarz's Bayesian Criterion	-50.6829
-2 Res Log Likelihood	92.3786

Solution for Fixed Effects

Effect	DRUG	DOSE	Estimate	Std Error	DF	t	Pr > |t|
INTERCEPT			17.13333333	1.12799601	5	15.19	0.0001
DRUG	BERENIL		7.15000000	0.90983820	5	7.86	0.0005
DRUG	SAMORIN		0.00000000	.	.	.	.
DOSE		H	4.35000000	0.83594590	10	5.20	0.0004
DOSE		L	0.00000000	.	.	.	.
DRUG*DOSE	BERENIL	H	-3.18333333	1.18220604	10	-2.69	0.0226

Solution for Random Effects

Effect	DRUG	REGION	Estimate	SE Pred	DF	t	Pr > |t|
REGION		1	-1.85417931	1.17615928	10	-1.58	0.1460
REGION		2	2.84161208	1.17615928	10	2.42	0.0363
REGION		3	-1.13006195	1.17615928	10	-0.96	0.3593
REGION		4	0.91063244	1.17615928	10	0.77	0.4567
REGION		5	1.70057866	1.17615928	10	1.45	0.1788
REGION		6	-2.46858193	1.17615928	10	-2.10	0.0622
DRUG*REGION	BERENIL	1	-0.21908499	0.58155179	10	-0.38	0.7142
DRUG*REGION	BERENIL	2	0.32134437	0.58155179	10	0.55	0.5927
DRUG*REGION	BERENIL	3	0.03057673	0.58155179	10	0.05	0.9591
DRUG*REGION	BERENIL	4	0.32969795	0.58155179	10	0.57	0.5833
DRUG*REGION	BERENIL	5	-0.20688560	0.58155179	10	-0.36	0.7294
DRUG*REGION	BERENIL	6	-0.25564846	0.58155179	10	-0.44	0.6696
DRUG*REGION	SAMORIN	1	0.07977406	0.58155179	10	0.14	0.8936
DRUG*REGION	SAMORIN	2	-0.10784419	0.58155179	10	-0.19	0.8566
DRUG*REGION	SAMORIN	3	-0.11548221	0.58155179	10	-0.20	0.8466
DRUG*REGION	SAMORIN	4	-0.26127897	0.58155179	10	-0.45	0.6628
DRUG*REGION	SAMORIN	5	0.33465598	0.58155179	10	0.58	0.5777
DRUG*REGION	SAMORIN	6	0.07017532	0.58155179	10	0.12	0.9063

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
DRUG	1	5	64.58	0.0005
DOSE	1	10	21.78	0.0009
DRUG*DOSE	1	10	7.25	0.0226

Dr Firas

مثال (128): اختيرت مجموعة من الابقار من سلالتين وبصورة عشوائية وقيست فيها قيم PCV خلال 5 فترات، هل للسلالة والتداخل بين السلالة والفترة تأثير على قيم PCV؟

التأثير العشوائي هنا سيكون الحيوان ضمن السلالة.

Breed	Animal	Days following infection				
		0	2	4	7	9
Boran	BO1	36.2	35.9	35.3	35.4	35.4
	BO2	35.9	38.5	35.9	36.0	36.3
	B03	29.5	33.3	29.2	29.9	29.0
N'Dama	ND1	25.4	27.0	20.3	27.9	20.6
	ND2	29.5	21.8	26.5	29.7	18.7
	ND3	22.4	25.4	29.7	24.2	28.5

```
data ex33;
input anim $ breed $ time PCV;
cards;
BO1   BO  0  36.2
BO1   BO  2  35.9
BO1   BO  4  35.3
BO1   BO  7  35.4
BO1   BO  9  35.4
BO2   BO  0  35.9
BO2   BO  2  38.5
BO2   BO  4  35.9
BO2   BO  7  36.0
BO2   BO  9  36.3
BO3   BO  0  29.5
BO3   BO  2  33.3
BO3   BO  4  29.2
BO3   BO  7  29.9
BO3   BO  9  29.0
ND1   ND  0  25.4
ND1   ND  2  27.0
ND1   ND  4  20.3
ND1   ND  7  27.9
ND1   ND  9  20.6
ND2   ND  0  29.5
ND2   ND  2  21.8
ND2   ND  4  26.5
ND2   ND  7  29.7
ND2   ND  9  18.7
ND3   ND  0  22.4
ND3   ND  2  25.4
ND3   ND  4  29.7
ND3   ND  7  24.2
ND3   ND  9  28.5
proc mixed data=ex33;
class breed anim;
model pcv=breed breed*time/solution;
random anim(breed) / solution;
run;
```

Dr. Firas

The SAS System
The MIXED Procedure

Class Level Information

Class	Levels	Values
BREED	2	BO ND
ANIM_ID	6	BO1 BO2 BO3 ND1 ND2 ND3

REML Estimation Iteration History

Iteration	Evaluations	Objective	Criterion
0	1	106.84710407	
1	1	103.27062957	0.00000000

Convergence criteria met.

Covariance Parameter Estimates (REML)

Cov Parm	Estimate
ANIM(BREED)	4.54021112
Residual	8.82461107

Model Fitting Information for PCV

Description	Value
Observations	30.0000
Res Log Likelihood	-75.5277
Akaike's Information Criterion	-77.5277
Schwarz's Bayesian Criterion	-78.7858
-2 Res Log Likelihood	151.0554

Solution for Fixed Effects

Effect	BREED	Estimate	Std Error	DF	t	Pr > |t|
INTERCEPT		25.84160401	1.78105732	4	14.51	0.0001
BREED	BO	8.79774436	2.51879542	4	3.49	0.0251
BREED	ND	0.00000000	.	.	.	.
TIME*BREED	BO	-0.11954887	0.23514261	22	-0.51	0.6162
TIME*BREED	ND	-0.15187970	0.23514261	22	-0.65	0.5250

The SAS System

Solution for Random Effects

Effect	BREED	ANIM_ID	Estimate	SE Pred	DF	t	Pr > |t|
ANIM(BREED)	BO	BO1	1.09932473	1.53644476	22	0.72	0.4818
ANIM(BREED)	BO	BO2	1.73299662	1.53644476	22	1.13	0.2715
ANIM(BREED)	BO	BO3	-2.83232135	1.53644476	22	-1.84	0.0788
ANIM(BREED)	ND	ND1	-0.67207625	1.53644476	22	-0.44	0.6661
ANIM(BREED)	ND	ND2	0.04800545	1.53644476	22	0.03	0.9754
ANIM(BREED)	ND	ND3	0.62407081	1.53644476	22	0.41	0.6885

Tests of Fixed Effects

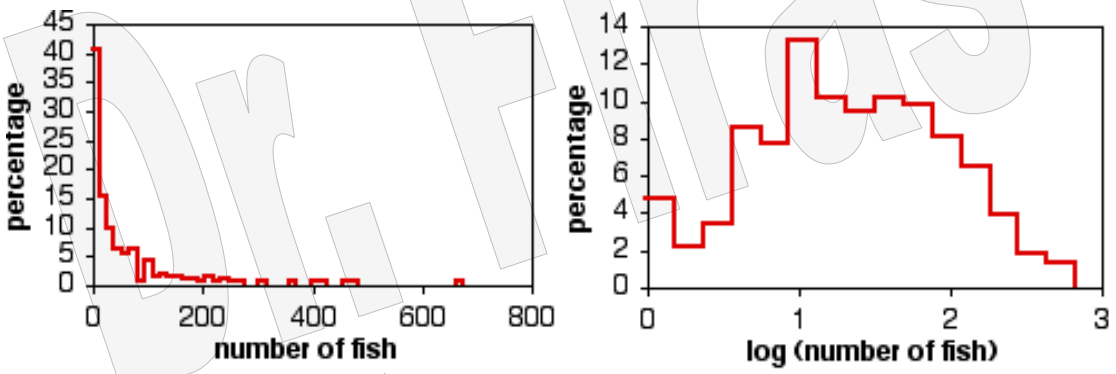
Source	NDF	DDF	Type III F	Pr > F
BREED	1	4	12.20	0.0251
TIME*BREED	2	22	0.34	0.7169

Dr Firas

الفصل التاسع

تحويل البيانات Data transformation

يلاحظ بأن هناك الكثير من المتغيرات البايولوجية لاتتفق مع الاختبارات الاحصائية بسبب كون المتغيرلايتوزع توزيعا طبيعيا او بسبب عدم تجانس التباين او لكلا السببين. لذا فأن الاختبار في بعض العمليات الاحصائية (جدول تحليل التباين او الانحدار) لايعطي نتائج دقيقة. لذا فأن تحويل البيانات في بعض الحالات سيؤدي الى مطابقة البيانات بعد التحويل للفرضيات الاحصائية. فمثلا ان انتشار الاسماك نوع (Eastern mudminnow) *Umbra pygmaea* في بحيرات ماريلاند لايكون موزع توزيع طبيعي اذ ان هناك عدد كبير من البحيرات كثافة الاسماك فيها قليلة وعدد قليل من البحيرات كثافة الاسماك فيها عالية ولكن هذه البيانات سيكون توزيعها طبيعي بعد اجراء تحويل للبيانات (شكل 1) .



البيانات قبل التحويل البيانات بعد التحويل

9 – 1 التحويل اللوغاريتمي

وفيه يتم تحويل البيانات الى اللوغارتيم الطبيعي او اللوغارتيم الآساس 10. ويشترط ان تكون جميع القيم موجبة ويستعمل هذا التحويل لتقريب البيانات التضاعفية الى النماذج الخطية كما يستعمل لجعل تباين الخطأ ثابتا ومتجانسا.

9 – 2 التحويل الجذري

ويستعمال هذا التحويل لجعل تباين الخطأ ثابتا ومتجانسا وهو ملائم للبيانات التي تتبع توزيع بواسون اذ ان هذا التوزيع له تطبيقات عملية في التجارب والتي تكون المشاهدات فيها تقاس على ضوء وحدة الزمان او المكان مثال ذلك عدد المكالمات التليفونية المستلمة في الدقيقة في دائرة البريد او عدد حالات الاصابة بحروق في يوم واحد في مستشفى.

مثال (129): البيانات في المثال ادناه تعود الى انتشار اسماك نوع *Umbra pygmaea* في عدة بحيرات المطلوب اجراء تحويل للبيانات ؟

```
data s;
input x;
x1=log10(x);
x2=sqrt(x);
cards;
38
1
13
2
13
20
50
9
28
6
4
43
proc print;
run;
proc univariate normal plot;
var x x1 x2;
run;
```

Dr. Firas

```
                                                The SAS System

            OBS    X      X1        X2

              1   38    1.57978   6.16441
              2    1    0.00000   1.00000
              3   13    1.11394   3.60555
              4    2    0.30103   1.41421
              5   13    1.11394   3.60555
              6   20    1.30103   4.47214
              7   50    1.69897   7.07107
              8    9    0.95424   3.00000
              9   28    1.44716   5.29150
             10    6    0.77815   2.44949
             11    4    0.60206   2.00000
             12   43    1.63347   6.55744
                  The SAS
                Univariate Procedure
Variable=X
                    Moments

          N             12  Sum Wgts       12
          Mean       18.91667  Sum          227
          Std Dev    16.94622  Variance   287.1742
          Skewness   0.767652  Kurtosis   -0.81617
          USS            7453  CSS       3158.917
          CV         89.58352  Std Mean   4.891951
          T:Mean=0   3.866896  Pr>|T|      0.0026
          Num ^= 0         12  Num > 0        12
          M(Sign)           6  Pr>=|M|     0.0005
          Sgn Rank         39  Pr>=|S|     0.0005
          W:Normal   0.890905  Pr<W        0.1158

                 Quantiles(Def=5)

          100% Max       50      99%       50
          75% Q3         33      95%       50
          50% Med        13      90%       43
          25% Q1          5      10%        2
           0% Min         1      5%         1
                          1%         1
          Range          49
          Q3-Q1          28
```

Mode 13

Extremes

Lowest	Obs	Highest	Obs
1(	2)	20(	6)
2(	4)	28(	9)
4(	11)	38(	1)
6(	10)	43(	12)
9(	8)	50(	7)

```
Stem Leaf              #          Boxplot
 5 0                   1          |
 4 3                   1          |
 3 8                   1          +-----+
 2 08                  2          |     |
 1 33                  2          *--+--*
 0 12469               5          +-----+
   ----+----+----+----+
```

Multiply Stem.Leaf by 10**+1

The SAS System
Univariate Procedure

Variable=X

Normal Probability Plot

```
55+                          * ++++++
 |                       * ++++++
 |                    +*++++
 |                ++*+*+
 |             ++++*+*
 5+      *   *++*++* *
  +----+----+----+----+----+----+----+----+----+----+
    -2     -1      0     +1     +2
```

The SAS System
Univariate Procedure

Variable=X1

Moments

N	12	Sum Wgts	12
Mean	1.043648	Sum	12.52378
Std Dev	0.540865	Variance	0.292535
Skewness	-0.6396	Kurtosis	-0.438
USS	16.28831	CSS	3.21789
CV	51.82449	Std Mean	0.156134
T:Mean=0	6.684295	Pr>|T|	0.0001
Num ^= 0	11	Num > 0	11
M(Sign)	5.5	Pr>=|M|	0.0010
Sgn Rank	33	Pr>=|S|	0.0010
W:Normal	0.94425	Pr<W	0.5156

Quantiles(Def=5)

100% Max	1.69897	99%	1.69897
75% Q3	1.513471	95%	1.69897
50% Med	1.113943	90%	1.633468
25% Q1	0.690106	10%	0.30103
0% Min	0	5%	0
		1%	0
Range	1.69897		
Q3-Q1	0.823365		
Mode	1.113943		

Dr. Firas
Dr. Firas

Extremes

Lowest	Obs	Highest	Obs
0(	2)	1.30103(	6)
0.30103(	4)	1.447158(	9)
0.60206(	11)	1.579784(	1)
0.778151(	10)	1.633468(	12)
0.954243(	8)	1.69897(	7)

The SAS System
Univariate Procedure

Variable=X1

```
Stem Leaf              #          Boxplot
 16 30                 2            |
 14 58                 2          +-----+
 12 0                  1          |     |
 10 11                 2          *--+--*
  8 5                  1          |     |
  6 08                 2          +-----+
  4                               |
  2 0                  1            |
  0 0                  1            |
   ----+----+----+----+
Multiply Stem.Leaf by 10**-1
```

Normal Probability Plot

```
1.7+                        *++  *
   |                    * +*++
   |                   *+++
   |                *+*+
0.9+               +*++
   |            *++*
   |          +++
   |       ++++*
0.1+    +++*
   +----+----+----+----+----+----+----+----+----+----+
     -2      -1       0      +1      +2
```

Univariate Procedure
Variable=X2

Moments

N	12	Sum Wgts	12
Mean	3.885947	Sum	46.63136
Std Dev	2.040343	Variance	4.162999
Skewness	0.213679	Kurtosis	-1.22327
USS	227	CSS	45.79298
CV	52.50568	Std Mean	0.588996
T:Mean=0	6.597576	Pr>|T|	0.0001
Num ^= 0	12	Num > 0	12
M(Sign)	6	Pr>=|M|	0.0005
Sgn Rank	39	Pr>=|S|	0.0005
W:Normal	0.954596	Pr<W	0.6534

Quantiles(Def=5)

100% Max	7.071068	99%	7.071068
75% Q3	5.727958	95%	7.071068
50% Med	3.605551	90%	6.557439
25% Q1	2.224745	10%	1.414214
0% Min	1	5%	1
		1%	1

Range	6.071068
Q3-Q1	3.503213
Mode	3.605551

Extremes

Dr. Firas

```
            Lowest    Obs    Highest    Obs
               1(      2) 4.472136(      6)
          1.414214(      4) 5.291503(      9)
               2(     11) 6.164414(      1)
           2.44949(     10) 6.557439(     12)
               3(      8) 7.071068(      7)


      Stem Leaf              #          Boxplot
        7 1                  1            |
        6 26                 2            |
        5 3                  1          +-----+
        4 5                  1          |     |
        3 066                3          *--+--*
        2 04                 2          +-----+
        1 04                 2            |
         ----+----+----+----+

                Univariate Procedure

Variable=X2

                Normal Probability Plot
      7.5+                          *++++
        |                       *  *++++
        |                     *+++++
      4.5+                  ++*++
        |               +*+*+
        |            +*++*
      1.5+       * +++*+
         +----+----+----+----+----+----+----+----+----+----+
           -2     -1      0     +1     +2
```

نلاحظ ان قيمة معامل الالتواء والتفلطح كانتا اقل عند x1 مقارنة مع قيم x او x2 كما نلاحظ ان قيمة المتوسط والوسيط في الصندوق اصبحت متساوية مما يعني ان توزيع البيانات اصبح توزيعا طبيعيا.

وعند حساب المتوسط العام للبيانات بأستعمال الآمر proc means نجد ان المتوسطات مختلفة

```
data s;
input x;
x1=log10(x);
x2=sqrt(x);
cards;
38
1
13
2
13
20
50
9
28
6
4
43
proc means;var x x1 x2;
run;
```

```
                    The SAS System

      Variable  N        Mean      Std Dev      Minimum      Maximum
      ------------------------------------------------------------------
      X        12   18.9166667   16.9462162    1.0000000   50.0000000
```

Dr. Firas

```
X1      12    1.0436484    0.5408654          0    1.6989700
X2      12    3.8859471    2.0403428    1.0000000    7.0710678
---------------------------------------------------------------
```

```
data s;
input x;
x1=log10(x);
x2=10**x1;
x3=log(x);
x4=exp(x3);
x5=sqrt(x);
x6=x5**2;
cards;
38
1
13
2
13
20
50
9
28
6
4
43
proc means;var x x1 x2 x3 x4 x5 x6;
run;
```

```
                    The SAS System
Variable   N        Mean     Std Dev      Minimum      Maximum
---------------------------------------------------------------
X         12   18.9166667   16.9462162    1.0000000   50.0000000
X1        12    1.0436484    0.5408654            0    1.6989700
X2        12   18.9166667   16.9462162    1.0000000   50.0000000
X3        12    2.4030892    1.2453887            0    3.9120230
X4        12   18.9166667   16.9462162    1.0000000   50.0000000
X5        12    3.8859471    2.0403428    1.0000000    7.0710678
X6        12   18.9166667   16.9462162    1.0000000   50.0000000
---------------------------------------------------------------
```

بعد اجراء عملية التحويل واكمال الاختبار على القيم المحولة فأن من غير المناسب كتابة التقديرات كالمتوسط والخطأ القياسي او الانحراف وغيرها بالقيم المحولة وانما يفضل اعادة التقديرات الى قيمها الحقيقية بمعنى وضع معادلة رياضية تعكس القيم المحولة لترجعها الى القيم الحقيقية .

```
data s;
input x1 x2;
x3=10**1.044;
x4=x2**2;
cards;
1.04 3.88
;
proc means mean;var  x3 x4;
run;
```

```
                    The SAS System

                     Variable        Mean
```

```
----------------------
X3        11.0662378
X4        15.0544000
----------------------
```

مثال (130) :البيانات ادناه تمثل متغيران الاول t يمثل عدد حقول الدواجن والثاني g يمثل عدد الدجاج جد معادلة انحدار عدد الدجاج على عدد الحقول بعد اختبار تجانس البيانات؟

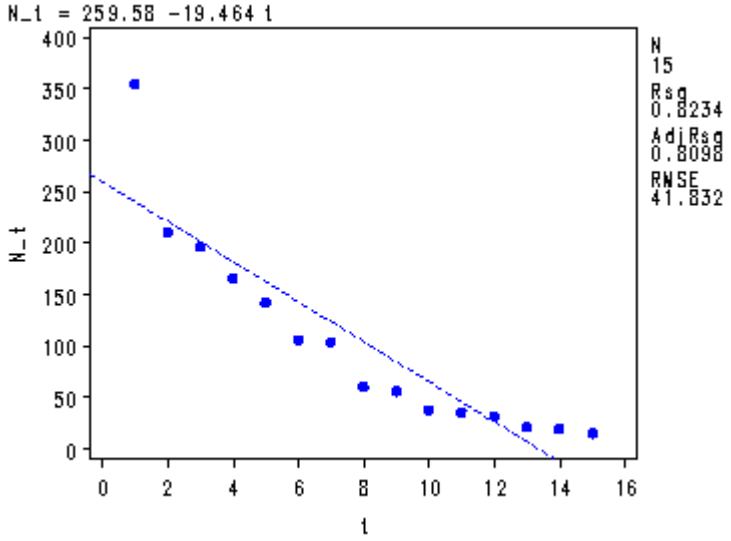
لغرض معرفة انتشار البيانات نقوم بتنفيذ برنامج الرسم البياني للمتغير الثاني بقيمه الحقيقية وقيمه بعد التحويل وكالاتي:

```
data f;
 input t g;
m=log(g);
cards;
1 355
2 211
3 197
4 166
5 142
6 106
7 104
8 60
9 56
10 38
11 36
12 32
13 21
14 19
15 15
option nodate nonumber;
proc reg data = f;
 model g = t;
 plot g*t;
run;
proc reg data = f;
 model m = t;
 plot m*t;
run;
```

```
                    The SAS System
Model: MODEL1
Dependent Variable: G
                  Analysis of Variance

                    Sum of       Mean
        Source       DF     Squares      Square     F Value      Prob>F

        Model         1 106080.35714 106080.35714      60.619      0.0001
        Error        13  22749.37619   1749.95201
        C Total      14 128829.73333

          Root MSE     41.83243    R-square      0.8234
          Dep Mean    103.86667    Adj R-sq      0.8098
          C.V.         40.27512

                  Parameter Estimates

              Parameter     Standard    T for H0:
```

Dr. Firas

Variable	DF	Estimate	Error	Parameter=0	Prob > |T|
INTERCEP	1	259.580952	22.72999119	11.420	0.0001
T	1	-19.464286	2.49996572	-7.786	0.0001



The SAS System

Model: MODEL1
Dependent Variable: M

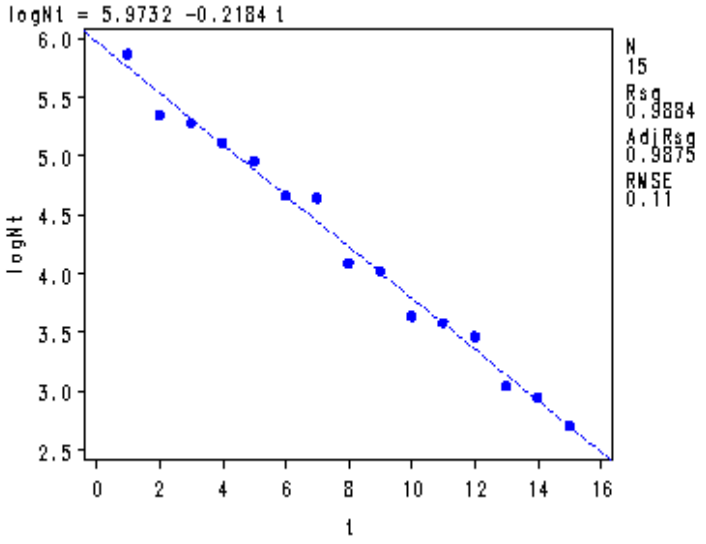
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	13.35869	13.35869	1103.702	0.0001
Error	13	0.15735	0.01210		
C Total	14	13.51603			

Root MSE	0.11002	R-square	0.9884
Dep Mean	4.22576	Adj R-sq	0.9875
C.V.	2.60346		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	5.973160	0.05977810	99.922	0.0001
T	1	-0.218425	0.00657471	-33.222	0.0001



Dr.Firas

يلاحظ من الرسوم البياني الاول ان هناك قيمة واحدة متطرفة ويمكن ان نجري اختبار اخر باستعمال الايعاز التالي :

```
PROC REG DATA=f;
 MODEL g= t / clm r influence;
RUN;

 PROC REG DATA=f;
 MODEL m= t / clm r influence;
RUN;
```

ويمكن تحديد القيم الشاذة او المتطرفة عن طريق قيم Cook's Distance او قيم Dfits .

The SAS System

Model: MODEL1
Dependent Variable: G

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	106080.35714	106080.35714	60.619	0.0001
Error	13	22749.37619	1749.95201		
C Total	14	128829.73333			

Root MSE 41.83243 R-square **0.8234**
Dep Mean 103.86667 Adj R-sq 0.8098
C.V. 40.27512

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	259.580952	22.72999119	11.420	0.0001
T	1	-19.464286	2.49996572	-7.786	0.0001

Obs	Dep Var G	Predict Value	Std Err Predict	Lower95% Mean	Upper95% Mean	Residual	Std Err Residual	Student Residual
1	355.0	240.1	20.565	195.7	284.5	114.9	36.429	3.154
2	211.0	220.7	18.484	180.7	260.6	-9.6524	37.527	-0.257
3	197.0	201.2	16.520	165.5	236.9	-4.1881	38.432	-0.109
4	166.0	181.7	14.719	149.9	213.5	-15.7238	39.157	-0.402
5	142.0	162.3	13.150	133.9	190.7	-20.2595	39.712	-0.510
6	106.0	142.8	11.902	117.1	168.5	-36.7952	40.103	-0.918
7	104.0	123.3	11.087	99.3798	147.3	-19.3310	40.337	-0.479
8	60.00	103.9	10.801	80.5323	127.2	-43.8667	40.414	-1.085
9	56.00	84.4024	11.087	60.4512	108.4	-28.4024	40.337	-0.704
10	38.00	64.9381	11.902	39.2249	90.6513	-26.9381	40.103	-0.672
11	36.00	45.4738	13.150	17.0658	73.8818	-9.4738	39.712	-0.239
12	32.00	26.0095	14.719	-5.7898	57.8089	5.9905	39.157	0.153
13	21.00	6.5452	16.520	-29.1440	42.2344	14.4548	38.432	0.376
14	19.00	-12.9190	18.484	-52.8512	27.0132	31.9190	37.527	0.851
15	15.00	-32.3833	20.565	-76.8106	12.0439	47.3833	36.429	1.301

Obs	-2-1-0 1 2	Cook's D	Rstudent	Hat Diag H	Cov Ratio	INTERCEP Dffits	T Dfbetas	Dfbetas
1	|******|	1.585	6.2508	0.2417	0.0854	3.5287	3.5228	-3.0028
2	| |	0.008	-0.2478	0.1952	1.4435	-0.1220	-0.1210	0.0990
3	| |	0.001	-0.1047	0.1560	1.3879	-0.0450	-0.0440	0.0341
4	| |	0.011	-0.3882	0.1238	1.3064	-0.1459	-0.1381	0.0991
5	*| |	0.014	-0.4951	0.0988	1.2507	-0.1639	-0.1463	0.0935
6	*| |	0.037	-0.9115	0.0810	1.1170	-0.2705	-0.2167	0.1136

Dr Firas

```
7 |     |     |    0.009  -0.4646 0.0702   1.2181  -0.1277  -0.0844   0.0288
8 |    **|     |    0.042  -1.0936 0.0667   1.0398  -0.2923  -0.1389   0.0000
9 |     *|     |    0.019  -0.6898 0.0702   1.1678  -0.1896  -0.0502  -0.0428
10 |    *|     |    0.020  -0.6569 0.0810   1.1899  -0.1949  -0.0120  -0.0819
11 |     |     |    0.003  -0.2297 0.0988   1.2909  -0.0761   0.0085  -0.0434
12 |     |     |    0.002   0.1471 0.1238   1.3346   0.0553  -0.0138   0.0376
13 |     |     |    0.013   0.3633 0.1560   1.3604   0.1562  -0.0555   0.1182
14 |     |*    |    0.088   0.8409 0.1952   1.3005   0.4142  -0.1807   0.3361
15 |     |**   |    0.270   1.3399 0.2417   1.1710   0.7564  -0.3776   0.6437

Sum of Residuals                0
Sum of Squared Residuals    22749.3762
Predicted Resid SS (Press)  35910.8512
```

The SAS System

Model: MODEL1
Dependent Variable: M

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	13.35869	13.35869	1103.702	0.0001
Error	13	0.15735	0.01210		
C Total	14	13.51603			

Root MSE	0.11002	R-square	**0.9884**
Dep Mean	4.22576	Adj R-sq	0.9875
C.V.	2.60346		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	5.973160	0.05977810	99.922	0.0001
T	1	-0.218425	0.00657471	-33.222	0.0001

Obs	Dep Var M	Predict Value	Std Err Predict	Lower95% Mean	Upper95% Mean	Residual	Std Err Residual	Student Residual
1	5.8721	5.7547	0.054	5.6379	5.8716	0.1174	0.096	1.225
2	5.3519	5.5363	0.049	5.4313	5.6413	-0.1845	0.099	-1.869
3	5.2832	5.3179	0.043	5.2240	5.4117	-0.0347	0.101	-0.343
4	5.1120	5.0995	0.039	5.0158	5.1831	0.0125	0.103	0.122
5	4.9558	4.8810	0.035	4.8063	4.9557	0.0748	0.104	0.716
6	4.6634	4.6626	0.031	4.5950	4.7302	0.0008	0.105	0.008
7	4.6444	4.4442	0.029	4.3812	4.5072	0.2002	0.106	1.887
8	4.0943	4.2258	0.028	4.1644	4.2871	-0.1314	0.106	-1.236
9	4.0254	4.0073	0.029	3.9443	4.0703	0.0180	0.106	0.170
10	3.6376	3.7889	0.031	3.7213	3.8565	-0.1513	0.105	-1.435
11	3.5835	3.5705	0.035	3.4958	3.6452	0.0130	0.104	0.125
12	3.4657	3.3521	0.039	3.2684	3.4357	0.1137	0.103	1.104
13	3.0445	3.1336	0.043	3.0398	3.2275	-0.0891	0.101	-0.882
14	2.9444	2.9152	0.049	2.8102	3.0202	0.0292	0.099	0.296
15	2.7081	2.6968	0.054	2.5799	2.8136	0.0113	0.096	0.118

Obs	-2-1-0 1 2	Cook's D	Rstudent	Hat Diag H	Cov Ratio	INTERCEP Dffits	T Dfbetas	Dfbetas
1	|** |	0.239	1.2516	0.2417	1.2108	0.7066	0.7054	-0.6013
2	|***|	0.424	-2.0997	0.1952	0.7799	-1.0342	-1.0256	0.8393
3	| |	0.011	-0.3312	0.1560	1.3654	-0.1423	-0.1390	0.1077
4	| |	0.001	0.1170	0.1238	1.3364	0.0440	0.0416	-0.0299
5	|* |	0.028	0.7020	0.0988	1.2016	0.2325	0.2074	-0.1326
6	| |	0.000	0.0076	0.0810	1.2770	0.0022	0.0018	-0.0009
7	|*** |	0.135	2.1281	0.0702	0.6653	0.5849	0.3868	-0.1319
8	|**|	0.055	-1.2646	0.0667	0.9791	-0.3380	-0.1606	0.0000

Dr. Firas
Dr. Firas

```
9 |    |    | 0.001   0.1634   0.0702   1.2567   0.0449   0.0119   0.0101
10 |   **|    | 0.091  -1.5025   0.0810   0.9046  -0.4459  -0.0275  -0.1873
11 |    |    | 0.001   0.1200   0.0988   1.2992   0.0397  -0.0044   0.0227
12 |    |**  | 0.086   1.1141   0.1238   1.1001   0.4188  -0.1043   0.2845
13 |   *|    | 0.072  -0.8736   0.1560   1.2292  -0.3755   0.1333  -0.2841
14 |    |    | 0.011   0.2855   0.1952   1.4387   0.1406  -0.0614   0.1141
15 |    |    | 0.002   0.1131   0.2417   1.5443   0.0638  -0.0319   0.0543

Sum of Residuals               0
Sum of Squared Residuals       0.1573
Predicted Resid SS (Press)     0.2087
```

يلاحظ من النتائج ان القيم قبل التحويل كان فيها قيمة واحدة متطرفة لذل فأن حذف القيمة وبدون تحويل سيعطينا النتائج التالية:

```
                                The SAS System

Model: MODEL1
Dependent Variable: G

                    Analysis of Variance

                      Sum of       Mean
      Source          DF     Squares     Square    F Value     Prob>F

      Model            1  55911.74615  55911.74615    125.523     0.0001
      Error           12   5345.18242   445.43187
      C Total         13  61256.92857

        Root MSE     21.10526    R-square     0.9127
        Dep Mean     85.92857    Adj R-sq     0.9055
        C.V.         24.56140

                    Parameter Estimates

              Parameter     Standard   T for H0:
      Variable  DF     Estimate       Error   Parameter=0   Prob > |T|

      INTERCEP   1   219.182418  13.16350605      16.651      0.0001
      T          1   -15.676923   1.39926491     -11.204      0.0001
```

نلاحظ من مجمل النتائج ان قيم R^2 قد بلغت في القيم قبل التحويل 0.83 وفي القيم بعد ازالة القيمة الشاذة 0.91 وفي القيم المحولة 0.98 مما يعني ان التحويل كانت نتائجه الافضل.

9 – 3 التحويل الزاوي

يستعمل هذا التحويل لغرض تثبيت التباين والارقام التي يمكن تحويلها تتراوح من -1 الى 1 اي ان التحويل يستعمل عندما تكون الارقام نسب.

مثال (131):

```
data s;
input    d ;
d1=arsin (sqrt(d));
d2=sin(d1)**2;
cards;
 0.36
 0.49
```

Dr. Firas

```
0.64
0.11
0.56
0.02
  proc print;var d1 d2 ;
run;
```

يمكن الحصول على تحويل زاوي الذي يمثله d1 ومن ثم ارجاع القيم المحولة الى القيم الحقيقية

```
The SAS System

 OBS     D1      D2

  1    0.64350   0.36
  2    0.77540   0.49
  3    0.92730   0.64
  4    0.33807   0.11
  5    0.84554   0.56
  6    0.14190   0.02
```

الفصل العاشر

اختبارات التشخيص الطبية Medical Diagnostic Tests

اختبار التشخيص هواي طريقة تستعمل لجمع معلومات سريرية لغرض اتخاذ قرار طبي بصدد حالة فرد ما. ومن الناحية الاحصائية فأن القرار يستند على الاحتمالية ، كما ان الهدف من اختبار التشخيص هو تحريك احتمالية المرض الى احدى النهايتين في مقياس الاحتمالات اما 0 وتعني غير مصاب او 1 وتعني مصاب.

10 – 1 تحليل اختبارات التشخيص Analysis of Diagnostic Tests

لو اجري اختبار على مجموعة اشخاص بعضهم مصاب والآخر غير مصاب وكانت النتائج معبر عنها بأحتمالين هما موجب او سالب فأننا سنحصل على اربع مجاميع وكما موضح في الجدول ادناه:

	Disease	No Disease
Test +	a موجب حقيقي	b موجب كاذب

Dr. Firas

Test –	c سالب كاذب	d سالب حقيقي

a = الافراد المصابون بالمرض والذين اعطوا اختبار موجب

b = الافراد غير المصابين ولكن اعطوا اختبار موجب

c = الافراد المصابون ولكن اعطوا اختبار سالب

d = الافراد المصابون ولكن اعطوا اختبار سالب

a + c = العدد الكلي للمرضى

d + b = العدد الكلي لغير المصابين

Gold Standard هو الاختبار الذي يستعمل لتحديد الحالات المرضية الصحيحة ويستعمل للمقارنة مع اختبارات جديدة لتقدير كفائتها.

ومن الامثلة عن هذه الاختبارات:

Gold Standard	Target Disorder
Exisional biopsy	Breast cancer
Transrectal biopsy	Prostate cancer
Coronary angiography	Coronary stenosis
Catteterization	Myocardial infection
Throat culture	Strep throat

10 – 2 وصف اختبارات التشخيص

هناك عدة مقاييس تستعمل لغرض وصف مستوى اختبار تشخيص معين نسبة الى الاختبار القياسي (Gold Standard) وهي بذلك توضح مدى نجاح ذلك الاختبار.

Sensitivity : هو احتمال ان يكون الفرد مصاب بمرض ما واعطى اختبار موجب عند استعمال الاختبار الجديد وتحسب وفق المعادلة : a/(a+c) .

Specificity : احتمال الفرد ان يكون غير مصاب واعطى اختبار سالب عند استعمال الاختبار الجديد وتحسب وفق المعادلة: d/ (b+d) .

Accurecy : هو احتمال ان يعطي الاختبار الجديد اختبار صحيح من مجموع الاحتمالات وتحسب وفق المعادلة: (a+d) / (a+b+c+d).

مثال (132): في دراسة لمقارنة كفاءة الاختبار FNA بالمقارنة مع اختبار Gold Standard على 114 امرأة اذ اشار الفحص الفيزياوي .Physical exam بعدم وجود ورم ظاهري في

Dr. Firas

الثدي الا ان Mammograms كان غير طبيعي ، جرى اختبار النساء وفق اختبار FNA وسجلت الحالات المصابة (22 حالة) وغير المصابة (92 حالة) واخضعت النساء الى عمليات جراحية فوجد ان 8 حالات من الـ 22 حالة كانت موجبة كاذبة وحالة واحدة من الـ 92 حالة كانت سالبة كاذبة ، ويمكن ترتيب جدول النتائج كالاتي:

	Cancer	No Cancer
FNA +	14	8
FNA –	1	91

Sensitivity = 14/ 15 = 0.93 أو 93%

Specificity = 91/ 99 = 0.92 او 92%

Accurecy = (14 + 91) / (14 + 8 + 1 + 9) = 0.92 او 92%

10 – 3 تقدير احتمالية الاصابة

ان كل من Sensitivily و Specificity يمثلان مقياس لكفاءة الاختبار،واجراء الاختبار مهم ليتم على ضوئه تقرير حالة الفرد، ونحن نسعى من اجل ان نتمكن من تقدير احتمالية تشخيص المرض اعتمادا على فحص تشخيصي واحد او اكثر. وهناك عدة مقاييس لذلك.

Prevalence: هو احتمالية ان يكون الفرد مريض او مايسمى بأسبقية احتمال مرض الفرد وتحسب من المعادلة:

Prevalence = (a + c) / (a + b+ c+ d.

Positive Prediction Value (PV+) = a / (a+b)

Negative Prediction Value (PV –) = d / (c+d)

وفي مثالنا السابق عن اصابة النساء بسرطان الثدي فأن المقاييس تساوي:

Prevalence= 15/114 = 0.13

PV + = 14/ (14 + 8) = 0.64

PV – = 91/ (1 + 91) = 0.99

النساء لهن اسبقية احتمال اصابة قدرها 0.13 اذا كان Mammograms غير طبيعي ولكن الاسبقية تزداد اذا كانت نتائج اختبار FNA موجبة اذ تصبح 0.64، فيما تكون الاسبقية لاحتمالية عدم وجود المرض في النساء هو 0.87 (99\114) وهذه النسبة تزداد الى 0.99 اذا كان اختبار FNA سالب.

يمكن تقدير قيمتا التنبوء الموجبة والسالبة بأعتماد المعادلتين التاليتين:

Dr.Firas

$$PV+=\frac{\text{Prevalence}\times\text{Sensitivity}}{(\text{Prevalence}\times\text{Sensitivity})+\{(1-\text{Prevalence})\times(1-\text{Specificity})\}}$$

$$PV-=\frac{(1-\text{Prevalence})\times\text{Specificity}}{\{(1-\text{Prevalence})\times\text{Specificity}\}+\{\text{Prevalence}\times(1-\text{Sensitivity})\}}$$

PV+ = (0.13)(0.93)/{(0.13)(0.93) + (0.87)(0.08)} = 0.64

PV- = (0.87)(0.92)/{(0.87)(0.92) + (0.13)(0.07)} = 0.99

وهي ايضا يمكن الحصول عليها من الجدول مباشرة كما اوضحنا سابقا

PV + = 14/ (14 + 8) = 0.64

PV – = 91/ (1 + 91) = 0.99

هناك اختبار آخر نجده في بعض البحوث الطبية ويسمى Likelihood ratio وهو يمثل احتمال نتيجة اختبار معين (موجب او سالب) في مجموعة من المرضى مقسوما على احتمال نتيجة مجموعة من الافراد بدون مرض. ويعبر عن الاختبار الموجب +LR والاختبار السالب – LR وتحسب وفق المعادلتين الآتية:

LR+ = Sensitivity/(1 - Specificity)

LR – = (1 - Sensitivity)/Specificity

وعند تطبيق المعادلتين على مثالنا السابق

LR+ = 0.933/0.081 = 11.52

LR – = 0.067/0.919 = 0.07

ان نتيجة الاختبار FNA الموجب تعني ان 11.52 مرة يكون اكثر مشابها لدى النساء المصابة بالسرطان مقارنة بغير المصابة والنتيجة السالبة لنفس الاختبار تعني 0.07 مرة يكون مشابها للنساء المصابة مقارنة بغير المصابة.

10 – 4 مقارنة اختبارين للتشخيص

لو فرضنا اننا اردنا اجراء مقارنة في Sensitivity و Specificity بين اختبارين ، فأذا اجرينا الاختبارين على مجموعتين مستقلتين من المرضى فأن المقارنة بينهما تتم وفق مربع كاي او اختبار فيشر اما اذا اجري الاختباران على نفس المجموعة فلابد من استعمال اختبار McNemars .

مثال (133): اجري اختباران على مجموعتين من المرضى وقد تبين من الاختبار الاول ان هناك 112 حالة موجبة و 88 حالة سالبة وعند اجراء اختبار Gold standard وجد ان هناك 30 حالة موجبة كاذبة و 18 حالة سالبة كاذبة في ذلك الاختبار، بمعنى ان عدد الحالات

الموجبة الصحيحة هو 112 – 30 = 82 حالة وعدد الحالات السالبة الحقيقية هو 88 – 18 = 70 حالة وبذلك يمكن تمثيل النتائج بالجدول التالي:

	Disease	No Disease
+	82	30
–	18	70

اما الاختبار الثاني فوجد ان الحالات الموجبة 150 والحالات السالبة 150 وعند اجراء الاختبار القياسي وجد ان 10 حالات كانت موجبة كاذبة اي ان عدد الحالات الموجبة الصحيحة 140 كما وجد 60 حالة سالبة كاذبة اي ان عدد الحالات السالبة الصحيحة 90 ، وبذلك فأن الجدول كألاتي:

	Disease	No Disease
+	140	10
–	60	90

قدر حساسية كل من الاختبارين؟

Sensitivity للأختبار الاول = 100/82 = 0.82

Sensitivity للأختبار الثاني = 140/ 200 = 0.70

مثال (134) :لو اردنا المقارنة بين الاختبارين في المثال السابق ؟

```
data d;
input test positive count;
cards;
1 1 82
1 2 18
2 1 140
2 2 60
proc freq;
tables positive*test/chisq;
exact chisq;
weight count;
 run;
```

```
          TABLE OF POSITIVE BY TEST
POSITIVE    TEST

Frequency,
Percent ,
Row Pct ,
Col Pct ,     1,     2, Total
ƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
     1 ,   82 ,  140 ,  222
       , 27.33 , 46.67 , 74.00
       , 36.94 , 63.06 ,
       , 82.00 , 70.00 ,
ƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
     2 ,   18 ,   60 ,   78
       ,  6.00 , 20.00 , 26.00
       , 23.08 , 76.92 ,
       , 18.00 , 30.00 ,
ƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
```

Dr Firas

```
Total       100     200     300
            33.33   66.67   100.00

      STATISTICS FOR TABLE OF POSITIVE BY TEST

Statistic                  DF    Value      Prob
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Chi-Square                  1    4.990     0.026
Likelihood Ratio Chi-Square   1    5.210      0.022
Continuity Adj. Chi-Square    1    4.385      0.036
Mantel-Haenszel Chi-Square    1    4.973      0.026
Fisher's Exact Test (Left)              0.992
              (Right)               0.017
              (2-Tail)              0.026
Phi Coefficient                 0.129
Contingency Coefficient           0.128
Cramer's V                      0.129

Sample Size = 300
```

من نتائج التحليل تبين بأن اختبار مربع كاي وفيشر (0.026) يشيران الى وجود فرق معنوي في حساسية الاختبارين وان الاختبار الاول افضل من الثاني.

مثال (135) : اجريت دراسة للمقارنة بين اختبارين اجريا على مجموعة واحدة وكانت نتائج الاختبار الاول 65 حالة موجبة و 35 حالة سالبة والاختبار الثاني 53 حالة موجبة و 47 حالة سالبة وعند مقارنة نتائج الاختبارين بالاختبار القياسي وجد ان عدد الحالات الموجبة كانت 35 حالة موجبة كاذبة في كلا الاختبارين و 23 حالة سالبة كاذبة في كلا الاختبارين اي ان عدد الاختبارات الموجبة الصحيحة في الاختبار الاول 65 – 35 = 30 حالة وفي الاختبار الثاني 53 – 23 = 30 حالة وفي الاختبارات السالبة فأن عددها في الاختبار الاول 35 – 23 = 12 وفي الاختبار الثاني 47 – 35 = 12 وبذلك يمكن التعبيرعنها بالجدول الآتي :

	الاختبار الثاني	
	+	–
الاختبار الاول +	30	35
–	23	12

هنا سنستخدم اختبار McNemar

```
data compar;
input test1 test2 count;
cards;
1 1 30
1 2 35
2 1 23
2 2 12
proc freq data=compar;
tables test1*test2/agree;
weight count;
exact McNem;
run;
```

```
                The SAS System
              TABLE OF TEST1 BY TEST2


          TEST1     TEST2
```

```
Frequency,
Percent ,
Row Pct ,
Col Pct ,     1,     2, Total
ƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
    1 ,   30 ,   35 ,   65
      , 30.00 , 35.00 , 65.00
      , 46.15 , 53.85 ,
      , 56.60 , 74.47 ,
ƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
    2 ,   23 ,   12 ,   35
      , 23.00 , 12.00 , 35.00
      , 65.71 , 34.29 ,
      , 43.40 , 25.53 ,
ƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total       53     47    100
          53.00   47.00  100.00

STATISTICS FOR TABLE OF TEST1 BY TEST2

          McNemar's Test
          --------------
Statistic = 2.483   DF = 1   Prob (Asymptotic) = 0.115
                    Prob (Exact)     = 0.148
        Simple Kappa Coefficient
        ------------------------
                     95% Confidence Bounds
Kappa = -0.181   ASE = 0.094      -0.366   0.004
Sample Size = 100
```

الاختلافات في حساسية الاختبارين غير معنوية اي ان لهما نفس الكفاءة

مثال (136) :في دراسة لمعرفة تأثير عقار على شفاء المرضى اخذت عينة من المرضى في مركزين وبواقع 90 مريض لكل مركز ثم قسمت العينة بصورة عشوائية الى مجموعتين اعطيت احد المجموعتين العقار المراد دراسة تأثيره فيما اعطيت المجموعة الثانية Placebo وكانت الفرضية هي عدم وجود علاقة بين العقار ونسبة الشفاء فيما كانت النظرية البديلة ان العقار والشفاء عاملان غير مستقلان.

وقد بلغ عدد الافراد الذين استجابوا للعلاج في المركز الاول 29 ومن لم يستجب 16 فيما كان عدد الذين تم شفائهم في الجموعة الثانية 14 والذين لم يتم شفائهم 31 وكانت التقديرات المناظرة لها في المركز الثاني 37 و 8 و 24 و 21 على التوالي.

```
data respire;
   input center treat $ response $ count @@;
   datalines;
 1 test    y 29 1 test    n 16
 1 placebo y 14 1 placebo n 31
 2 test    y 37 2 test    n  8
 2 placebo y 24 2 placebo n 21
 proc freq order=data;
   weight count;
   tables center*treat*response /
      nocol nopct chisq cmh;
 run;
```

Dr.Firas

The SAS System
TABLE 1 OF TREAT BY RESPONSE
CONTROLLING FOR CENTER=1

TREAT RESPONSE

Frequency,
Row Pct ,y ,n , Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
test , 29 , 16 , 45
, 64.44 , 35.56 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
placebo , 14 , 31 , 45
, 31.11 , 68.89 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 43 47 90

STATISTICS FOR TABLE 1 OF TREAT BY RESPONSE
CONTROLLING FOR CENTER=1

Statistic	DF	Value	Prob
Chi-Square	1	10.020	0.002
Likelihood Ratio Chi-Square	1	10.216	0.001
Continuity Adj. Chi-Square	1	8.728	0.003
Mantel-Haenszel Chi-Square	1	9.908	0.002
Fisher's Exact Test (Left)			1.000
(Right)			1.46E-03
(2-Tail)			2.92E-03
Phi Coefficient			0.334
Contingency Coefficient			0.317
Cramer's V			0.334

Sample Size = 90

تبين وجود تأثير معنوي للعلاج على الاستجابة في المركز الاول

The SAS System
TABLE 2 OF TREAT BY RESPONSE
CONTROLLING FOR CENTER=2

TREAT RESPONSE

Frequency,
Row Pct ,y ,n , Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
test , 37 , 8 , 45
, 82.22 , 17.78 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
placebo , 24 , 21 , 45
, 53.33 , 46.67 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 61 29 90

STATISTICS FOR TABLE 2 OF TREAT BY RESPONSE
CONTROLLING FOR CENTER=2

Statistic	DF	Value	Prob
Chi-Square	1	8.598	0.003
Likelihood Ratio Chi-Square	1	8.832	0.003
Continuity Adj. Chi-Square	1	7.326	0.007
Mantel-Haenszel Chi-Square	1	8.503	0.004
Fisher's Exact Test (Left)			0.999
(Right)			3.14E-03
(2-Tail)			6.28E-03
Phi Coefficient			0.309
Contingency Coefficient			0.295

Dr. Firas

Cramer's V 0.309

Sample Size = 90

تبين وجود تأثير معنوي للعلاج على الاستجابة في المركز الثاني

The SAS System
SUMMARY STATISTICS FOR TREAT BY RESPONSE
CONTROLLING FOR CENTER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	18.411	0.001
2	Row Mean Scores Differ	1	18.411	0.001
3	General Association	1	18.411	0.001

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Bounds	
Case-Control	Mantel-Haenszel	4.029	2.132	7.614
(Odds Ratio)	Logit	4.029	2.106	7.707
Cohort	Mantel-Haenszel	1.737	1.350	2.235
(Col1 Risk)	Logit	1.676	1.294	2.170
Cohort	Mantel-Haenszel	0.462	0.324	0.657
(Col2 Risk)	Logit	0.474	0.326	0.688

The confidence bounds for the M-H estimates are test-based.

ارجحية الاستجابة من العلاج للمرضى الذين تناولوا الدواء تساوي 4 مرات مقارنة بالمرضى الذين لم يأخذوا الدواء

Breslow-Day Test for Homogeneity of the Odds Ratios

Chi-Square = 0.000 DF = 1 Prob = 0.990

Total Sample Size = 180

اختبار بريسلو يشير الى تجانس تقديرات نسبة الارجحية في كلا المركزين

يلاحظ من النتائج ان العقار والشفاء عاملان غير مستقلان عن بعضهما وفي كلا المركزين حسب اختبار مربع كاي وعند تقدير نسبة الارجحية وجد ان المرضى الذين اخذوا العلاج كان ارجحية شفائهم هي 4.029 مرة مقارنة بالمجموعة الثانية وبمجال ثقة 2.132 الى 7.62 فيما بلغت خطورة عدم الشفاء في المجموعة الثانية مقارنة بالاولى 1.737 وبحدود ثقة من 1.35 الى 2.17 . واوضحت نتائج اختبار بريسلو عدم وجود فروق معنوية بين تقديرات نسب الارجحية للمجاميع المقارنة. ويمكن تقدير ارجحية الشفاء لكل مركز بأضافة كلمتي by center الى نفس الايعاز السابق وكالآتي :

```
proc freq order=data;
    weight count; by center;
    tables treat*response /
      nocol nopct chisq cmh;
```

Dr Firas

run;

مثال (137): في دراسة لمعرفة العلاقة بين المرض وجنس المريض (ذكر او انثى) وعلاقة المرض بقراءة EGC شملت الدراسة 33 امرأة و35 رجل وقد وجد ان قراءة EGC عندما كانت اقل من 0.1 كان عدد النساء المريضات 4 وغير المريضة 11 وعندما كانت القراءة اكبر او تساوي 0.1 كان عدد النساء المريضة 8 وغير المريضة 10 وبلغت التقديرات المناظرة لها في الرجال 9 ، 9 ، 21 ، 6 على التوالي.

```
data ca;
    input gender $ ECG $ disease $ count;
    datalines;
  female <0.1  yes   4
  female <0.1  no   11
  female >=0.1 yes   8
  female >=0.1 no   10
  male   <0.1  yes   9
  male   <0.1  no    9
  male   >=0.1 yes  21
  male   >=0.1 no    6
  proc freq;
    weight count;
    tables gender*disease / nocol nopct chisq;
    tables gender*ECG*disease / nocol nopct cmh chisq measures;
  run;
```

```
                The SAS System
              TABLE OF GENDER BY DISEASE

            GENDER    DISEASE

            Frequency,
            Row Pct  ,no     ,yes    , Total
            ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
            female   ,   21 ,   12 ,   33
                 , 63.64 , 36.36 ,
            ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
            male     ,   15 ,   30 ,   45
                 , 33.33 , 66.67 ,
            ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
            Total       36     42     78

          STATISTICS FOR TABLE OF GENDER BY DISEASE

        Statistic                 DF     Value      Prob
        ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
        Chi-Square                 1     7.035      0.008
        Likelihood Ratio Chi-Square  1     7.121      0.008
        Continuity Adj. Chi-Square   1     5.868      0.015
        Mantel-Haenszel Chi-Square   1     6.944      0.008
        Fisher's Exact Test (Left)                    0.998
                  (Right)                  7.51E-03
                  (2-Tail)                    0.011
        Phi Coefficient                    0.300
        Contingency Coefficient            0.288
        Cramer's V                         0.300

        Sample Size = 78
```

اختبار مربع كاي يشير الى وجود علاقة بين المرض وجنس المريض

```
                The SAS System
```

Dr. Firas

TABLE 1 OF ECG BY DISEASE
CONTROLLING FOR GENDER=female

ECG DISEASE

Frequency,
Row Pct ,no ,yes , Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
<0.1 , 11 , 4 , 15
, 73.33 , 26.67 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
>=0.1 , 10 , 8 , 18
, 55.56 , 44.44 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 21 12 33

The SAS System
STATISTICS FOR TABLE 1 OF ECG BY DISEASE
CONTROLLING FOR GENDER=female

Statistic	DF	Value	Prob
Chi-Square	1	1.117	0.290
Likelihood Ratio Chi-Square	1	1.134	0.287
Continuity Adj. Chi-Square	1	0.481	0.488
Mantel-Haenszel Chi-Square	1	1.084	0.298
Fisher's Exact Test (Left)			0.923
(Right)			0.245
(2-Tail)			0.469
Phi Coefficient			0.184
Contingency Coefficient			0.181
Cramer's V			0.184

لاتوجد علاقة معنوية بين المرض وقراءة ECG لدى النساء

Statistic	Value	ASE
Gamma	0.375	0.323
Kendall's Tau-b	0.184	0.169
Stuart's Tau-c	0.176	0.162
Somers' D C|R	0.178	0.164
Somers' D R|C	0.190	0.174
Pearson Correlation	0.184	0.169
Spearman Correlation	0.184	0.169
Lambda Asymmetric C|R	0.000	0.000
Lambda Asymmetric R|C	0.067	0.295
Lambda Symmetric	0.037	0.167
Uncertainty Coefficient C|R	0.026	0.048
Uncertainty Coefficient R|C	0.025	0.046
Uncertainty Coefficient Symmetric	0.026	0.047

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Bounds	
Case-Control	2.200	0.504	9.611
Cohort (Col1 Risk)	1.320	0.790	2.206
Cohort (Col2 Risk)	0.600	0.224	1.607

Sample Size = 33

ارجحية المرض تساوي 2.2 مرة لدى النساء اللاتي كانت قراءة ECG لهن اكبر اوتساوي 0.1

The SAS System
TABLE 2 OF ECG BY DISEASE

Dr Firas

CONTROLLING FOR GENDER=male

ECG DISEASE

Frequency,
Row Pct ,no ,yes , Total
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
<0.1 , 9 , 9 , 18
, 50.00 , 50.00 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
>=0.1 , 6 , 21 , 27
, 22.22 , 77.78 ,
ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
Total 15 30 45

The SAS System
STATISTICS FOR TABLE 2 OF ECG BY DISEASE
CONTROLLING FOR GENDER=male

Statistic	DF	Value	Prob
Chi-Square	1	3.750	0.053
Likelihood Ratio Chi-Square	1	3.729	0.053
Continuity Adj. Chi-Square	1	2.604	0.107
Mantel-Haenszel Chi-Square	1	3.667	0.056
Fisher's Exact Test (Left)			0.988
(Right)			0.054
(2-Tail)			0.105
Phi Coefficient			0.289
Contingency Coefficient			0.277
Cramer's V			0.289

لاتوجد علاقة معنوية بين المرض وقراءة ECG لدى الرجال

Statistic	Value	ASE
Gamma	0.556	0.228
Kendall's Tau-b	0.289	0.146
Stuart's Tau-c	0.267	0.138
Somers' D C|R	0.278	0.142
Somers' D R|C	0.300	0.152
Pearson Correlation	0.289	0.146
Spearman Correlation	0.289	0.146
Lambda Asymmetric C|R	0.000	0.283
Lambda Asymmetric R|C	0.167	0.196
Lambda Symmetric	0.091	0.209
Uncertainty Coefficient C|R	0.065	0.066
Uncertainty Coefficient R|C	0.062	0.063
Uncertainty Coefficient Symmetric	0.063	0.064

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Bounds	
Case-Control	3.500	0.959	12.778
Cohort (Col1 Risk)	2.250	0.968	5.230
Cohort (Col2 Risk)	0.643	0.388	1.064

Sample Size = 45

ارجحية المرض تساوي 3.5 مرة لدى الرجال الذين كانت قراءة ECG لديهم اكبر او تساوي 0.1.

The SAS System
SUMMARY STATISTICS FOR ECG BY DISEASE
CONTROLLING FOR GENDER

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob

Dr.Firas

```
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
  1      Nonzero Correlation       1      4.503    0.034
  2      Row Mean Scores Differ    1      4.503    0.034
  3      General Association       1      4.503    0.034
```

توجد علاقة معنوية بين المرض بغض النظر عن الجنس وقراءة ECG

```
    Estimates of the Common Relative Risk (Row1/Row2)
                                95%
Type of Study   Method            Value   Confidence Bounds
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Case-Control    Mantel-Haenszel   2.847     1.083    7.482
 (Odds Ratio)   Logit             2.859     1.081    7.565

Cohort          Mantel-Haenszel   1.641     1.039    2.594
 (Col1 Risk)    Logit             1.525     0.983    2.365

Cohort          Mantel-Haenszel   0.630     0.411    0.965
 (Col2 Risk)    Logit             0.634     0.405    0.993

The confidence bounds for the M-H estimates are test-based.
    Breslow-Day Test for Homogeneity of the Odds Ratios
```

ارجحية الاصابة تكون 2.8 مرة للمرضى الذين قراءة ECG لهم اكبر او تساوي 0.1

```
Chi-Square =   0.215        DF =   1        Prob = 0.643

Total Sample Size = 78
```

مثـال (138) فـي دراسـة عـن تـأثير العـلاج بعقـار جديـد مـع placebo علـى الصـداع سـجلت الاسـتجابة بحـالتين (أفضـل او لاتغييـر) وشـملت رجـال ونسـاء المطلـوب اختبـار الاقتـران بـين المعاملة والاستجابة من العلاج؟

```
data Migraine;
   input Gender $ Treat $ Response $ Count @@;
   datalines;
 female Active  Better 16   female Active  Same 11
 female Placebo Better  5   female Placebo Same 20
 male   Active  Better 12   male   Active  Same 16
 male   Placebo Better  7   male   Placebo Same 19
   proc freq data=Migraine;
   weight Count;
   tables Gender*Treat*Response / cmh noprint;
 run;
```

```
                                 Clinical Trial for Treatment of Pain
       SUMMARY STATISTICS FOR TREAT BY RESPONSE
            CONTROLLING FOR GENDER

 Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic  Alternative Hypothesis    DF      Value     Prob
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
  1      Nonzero Correlation       1      8.305    0.004
  2      Row Mean Scores Differ    1      8.305    0.004
  3      General Association       1      8.305    0.004
```

ان نتائج اختبارات CMH (P=0.004) تشير الى وجود اقتران بين المعاملة والاستجابة عند التعديل للجنس.

```
    Estimates of the Common Relative Risk (Row1/Row2)
                                95%
Type of Study   Method            Value   Confidence Bounds
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Case-Control    Mantel-Haenszel   3.313     1.467    7.483
```

Dr. Firas

```
(Odds Ratio)  Logit            3.294    1.418   7.652

Cohort        Mantel-Haenszel  2.164    1.280   3.657
 (Col1 Risk)  Logit            2.106    1.195   3.711

Cohort        Mantel-Haenszel  0.642    0.475   0.868
 (Col2 Risk)  Logit            0.661    0.485   0.901
```

احتمال شفاء المرضى من الصداع عند تناولهم للعقار هو 3 أضعاف احتمال الشفاء لمجموعة السيطرة.

```
      Breslow-Day Test for Homogeneity of the Odds Ratios

Chi-Square =  1.493      DF =  1      Prob = 0.222
Total Sample Size = 106
```

نتائج اختبار بريسلو لتجانس نسبة الارجحية تشير الى عدم اختلافها بين الرجال والنساء.

مثال (139) في دراسة لمعرفة تأثير الغذاء الغني بالدهون والمنخفض الدهون على الاصابة بامراض الشرايين التاجية للقلب ، سجلت البيانات على مجموعة من الافراد ورتبت البيانات لغرض تحليلها وكانت قيم الغذاء عالي الكوليستيرول = 1 ومنخفض الكوليستيرول = 0 ، والمصاب = 1 وغير المصاب = 0.

```
data FatComp;
input Exposure Response Count;
cards;
0 0  6
0 1  2
1 0  4
1 1 11
proc freq data=FatComp order=data;
 weight Count;
 tables Exposure*Response / chisq relrisk;
 exact pchi or;
 run;
```

```
                TABLE OF EXPOSURE BY RESPONSE

                EXPOSURE    RESPONSE

                Frequency,
                Percent ,
                Row Pct ,
                Col Pct ,     0,     1, Total
                ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
                     0 ,    6 ,    2 ,    8
                       , 26.09 ,  8.70 , 34.78
                       , 75.00 , 25.00 ,
                       , 60.00 , 15.38 ,
                ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
                     1 ,    4 ,   11 ,   15
                       , 17.39 , 47.83 , 65.22
                       , 26.67 , 73.33 ,
                       , 40.00 , 84.62 ,
                ƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒˆ
                Total       10     13     23
                       43.48   56.52  100.00

        STATISTICS FOR TABLE OF EXPOSURE BY RESPONSE
                                    Prob      Prob
Statistic               DF    Value  (Asymptotic) (Exact)
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
Chi-Square               1    4.960    0.026     0.039
Likelihood Ratio Chi-Square   1    5.098    0.024
Continuity Adj. Chi-Square    1    3.188    0.074
Mantel-Haenszel Chi-Square    1    4.744    0.029
Fisher's Exact Test (Left)                    0.997
```

Dr. Firas

	(Right)	0.037
	(2-Tail)	0.039
Phi Coefficient	0.464	
Contingency Coefficient	0.421	
Cramer's V	0.464	

WARNING: 50% of the cells have expected counts less than 5.

بسبب ان القيم المتوقعة في بعض الخلايا تقل عن خمسة فسيظهر تحذير في البرنامج يشير الى عدم صلاحية اختبار مربع كاي ، لذا فأن اختبار فيشر المضبوط هو المطلوب وبما ان الفرضية البديلة تفترض ان الاصابة بامراض القلب تزداد مع زيادة الكوليستيرول في الغذاء المتناول لذا فأن اختبار فيشر المضبوط لطرف واحد (الى اليمين هو المطلوب) والذي يساوي 0.03 مما يدعم الفرضية البديلة.

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Bounds (Asymptotic)		95% Confidence Bounds (Exact)	
Case-Control	**8.250**	1.154	59.003	0.868	105.549
Cohort (Col1 Risk)	2.813	1.110	7.126		
Cohort (Col2 Risk)	0.341	0.099	1.176		

Sample Size = 23

نلاحظ من النتائج ان نسبة الارجحية لاحتمال الاصابة بامراض القلب للافراد الذين يتناولون غذاء عالي الكوليستيرول يساوي 8.25 مرة احتمال الاصابة للافراد الذين يتناولون غذاء منخفض الكوليستيرول.

Dr Firas

الفصل الحادي عشر

تقدير العمر

يعد العمر من الصفات المهمة في الانسان والحيوان اذ ان زيادة معدل العمرفي الانسان يعكس مدى رقي وتقدم البلدان اذ ترتفع في البلدان المتقدمة فيما تنخفض في البلدان النامية ويكتسب العمر اهمية اقتصادية في الحيوانات مثل الابقار والاغنام اذ ان ارتفاع معدل الاعمار هو مؤشر عن تحسن صحة الحيوان وادائه التناسلي وزيادة انتاجه.

ان تقدير العمر يمكن ان يجرى باعتماد نماذج خطية الا ان الطريقة الحديثة والتي تعتمد اليوم لدى العديد من دول العالم هي طريقة تحليل البقاء (Survival analysis) وهي طريقة احصائية لاخطية يستعمل فيها انموذج الخطورة النسبية (Prportional hazard mdel) وقد استعملت في البداية في الابحاث الطبية والهندسية ، ويعد Famula اول من اقترح استعمالها في الحيوانات حتى اصبحت الطريقة المعتمدة لدى معظم دول اوربا بسبب الخصائص المميزة لها.

11 – 1 تقدير العمر بأستعمال طريقة تحليل البقاء Survival Analysis Method

تتصف البيانات الخاصة بالعمر بكونها تتضمن بيانات عن افراد ميتة (غير مراقبة) (uncensord) علاوة على بيانات افراد حية (افراد مراقبة) (censored) وفي مثل هذه البيانات يمكننا معرفة عمر الافراد التي تصل الى قيمة معينة فيما تبقى القيمة الحقيقية للعمر غير معروفة اذ ان شمول الافراد الميتة بالتحليل واهمال الافراد الحية سيؤدي الى اعطاء نتائج متحيزة اذ ان الافراد ذات العمر الطويل تكون عادة غير مشمولة بالتحليل في النماذج الخطية لذا فأن التحليل الاكثر دقة هوالذي يشمل كلا البيانات في التحليل وهو مايمكن تطبيقه باستعمال طريقة تحليل البقاء. والخطوة الاولى في هذه الطريقة هو تقدير توزيع اوقات البقاء وتستعمل دالة توزيع البقاء Survival Distribution Function (SDF) لوصف عمر العشيرة المدروسة.

توفر طريقة Lifetest امكانية حساب التقديرات اللامعلمية لدالة البقاء اما بطريقة product-limit او ماتسمى بطريقة Kaplan-Meier او بطريقة جدول الحياة Life Table.

تمكن طريق تحليل البقاء من اجراء مقارنة بين منحنيات البقاء ، اذ ان من المفيد الاقرار بأن عينتين او اكثر لهما نفس دالة البقاء. توفر طريقة Lifetest اختبارين للرتبة (Savage و Wilicoxon) واختبار التجانس (Likelihood Ratio) لدوال البقاء خلال الطبقات (Strata).

مثال (140): في دراسة عن 40 فأر تعرضوا الى carcinogen وقسمت الى مجموعتين واعطي علاج لكل مجموعة وقد عبر عن الاستجابة بفترة بقاء الافراد منذ التعرض لغاية موتها وقد هلكت اربعة افراد لاسباب لاتتعلق بالسرطان لذا اعتبرت افراد مراقبة ، المطلوب معرفة هل ان توزيع البقاء بين المجموعتين مختلف؟

تتضمن البيانات اربعة متغيرات 1- Days (فترة بقاء بالايام من العلاج لغاية موتها).2- الحالة Status فاذا كانت مراقبة =0 واذا كانت غير مراقبة = 1 ، 3- المعاملة ، 4- جنس الحيوان اذ ان F للانثى و M للذكر.

```
data Exposed;
input Days Status Treat Sex $ @@;
datalines;
179 1 1 F 378 0 1 M
256 1 1 F 355 1 1 M
262 1 1 M 319 1 1 M
256 1 1 F 256 1 1 M
255 1 1 M 171 1 1 F
224 0 1 F 325 1 1 M
225 1 1 F 325 1 1 M
287 1 1 M 217 1 1 F
319 1 1 M 255 1 1 F
264 1 1 M 256 1 1 F
237 0 2 F 291 1 2 M
156 1 2 F 323 1 2 M
270 1 2 M 253 1 2 M
257 1 2 M 206 1 2 F
242 1 2 M 206 1 2 F
157 1 2 F 237 1 2 M
249 1 2 M 211 1 2 F
180 1 2 F 229 1 2 F
226 1 2 F 234 1 2 F
268 0 2 M 209 1 2 F
proc lifetest data=Exposed plots=(s,ls,lls);
time Days*Status(0);
strata Treat;
run;
```

ان وظيفة Plots تستعمل لغرض رسم العلاقة البيانية بين دالة البقاء المقدرة والزمن (تحدد بـ S) ، والعلاقة البيانية بين اللوغاريتيم السالب لدالة البقاء المقدرة والزمن (تحدد بـ LS) ، والعلاقة البيانية بين لوغاريتيم اللوغاريتيم السالب لدالة البقاء المقدرة ولوغاريتيم الزمن(يحدد بـ LLS). وتوفر كل من LS و LLS اختبار عملي لملائمة النموذج الاسي او نموذج Weibull. فاذا كان النموذج الاسي ملائم فأن منحنى LS يجب ان يكون خطي تقريبا ومن نقطة الاصل. واذا كان نموذج Weibull ملائم فأن منحنى LLS يجب ان يكون خطي تقريبا.

Dr. Firas

ونظرا لوجود عدة طبقات (Stratum) فأن منحنى LLS يمكن ان يستعمل للتأكد من صحة فرضية نموذج الخطورة النسبية (proportional hazards model) التي تفترض ان منحنى LLS يجب ان يكون متماثل خلال الطبقات (Strata).

The SAS System

The LIFETEST Procedure
Product-Limit Survival Estimates
TREAT = 1

DAYS	Survival	Survival Standard Failure	Number Error	Number Failed	Number Left
0.000	1.0000	0	0	0	20
171.000	0.9500	0.0500	0.0487	1	19
179.000	0.9000	0.1000	0.0671	2	18
217.000	0.8500	0.1500	0.0798	3	17
224.000*	.	.	.	3	16
225.000	0.7969	0.2031	0.0908	4	15
255.000	.	.	.	5	14
255.000	0.6906	0.3094	0.1053	6	13
256.000	.	.	.	7	12
256.000	.	.	.	8	11
256.000	.	.	.	9	10
256.000	0.4781	0.5219	0.1146	10	9
262.000	0.4250	0.5750	0.1135	11	8
264.000	0.3719	0.6281	0.1111	12	7
287.000	0.3188	0.6813	0.1071	13	6
319.000	.	.	.	14	5
319.000	0.2125	0.7875	0.0942	15	4
325.000	.	.	.	16	3
325.000	0.1063	0.8938	0.0710	17	2
355.000	0.0531	0.9469	0.0517	18	1
378.000*	.	.	.	18	0

* Censored Observation

النتائج اعلاه توضح دالة البقاء للطبقة الاولى (معاملة 1) وتتضمن بيانات عن الزمن وتقدير البقاء ونسبة الهلاك والخطأ القياسي وعدد الافراد الميتة وعدد الافراد الباقية.

Summary Statistics for Time Variable DAYS

Quantile	Point Estimate	95% Confidence Interval [Lower, Upper)	
75%	319.000	262.000	325.000
50%	256.000	255.000	319.000
25%	255.000	217.000	256.000

Mean 271.131 Standard Error 11.877

يمثل الجدول ملخص احصاءات متغير الزمن مثل الوسيط (256 يوم) والمتوسط (271.13 يوم) ويلاحظ ان التقديرات منخفضة وذلك لان اعلى قيمة للبقاء هي مراقبة اذ ان التقديرات تتحدد بأعلى قيمة للحدث غير المراقب.

The SAS System

The LIFETEST Procedure
Product-Limit Survival Estimates
TREAT = 2

Survival

DAYS	Survival	Standard Failure	Number Error	Number Failed	Number Left
0.000	1.0000	0	0	0	20
156.000	0.9500	0.0500	0.0487	1	19
157.000	0.9000	0.1000	0.0671	2	18
180.000	0.8500	0.1500	0.0798	3	17
206.000	.	.	.	4	16
206.000	0.7500	0.2500	0.0968	5	15
209.000	0.7000	0.3000	0.1025	6	14
211.000	0.6500	0.3500	0.1067	7	13
226.000	0.6000	0.4000	0.1095	8	12
229.000	0.5500	0.4500	0.1112	9	11
234.000	0.5000	0.5000	0.1118	10	10
237.000	0.4500	0.5500	0.1112	11	9
237.000*	.	.	.	11	8
242.000	0.3937	0.6063	0.1106	12	7
249.000	0.3375	0.6625	0.1082	13	6
253.000	0.2812	0.7188	0.1038	14	5
257.000	0.2250	0.7750	0.0971	15	4
268.000*	.	.	.	15	3
270.000	0.1500	0.8500	0.0891	16	2
291.000	0.0750	0.9250	0.0693	17	1
323.000	0	1.0000	0	18	0

* Censored Observation

النتائج اعلاه توضح دالة البقاء للطبقة الثانية (معاملة 2) وتتضمن بيانات عن الزمن وتقدير البقاء ونسبة الهلاك والخطأ القياسي وعدد الافراد الميتة وعدد الافراد المتبقية.

Summary Statistics for Time Variable DAYS

Quantile	Point Estimate	95% Confidence Interval [Lower, Upper)	
75%	257.000	237.000	291.000
50%	235.500	209.000	253.000
25%	207.500	180.000	234.000

Mean 235.156 Standard Error 10.211

تقديرات الوسيط (235.5 يوم) والمتوسط (235.15 يوم) للمعاملة الثانية.

The SAS System
The LIFETEST Procedure

Summary of the Number of Censored and Uncensored Values

TREAT	Total	Failed	Censored	%Censored
1	20	18	2	10.0000
2	20	18	2	10.0000
Total	40	36	4	10.0000

ملخص بالاحداث المراقبة وغير المراقبة والخاصة بكل طبقة والنسبة المؤية للاحداث المراقبة.

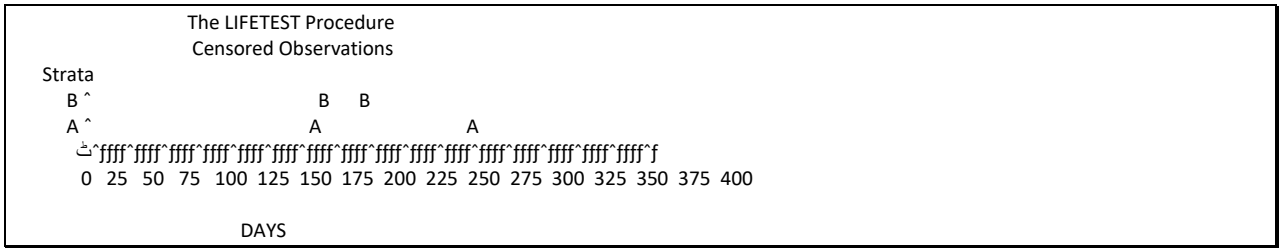
The LIFETEST Procedure

Survival Function Estimates

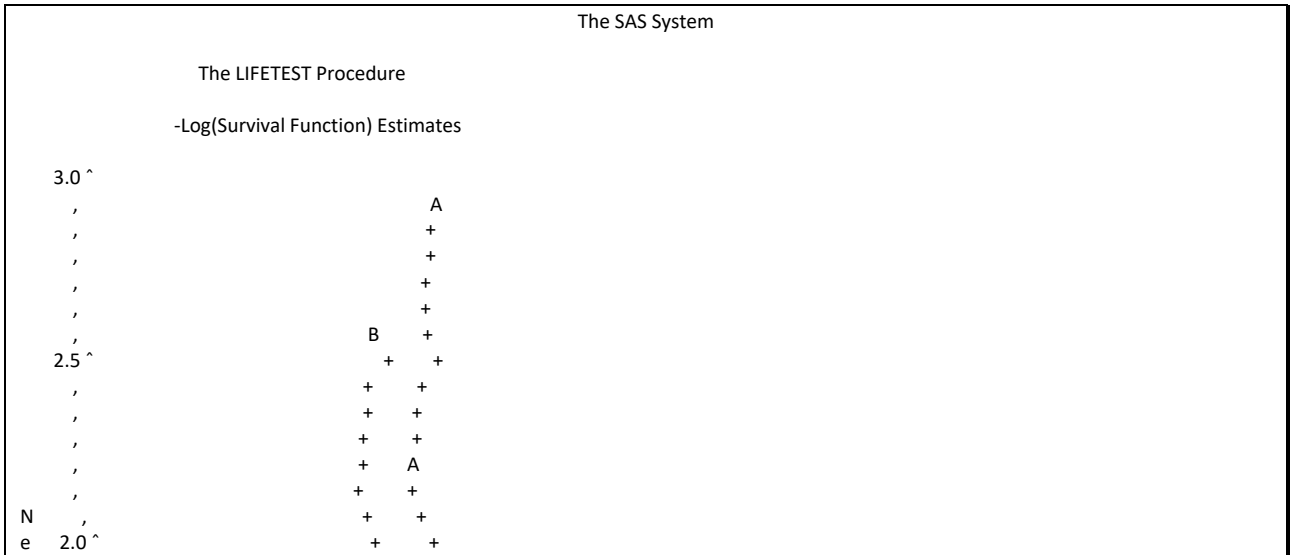
SDF ,
,
,
,
1.0 ^ *ffffffffffffffffffffffffffffffff*ffA
, , ,
S , B AfA
u , Bffff*ffffffA
r , , ,
v , BffffB AfA
i , , ,

Dr.Firas

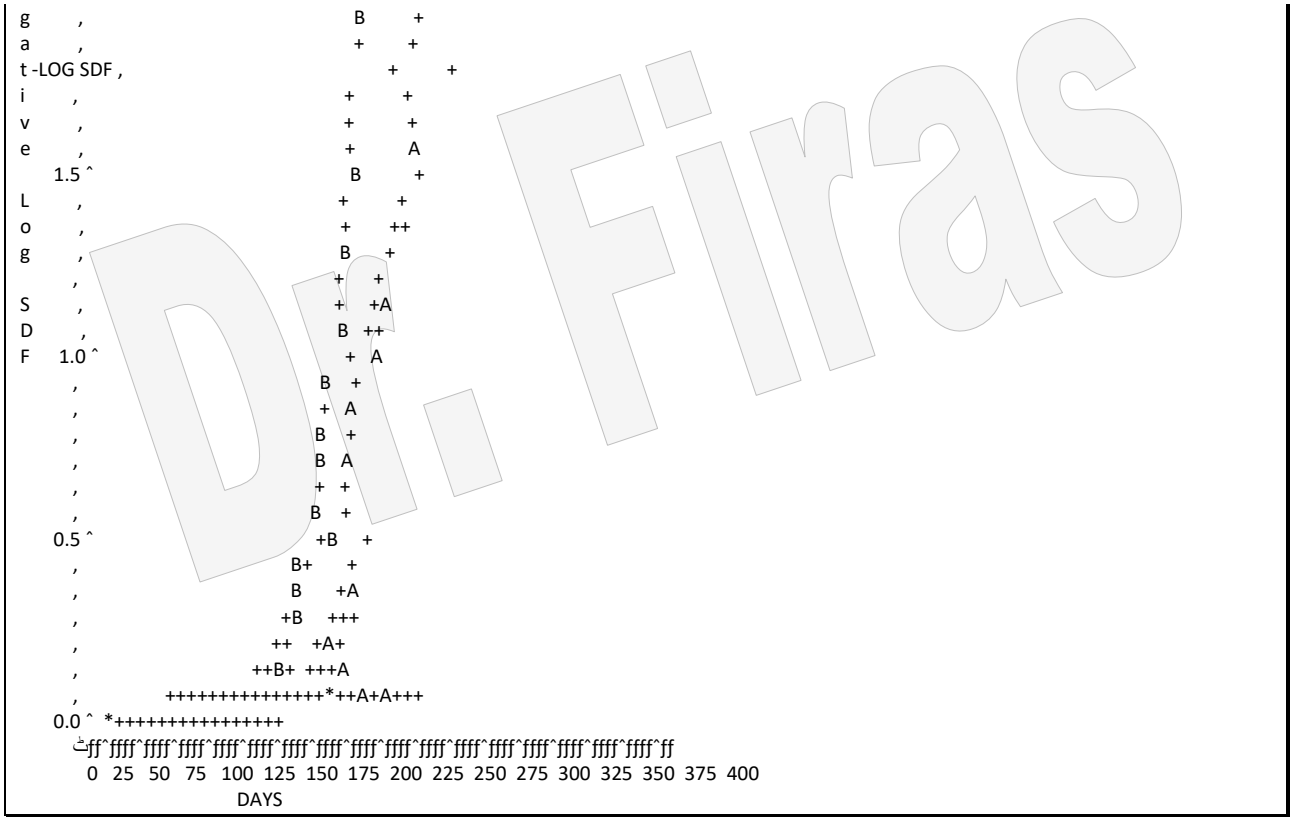
الشكل البياني اعلاه يوضح العلاقة بين دالة البقاء والزمن للمجموعتين ويلاحظ وجود اختلاف بينهما.



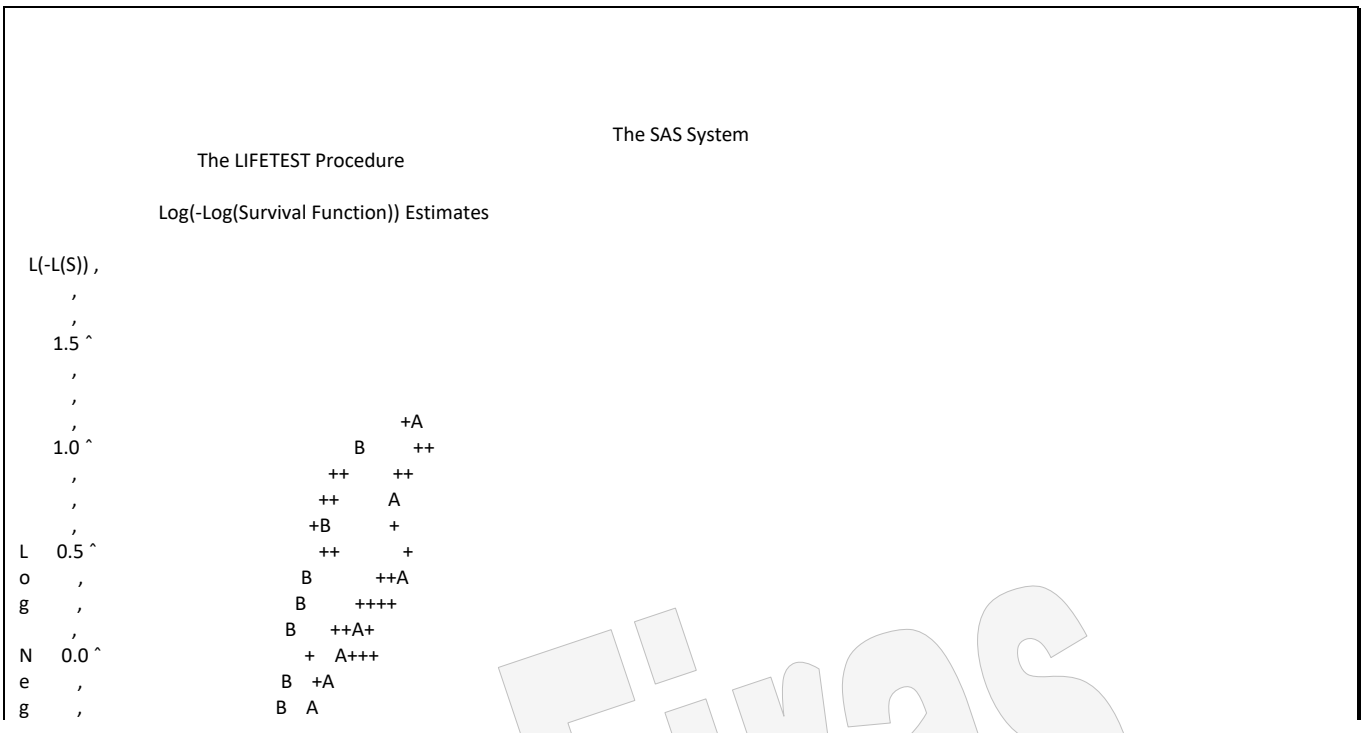
الشكل البياني اعلاه يوضح عدد البيانات المراقبة لكل معاملة.

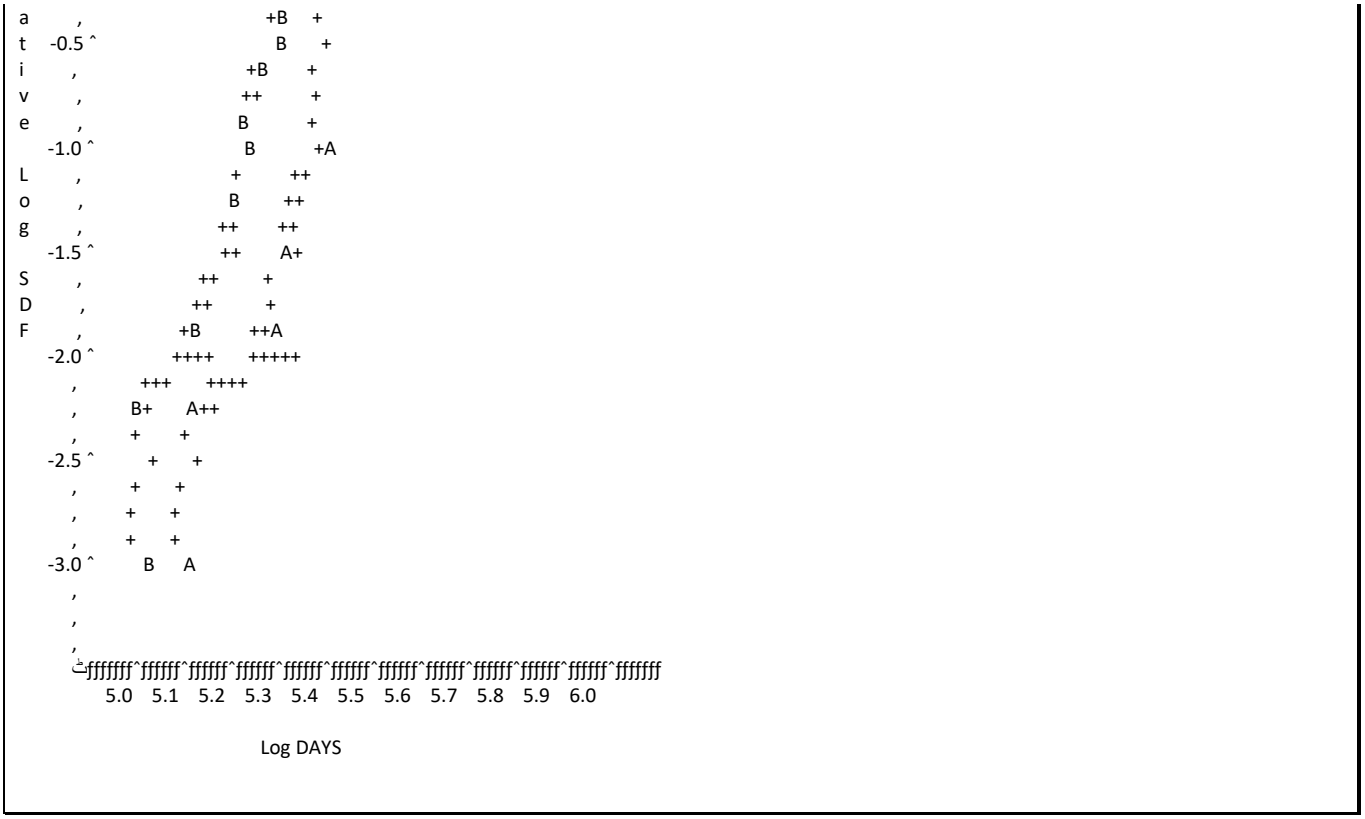


Dr. Firas



الشكل اعلاه يوضح العلاقة بين دالة التوزيع اللوغاريتمي وزمن البقاء للمعاملتين ويلاحظ ان كلا المنحنيين غير مستقيمان من نقطة الاصل ، لذا فأن النموذج الاسي غير ملائم لبيانات القدرة على الحياة.





الرسم البياني اعلاه يمثل العلاقة بين اللوغاريتيم السالب لدالة البقاء اللوغاريتمي وزمن البقاء للمعاملتين.

Test of Equality over Strata

Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	5.6485	1	0.0175
Wilcoxon	5.0312	1	0.0249
-2Log(LR)	0.1983	1	0.6561

الجدول اعلاه يوضح نتائج اختبار المقارنة بين منحني البقاء للمعاملتين اذ ان اختبارات الرتبة للتجانس تشير الى وجود اختلافات معنوية بين المعاملتين (p=0.0175 بالنسبة لـ log rank و p=0.0249 لـ Wilcoxon) وهذا يعني ان افراد المعاملة 1 عاشوا لفترة اطول معنويا من افراد المعاملة 2.

وبسبب ان النموذج الاسي غير ملائم لهذه البيانات كما اتضح سابقا لذا فأن اختبار Likelihood Ratio يجب ان يهمل.

لوفرضنا بأن جنس الفرد له علاقة بزمن البقاء لذا فأن من المهم معرفة تأثير المعاملة بعد التعديل لتأثير الجنس لذا يمكن تنفيذ ذلك كما يلي:

```
proc lifetest data=Exposed notable;
time Days*Status(0);
strata Sex;
test Treat;
run;
```

نلاحظ وجود كلمة notable حتى نتجنب تقدير منحنى البقاء لكل جنس ، كما نضع كلمة sex بعد strata.

The SAS System

Dr. Firas

The LIFETEST Procedure

Rank Tests for the Association of DAYS with Covariates
Pooled over Strata

Univariate Chi-Squares for the WILCOXON Test

Variable	Test Statistic	Standard Deviation	Chi-Square	Pr > Chi-Square
TREAT	-4.2372	1.7371	5.9503	0.0147

Univariate Chi-Squares for the LOG RANK Test

Variable	Test Statistic	Standard Deviation	Chi-Square	Pr > Chi-Square
TREAT	-6.8021	2.5419	7.1609	0.0075

يتبين من النتائج ان قيمة اختبار Wilicoxon بلغت p=0.0147 فيما بلغت قيمة اختبار Log rank 0.0076 ومقارنة بالنتائج قبل التعديل فأن تأثير المعاملة اصبح اعلى.

ولغرض التعرف على نسبة الخطورة يمكن استعمال الايعاز التالي:

```
proc phreg data=Exposed;
model days*status(0)=treat;
run;
```

ان نسبة الخطورة التي نحصل عليها من الايعاز اعلاه تكون صحيحة اذا لم تتغير خلال زمن التجربة اما اذا حصل فيها تغيير خلال زمن التجربة بمعنى ان تزيد في المراحل اللاحقة او تقل فهنا لابد من استعمال ايعاز اخر لغرض تقدير تأثير المتغير x وهو متغير تابع للزمن ويوجد ضمن برنامج SAS وكما يلي:

```
proc phreg data=Exposed;
    model days*status(0)=treat x;
    x=treat*(log(days));
  run;
```

11 – 2 تقدير الزمن في البيانات المراقبة

كان من المفترض ان تذكر هذه الطريقة في الفصل الخاص بالانحدار لكونها احدى طرق الانحدار التي يجري فيها مطابقة النموذج للبيانات الا انه وبسبب ان هذه الطريقة تخص البيانات المراقبة وتتبع توزيع Weibull فقد وجدنا من المناسب ان نذكرها هنا بعد ان يكون القارىء قد عرف معنى البيانات المراقبة وغير المراقبة في هذا الفصل والمثال التالي سيوضح لنا كيف نستعمل طريقة lifereg لمطابقة النموذج مع بيانات وقت الفشل المراقبة يمينا (right-censored failure time).

مثال (141) : نفذت تجربة لعقارين ضد الصداع على مجموعة من المرضى قسموا عشوائيا الى مجموعتين لغرض التعرف عاى افضلية اي من العقارين وقد سجل الوقت المستغرق

بالدقائق لكل مريض زال عنه الصداع ولكن بعض المرضى لم يشيروا الى زوال الصداع خلال مدة التجربة هؤلاء لايمكن تحديد موقفهم ولايفضل استبعادهم من التحليل لذا ستعد بياناتهم مراقبة.

البيانات تشمل متغير الوقت وهو الزمن المستغرق بالدقائق لزوال الصداع للمرضى ومتغير المجموعة ومتغير اخر ثنائي (1 او 0) يسمى مراقب اذ ان القيمة 0 تعني غير مراقب و1 تعني مراقب.

```
data headache;
input minutes group censor;
datalines;
11 1 0
12 1 0
19 1 0
19 1 0
19 1 0
19 1 0
21 1 0
20 1 0
21 1 0
21 1 0
20 1 0
21 1 0
20 1 0
21 1 0
25 1 0
27 1 0
30 1 0
21 1 1
24 1 1
14 2 0
16 2 0
16 2 0
21 2 0
21 2 0
23 2 0
23 2 0
23 2 0
23 2 0
25 2 1
23 2 0
24 2 0
24 2 0
26 2 1
32 2 1
30 2 1
30 2 0
32 2 1
20 2 1
;
proc lifereg;
class group;
model minutes*censor(1)=group;
```

Dr. FiraS

```
output out=new cdf=prob;
run;

proc sort data=new;
by prob;
proc plot data=new;
plot prob*minutes=group;
run;
```

الايعازالمستعمل اعلاه يتضمن الصف class وهو يشير الى المجموعة فيما يشير النموذج الى ان المتغير التابع هو الزمن وهو يعد مراقب اذا كانت قيمته تساوي 1. وتعمل هذه الطريقة على تحويل الزمن الى لوغاريتيم لتكون مطابقة لتوزيع Weibull. اما الايعاز output فهو لغرض استحداث بيانات جديدة وان الايعاز cdf يعني تقدير الاحتمالات بعد التعديل لتأثيرالزمن المراقب وغير المراقب لغرض رسمها على شكل رسم بياني.

```
                    The SAS System
                  Lifereg  Procedure
                 Class Level Information

                 Class    Levels    Values

                 GROUP       2    1 2

              Number of observations used = 38

                    The SAS System
                  Lifereg  Procedure

Data Set         =WORK.HEADACHE
Dependent Variable=Log(MINUTES)
Censoring Variable=CENSOR
Censoring Value(s)=     1
Noncensored Values=    30  Right Censored Values=     8
Left Censored Values=   0  Interval Censored Values=   0


Log Likelihood for WEIBULL  -9.37930239
                    The SAS System
                  Lifereg  Procedure

    Variable  DF   Estimate  Std Err ChiSquare  Pr>Chi Label/Value

    INTERCPT   1 3.30911843 0.058851    3161.7  0.0001 Intercept

    GROUP      1                      6.054016  0.0139
           1 -0.1933025 0.078563  6.054016  0.0139          1
           0       0      0      .  .              2

    SCALE      1 0.21218765 0.030356             Extreme value scale parameter
```

Dr. Firas

يتضح من النتائج اعلاه متوسط مدة زوال الصداع في كلا المجموعتين

Log(time) = 3.30911843 – 0.1933025 للمجموعةالاولى = exp(3.11581593)=22.55

Log(time) = 3.30911843 للمجموعةالثانية = exp(3.3091184) = 27.36

وتشير النتائج الى ان الفرق بين المجموعتين معنوي عند مستوى 5%.

The SAS System

Plot of PROB*MINUTES. Symbol is value of GROUP.

NOTE: 19 obs hidden.

الرسم اعلاه يوضح دالة التوزيع التجميعي لكل مجموعة.

Dr. Firas

الفصل الثاني عشر

تقدير LD50

هو مقياس للتعبير عن السمية ويمثل تركيز المادة التي تؤدي الى موت نصف المجموعة من الكائنات الحية.

مثال (142): البيانات ادناه تمثل تراكيز مختلفة من مادة كيمياوية ذات تأثير سمي (معبر عنها باللوغاريتيم) المطلوب تقدير قيمة LD50؟

يمكن ايجاد معادلة الانحدار اللوجستيكي حسب الايعاز التالي:

```
data N;
input logdose x N;
p=x/N;
cards;
2.079442 18 22
2.197225 17 34
2.302585 8 26
2.397895 3 25
2.484907 6 25
2.564949 0 23
proc probit data=n inversecl;← يستعمل هذا الايعاز لغرض الحصول على الاحتمالية
model x/N = logDose/dist=logistic;
output out=new1 p=p_hat;
run;
```

```
                The SAS System
                   Probit Procedure

Data Set        =WORK.N
Dependent Variable=X
Dependent Variable=N
Number of Observations=  6
Number of Events    =    52   Number of Trials =    155


Log Likelihood for LOGISTIC -77.10867963

                The SAS System
                Probit Procedure

   Variable  DF   Estimate  Std Err ChiSquare  Pr>Chi Label/Value

   INTERCPT   1 18.2208319 3.339292  29.77334  0.0001 Intercept
   LOGDOSE    1 -8.2180463 1.465612  31.44121  0.0001

       Probit Model in Terms of Tolerance Distribution

                    MU       SIGMA
                 2.217173    0.121683


       Estimated Covariance Matrix for Tolerance Parameters
```

	MU	SIGMA
MU	0.000686	-0.000222
SIGMA	-0.000222	0.000471

The SAS System

Probit Procedure
Probit Analysis on LOGDOSE

Probability	LOGDOSE	95 Percent Fiducial Limits	
		Lower	Upper
0.01	2.77632	2.64005	3.05232
0.02	2.69074	2.57540	2.92199
0.03	2.64016	2.53693	2.84520
0.04	2.60389	2.50918	2.79032
0.05	2.57546	2.48730	2.74743
0.06	2.55199	2.46912	2.71213
0.07	2.53193	2.45349	2.68206
0.08	2.51437	2.43971	2.65582
0.09	2.49870	2.42734	2.63251
0.10	2.48454	2.41608	2.61150
0.15	2.42825	2.37029	2.52902
0.20	2.38586	2.33418	2.46858
0.25	2.35086	2.30257	2.42043
0.30	2.32028	2.27300	2.38032
0.35	2.29250	2.24405	2.34599
0.40	2.26651	2.21485	2.31597
0.45	2.24159	2.18490	2.28915
0.50	2.21717	2.15384	2.26458
0.55	2.19275	2.12136	2.24142
0.60	2.16783	2.08707	2.21893
0.65	2.14185	2.05037	2.19642
0.70	2.11407	2.01037	2.17313
0.75	2.08349	1.96567	2.14816
0.80	2.04848	1.91389	2.12019
0.85	2.00610	1.85060	2.08691
0.90	1.94981	1.76587	2.04338
0.91	1.93564	1.74446	2.03252
0.92	1.91998	1.72076	2.02053
0.93	1.90242	1.69415	2.00713
0.94	1.88236	1.66372	1.99186
0.95	1.85888	1.62806	1.97403
0.96	1.83046	1.58482	1.95251
0.97	1.79419	1.52958	1.92512
0.98	1.74360	1.45241	1.88703
0.99	1.65802	1.32165	1.82281

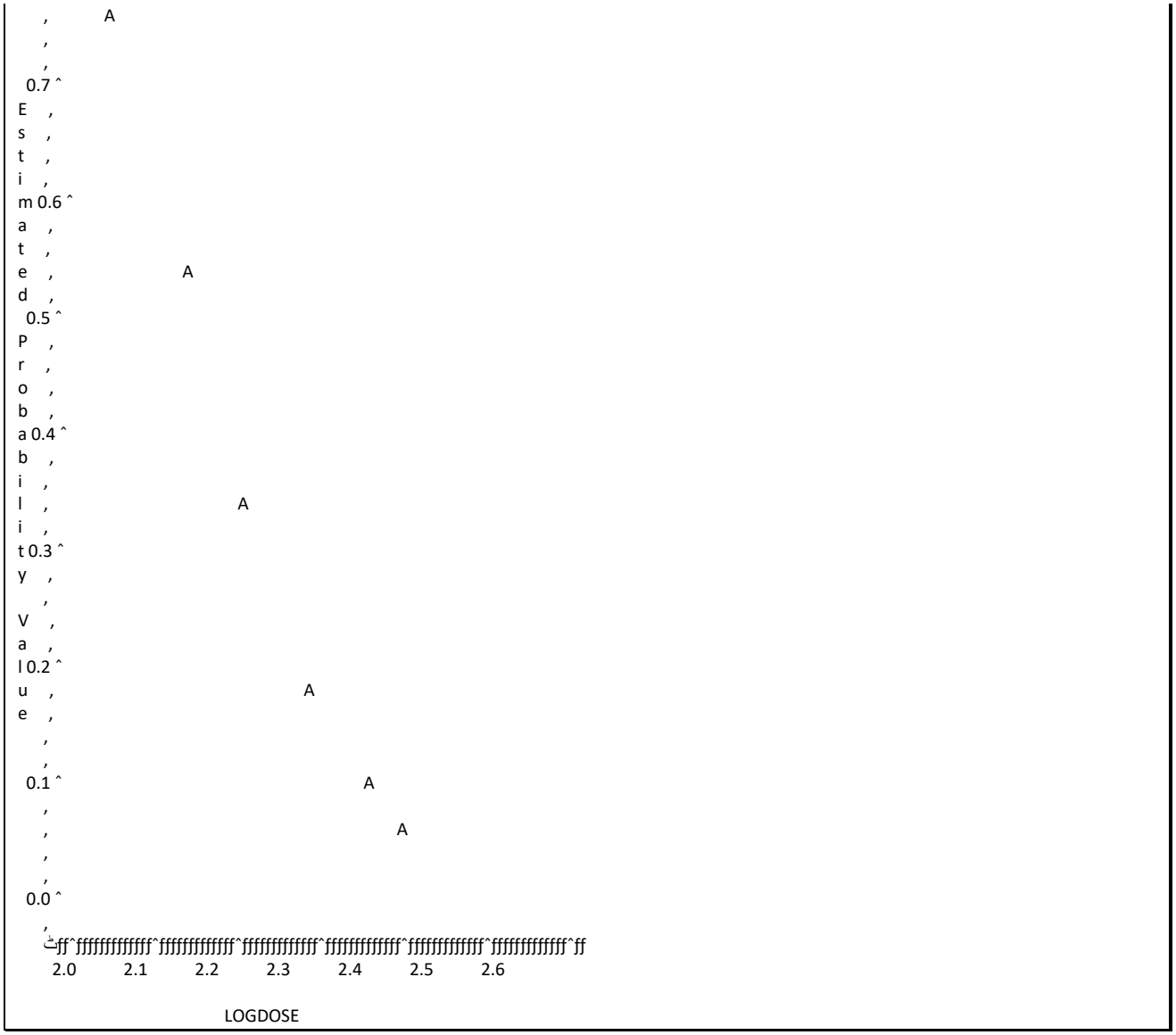
ولغرض تقدير قيمة LD50 نعتمد المعادلة الآتية:

$$\log \hat{} LD50 = - [18.22 / (-8.22) = - (2.22)]$$

ولغرض رسم العلاقة بين لوغاريتيم الجرعة والاحتمالية نستعمل الايعاز الاتي:

```
proc plot data=new1;
plot p_hat *logdose;
run;
```

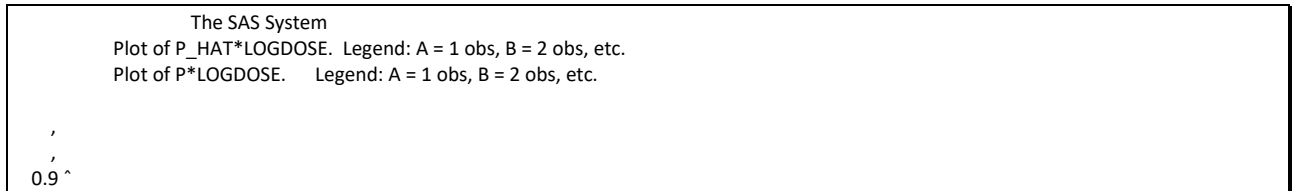
```
  ,
0.8 ^
  ,
```

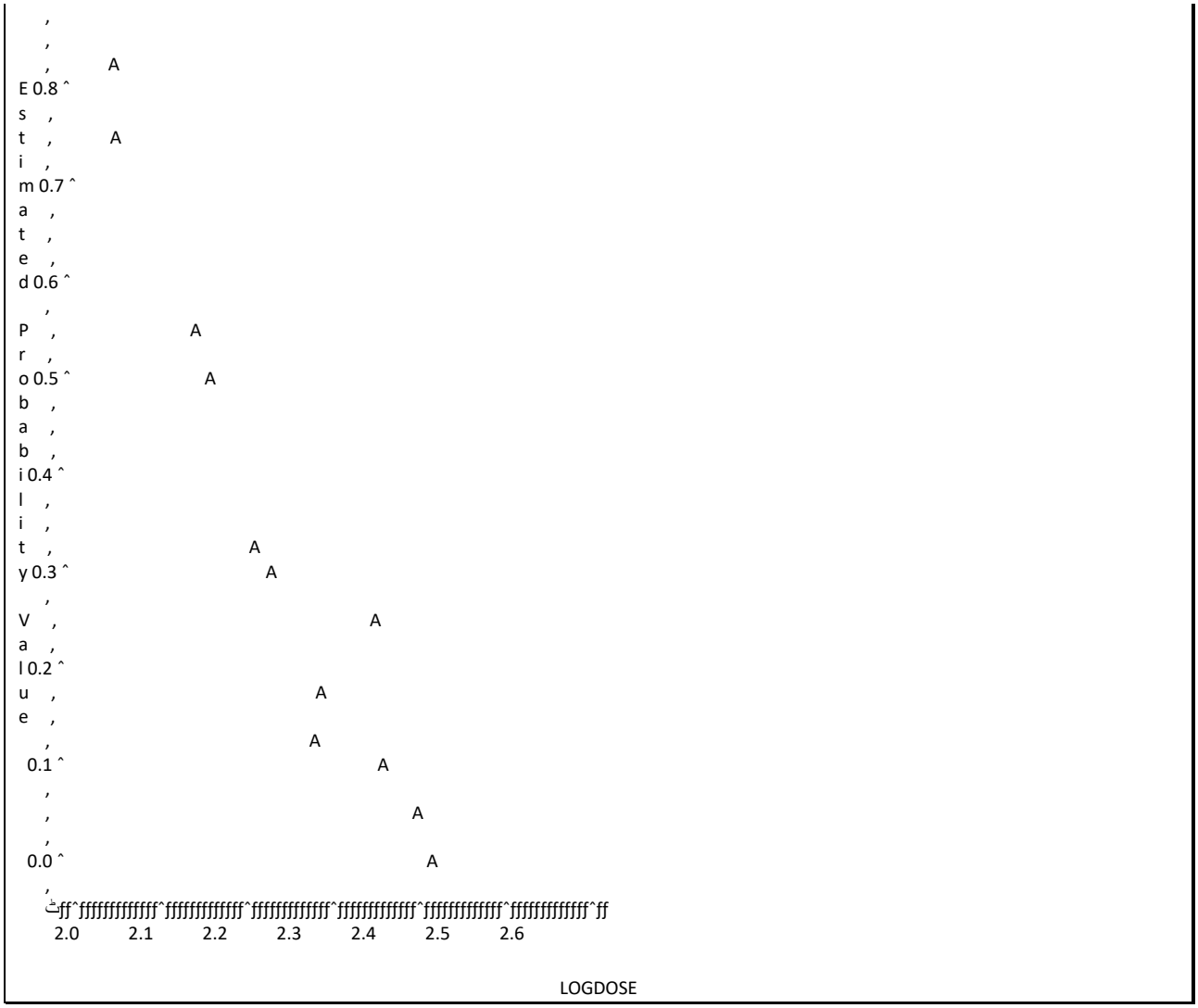


الرسم اعلاه يوضح العلاقة بين لوغاريتيم الجرعة والاحتمالية المحسوبة لنسبة الموت الى العدد الكلي.

ولغرض رسم العلاقة بين الجرعة والقيم المشاهدة والمتوقعة لنسبة الافراد الهالكة نستعمل الايعاز:

```
proc plot data=new1;
plot (p_hat p) *logdose/overlay; ← يستعمل هذا الايعاز لرسم لوغاريتيم الجرعة مع عدة عوامل
run;
```





الرسم يوضح العلاقة بين لوغاريتيم الجرعة والاحتمالية المحسوبة والمتوقعة للتعرف على مدى مطابقة النموذج للبيانات.

كما يمكن تقدير LD50 بأستعمال نموذج probit ويكون شكل المنحني مقارب لشكله في الانحدار اللوجستيكي الا ان معادلة الانحدار مختلفة.

```
data N;
input logdose x N ;
p=x/N;
cards;
2.079442 18 22
2.197225 17 34
2.302585 8 26
2.397895 3 25
2.484907 6 25
2.564949 0 23
proc probit data=n inversecl;
model x/N = logdose;
output out=new2 p=p_hat;
run;

proc plot data=new1;
```

```
plot p_hat *logdose;
run;
proc plot data=new2;
plot (p_hat p) *logdose/overlay;
run;
```

The SAS System
Probit Procedure

Data Set =WORK.N
Dependent Variable=X
Dependent Variable=N
Number of Observations= 6
Number of Events = 52 Number of Trials = 155

Log Likelihood for NORMAL -77.14780518

The SAS System
Probit Procedure

Variable	DF	Estimate	Std Err	ChiSquare	Pr>Chi	Label/Value
INTERCPT	1	**10.7380966**	1.850525	33.67168	0.0001	Intercept
LOGDOSE	1	-4.8389147	0.805523	36.08613	0.0001	

Probit Model in Terms of Tolerance Distribution

MU	SIGMA
2.219113	0.206658

Estimated Covariance Matrix for Tolerance Parameters

	MU	SIGMA
MU	0.000724	-0.000422
SIGMA	-0.000422	0.001183

The SAS System

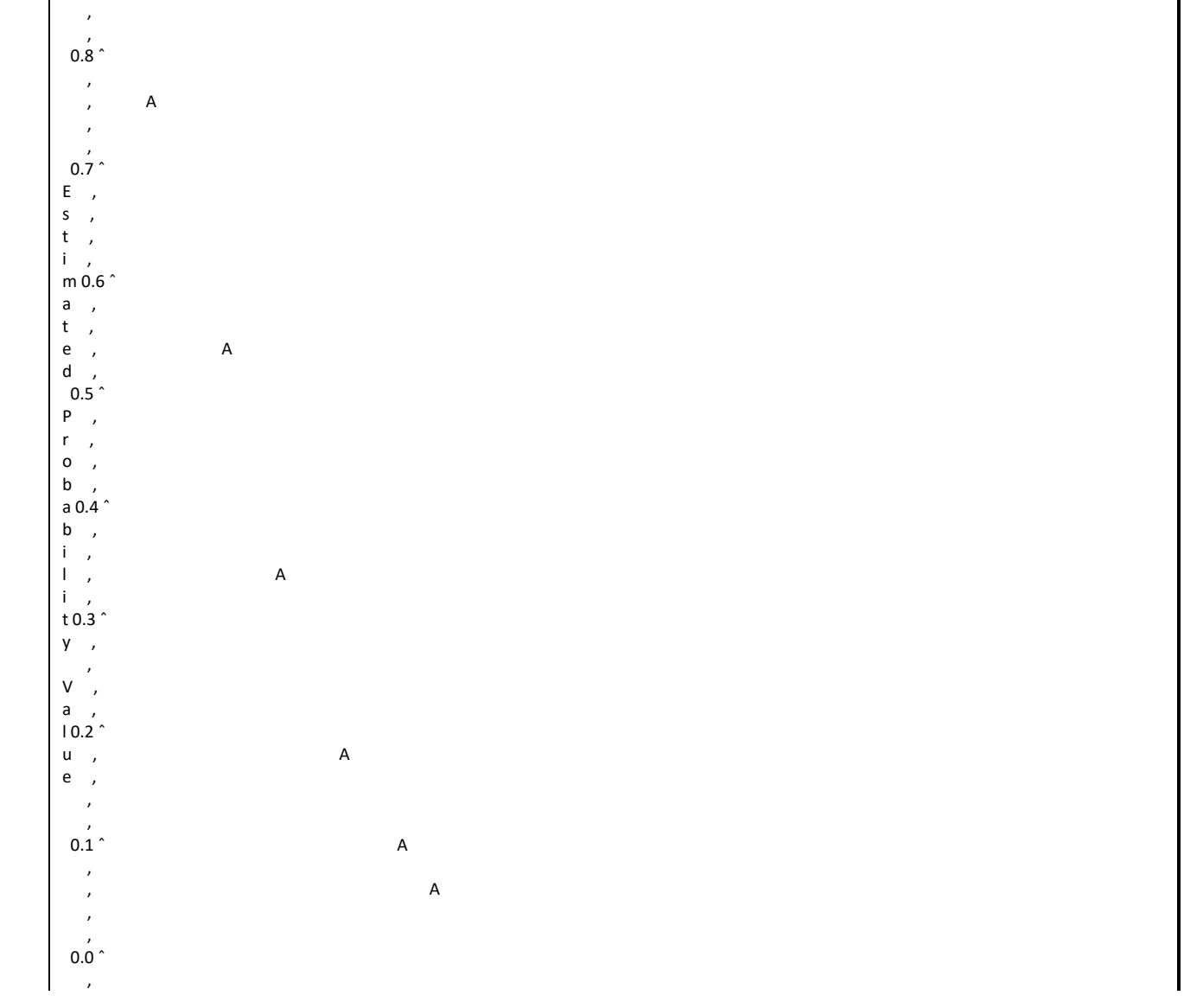
Probit Procedure
Probit Analysis on LOGDOSE

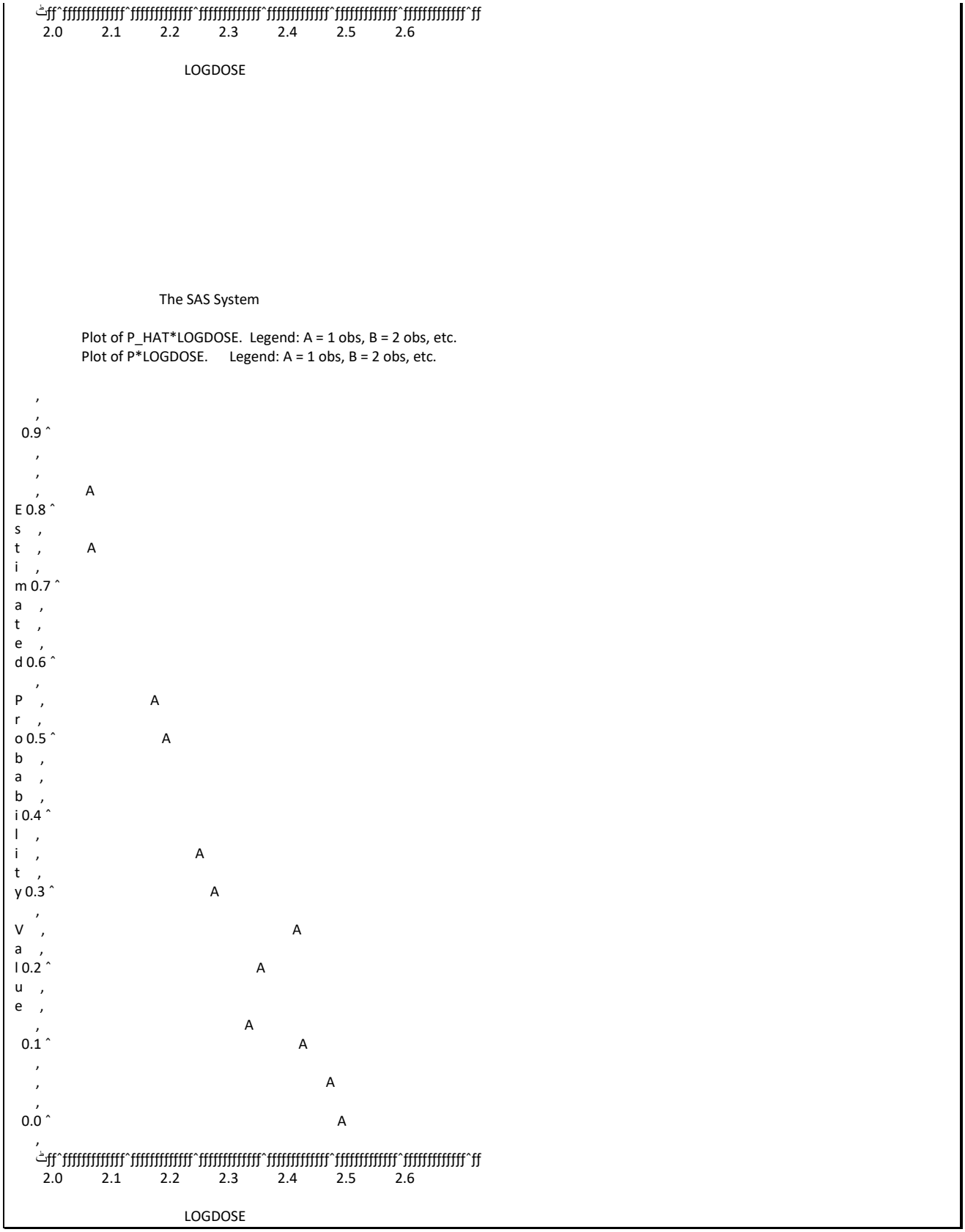
Probability	LOGDOSE	95 Percent Fiducial Limits	
		Lower	Upper
0.01	2.69987	2.59165	2.90509
0.02	2.64354	2.54796	2.82269
0.03	2.60779	2.52004	2.77060
0.04	2.58091	2.49890	2.73156
0.05	2.55903	2.48159	2.69992
0.06	2.54042	2.46677	2.67308
0.07	2.52410	2.45368	2.64963
0.08	2.50948	2.44189	2.62871
0.09	2.49619	2.43109	2.60976
0.10	2.48396	2.42107	2.59239
0.15	2.43330	2.37864	2.52144
0.20	2.39304	2.34330	2.46667
0.25	2.35850	2.31119	2.42147
0.30	2.32748	2.28038	2.38286
0.35	2.29874	2.24975	2.34915
0.40	2.27147	2.21867	2.31918
0.45	2.24508	2.18678	2.29201
0.50	2.21911	2.15387	2.26680
0.55	2.19314	2.11972	2.24283
0.60	2.16676	2.08404	2.21944
0.65	2.13948	2.04639	2.19604
0.70	2.11074	2.00608	2.17202

0.75	2.07972	1.96205	2.14663
0.80	2.04518	1.91254	2.11883
0.85	2.00493	1.85438	2.08688
0.90	1.95427	1.78073	2.04715
0.91	1.94203	1.76288	2.03761
0.92	1.92874	1.74347	2.02727
0.93	1.91413	1.72210	2.01593
0.94	1.89781	1.69821	2.00328
0.95	1.87919	1.67094	1.98889
0.96	1.85732	1.63886	1.97202
0.97	1.83043	1.59937	1.95133
0.98	1.79469	1.54681	1.92388
0.99	1.73835	1.46385	1.88075

The SAS System

Plot of P_HAT*LOGDOSE. Legend: A = 1 obs, B = 2 obs, etc.





يمكن تقدير قيمة LD50 بأستعمال المعادلة الآتية:

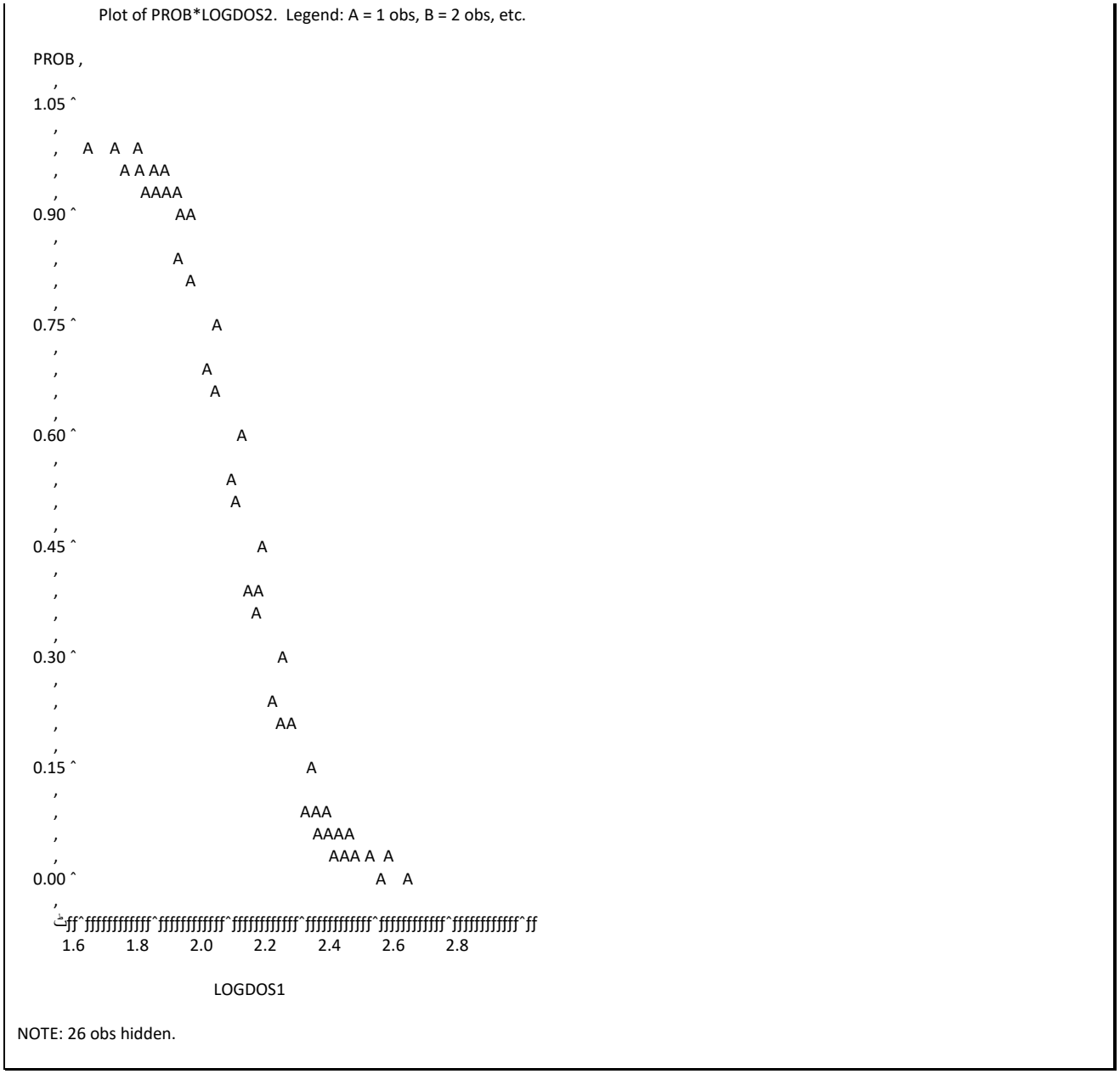
$$\log\hat{}LD50 = -\left[10.738/(-4.838) = -(2.22)\right]$$

يمكن مقارنة مدى مطابقة الطريقتين بأستعمال الايعاز الآتي:

```
data d;
 input OBS   prob logdos1 logdos2;
 cards;

 1    0.01    2.77632   2.69987
 2    0.02    2.69074   2.64354
 3    0.03    2.64016   2.60779
 4    0.04    2.60389   2.58091
 5    0.05    2.57546   2.55903
 6    0.06    2.55199   2.54042
 7    0.07    2.53193   2.52410
 8    0.08    2.51437   2.50948
 9    0.09    2.49870   2.49619
10    0.10    2.48454   2.48396
11    0.15    2.42825   2.43330
12    0.20    2.38586   2.39304
13    0.25    2.35086   2.35850
14    0.30    2.32028   2.32748
15    0.35    2.29250   2.29874
16    0.40    2.26651   2.27147
17    0.45    2.24159   2.24508
18    0.50    2.21717   2.21911
19    0.55    2.19275   2.19314
20    0.60    2.16783   2.16676
21    0.65    2.14185   2.13948
22    0.70    2.11407   2.11074
23    0.75    2.08349   2.07972
24    0.80    2.04848   2.04518
25    0.85    2.00610   2.00493
26    0.90    1.94981   1.95427
27    0.91    1.93564   1.94203
28    0.92    1.91998   1.92874
29    0.93    1.90242   1.91413
30    0.94    1.88236   1.89781
31    0.95    1.85888   1.87919
32    0.96    1.83046   1.85732
33    0.97    1.79419   1.83043
34    0.98    1.74360   1.79469
35    0.99    1.65802   1.73835

proc plot;
plot prob* (logdos1 logdos2)/overlay;
run;
```

The SAS System

Plot of PROB*LOGDOS1. Legend: A = 1 obs, B = 2 obs, etc.



يلاحظ بأن الطريقتين كانت متقاربة في تقدير احتمالية نسبة الموت.

امثلة متنوعة

ان اكتساب مهارة التحليل الاحصائي تستلزم اجراء المزيد من التحليلات اذ ان ذلك سيساهم في زيادة امكانية الباحث في اختيار النموذج والاختبار المناسب ، ولاظير في ان يعمل على تنظيم بياناته بالاعتماد على الامثلة.

مثال (143): قدر ارجحية حصول تضخم البروستات بسبب بعض العوامل المذكورة في قطعة البرنامج ادناه بأستعمال عدة طرق؟

```
data prostate;
  input case age acid xray size grade nodalinv @@;
  lacd=log(acid);
  datalines;
1  66  .48 0 0 0 0  2  68  .56 0 0 0 0  3  66  .50 0 0 0 0
4  56  .52 0 0 0 0  5  58  .50 0 0 0 0  6  60  .49 0 0 0 0
7  65  .46 1 0 0 0  8  60  .62 1 0 0 0  9  50  .56 0 0 1 1
10 49  .55 1 0 0 0  11 61  .62 0 0 0 0  12 58  .71 0 0 0 0
13 51  .65 0 0 0 0  14 67  .67 1 0 1 1  15 67  .47 0 0 1 0
16 51  .49 0 0 0 0  17 56  .50 0 0 1 0  18 60  .78 0 0 0 0
19 52  .83 0 0 0 0  20 56  .98 0 0 0 0  21 67  .52 0 0 0 0
22 63  .75 0 0 0 0  23 59  .99 0 0 1 1  24 64 1.87 0 0 0 0
25 61 1.36 1 0 0 1  26 56  .82 0 0 0 1  27 64  .40 0 1 1 0
28 61  .50 0 1 0 0  29 64  .50 0 1 1 0  30 63  .40 0 1 0 0
31 52  .55 0 1 1 0  32 66  .59 0 1 1 0  33 58  .48 1 1 0 1
34 57  .51 1 1 1 1  35 65  .49 0 1 0 1  36 65  .48 0 1 1 0
37 59  .63 1 1 1 0  38 61 1.02 0 1 0 0  39 53  .76 0 1 0 0
40 67  .95 0 1 0 0  41 53  .66 0 1 1 0  42 65  .84 1 1 1 1
43 50  .81 1 1 1 1  44 60  .76 1 1 1 1  45 45  .70 0 1 1 1
46 56  .78 1 1 1 1  47 46  .70 0 1 0 1  48 67  .67 0 1 0 1
49 63  .82 0 1 0 1  50 57  .67 0 1 1 1  51 51  .72 1 1 0 1
52 64  .89 1 1 0 1  53 68 1.26 1 1 1 1
;
 proc logistic data=prostate descending;
  model nodalinv=lacd xray size age grade;
run;

proc logistic data=prostate descending;
  model nodalinv=age lacd xray size grade / selection=stepwise
                                start=5;
 run;

 proc logistic data=prostate descending;
  model nodalinv = age lacd xray size grade / selection=score
                               best=2;
run;
```

مثال (144): قدر ارجحية الاستجابة لكل من الفايبرين والكلوبيولين ؟

```
data esr;
  input id fibrin globulin response @@;
  cards;
1  2.52 38 0   2  2.56 31 0   3  2.19 33 0   4  2.18 31 0
5  3.41 37 0   6  2.46 36 0   7  3.22 38 0   8  2.21 37 0
9  3.15 39 0   10 2.60 41 0   11 2.29 36 0   12 2.35 29 0
13 5.06 37 1   14 3.34 32 1   15 2.38 37 1   16 3.15 36 0
```

```
17 3.53 46 1   18 2.68 34 0   19 2.60 38 0   20 2.23 37 0
21 2.88 30 0   22 2.65 46 0   23 2.09 44 1   24 2.28 36 0
25 2.67 39 0   26 2.29 31 0   27 2.15 31 0   28 2.54 28 0
29 3.93 32 1   30 3.34 30 0   31 2.99 36 0   32 3.32 35 0
;
proc logistic data=esr descending;
  model response=fibrin globulin;
run;
```

مثال (145): في دراسة عن العلاقة بين الاصابة بضغط الدم وقدرة الانسان على العيش جمعت بيانات عن الذين لديهم ضغط دم منخفض وعالي ، اختبر العلاقة بين العاملين (عاملان مستقلان ام يؤثر الاول على الثاني) وقدر الخطورة النسبية ونسبة الارجحية؟

مستوى ضغط الدم	عدد الافراد المتوفون	عدد الاشخاص على قيد الحياة
واطيء	21	2655
عالي	55	3283

```
DATA e;
 INPUT bpr  status  freq;
cards;
0 0 55
0 1 3238
1 0 21
1 1 2655
 proc freq data=e;
 tables bpr*status/chisq;
weight freq;
run;
proc freq;
 weight freq;
 tables bpr*status/cl riskdiff relrisk rdiff;
 run;
```

مثال (146):قدر متوسط طول العمر بالاسابيع لمجموعة من المرضى اعتمادا على متوسط الاحياء والاموات؟

```
data melanoma;
 input weeks status$;
 censor=status='alive';
cards;
12.8   dead
15.6   dead
24.0   alive
26.4   dead
29.2   dead
30.8   alive
39.2   dead
42.0   dead
58.4   alive
72.0   alive
77.2   dead
82.4   dead
87.2   alive
```

Dr.Firas

```
94.4    alive
97.2    alive
106.0   alive
114.8   alive
117.2   alive
140.0   alive
168.0   alive
proc lifetest data=melanoma plots=(s);
 time weeks*censor(0);
run;
```

مثال (147): في دراسة لمعرفة تأثير نوع من العقار على الاستجابة بين الرجال والنساء اختبر تأثير العقار؟

```
data arth;
  input gender $ treat $ response $ count @@;
  cards;
female test    none 6  female test    some 5  female test    marked 16
female placebo none 19 female placebo some 7  female placebo marked 6
male   test    none 7  male   test    some 2  male   test    marked 5
male   placebo none 10 male   placebo some 0  male   placebo marked 1
;
proc freq data=arth order=data;
  weight count;
  tables treat*response / chisq nocol nopct;
run;
proc freq data=arth order=data;
  weight count;
  tables gender*treat*response / cmh nocol nopct;
run;
```

مثال (148): اختبر العلاقة بين العقار المستعمل والاستجابة وهل تختلف المجاميع في الاستجابة؟

```
data pain;
  input dgnosis $ treatmnt $ response $ count @@;
  cards;
  I placebo  no 26 I  placebo yes  6
  I dosage1  no 26 I  dosage1 yes  7
  I dosage2  no 23 I  dosage2 yes  9
  I dosage3  no 18 I  dosage3 yes 14
  I dosage4  no  9 I  dosage4 yes 23
  II placebo no 26 II placebo yes  6
  II dosage1 no 12 II dosage1 yes 20
  II dosage2 no 13 II dosage2 yes 20
  II dosage3 no  1 II dosage3 yes 31
  II dosage4 no  1 II dosage4 yes 31
proc freq order=data;
  weight count;
  tables treatmnt*response /chisq;
  tables dgnosis*treatmnt*response/
         chisq cmh;
  tables dgnosis*treatmnt*response/
         scores=modridit cmh;
run;
```

Dr. Firas

مثال (149): اختبر العلاقة بين تركيز الجرعة والاستجابة لمجموعة من المرضى؟

```
data pain;
  input treatmnt $ hours count @@;
  cards;
  placebo  0 6 placebo  1  9 placebo   2 6 placebo  3 3 placebo  4 1
  standard 0 1 standard 1  4 standard  2 6 standard 3 6 standard 4 8
  test    0 2 test    1  5 test     2 6 test    3 8 test    4 6
proc freq;
  weight count;
  tables treatmnt*hours/ cmh nocol nopct;
run;
```

مثال (150): مجموعة من الطلبة اعطيت سؤالين لغرض الاجابة عنهما وكان عدد الذين اجابوا عن السؤال الاول 172 وعدد الذين لم يجيبوا عن السؤال 6 فيما كانت الاعداد بالنسبة للسؤال الثاني 160 و 18. اختبر الاختلافات في الاختبارين؟

بما ان الاختبار اجري على نفس الطلاب فيفضل استعمال اختبار ماكنيمار.

```
data set1;
 input Q1 Q2 students;
 datalines;
 1 1 172
 0 1 6
 1 0 160
 0 0 18
run;
proc freq data=set1;
 table Q1*Q2;
 exact mcnem;
 weight students;
run;
```

مثال (151): اجريت تجربة على دجاج اللحم شملت ثلاثة معاملات لدراسة تأثير اضافة علفية على (1) معدل وزن الافراخ لكل اسبوع ولغاية الاسبوع السادس (2) معدل الوزن العام (3) معدل الزيادة الوزنية اليومية (4) المعدل العام للزيادة الوزنية اليومية ؟

وتمثل trt المعاملات و obs المشاهدات في كل معاملة وان x1 – x6 تمثل وزن الافراخ من الاسبوع الاول لغاية السادس xx الهلاكات (1= حي ، 0= هلاك) و w وزن الافراخ بعمر يوم.

```
data s;
input trt obs x1 x2 x3 x4 x5 x6 xx w;
wt1=x1-w;
wt2=x2-x1;
wt3=x3-x2;
wt4=x4-x3;
wt5=x5-x4;
wt6=x6-x5;
TWT=wt1+wt2+wt3+wt4+wt5+wt6;
```

Dr. Firas

```
DG1=wt1/7;
DG2=wt2/7;
DG3=wt3/7;
DG4=wt4/7;
DG5=wt5/7;
DG6=wt6/7;
TDG=(wt1+wt2+wt3+wt4+wt5+wt6)/42;
xr=(1-xx)*100;
cards;
1 1  140 374 900  1110 1800 2100 1 44
1 2  145 376 990  1110 1800 2200 1 43
1 3  130 287 950  1150 1750 2200 1 43
1 4  160 288 590  1010 1800 .    0 44
1 5  165 470 800  1100 2000 2100 1 42
2 1  166 387 725  1483 2050 2700 1 41
2 2  167 388 850  1483 2050 2450 1 43
2 3  168 420 800  1484 2150 2550 0 44
2 4  166 400 850  1458 .    .    0 43
2 5  166 450 850  1458 2280 2700 1 43
3 1  150 360 900  1350 1850 2700 1 42
3 2  150 360 900  1350 1850 2700 1 43
3 3  150 400 950  1350 1850 .    0 44
3 4  174 400 650  1406 .    .    0 42
3 5  173 410 850  1407 1800 .    0 42
;proc GLM;classes trt;
model w wt1 wt2 wt3 wt4 wt5 wt6 twt dg1 dg2 dg3 dg4 dg5 dg6 tdg x1 x2 x3 x4 x5 x6 xr =trt;
means trt/LSD DUNCAN;
proc means mean n stderr  std;by trt;var w wt1 wt2 wt3 wt4 wt5 wt6 twt dg1 dg2 dg3 dg4 dg5 dg6
tdg x1 x2 x3 x4 x5 x6 xr;
run;
```

في تجارب الدواجن من الضروري ان تكون هناك مكررات لغرض امكانية اجراء التحليل الاحصائي بصورة صحيحة اذ ان وجود المكررات يعني وجود تباين وهو الاساس الذي يقوم عليه التحليل الاحصائي لاننا في الاحصاء نبحث على التباينات لدراستها وغياب التباين ينفي الحاجة الى الاحصاء.

كيف نحسب كمية العلف الاسبوعي المستهلك:

لو فرضنا ان افراخ استهلكت 1500 غرام خلال اسبوع وخلال الاسبوع حدثت هلاكات لثلاثة افراخ في اليوم الثاني هلاك واحد وفي اليوم الرابع هلاكان كيف تقدر كمية العلف المستهلك لكل فرخ؟

مجموع الايام التي تغذت فيه الطيور الحية والهالكة= عدد الطيور الحية نهاية الاسبوع + (عدد الطيور الهالكة × عدد الايام التي تغذت فيها) وفي مثالنا:

$$14 =(3\times2 + 1\times1) + 7)$$

لو فرضنا ان الافراخ لم يحصل فيها هلاك فأن العشرة افراخ تتغذى كل يوم من الاسبوع فيكون:

$7\times10= 70$ فسيكون معدل استهلاك الفرخ الواحد:

Dr Firas

70/1500= 21.42 غرام وهو تقدير غير صحيح والتقدير الصحيح هو

70 – 14 = 56

56/1500=26.78 وللتوضيح اكثر نرتب ايام الاسبوع وعدد الطيور الحية:

1	10
2	9
3	9
4	7
5	7
6	7
7	7
	56

هنا يتضح اهمية المكررات لكل معاملة اذ ان كل معاملة تعطينا تقدير واحد فيما نحتاج على الاقل تقديرين وهذا يعني عدم امكانية اجراء التحليل الاحصائي الصحيح.

ملاحظة : من المفيد التمييز بين طريقتين لتقدير كفاءة التحويل الغذائي واعتماد الطريقة الاكثر دقة لاسيما وان الطريقتان تستعملان في العديد من الرسائل والاطاريح. ولغرض اعطاء صورة واضحة عن ذلك نفترض المثال الآتي:

البيانات ادناه تمثل كمية العلف المستهلك (غم) للاسبوع الاول والثاني والثالث اضافة الى مقدار الزيادة الوزنية لكل اسبوع لعدد من الافراخ المطلوب تقدير كفاءة التحويل الغذائي الاسبوعي والكلي؟

	الاسبوع الاول	الاسبوع الثاني	الاسبوع الثالث	الكلي
العلف المستهلك	400	900	1200	2500
الزيادةالوزنية	100	300	600	1000
كفاءة التحويل	4=100/400	3=300/900	2=600/1200	2.5 3

الطريقة الاولى لتقدير كفاءة التحويل الغذائي الكلية = (4 + 3 + 2) / 3 = 3 والذي يمثل كفاءة التحويل الغذائي لمعدل الكفاءات خلال الاسابيع الثلاثة (الكفاءة الكلية).

الطريقة الثانية = 1000/2500 = 2.5

Dr. Firas

تقدير الطريقة الثانية هو الاصح لانه يمثل القيمة الحقيقية ونلاحظ ان القيمة لايوجد لها تباين ولاانحراف قياسي ولا خطأ قياسي اما الطريقة الثانية فهي تمثل تقدير غير متحيز للتقدير الحقيقي اذ ان لها تباين وانحراف قياسي وخطأ قياسي وهذا يعني ان التقدير يكون اكثر دقة كلما زاد عدد الاسابيع اذ ان حجم العينة كلمل زاد ازدادت الدقة (انخفاض الخطأ القياسي).

مثال (152): في دراسة شملت 9 مصابين بالسكر و 8 اشخاص غير مصابين حيث قدرضغط الدم الانقباضي Systolic blood pressure في كليهما وكانت النتائج كما موضح في قطعة البرنامج اختبر الاختلافات بين المجموعتين؟

```
data react;
input trt response;
cards;
1 114
1 120
1 120
1 128
1 130
1 135
1 138
1 140
1 141
2 110
2 112
2 112
2 118
2 120
2 122
2 125
2 130
;option nodate nonumber;
proc npar1way wilcoxon data=react;
class trt;
var response;
run;
```

مثال (153): اخذت ثلاث عينات من الطلاب من ثلاثة صفوف وكانت درجاتهم في نفس الاختبار كما موضح ادناه فهل توجد اختلافات بين المجموعات؟

```
data k;
 input group score @@;
 cards;
1 55  1 0  1 1  1 0  1 50  1 60  1 44
2 73  2 85  2 51  2 63  2 85  2 81
2 66  2 69  3 61  3 54  3 80  3 47
proc npar1way wilcoxon;
 class group;
 var score;
 run;
```

Dr. Firas

مثـال (154): فـي دراسـة شـملت 59 مـريض خضـعوا الـى برنـامج علاجـي اذ قسـموا الـى مجموعتين اعطيت الاولى عقار والثانية اعطيت placebo لازالـة التـأثير النفسـي وقد عبر عن الاستجابة بخمسة مستويات اختبر الفرق بين استجابة المجموعتين؟

```
data Arthriti;
input Treatmen $ Response Freq @@;
datalines;
Active 5 5 Active 4 11 Active 3 5 Active 2 1 Active 1 5
Placebo 5 2 Placebo 4 4 Placebo 3 7 Placebo 2 7 Placebo 1 12
data B;
set arthriti;
do i=1 to freq;
output B;
end;
run;
proc npar1way wilcoxon median edf data=B;
class treatmen;
var response;
exact;
run;
```

ملاحظة: ان العبارة do i=1 to freq هو لغرض اعلام البرنـامج بأننـا نريد ان نكرر كل قيم treatmen و Response بقيمـة Freq اي ان اول قيمـة (5 Active) سـتكرر 5 مـرات وان القيمة (4 Active) تكرر 11 مرة ولغاية القيمة (1 Placebo) تكرر 12 مرة.

بعد التنفيذ سنجد ان قيمة الاحتمالية p المقدرة بـ z ستكون 0.0032 على جانبي التوزيع وعند تقسيمها على 2 سنحصل على 0.0016 وهذا يعني رفض نظرية العدم بخصوص تساوي الاستجابة بين المجموعتين.

وفي اختبار الوسيط لعينتين سنجد الاحتمالية p المقدرة بـ z ستكون 0.0011 على جانبي التوزيع وعن تقسيمها على 2 سنحصل على 0.0005 وهذا يعني رفض نظرية العدم بخصوص تساوي الاستجابة بين المجموعتين.

وفي اختبار كولموجروف وسيرنوف سنجد ان احتمالية التماثل بين المجموعتين 0.0164 مما يعني رفض فرضية العدم بخصوص تطابق توزيع الاستجابة لكلا المجموعتين.

مثال (155): في دراسة شملت مجموعة من المرضى (رجال ونساء) قسموا الى مجموعتين واعطيت مجموعة السيطرة Placebo فيما اعطيت المجموعة الثانية عقار جديد وسجلت حالات المرضى حسب التصنيف تحسن ، تحسن قليل ، ولاتغير، المطلوب اختبار هل ان للعقار تأثير على حالة المريض؟

```
data a;
input sex $ treat $ improve $ count @@ ;

cards ;
female active marked 16 female active some 5 female active none 6
```

Dr. Firas

```
female placebo marked 6 female placebo some 7 female placebo none 19
male active marked 5 male active some 2 male active none 7
male placebo marked 1 male placebo some 0 male placebo none 10
proc freq order=data;
weight count;
tables sex*treat*improve / nocol nopct chisq exact;
run;
```

وعند تنفيذ البرنامج سنجد ان تأثير العقار كان معنويا في النساء فقط ولكن عند ازالة تأثير الجنس من قطعة البرنامج سنجد ان قيمة مربع كاي تشير الى ان تأثير العقار معنوي.

مثال (156): في تجربة لدراسة تأثير خمسة مستويات من الكوسيبول (مكون اساسي لغلاف بذور القطن) على الزيادة الوزنية في مجموعة من الحيوانات اختبر تأثير الكوسيبول؟

```
data Gossypol;
input Dose n;
do i=1 to n;
input Gain @@;
output;
end;
datalines;
0 16
228 229 218 216 224 208 235 229 233 219 224 220 232 200 208 232
.04 11
186 229 220 208 228 198 222 273 216 198 213
.07 12
179 193 183 180 143 204 114 188 178 134 208 196
.10 17
130 87 135 116 118 165 151 59 126 64 78 94 150 160 122 110 178
.13 11
154 130 130 118 118 104 112 134 98 100 104
proc npar1way data=Gossypol;
class Dose;
var Gain;
                                                    run;
```

يمكن ترتيب البيانات بالشكل التالي:

```
data Gossypol;
input dose rep Gain ;
cards;
0 1 228
0 2 229
0 3 218
0 4 216
0 5 224
0 6 208
0 7 235
0 8 229
0 9 233
0 10 219
0 11 224
0 12 220
0 13 232
0 14 200
```

Dr. Firas

0 15 208
0 16 232
.04 1 186
.04 2 229
.04 3 220
.04 4 208
.04 5 228
.04 6 198
.04 7 222
.04 8 273
.04 9 216
.04 10 198
.04 11 213
.07 1 179
.07 2 193
.07 3 183
.07 4 180
.07 5 143
.07 6 204
.07 7 114
.07 8 188
.07 9 178
.07 10 134
.07 11 208
.07 12 196
.10 1 130
.10 2 87
.10 3 135
.10 4 116
.10 5 118
.10 6 165
.10 7 151
.10 8 59
.10 9 126
.10 10 64
.10 11 78
.10 12 94
.10 13 150
.10 14 160
.10 15 122
.10 16 110
.10 17 178
.13 1 154
.13 2 130
.13 3 130
.13 4 118
.13 5 118
.13 6 104
.13 7 112
.13 8 134
.13 9 98
.13 10 100
.13 11 104

```
proc npar1way data=Gossypol;
class Dose;
var Gain;
run;
```

Dr. Firas

بسبب عدم تحديد الاختبار سنحصل على عدة اختبارات WILCOXON, MEDIAN, VW, SAVAGE, and EDF . كما ان النتائج تتضمن اختبار F لتأثير الجرعة (p=0.001) وهي نفس النتيجة التي يمكن الحصول عليها بأستعمال الايعاز الخاص بجدول تحليل التباين الاتي:

```
proc anova;classes dose;
model gain=dose;
means dose/lsd duncan;
                                    run;
```

مثال (157): في دراسة شملت 43 طالبا خضعوا لبرنامجين للتقييم يضمان اربعة فئات المطلوب اختبار التوافق بين درجات البرنامجين؟ ويمكن اختبارالتوافق باعتماد معامل كابا= صفر.

```
data fitness;
input score1 $ score2 $ count;
datalines;
poor poor 5
average average 4
good good 4
superior superior 3
poor average 3
average poor 1
average good 6
good average 5
good superior 1
superior average 10
superior good 1
proc freq;
weight count;
tables score1 * score2 / agree;
exact kappa;
run;
```

بعد تنفيذ البرنامج سنجد ان الاحتمال في اختبار التماثل (Test of symmetry) يساوي 0.086 واختبار معامل كابا يساوي 0.055 مما يعني ان قيمة كابا لاتختلف عن الصفر وان برنامجي التقييم متماثلين. من جهة اخرى نجد ان قيمة p في .asymptotic test = 0.038 وهي تعطي استنتاج معاكس لقيمة اختبار معامل كابا.

مثال (158): في دراسة لاختبار فرضية ان توزيع درجات الطلاب في امتحان الاحصاء متشابهه في عدة كليات ، سجلت البيانات ادناه ، اختبر التوزيع بين المجاميع وبين كل مجموعتين؟

نستعمل اختبار كرسكال لوجود اكثر من عينتين ولانها عينات مستقلة. ثم نجري اختبار فريدمان الخاص بعينتين مستقلة لتوليفات المجاميع (1 و2 ، 1 و3 ، 2 و 3).

```
data kruskal;
 input group score @@;
```

Dr. Firas

```
 cards;
1 55  1 0  1 1  1 0  1 50  1 60  1 44
2 73  2 85  2 51  2 63  2 85  2 81
2 66  2 69  3 61  3 54  3 80  3 47
proc npar1way wilcoxon;
 class group;
 var score;
 run;
data kr12;
 set kruskal;
 if group < 3;
proc npar1way wilcoxon; class group;
 exact; var score;
 run;
data kr23;
set kruskal; if group > 1;
proc npar1way wilcoxon; class group;
 exact; var score;
 run;
data kr13;
 set kruskal; if group NE 2;
proc npar1way wilcoxon; class group;
 exact; var score;
 run;
```

واذا وجد القارىء ان قطعة البرنامج صعبة فيمكن الغاء جزء من قطعة البرنامج الذي يلي اول run في البرنامج ، ولكي نحصل على اختبارات زوجية نستعمل العبارة التالية بعد عبارة input If group =2 then delete وذلك للمقارنة بين 1 و 3 ، ثم ننفذ البرنامج. ثم نغير رقم group=3 وننفذ فنحصل على مقارنة بين 1 و 2. واخيرا نغير group=1 لنحصل على مقارنة بين 2 و 3.

```
data kruskal;
 input group score @@;
if group=3 then delete;
 cards;
1 55  1 0  1 1  1 0  1 50  1 60  1 44
2 73  2 85  2 51  2 63  2 85  2 81
2 66  2 69  3 61  3 54  3 80  3 47
proc npar1way wilcoxon;
 class group;
 var score;
 run;
```

مثال (159): في دراسة عن العوامل المؤثرة على طول عمر الافراد عند زراعة كلية لهم سجلت البيانات التالية TIME الوقت المستغرق من اجراء العملية لغاية موت المريض ، DEATH حالة الشخص 0=حي و 1= ميت ، GENDER اذ ان 1= ذكر و2= انثى ، RACE اذ ان 1 = ابيض و2= اسود ، AGE = العمر بالسنوات.

```
data u;
input obs time death gender race age;
cards;
  1     1     0     1     1     46
```

Dr Firas

2	5	0	1	1	51
3	7	1	1	2	55
4	9	0	1	2	57
5	13	0	1	1	45
6	13	0	2	2	43
7	17	1	1	1	47
8	20	0	1	1	65
9	26	1	2	2	55
10	26	1	2	1	44
11	28	1	1	1	49
12	32	0	1	2	52
13	32	0	2	1	31
14	43	0	1	2	63
15	43	1	2	1	55
16	44	1	2	2	50
17	51	0	1	1	49
18	51	0	2	2	35
19	51	0	1	1	23
20	56	1	1	1	57
21	57	1	2	2	45
22	59	1	2	2	41
23	62	1	1	2	38
24	66	0	1	1	55
25	66	0	1	2	45
26	67	0	1	2	44
27	68	1	1	1	45
28	69	1	1	2	65
29	79	0	1	2	66
30	79	1	1	1	57

```
proc lifetest data=u plots=(s);
 time time*age(0);
 strata gender;
run;
 proc phreg data=u;
 model time*death(0) = age;
run;
```

مثال (160): في دراسة لمقارنة نتائج اختبار فحص نسبة السكر في الدم قبل الافطار ووجد ان عدد الافراد الذين ارتفعت لديهم نسبة السكر هو 53 حالة موجبة و47 حالة سالبة فيما اعطى الاختبارنفسه بعد الافطار 65 حالة موجبة و35 حالة سالبة علما بان كلا الاختبارين اشتركا في الكشف عن 30 حالة موجبة و12 حالة سالبة فهل هناك فرق بين نتائج الاختبارين؟

```
data compar;
input test1 test2 count;
cards;
1 1 30
1 2 35
2 1 23
2 2 12
proc freq data=compar;
tables test1*test2/agree;
weight count;
exact McNem;
run;
```

Dr. Firas

مثال (161): في دراسة تضمنت مجموعتين من الفئران المجموعة الاولى عرضت الى مسببات مرضية لمدة ساعة والثانية لمدة 24 ساعة ثم قسمت كل مجموعة الى ثلاثة اقسام بصورة عشوائية واعطي لكل من هذه المجاميع ثلاثة انواع من العقار وسجلت عدد ايام بقائها لمدة 18 يوم (1= ميت و0=حي) المطلوب معرفة هل ان فترة بقائها بين المجموعتين مختلفة؟

```
data g;
input trt drug$ days status;
cards;
1   a1 1   1
1   a1 4   1
1   a1 0   1
1   a1 7   1
1   a2 13  1
1   a2 5   1
1   a2 7   1
1   a2 15  1
1   a3 9   1
1   a3 16  1
1   a3 18  0
1   a3 13  1
2   a1 15  1
2   a1 6   1
2   a1 10  1
2   a1 13  1
2   a2 6   1
2   a2 18  0
2   a2 9   1
2   a2 15  1
2   a3 14  1
2   a3 7   1
2   a3 6   1
2   a3 13  1
proc lifetest data=g plots=(s,ls,lls);
time days*Status(0);
strata drug;
test trt;
run;
```

اما اذا اردنا اختبار تأثير الدواء فيمكن عمل ذلك مع اجراء تحوير بسيط للبيانات وكالاتي:

```
data g;
input trt drug days status;
cards;
1   1 1   1
1   1 4   1
1   1 0   1
1   1 7   1
1   2 13  1
1   2 5   1
1   2 7   1
1   2 15  1
1   3 9   1
1   3 16  1
1   3 18  0
1   3 13  1
```

Dr. Firas

```
2    1  15   1
2    1  6    1
2    1  10   1
2    1  13   1
2    2  6    1
2    2  18   0
2    2  9    1
2    2  15   1
2    3  14   1
2    3  7    1
2    3  6    1
2    3  13   1
proc sort; by drug;
run;
proc lifetest data=g plots=(s,ls,lls);
time days*Status(0);
strata trt;
test drug;
run;
```

مثال (162): اخذت عينة شملت 45 عائلة من الآباء والامهات وخضعت لبرنامج حمية لغرض تخفيف اوزانهم وكان عدد الازواج (زوج وزوجته) الذين استجابوا 20 وعدد الآباء الذين استجابوا دون زوجاتهم 5 ، وبلغ عدد الأمهات اللاتي استجبن دون ازواجهن 10 فيما كان عدد الازواج (زوج وزوجته) الذين لم يستجيبوا 10 فهل كان للبرنامج تأثير على العينة المدروسة؟

```
data approval;
  input hus $ wif $ count ;
  cards;
  yes yes 20
  yes no   5
  no yes  10
  no no   10
proc freq order=data;
  weight count;
  tables hus *wif / agree;
run;
```

مثال (163): في دراسة عن تأثير علاج جديد لمرضى الربو شملت مركزين طبيين و 180 مريضا وجد ان عدد الذين استجابوا للعلاج هو 29 ومن لم يستجب 16 فيما كان عدد الذين تناولوا placebo واستجابوا 14 ومن لم يستجب 31 في المركز الاول وبلغت الاعداد المناظرة في المركز الثاني 37 ، 8 ، 24 ، 21 ، فهل ان العلاج كان ناجحا؟

```
data respire;
  input center treatmnt $ response $ count @@;
  cards;
  1 test    y 29 1 test    n 16
  1 placebo y 14 1 placebo n 31
  2 test    y 37 2 test    n  8
  2 placebo y 24 2 placebo n 21
proc freq order=data;
    weight count;
    tables center*treatmnt*response /
```

Dr. Firas

```
      nocol nopct chisq cmh;
run;
```

مثال (164) : اجريت دراسة على 322 مريضا لمقارنة تأثير اربعة عقارات مع السيطرة (placebo) على استجابة المرضى للعلاج بأعتماد طريقتين مختلفتين في التشخيص فهل كان هناك تأثير للعقار على الاستجابة وهل ان هذه العلاقة تختلف باختلاف طريقة التشخيص؟

```
data pain;
    input diagnos $ treat $ response $ count @@;
    datalines;
 I placebo  no 26 I  placebo yes  6
 I dosage1  no 26 I  dosage1 yes  7
 I dosage2  no 23 I  dosage2 yes  9
 I dosage3  no 18 I  dosage3 yes 14
 I dosage4  no  9 I  dosage4 yes 23
 II placebo no 26 II placebo yes  6
 II dosage1 no 12 II dosage1 yes 20
 II dosage2 no 13 II dosage2 yes 20
 II dosage3 no  1 II dosage3 yes 31
 II dosage4 no  1 II dosage4 yes 31
 proc freq order=data;
   weight count;
   tables treat*response / chisq;
   tables diagnos*treat*response / chisq cmh;
 run;
```

مثال (165): البيانات ادناه تمثل اوزان مجموعتين من الحيوانات (ذكور واناث) ، وقد غذيت على عليقتين مختلفتين ، المطلوب معرفة تأثير المعاملة والجنس والتداخل بينهما على اوزان الحيوانات ؟

```
DATA weight;
INPUT TREAT $ SEX $ WT  @@;
DATALINES;
A F 110  A F 101  A F 124  A F 120  A F 111
A F 117  A F 120  A F 131  A M 185  A M 181
A M 173  A M 190  A M 181  A M 202  A M 175
B F 121  B F 116  B F 144  B F 125  B F 115
B F 118  B F 127  B M 205  B M 193  B M 196
B M 189  B M 180  B M 193  B M 210  B M 189
B M 179
PROC GLM DATA = weight;
   CLASSES TREAT SEX;
   MODEL WT = TREAT SEX TREAT*SEX / SS3;
means treat sex treat*sex/lsd;
RUN;
```

مثال (166): البيانات ادناه تمثل الجيل الثاني من نبات الشعير وهو باربعة اشكال مظهرية اختبر فرضية ان الاعداد تتبع النسبة المندلية 1:3:3:9 ، علما بأن الاحتمالية اذا كانت = 0.05 او اقل فأن ذلك يعني انها لاتتبع النسبة المندلية.

```
data f2;
input pheno count ;
```

Dr. Firas

```
cards;
1 1178
2 291
3 273
4 156
proc freq;
weight count;
tables pheno / testp = (0.5625, 0.1875, 0.1875, 0.0625);
run;
```

مثال (167): البيانات ادناه تعود الى مجموعة من المرضى المصابين بالسرطان وقد استعمل في علاج المجموعة الاولى عقار جديد واستعمل مع المجموعة الثانية نفس العقار علاوة على العلاج بالاشعاع وسجل عدد ايام بقائهم بالايام على قيد الحياة من تاريخ بدء العلاج وقد عبر عن الموت = 1 وعلى قيد الحياة =0، قدر متوسط العمر في كلا المجموعتين ووضح هل كان للاشعاع تأثير على عمر المرضى معبرا عن ذلك بنسبة الخطورة؟

```
data s;
input time death group;
cards;
 19  1 1
 25  1 1
 30  1 1
 34  1 1
 37  1 1
 46  1 1
 47  1 1
 51  1 1
 56  1 1
 57  1 1
 61  1 1
 66  1 1
 67  1 1
 74  1 1
 78  1 1
 62  1 2
 83  1 2
 38  1 2
 42  1 2
 73  1 2
 77  1 2
 89  1 2
 55  1 2
 44  0 2
 ; proc lifetest data=s plots=(s);
  time time*death(0);
  strata group ;
run;
 proc phreg data=s;
  model time*death(0) = group;
run;
```

Dr. Firas

مثال (168): في دراسة شملت مجموعة من الاطفال في منطقتين مختلفة من اوربا سجلت الوان العيون والشعربحيث بلغت عدد التوليفات 15 توليفة جد تكرارات كل من الوان العيون والشعر وجد عدد افراد كل توليفة والنسبة المؤية لكل منها؟ وهل ان لون العيون له علاقة بلون الشعر؟

```
data Color;
   input Region Eyes $ Hair $ Count @@;

     cards;
 1 blue  fair   23  1 blue  red    7  1 blue  medium 24
 1 blue  dark   11  1 green fair   19  1 green red    7
 1 green medium 18  1 green dark   14  1 brown fair   34
 1 brown red    5  1 brown medium 41  1 brown dark   40
 1 brown black   3  2 blue  fair   46  2 blue  red    21
 2 blue  medium 44  2 blue  dark   40  2 blue  black   6
 2 green fair   50  2 green red    31  2 green medium 37
 2 green dark   23  2 brown fair   56  2 brown red    42
 2 brown medium 53  2 brown dark   54  2 brown black  13
  proc freq data=Color;
    weight Count;
    tables Eyes Hair Eyes*Hair/out=FreqCnt outexpect sparse;
 run;

 proc print data=FreqCnt noobs;
 run;
 proc freq  order=data;
    weight Count;
    tables Eyes*Hair / chisq expected cellchi2 norow nocol;
    output out=ChiSqDat pchi lrchi n nmiss;
 run;
```

مثال (169): في دراسة لمقارنة كفاءة اختبار بالمقارنة مع اختبار Gold Standard على 344 شخص اذ اشار الفحص الجديد الى وجود 263 حالة اصابة و81 حالة سليمة وعند مقارنة نتائج الاختبار الجديد مع الاختبار القياسي وجد ان من بين النتائج الموجبة كانت 32 كاذبة ومن بين النتائج السالبة 27 حالة كاذبة جد مايلي:

Sensitivity ، Specificity ، positive predictive value (PPV)، Negative predictive value (NPV) اذا علمت ان Prevalence = 75%؟

		نتيجة الاختبار القياسي		الكلي
		+	–	
نتيجة الاختبار الجديد	+	231	32	263
	–	27	54	81
		258	86	344

Sens = 231 /258 = 0.90
Spec = 54 / 86 = 0.63
PPV = 231 / 263 = 0.88

$$NPV = 54 / 81 = 0.67$$

مثال (170): في دراسة عن مجموعة من الفئران عرضت الى 240 راد من اشعة كاما وسجل عدد ايام بقائها من تاريخ التعرض قدرمعدل الخطورة باستعمال طريقة تحليل البقاء؟

```
Data mice;
 input lifetime freq @@;
 cards;
  40 1  48 1  50 1  54 1  56 1  59 1  62 1  63 1  67 2  69 1
  70 1  71 1  73 2  76 1  77 1  80 1  81 2  82 1  83 1  84 1
  86 2  87 1  88 5  89 1  90 2  91 1  93 1  94 1  95 1  96 1
  97 2  98 1  99 2 100 4 101 3 102 2 103 5 104 3 105 2 106 3
 107 1 108 1 109 2 110 3 111 3 112 1 113 2 114 2 115 1 116 2
 117 1 118 3 119 2 120 3 121 2 123 2 124 3 125 2 126 5 127 4
 128 4 129 6 130 4 131 2 132 1 133 3 134 4 135 3 136 4 137 3
 138 1 139 2 140 2 141 5 142 1 144 5 145 2 146 4 147 4 148 4
 149 1 150 1 151 4 152 2 153 1 155 1 156 1 157 1 158 2 160 1
 161 1 162 2 163 2 164 1 165 2 166 1 168 1 169 1 171 2 172 2
 174 1 177 2
 proc lifetest method = life
   intervals = (30 to 180 by 15)
   plots = (pdf, hazard);
   time lifetime;
   freq freq;
run;
```

مثال(171): اجريت تجربة لدراسة تأثير عقار على مجموعتين من المرضى وسجلت الاستجابة اذا كانت موجبة (f) وعم زجود استجابة (u) المطلوب اختبار ان العقار له تأثير في الاستجابة ام لا؟

```
data respire;
input treat $ outcome $ @@ ;
cards;
placebo f placebo f placebo f
placebo f placebo f
placebo u placebo u placebo u
placebo u placebo u placebo u
placebo u placebo u placebo u
placebo u
test f test f test f
test f test f test f
test f test f
test u test u test u
test u test u test u
test u test u test u
test u test u test u
test u test u test u
test u test u test u
test u test u
;
proc freq;
tables treat*outcome/cl riskdiff relrisk rdiff;
run;
```

Dr. Firas

يمكن كتابة البيانات بشكل آخر:

```
DATA e;
 INPUT bpr  status  freq;
cards;
0 0 5
0 1 10
1 0 8
1 1 20
proc freq;
 weight freq;
 tables bpr*status/cl riskdiff relrisk rdiff;
 run;
```

مثال (172): في دراسة عن تأثير بعض العوامل على عدد المواليد الهالكة عند الولادة تضمنت طول مدة الحمل 261 يوم فأكثر= 1 وأقل = 0 ، وعمر الام أقل من 30 سنة فأقل =1 واكثر = 0 ، عدد السكاير المتناولة في اليوم الواحد لاتزيد عن 6 سكاير= 1 وأكثر = 0 ، المولود ميت عند الولادة =1 او بعد اقل من 4 ساعات = 0 المطلوب تحديد تأثير العوامل ؟

الصفة عددية لذا نستعمل طريقة genmod والتوزيع بواسوني (D=poisson).

```
data f;
 input gest age cigs vita count @@;
 cards;
0 0 0 0  50     0 0 0 1 315     0 0 1 0   9
0 0 1 1  40     0 1 0 0  41     0 1 0 1  147
0 1 1 0   4     0 1 1 1  11     1 0 0 0  24
1 0 0 1 4012     1 0 1 0   6     1 0 1 1  459
1 1 0 0  14     1 1 0 1 1494     1 1 1 0   1
1 1 1 1  124

proc genmod;
  model count = gest  age  cigs  vita / dist = Poisson;
  run;

proc genmod;
  model count = gest  age  cigs  vita  gest*age age*cigs
          gest*vita age*vita cigs*vita / dist = Poisson;
  run;
```

مثال (173): اجريت دراسة على مجموعتين من المرضى قسموا على اساس الحالة المرضية للتحري عن تأثير ثلاثة عقارات وسجلت عدد حالات الشفاء وعدد الحالات الكلي ، اختبر تأثير العاملين على نسبة الشفاء ؟

```
data uti2;
input diagnoss $  treat $   events trials;
cards;
compli    A   78   106
compli    B  101   112
compli    C   68   114
uncomp    A   40   45
uncomp    B   54   59
```

Dr Firas

```
uncomp    C  34  40
proc genmod;
  class diagnoss treat;
  model events/trials = diagnoss treat /
  link=logit dist=binomial type3;
  run;

proc probit;
  class diagnoss treat;
  model events/trials=diagnoss treat / d=logistic itprint;
  run;
```

بما ان بعض بيانات المثال اعلاه اسمية لذا فأن من الاسهل اعتماد طريقة genmod لانها تضع معادلة توقع وتعطي نتائج اختبار تحليل التباين. ويمكن استعمال طريقة probit ايضا للحصول على معادلة توقع اما اذا اردنا استعمال طريقة logistic فيجب اعطاء بعض الأوامر :

```
if diagnoss='compli' then d1=1; else d1=0;
if diagnoss='uncomp' then d2=1; else d2=0;
if treat='A' then T1=1; else T1=0;
if treat='B' then T2=1; else T2=0;
if treat='C' then T3=1; else T3=0;
```

فيما سيكون امر التنفيذ:

```
proc logistic;
  model events/trials=d1 d2 T1 T2 T3;
  run;
```

مثال(174): في دراسة تضمنت بيانات عن شراء احد المنتوجات للتحري عن تأثير عاملين هما المستوى العلمي ومستوى الدخل في شراء ذلك المنتوج وتضمن المستوى العلمي مستويان (بكالوريوس فأكثر او ادنى من البكالوريرس) ومستوى الدخل (عالي او واطىء) المطلوب دراسة تأثير العاملين والتداخل بينهما ووضع معادلة تنبوء اعتمادا على العوامل المؤثرة ؟

ملاحظة: في حالة كون التداخل غير معنوي يفضل رفعه ومن ثم الحصول على معادلة التنبوء.

ان المتغير no يأخذ حالتين نعم وتمثل شراء و كلا وتمثل عدم شراء وبما ان كلمة نعم هي الاولى في الترتيب فعند تطبيق القيم على معادلة التوقع سنحصل على احتمالية الشراء .

```
data family;
input Educ $ Income $ no $  wt;
 cards;
  high  high  yes     54
  high  high  no      23
  high  low   yes     41
  high  low   no      12
  low   high  yes     35
  low   high  no      42
  low   low   yes     19
  low   low   no       8
  ;
 proc catmod order=data;
```

Dr.Firas

```
  weight wt;
  response 1 0;
  model no=Educ Income educ|income/ pred ;
  run;
```

مثال (175): اجريت تجربة للمقارنة بين تأثير ثلاثة عقارات لعلاج الصداع على مجموعة من المرضى قسموا بصورة عشوائية الى ثلاث مجموعات وسجل الوقت المستغرق من تناول الدواء لغاية زوال الصداع المطلوب اختبار الاختلافات بين المجاميع بسبب نوع العقار؟

قبل اختيار نوع الاختبار لابد من اختبار التوزيع الطبيعي للبيانات واختبار تجانس التباين.

```
DATA ACHE;
INPUT BRAND RELIEF;
CARDS;
1 24.5
1 23.5
1 26.4
1 27.1
1 29.9
2 28.4
2 34.2
2 29.5
2 32.2
2 30.1
3 26.1
3 28.3
3 24.3
3 26.2
3 27.8
;
PROC ANOVA DATA=ACHE;
   CLASSES BRAND;
   MODEL RELIEF=BRAND;
   MEANS BRAND/hovtest=levene welch;
    MEANS BRAND/TUKEY;
run;
PROC PLOT;
   PLOT RELIEF*BRAND;
run;
proc univariate normal plot; var relief;run;
```

مثال (176): لو كانت لدينا البيانات ادناه والتي تعود الى صفة ما قيست في مجموعتين فكيف تحدد الاختبار المناسب لاختبار الاختلافات بين متوسطي المجموعتين؟

```
data  d;
input Obs GROUP $ no;
cards;
1  A 1202.6
2  A 830.1
3  A 372.4
4  A 345.5
5  A 321.2
6  A 244.3
```

Dr. Firas

7 A 163.0
8 A 147.8
9 A 95.0
10 A 87.0
11 A 81.2
12 A 68.5
13 A 47.3
14 A 41.1
15 A 36.6
16 A 29.0
17 A 28.6
18 A 26.3
19 A 26.1
20 A 24.4
21 A 21.7
22 A 17.3
23 A 11.5
24 A 4.9
25 A 4.9
26 A 1.0
25 A 4.9
26 A 1.0
27 B 2745.6
28 B 1697.8
29 B 1656.0
30 B 978.0
31 B 703.4
32 B 489.1
33 B 430.0
34 B 334.1
35 B 302.8
36 B 274.7
37 B 274.7
38 B 255.0
39 B 242.5
40 B 200.7
41 B 198.6
42 B 129.6
43 B 119.0
44 B 118.3
45 B 115.3
46 B 92.4
47 B 40.6
48 B 32.7
49 B 31.4
50 B 17.5
51 B 7.7
52 B 4.1

قبل تحديد نوع الاختبار اللازم (معلمي او غير معلمي) لابد من اجراء اختبار ان البيانات تتوزع توزيع طبيعي وان تباينهما متساوي.

```
proc univariate normal plot;
by group;
var no;
run;
 proc ttest;
```

Dr. Firas

```
class group;
var no ;
run;
```

سنلاحظ من النتائج ان توزيع البيانات غير طبيعي كما سنجد بأت تباينهما غير متساوي لذا لابد من اجراء اختبار لامعلمي.

```
PROC NPAR1WAY WILCOXON;
CLASS GROUP;
VAR no;
EXACT WILCOXON;
Run;
```

مثال(177): لغرض معرفة تأثير وقت الاجابة على الدرجة اجرى احد الاساتذة اختبارا لمجموعة من الطلاب لهم نفس درجة الذكاء بعد ان قسمهم الى مجموعتين اختبر الاولى صباحا واختبرالثانية مساءا المطلوب معرفة هل ان الاختلافات بين المجموعتين معنوية؟ واي من الاختبارين يجب اعتماده T او WILCOXIN ولماذا؟

واذا اعتيرنا القيمة 83.9 ضمن المجموعة B هل تتغير النتائج والاستنتاجات؟

```
data f;
input group $ texam;
cards;
A 89.8
A 90.2
A 98.1
A 91.2
A 88.9
A 90.3
A 99.2
A 94.0
A 88.7
A 83.9
B 57.3
B 87.6
B 87.3
B 91.8
B 86.4
B 86.4
B 93.1
B 89.2
B 90.1

;option nodate nonumber;
proc npar1way WILCOXIN data=F;
class GROUP;
var TEXAM;
exact;
run;
 proc ttest;
class GROUP;
var TEXAM;
run;
```

Dr. Firas

مثال (178): عرضت مجموعتان من المرضى الى محفز وسجل الوقت المستغرق لحصول الاستجابة المطلوب اختبار الاختلافات بين المجموعتين؟

لغرض تحديد نوع الاختبار نجري اختبار التوزيع الطبيعي لغرض تحديد نوع الاختبار معلمي ام لامعلمي.

```
Data react;
input stim time @@;
cards;
1 1.94 1 1.94 1 2.92 1 2.92
1 2.92 1 2.92 1 3.27 1 3.27
1 3.27 1 3.27 1 3.70 1 3.70
1 3.74 2 3.27 2 3.27 2 3.27
2 3.70 2 3.70 2 3.74
;option nodate nonumber;
proc univariate normal plot;
by stim;
var time;
run;
proc npar1way wilcoxon data=react;
class stim;
var time;
exact;
run;
```

مثال (179): في دراسة افتراضية عن تأثيرعدة جرع من عقار جديد على السيطرة على الألم سجل التأثير العكسي لزيادة الجرعة للعقار على الألم اذ استلمت الجموعة الاولى placebo فيما استلمت المجاميع الاخرى جرع اكبر(تصاعدي) المطلوب تحديد حجم الجرعة على زيادة او قلة الألم؟

```
data Pain;
input Dose Adverse $ Count;
cards;
0 No   26
0 Yes  6
1 No   26
1 Yes  7
2 No   23
2 Yes  9
3 No   18
3 Yes  14
4 No   9
4 Yes  23
proc freq data=Pain;
weight Count;
tables Dose*Adverse / trend  cl;
exact trend;
run;
```

استعمال الامر trend لغرض اختبار الاتجاه لقيم فئات متغير الجرعة بأستعمال اختبار كوكران ارميتاج.

Dr. Firas

مثال (180): البيانات ادناه تمثل عدد الجرائم مصنفة الى : احراق ممتلكات ، ارهاب ، عنف، سرقة ، تزوير ، احتيال ، وقد وزعت حسب حالة القائم بالعمل : مدمن على المخدرات ام لا المطلوب اختبار وجود علاقة ام لا بين نوع الجريمة وحالة القائم بها؟

نوع الجريمة	مدمن	غير مدمن
احراق ممتلكات	50	43
ارهاب	88	62
عنف	155	110
سرقة	379	300
تزوير	18	14
احتيال	63	144

هنا يعتمد اختبار مربع كاي .

```
data S;
input  status  type count;
cards;
1 1  50
1 2  88
1 3  155
1 4  379
1 5  18
1 6  63
2 1  43
2 2  62
2 3  110
2 4  300
2 5  14
2 6  144
;
proc freq data=s order=data;
weight count;
tables status*type / chisq;
run;
```

مثال (181): اجريت تجربة لدراسة العلاقة بين التعرض لمادة كيمياوية في احد المنظفات وحدوث الحساسية الجلدية ، اذ اخذت مجموعتين تضم كل منها 20 شخصا المجموعة الاولى سبق وان تعامل افرادها بتلك المنظفات فيما لم يسبق ان تعامل افراد المجموعة الثانية بتلك المنظفات وخضع الجميع للفحص لتحديد وجود الحساسية المطلوب تحديد الاختلافات بين المجموعتين وماهو المقياس الذي يفضل استعماله : نسبة الارجحية او الخطورة النسبية ولماذا؟

	وجود حساسية	عدم وجود حساسية	
التعرض	13	7	20

Dr.Firas

عدم التعرض	4	16	20

للمقارنة يستعمل مربع كاي.

```
data S;
input   status group  count;
cards;
1 1 13
1 2  4
2 1  7
2 2 16
;
proc freq;
weight count;
tables status*group / chisq;
run;
```

يفضل استعمال نسبة الارجحية لان الاحداث تتعلق بالماضي.

```
data S;
input   status group  count;
cards;
1 1 13
1 2  4
2 1  7
2 2 16
;
proc freq;
weight count;
tables status*group / cl riskdiff;
run;
```

اما اذا اردنا استعمال الخطورة النسبية في هذه الحالة يجب ان نأخذ مجموعتين الاولى تعرضت للخطر والثانية لم تتعرض للخطر ومن ثم نتابع هاتين المجموعتين لمدة من الزمن لغرض تسجيل من تحسس من كل مجموعة .

مثال (182): البيانات التالية تمثل جرع بتراكيز مختلفة (z) فيما تمثل x1 عدد الافراد الهالكة وx2 عدد الافراد الحية المطلوب تقدير قيمة LD50؟

(الجواب : LD50= 4.07).

```
DATA SET1;
INPUT Z X1 X2;
  ZL = LOG(Z);
  N = X1+x2;
CARDS;
49.057 6 53
52.991 13 47
56.911 18 44
60.842 28 28
64.759 52 11
68.691 53 6
72.611 61 1
```

```
76.542 60 0
proc probit data=set1 inversecl;
model x1/N = zl/dist=logistic;
output out=new1 p=p_hat;
run;
proc print;run;
proc plot;
plot p_hat*zl;run;


proc probit data=set1 inversecl;
model x1/N = zl;
output out=new2 p=p_hat;
run;
proc print data=new2;
proc plot data=new1;
plot p_hat *zl;
run;
```

يمكن تقديرها ايضا باستعمال الايعاز الآتي:

```
 PROC LOGISTIC DATA=SET1;
 MODEL X1/N = ZL;
 OUTPUT OUT=SETR1 L=LOWER95 P=PHAT
        U=UPPER95 / ALPHA=.05;
 RUN;
 proc print;run;
 proc plot;
 plot phat*zl;run;
```

ملاحظة: عندما يكون المتغير التابع نسبة ويراد تقدير قيمة LD50 فيجب استعمال طريقة الانحدار اللوجستيكي بدلا من الانحدار الخطي اذ ان الانحدار اللوجستيكي سيكون اكثر مطابقة لوصف البيانات. فمثلا لوحاولنا المقارنة بين الانحدار اللوجستيكي والخطي في وصف العلاقة بين نسبة الهلاكات ولوغريتيم الجرعة للمثال السابق سنجد ان الانحدار اللوجستيكي سيعطينا قيمة ادق لـ LD50 كما تكون القيم المتوقعة اكثر مطابقة للقيم المتنبأبها.

```
DATA SET1;
INPUT Z X1 X2;
 ZL = LOG(Z);
 N = X1+x2;
 TP=(x1/n);
CARDS;
49.057 6 53
52.991 13 47
56.911 18 44
60.842 28 28
64.759 52 11
68.691 53 6
72.611 61 1
76.542 60 0
proc probit data=set1 inversecl;
model x1/N = zl/dist=logistic;
output out=new1 p=p_hat;
```

```
run;
proc print;run;

 proc plot;
plot (p_hat TP) *zl/overlay;run;

proc reg data=SET1 ;
MODEL TP= ZL/P;
output out=new2 p=p_hat;
Run;

PROC PRINT DATA=NEW2;RUN;
proc plot DATA=NEW2;
plot (p_hat TP) *zl/overlay;run;
```

لابد من الاشارة الى ان قيمة LD50 والتي تحسب يدويا بالاعتماد على البيانات ممكنة في بعض الحالات كما في مثالنا هذا(الجرعة 60 اعطت 50% هلاكات) الا ان التقدير لايعتد به لانه لايأخذ بنظر الاعتبار توزيع المتغيرين (التابع والمستقل) وبذلك فأننا لانستطيع توقع تأثير مستوى قليل من الجرعة او مستوى عالي منها على نسبة الهلاكات.

Dr. Firas

المصادر

الراوي، خاشع محمود.1980.المدخل الى الاحصاء.

الراوي، خاشع محمود.1987. المدخل الى تحليل الانحدار.

الـراوي، خاشـع محمـود ، خلـف الله ، عبـدالعزيز محمـد. 1980. تصـميم وتحليـل التجـارب الزراعية.

المحمد، نعيم ثاني ؛ الراوي ، خاشع محمود؛ يونس، مؤيد أحمد و المراني، وليد خضير.1983. مبادىء الاحصاء.

غزال، نجيب توفيق.1982. الوراثة الكمية في تربية وتحسين الحيوان.

Ahlbom, A., 1993. Biostatistics for Epidemiologists. Lewis Publishers.

Allison, P.D. 1995. Survival Analysis Using the SAS System: A Practical Guide. SAS Institute Inc., Cary, NC, USA .

Alison, P.D. 1999. Logistic Regression Using the SAS System: Theory and Application Copyright(c). SAS Institute Inc., Cary, NC, USA

Cody, R.P. and Smith, J.K. 1997. Applied Statistics and the SAS Programming Language, Fourth Edition Copyright(c), Prentice-Hall, Inc. Upper Sadle River, New Jersey.

Der, G. and Everitt, B.2001. A Handbook of Statistical Analyses Using SAS, Second Edition. Copyright(c). Chapman & Hall / CRC Press.

Derr, E. R. 1998. Performing Exact Logistic Regression with the SAS System. SAS Institute Inc., Cary, NC.

Dmitrienko, A., Molenbergh, G., Chuang-Stein, C. and Offen W. 2005.

Analysis of Clinical Trials Using SAS: A Practical Guide. Copyright(c). SAS Institute Inc., Cary, NC, USA.

Hosmer, Jr., D. W. and Lemeshow, S., 1989. Applied Logistic Regression. John Wiley and Sons.

Littell, R., Milliken, G., Stroup, W., Wolfinger, R. and Schabenberger, O. 2006. SAS for Mixed Models, Second Edition Copyright(c). SAS Institute Inc., Cary, NC, USA.

Dr.Firas

Muller, K. and Fetterman, B.2002.Regression and ANOVA: An Integrated Approach Using SAS Software.Copyright(c). SAS Institute Inc., Cary, NC, USA.

O'Brien, R.G. (1986), "Using the SAS System to Perform Power Analyses for Log-linear Models," Proceedings of the Eleventh Annual SAS Users Group International Conference, Cary, NC: SAS Institute Inc., 778–782.

Sackett, D. L., Haynes, R. B., and Tugwell, P., 1985. Clinical Epidemiology - A Basic Science for Clinical Medicine. Little, Brown, and Company.

SAS Institute, Inc. (1990), SAS/STAT User's Guide, Volume 2, GLM-VARCOMP, Version 6, Fourth Edition, Cary, NC: SAS Institute, Inc.

SAS Institute, Inc. (1992), SAS Technical Report P-229, SAS/STAT Software: Changes and Enhancements, Release 6.07, Cary, NC: SAS Institute, Inc. 12.

SAS Institute Inc. (1999). SAS Language Reference: Dictionary, Version 6.12, Cary, NC: SAS Institute Inc.

Stokes, M. E.1995. Statistical Enhancements in Release 6.12 of the SAS System. SAS Institute Inc., Cary, NC

Stokes, M. E., Davis, C. S., and Koch, G. G. (1995). Categorical Data Analysis Using the SAS System, Cary: SAS Institute, Inc.

Stroup, W. and Freund, R. 2002. SAS System for Linear Models, Fourth Edition. Copyright(c) 2002 by SAS Institute, Cary, NC, USA.

Timm, N.H. and Mieczkowski, T.A. (1997), Univariate and Multivariate General Linear Models: Theory and Applications Using SASâ Software, Cary, NC: SAS Institute Inc.

Walker, G.2002.Common Statistical Methods for Clinical Research with SAS Examples Copyright(c). SAS Institute, Cary, NC, USA.

Dr Firas