## اختبار التجانس

مثال 1: في الجدول ا لتالي أعداد الطلاب (count) في مادة 207 كمي (برنامج بكالوريوس) عند الدكتور/ محمد، وكذلك يبين الجدول التقديرات (grade) والنسبة المتوقعة (exp%) لكل تقدير. وتم الحصول على النسبة المتوقعة بناء على بيانات تاريخية تمثل نتائج الطلاب في جميع الشعب.

أعداد الطلاب حسب تقديراتهم في الشعبة (عدة شعب)

| GRADE | count | EXP % |
|---|---|---|
| A | 10 | 10% |
| B | 17 | 15% |
| C | 21 | 25% |
| D | 32 | 30% |
| F | 15 | 20% |

والمطلوب هو اختبار الفرضية القائلة بأن توزيع أعداد الطلاب على التقديرات يختلف عن النسب المحدد.

الحل

يهدف هذا السؤال إلى معرفة ما إذا كان توزيع تقديرات الطلاب متجانس مع توزيع التقديرات بناء على البيانات التاريخية. أي أنه ما نسبته 10% (أو نسبة قريبة من 10%) من طلاب الدكتور محمد حصلوا على تقدير A، وما نسبته 15% ( أو نسبة قريبة من 15%) من طلاب الدكتور محمد حصلوا على تقدير B، وهكذا.

والفرضية التي يتم اختبارها هي:

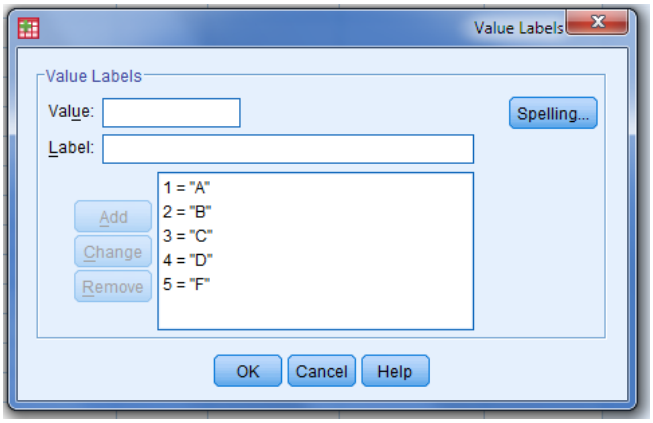
$H_0$: توزيع طلاب د. محمد على التقديرات متجانس مع النسب المتوقعة

$H_1$: توزيع طلاب د. محمد على التقديرات غير متجانس مع النسب المتوقعة

ويمكن إدخال البيانات بنفس الطريقة الموضحة في العرض الخاص بكاي المربع

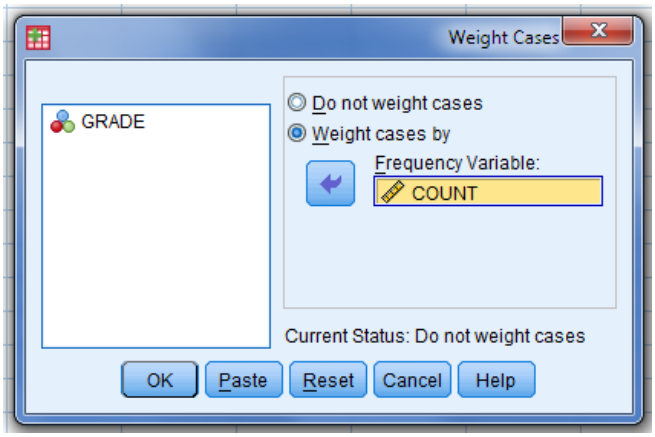
| | GRADE | COUNT | var |
|---|---|---|---|
| 1 | 1 | 10 | |
| 2 | 2 | 17 | |
| 3 | 3 | 21 | |
| 4 | 4 | 32 | |
| 5 | 5 | 15 | |
| 6 | | | |

ومن ثم تحديد التقدير المقابل لكل قيمة في المتغير GRADE على النحو التالي:



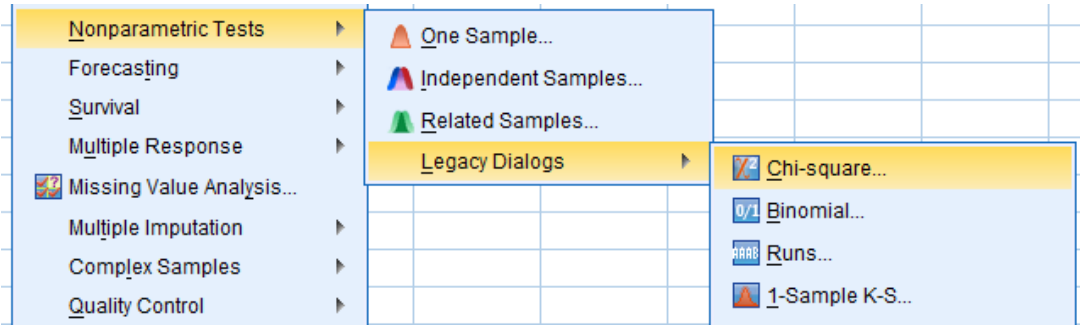

ثم تحديد المتغير COUNT على أنه يمثل تكرار أو عدد الطلاب لكل معدل، أي أنه يوجد 10 طلاب حصلوا على تقدير A و 17 طالب حصلوا على تقدير B، وهكذا. ويتم ذلك على النحو التالي:

DATA → WEIGHT CASES

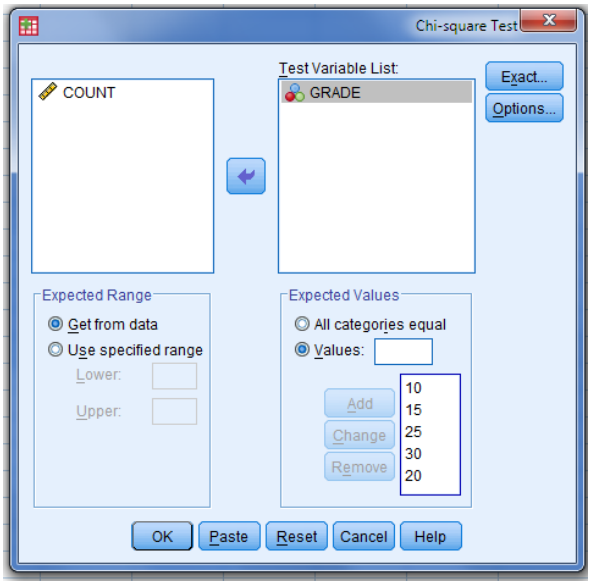



ولإجراء اختبار التجانس لكاي المربع نقوم بالتالي:

Analyze → Nonparametric Tests →Legacy dialogs → Chi-square




ويتم في مربع الحوار التالي تحديد المتغير الذي يمثل تقديرات الطلاب وتحديد النسب المقابلة لكل تقدير بناء على النسب المتوقعة.




وتكون نتائج الاختبار على النحو التالي:

**GRADE**

| | Observed N | Expected N | Residual |
|---|---|---|---|
| A | 10 | 9.5 | .5 |
| B | 17 | 14.3 | 2.8 |
| C | 21 | 23.8 | -2.8- |
| D | 32 | 28.5 | 3.5 |
| F | 15 | 19.0 | -4.0- |
| Total | 95 | | |

**Test Statistics**

| | GRADE |
|---|---|
| Chi-square | 2.147[a] |
| df | 4 |
| Asymp. Sig. | .709 |

ويوضح الجدول الأول التكرارات المشاهدة والتكرارات المتوقعة ، أما الجدول الثاني فيوضح قيمة اختبار كاي المربع وهي 2.147، و P-value = 0.709. وبمقارنة P-value بـ 0.05، فإنه لا يمكن رفض فرضية العدم، وبذلك فإن توزيع تقديرات الطلاب عند د. محمد متجانس مع البيانات التاريخية.

تحسب قيمة كاي المربع (إحصائية الاختبار) باستخدام العلاقة التالية:

$$\chi^2 = \sum_{i=1}^{5} \frac{(o_i - e_i)^2}{e_i}$$

ملاحظة (يجب تحديد جميع الخطوات الخمس لاختبار الفرضيات).

مثال 2:

أعداد الطلاب حسب تقديراتهم لدى أستاذين مختلفين

| الاستاذ | A | B | C | D | F |
|---|---|---|---|---|---|
| علي | 7 | 14 | 25 | 40 | 22 |
| إبراهيم | 10 | 17 | 21 | 32 | 15 |

اختبر الفرضية القائلة بأن توزيع التقديرات لدى الاستاذين غير متجانس. (توزيع التقديرات مختلف تباعا لاختلاف أستاذ المادة.

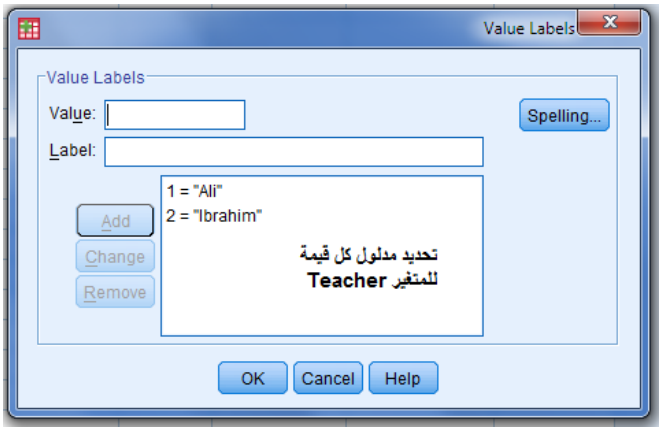
الفرضية

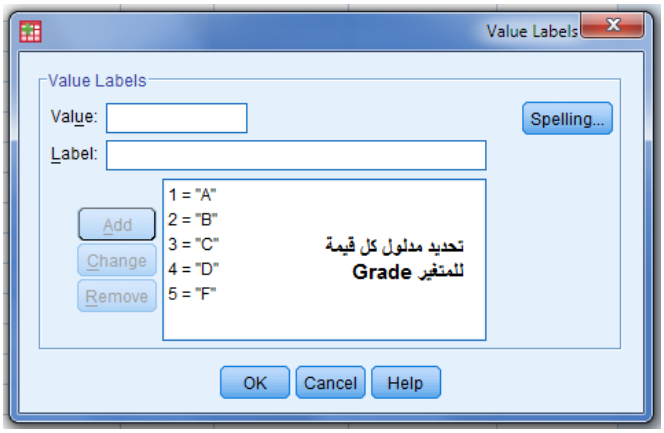
$H_0$: توزيع طلاب د. علي على التقديرات متجانس مع توزيع طلاب د. إبراهيم على التقديرات

$H_1$: توزيع طلاب د. علي على التقديرات غير متجانس مع توزيع طلاب د. إبراهيم على التقديرات

إدخال البيانات

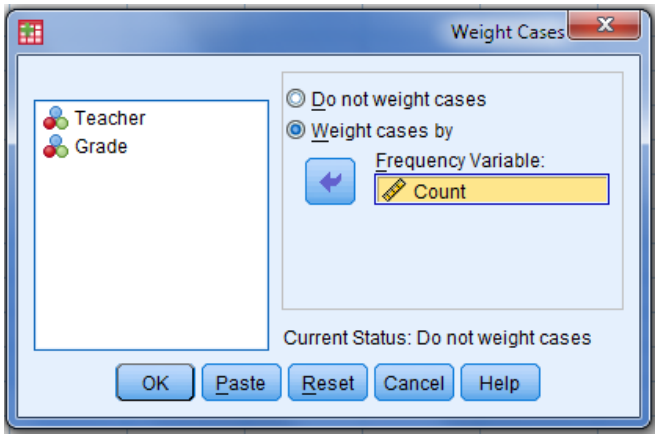
| | Teacher | Grade | Count |
|---|---|---|---|
| 1 | 1 | 1 | 7 |
| 2 | 1 | 2 | 14 |
| 3 | 1 | 3 | 25 |
| 4 | 1 | 4 | 40 |
| 5 | 1 | 5 | 22 |
| 6 | 2 | 1 | 10 |
| 7 | 2 | 2 | 17 |
| 8 | 2 | 3 | 21 |
| 9 | 2 | 4 | 32 |
| 10 | 2 | 5 | 15 |
| 11 | | | |





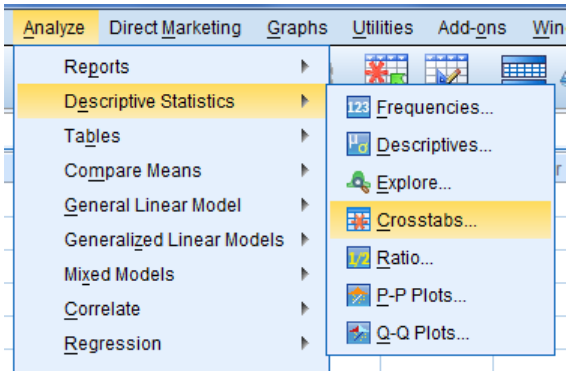

ومن ثم تحديد المتغير count على أنه يمثل تكرارات لكل تقدير.

Data→Weight Cases

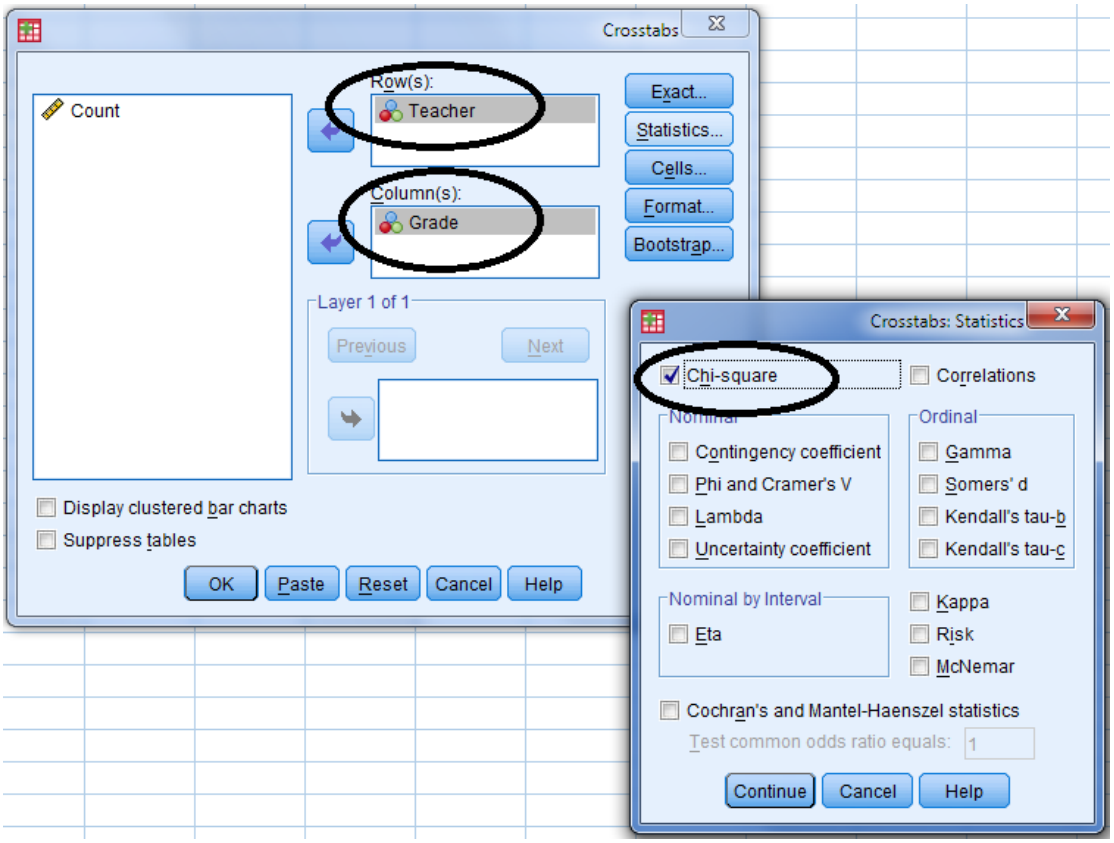


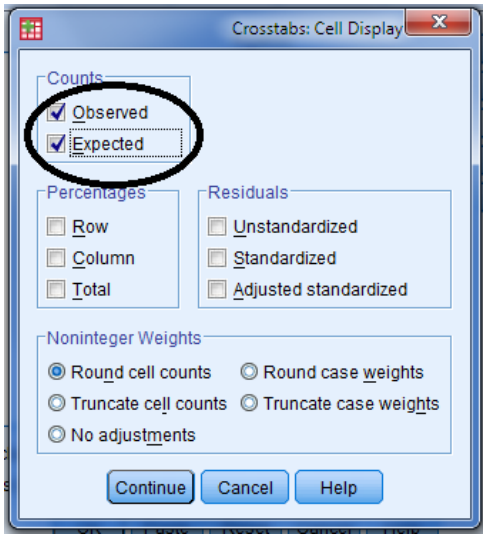

أي أن البرنامج يتعرف على أن عدد الأشخاص الحاصلين على تقدير A عند الدكتور علي يساوي 7 طلاب، وعدد الحاصلين على تقدير B عند الدكتور علي يساوي 14، وهكذا.

ولأن لدينا جدول تقاطعي، فإنه سيتم اختبار الفرضية على النحو التالي:



ومن ثم تحديد المتغيرات وتحديد الاختبار المطلوب من الخيار (Statistics) وكذلك تحديد النتائج في خلايا الجدول من الخيار (Cell) على النحو التالي:



والجداول التالية توضح نتائج اختبار كاي المربع.

**Teacher * Grade Crosstabulation**

| | | | Grade | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | A | B | C | D | F | Total |
| Teacher | Ali | Count | 7 | 14 | 25 | 40 | 22 | 108 |
| | | Expected Count | 9.0 | 16.5 | 24.5 | 38.3 | 19.7 | 108.0 |
| | Ibrahim | Count | 10 | 17 | 21 | 32 | 15 | 95 |
| | | Expected Count | 8.0 | 14.5 | 21.5 | 33.7 | 17.3 | 95.0 |
| Total | | Count | 17 | 31 | 46 | 72 | 37 | 203 |
| | | Expected Count | 17.0 | 31.0 | 46.0 | 72.0 | 37.0 | 203.0 |

**Chi-Square Tests**

| | Value | df | Asymp. Sig. (2-sided) |
|---|---|---|---|
| Pearson Chi-Square | 2.559[a] | 4 | .634 |
| Likelihood Ratio | 2.561 | 4 | .634 |
| Linear-by-Linear Association | 2.322 | 1 | .128 |
| N of Valid Cases | 203 | | |

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.96.

وبذلك فإن قيمة كاي المربع تساوي 2.559 و قيمة p-value = 0.634، وبذلك فإنه لا يمكن رفض فرضية العدم. مما يعني أن توزيع التقديرات عند الدكتور علي متجانس مع توزيع الدرجات عند الدكتور إبراهيم.

تحسب قيمة كاي المربع (إحصائية الاختبار) باستخدام العلاقة التالية:

$$\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{5}\frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

ملاحظة: يجب تحديد الخطوات الخمس لاختبار الفرضيات الاحصائية، وترفض فرضية العدم في حالة أن p-value أقل من 0.05.

مثال 3 : اختبار الاستقلالية

الجدول التالي يوضح أعداد الموظفين مصنفين حسب مستوى التعليم وحسب تواجد عادة التدخين لديهم.

| | غير متعلم | أقل من جامعي | جامعي أو أعلى | المجموع |
|---|---|---|---|---|
| مدخن | 10 | 9 | 5 | 24 |
| غير مدخن | 13 | 7 | 6 | 26 |
| المجموع | 23 | 16 | 11 | 50 |

ويمكن بناء على نتائج العينة معرفة ما إذا كانت مستوى التعليم يؤثر على تواجد عادة التدخين لدى الموظف أو أن تواجد عادة التدخين من عدمه يدل على مستوى التعليم لدى الموظف. أي أننا نرغب في معرفة ما أذا كان هناك تأثير متبادل بين التعليم وتواجد عادة التدخين أو أنهما متغيران مستقلان، أي أنهما لا يتأثران ببعضهما البعض.

ويمكن صياغة الفرضية على النحو التالي:

فرضية العدم : المتغيران مستقلان (مستوى التعليم ، والتدخين)

الفرضية البديلة : المتغيران غير مستقلان

ويمكن إدخال البيانات بنفس الطريقة المتبعة في المثال رقم 2.

وكذلك يمكن اختبار الاستقلالية، بنفس طريقة اختبار التجانس في مثال رقم 2.

إلا أنه يجب التنويه هنا أننا في حالة اختبار الاستقلالية، نهتم بطريقة توزيع المجموع الكلي لعناصر العينة على الخلايا، في حين أننا في اختبار التجانس نهتم بتوزيع المجاميع الهامشية للمتغير المستقل على مستويات التصنيف للمتغير التابع.

والجدول التالي يوضح الاعداد المشاهدة والاعداد المتوقعة للموظفين، أما الجدول الأخير فيوضح نتائج اختبار كاي المربع.

**Smoking * Education Crosstabulation**

| | | | Education | | | |
|---|---|---|---|---|---|---|
| | | | Uneducated | Average | High | Total |
| Smoking | NO | Count | 10 | 9 | 5 | 24 |
| | | Expected Count | 11.0 | 7.7 | 5.3 | 24.0 |
| | Yes | Count | 13 | 7 | 6 | 26 |
| | | Expected Count | 12.0 | 8.3 | 5.7 | 26.0 |
| Total | | Count | 23 | 16 | 11 | 50 |
| | | Expected Count | 23.0 | 16.0 | 11.0 | 50.0 |

**Chi-Square Tests**

| | Value | df | Asymp. Sig. (2-sided) |
|---|---|---|---|
| Pearson Chi-Square | .653[a] | 2 | .721 |
| Likelihood Ratio | .654 | 2 | .721 |
| Linear-by-Linear Association | .073 | 1 | .787 |
| N of Valid Cases | 50 | | |

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 5.28.

وبناء على نتائج الاختبار، فإنه لا يمكن رفض فرضية العدم، مما يدل على أن المتغيران مستقلان وأن التعليم وعادة التدخين لا يؤثران في بعضهما البعض.