

استخدام مربع كاي لحسن المطابقة و جدول التجانس

(١٢ - ١) مقدمة

استخدمنا في الفصول السابقة اختبار ص (التوزيع المعتدل) لاختبار تساوي وسطين أو تساوي نسبتين وذلك في العينات الكبيرة كما استخدمنا اختبار تي لاختبار تساوي وسطين للعينات الصغيرة، وذلك عندما تكون البيانات المدروسة كمية. ولكن إذا كان المطلوب اختبار البيانات لأكثر من مجموعتين أو إذا كانت بعض أو كل البيانات المدروسة وصفية فإنه لا يمكن استخدام الاختبارات السالفة الذكر. لذلك فإنه لا بد من استحداث بعض الاختبارات المناسبة لمعالجة مثل هذه الأوضاع.

في هذا الفصل سنتعرض لدراسة أحد الاختبارات المشهورة وهو المسمى اختبار مربع كاي. يعتبر اختبار مربع كاي من الاختبارات الإحصائية غير المعلمية لأنه لا يعتمد على طبيعة التوزيعات التي تتبعها البيانات المدروسة أو صيغ التوزيعات الاحتمالية التي تحكمها. وما يزيد أهمية استخدام مربع كاي في الإحصاء التطبيقي هو تحديد الصيغة الاحتمالية لتوزيع مربع كاي ووجود جداول رياضية لها مثل جدول (٤) في نهاية هذا الكتاب.

(١٢ - ١ - ١) فكرة توضيحية عن استخدام مربع كاي

لتكن لدينا تجربة لها الحوادث الشاملة ١، ٢، ٣، ... ، ن حيث إن التكرارات المشاهدة لهذه الحوادث هي مش_١، مش_٢، مش_٣، ... ، مش_ن والتكرارات المتوقعة لهذه

الحوادث هي مت_١، مت_٢، ... ، مت_ن على الترتيب، كما هو موضح في الجدول التالي.

التكرارات المشاهدة للحوادث والتكرارات المتوقعة لها

الحادثة	١	٢	ن
التكرار المشاهد (مش)	مش _١	مش _٢	مش _ن
التكرار المتوقع (مت)	مت _١	مت _٢	مت _ن

وغالبا ما تتركز الدراسة في معرفة ما إذا كانت التكرارات المشاهدة تختلف عن التكرارات المتوقعة حسب قيمة معنوية معينة. وتحسب قيمة مربع كاي التي يرمز لها بالرمز ك^٢ كما يلي:

$$\begin{aligned}
 & \frac{(مش_١ - مت_١)^2}{مت_١} + \dots + \frac{(مش_٢ - مت_٢)^2}{مت_٢} + \dots + \frac{(مش_ن - مت_ن)^2}{مت_ن} = ك^٢ \\
 & \text{مجم} = \frac{(مش - مت)^2}{مت} \dots \dots \dots (١)
 \end{aligned}$$

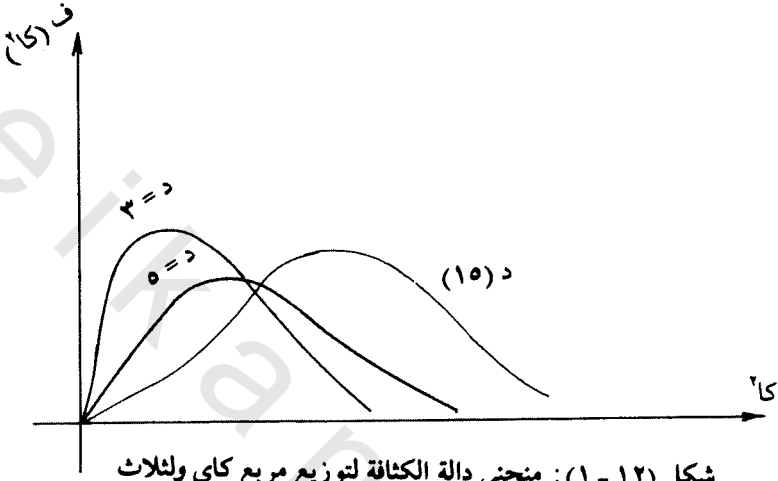
وهذه القيمة تحدد مدى التفاوت بين التكرارات المتوقعة والتكرارات المشاهدة، يلاحظ أن قيمة ك^٢ تساوي صفراً إذا تساوت كل قيمة مشاهدة بالقيمة المتوقعة المناظرة لها وتزداد قيمتها بازدياد الفرق بين القيم المشاهدة والقيم المتوقعة. إذا كان مجموع التكرارات الكلي يساوي ن فإن:

$$\text{مجم} = مش = مت = ن \dots \dots \dots (٢)$$

ويتلخص الاختبار بمقارنة القيمة المسحوبة بالعلاقة (١) السابقة مع القيمة المستخرجة من الجدول لتوزيع كاي الذي تعطى دالة الكثافة الاحتمالية بالصيغة التالية:

$$\text{ف (ك^٢) = ك (ك^٢)} \frac{٢-د}{٢} \text{ هـ} \frac{٢-ك^٢}{٢} ، ك^٢ < \text{صفر}$$

ويحدد المقدار (د) درجات الحرية، ك ثابت يعتمد على د بحيث تكون المساحة تحت منحنى الدالة ف (ك^٢) تساوي الوحدة، كما أن د تحدد شكل منحنى الدالة ف (ك^٢) كما هو موضح بالشكل التالي :



شكل (١٢ - ١) : منحنى دالة الكثافة لتوزيع مربع كاي ولثلاث

درجات حرية د = ٣ ، ٥ ، ١٥

وتحدد درجات الحرية كما يلي :

- ١ - تكون د = ن - ١ إذا لم نحتاج في حساب القيم المتوقعة إلى تقدير أية معالم من معالم المجتمع المدروس وقد طرحنا ١ من ن وذلك نظراً لوجود القيد (٢) الذي يعني أن معرفة ن - ١ من التكرارات المتوقعة يكفي لتحديد التكرار الباقي .
- ٢ - تكون درجات الحرية د = ن - ١ - م إذا كان لا يمكن حساب التكرارات المتوقعة إلا في حالة تقدير م من معالم المجتمع .

وسنوضح في المثالين التاليين طريقة استخدام اختبار مربع كاي لاختبار حسن المطابقة .

مثال (١)

رميت زهرة نرد تسعين مرة وكان التوزيع التكراري لظهور الأرقام من ١ إلى ٦

هي كما يلي :

تكرار ظهور أوجه زهرة النرد

الوجه الظاهر	١	٢	٣	٤	٥	٦
عدد مرات ظهوره	١٥	٧	١٥	٨	٣٠	١٥

والمطلوب فحص، إذا كانت زهرة النرد متزنة أم لا.

الحل

من البديهي أننا نتوقع أنه عند رمي زهرة نرد متزنة تسعين مرة فإن كل وجه يظهر بنفس الاحتمال وبالتالي بنفس عدد المرات أو ١٥ مرة وهو عدد المرات المتوقعة أو النظرية لظهور أي رقم. ولإجراء اختبار مربع كاي نجري الخطوات التالية:

- أولاً: نحدد عدد المرات المشاهدة (مش) لظهور أي رقم.
 ثانياً: نحدد عدد المرات المتوقعة (مت) أو النظرية لكل رقم.
 ثالثاً: نحسب الفرق بين القراءة المشاهدة والقراءة المتوقعة (مش - مت)، وكذلك مربع الفرق (مش - مت)^٢.
 رابعاً: نقسم (مش - مت)^٢ لكل رقم على مت المناظرة له.
 خامساً: نجمع المقادير الناتجة من «رابعاً» أي: $\frac{(مش - مت)^2}{مت}$
 ومن المثال الحالي نجد ذلك حسب الجدول التالي:

رقم حجر النرد	١	٢	٣	٤	٥	٦	المجموع
مش	١٥	٧	١٥	٨	٣٠	١٥	٩٠
مت	١٥	١٥	١٥	١٥	١٥	١٥	٩٠
(مش - مت)	٠	٨-	٠	٧-	١٥	٠	٠
$\frac{(مش - مت)^2}{مت}$	٠	٤,٣	٠	٣,٣	١٥	٠	٢٢,٦

وبالتالي فإن

$$\chi^2 = \frac{(\text{مش} - \text{مت})^2}{\text{مت}} = 22,6$$

ويتبع هذا المجموع توزيع مربع كاي أو χ^2 تحت الفرضية الأولية بأن حجر النرد متزن. ويعتمد كاي على ثابت أو معلمة يمكن تحديدها أو تمثيلها بعدد درجات الحرية. وفي هذه الحالة فإن القراءات الست ليست مستقلة تماماً عن بعضها، ففي القراءات المشاهدة يجب أن يكون مجموع الفروق بين القراءات ووسطها الحسابي مساوياً للصفر. وبالتالي سيكون لدينا $6 - 1 = 5$ أزواج من القراءات المستقلة أو درجات الحرية التي عن طريقها يمكن إيجاد قيمة كاي من الجدول.

في الواقع يمكن النظر إلى هذه المسألة على صورة تعبئة أو ملء ست خلايا تحت شرط واحد بأن مجموع قراءاتها يساوي تسعين، وبالتالي سيكون لدينا ستة خيارات مطروحة منها شرط واحد وتساوي 5 درجات للحرية.

وبذلك لا بد أن يكون $\chi^2 = \frac{(\text{مش} - \text{مت})^2}{\text{مت}} = \text{كاي} (5)$ أو مربع كاي بخمس درجات للحرية والمستوى معنوي 5% نأخذ أحد القرارين التاليين:

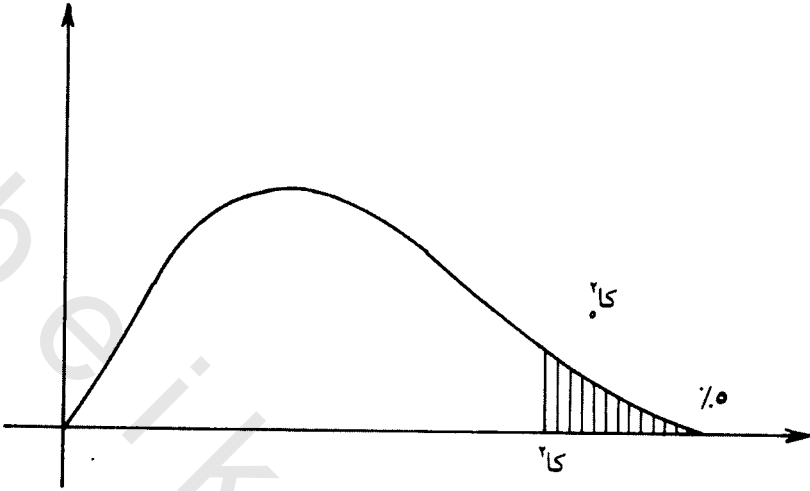
أولاً: نرفض الفرضية الأولية إذا كانت

$$\chi^2 \geq \text{كاي} (5) \dots (5)$$

ثانياً: نقبل الفرضية الأولية إذا كانت

$$\chi^2 < \text{كاي} (5) \dots (5)$$

تحدد المنطقة المحرجة عادة بنهاية المساحة تحت منحنى كاي كما في الشكل التوضيحي التالي:



شكل (١٢ - ٢): المنطقة المظللة تمثل منطقة الرفض لفرض العدم

ومن الجدول نجد أن $كأ = ١١,٠٧ = (٥)$ وهي أقل بكثير من $٢٢,٦$ وبالتالي فإن القيمة المحسوبة لمربع كاي عالية المعنوية وبالتالي فإن الفرضية الأولية مرفوضة أي أن حجر النرد غير متزن.

مثال (٢)

رميت قطعة نقدية مئة مرة. ظهرت صورة في ٦٠ مرة وكتابة في ٤٠ مرة والمطلوب اختبار إذا كانت القطعة النقدية متزنة تحت مستوى ٥% .

الحل

من المعروف أنه إذا كانت القطعة النقدية متزنة فإن:

$$ح (ص) = \frac{1}{٢} ، ح (ك) = \frac{1}{٢} ،$$

وبالتالي فإن عدد ظهور الصور أو الكتابة لا بد أن يساوي $١٠٠ \times \frac{1}{٢} = ٥٠$.

والآن نجري الحسابات التالية كما في الجدول

التكرارات المشاهدة لوجهي القطعة النقدية وتكراراتها المتوقعة

المجموع	ك	ص	
١٠٠	٤٠	٦٠	مش
١٠٠	٥٠	٥٠	مت
٠	١٠-	١٠	(مش - مت)
٤	٢	٢	$\frac{(مش - مت)^2}{مت}$

ومن ذلك نلاحظ أن :

$$كا^2 = \frac{(مش - مت)^2}{مت} = \frac{(٥٠ - ٦٠)^2}{٥٠} + \frac{(٥٠ - ٤٠)^2}{٥٠} = ٢ + ٢ = ٤$$

ولتحديد درجات الحرية فإن لدينا زوجين من القراءات والشرط الوحيد هو أن مجموعها ١٠٠ وبذلك فإننا نجد أن قيمة كا^٢... (١) المناظر لدرجة حرية واحدة، والتي تساوي من الجدول كا^٢... (١) = ٣,٨٤ وبالتالي فإننا نرفض الفرضية الأولى بأن القطعة النقدية متزنة.

(١٢ - ٢) اختبار حسن المطابقة لتوزيع ذي الحدين

ندرس في هذا الفصل استخدام اختبار كا^٢ لفحص ما إذا كانت الاحتمالات المشاهدة من توزيع ذي الحدين أم لا، وذلك عن طريق إيجاد القراءات المشاهدة والقراءات المتوقعة التي يمكن حسابها بضرب الاحتمال المتوقع في عدد المرات أو التكرارات، حيث إن دالة الثقل الاحتمالي لتوزيع ذي الحدين تعطى بالعلاقة

ح (س) = $\binom{n}{س} p^س q^{n-س}$ حيث س = صفر، ١، ٢، ...، ن

حيث إن ن عدد المحاولات أو التجارب، ح احتمال النجاح في كل محاولة أو تجربة،

ل = ١ - ح ولتوضيح استخدام اختبار كا^٢ في حالة التوزيع ذي الحدين نورد المثالين التاليين.

مثال (٣)

لنفرض أنه في أحد التجارب التي أعيدت مئة مرة كانت النتائج كما يلي:
تكرار المشاهدات لقيم متغير عشوائي س

المتغير س	صفر	١	٢	٣	٤	٥	٦	٧	المجموع
مش	١٦	صفر	٣٠	٢٠	٢٤	١٠	صفر	صفر	١٠٠

حيث إن القيم من صفر إلى ٧ للمتغير س هي عناصر فضاء العينة أو القيم الممكنة ظهورها. والمطلوب فحص ما إذا كان المتغير العشوائي الذي يحكم نتائج هذه التجربة يتبع توزيع ذي الحدين أم لا.

الحل

من الواضح أنه لا بد من حساب ح لتوزيع ذي الحدين حتى يمكن إيجاد القيم المتوقعة للقراءات، ولأن هذه القيم ليست معطاة في المثال فإننا نلجأ إلى حساب الوسط أولاً، حيث إنه في حالة توزيع ذي الحدين فإن $تو = ن ح$ الذي يمكن تقريبه بوسط العينة وهو $س$. ومن القراءات المعطاة في الجدول فإن:

$$\bar{س} = \frac{\text{مجموع س ك}}{\text{مجموع ك}}$$

وهذه يمكن حسابها كما يلي:

$$\bar{س} = \frac{٠ \times ٧ + ٠ \times ٦ + ١٠ \times ٥ + ٢٤ \times ٤ + ٢٠ \times ٣ + ٣٠ \times ٢ + ٠ \times ١ + ١٦ \times ٠}{٠ + ٠ + ١٠ + ٢٤ + ٢٠ + ٣٠ + ٠ + ١٦}$$

$$٢,٨٢ = \frac{٢٨٢}{١٠٠}$$

ومن ذلك نجد أن

$$\bar{c} = \frac{c}{n}$$

$$0,4 = \frac{2,82}{7} =$$

ومن ذلك نجد احتمال حدوث أي من القيم من صفر إلى ٧ من العلاقة التالية:

$$c(s) = \binom{7}{s} (0,4)^s (0,6)^{7-s}$$

حيث إن

$$s = 0, 1, 2, 3, 4, 5, 6, 7$$

والقيم المتوقعة لأي قيمة للمتغير s هي عبارة عن مجموع عدد المشاهدات في احتمال حدوثه أي ١٠٠ ح (س).

ومن ذلك نجد الجدول التالي

مت	ح (س)	س
٢,٨	٠,٠٢٧٩٩٣٦	٠
١٣,١	٠,١٣٠٦٣٦٧	١
٢٦	٠,٢٦١٢٧٣٦	٢
٢٩	٠,٢٩٠٣٠٤	٣
١٩,٤	٠,١٩٣٥٣٦	٤
٧,٧٤	٠,٠٧٧٤١٤٤	٥
١,٧٢	٠,٠١٧٢٠٣٢	٦
٠,١٦	٠,٠٠١٦٣٨٤	٧
١٠٠	١	المجموع

في الواقع لا بد أن يكون مجموع القراءات المتوقعة يساوي مئة ولكن لتقريب الكسور فإن المجموع الحالي يساوي ٩٩,٩٢ (≈ 100).

يلاحظ أن القراءة المتوقعة الأولى أقل من ٥ ، وكذلك بالنسبة للقيمتين المتوقعتين الأخيرتين لذا نضيف مثل هذه القيم إلى القيم المجاورة لها ونجري نفس الإضافة بالنسبة للقراءات المشاهدة لنحصل على الجدول التالي :

٧,٦٥	٤	٣	٢	١,٥	س
١٥	٢٤	٢٥	٣٥	١٦	مش
٩,٦٢	١٩,٤	٢٩	٢٦	١٥,٩	مت

وتكون قيمة مربع كاي المشاهدة هي :

$$K^2 = \frac{\sum (\text{مش} - \text{مت})^2}{\text{مت}}$$

وبالتعويض يكون :

$$K^2 = \frac{(9,62-15)^2}{9,62} + \frac{(19,4-24)^2}{19,4} + \frac{(29-25)^2}{29} + \frac{(26-35)^2}{26} + \frac{(15,9-16)^2}{15,9} = 4,51$$

ومن جدول مربع كاي تحت مستوى ٥٪ حيث إن درجات الحرية هي عدد أزواج القراءات مطروحاً منها عدد الشروط المفروضة على المتغير العشوائي المراد اختباره . أصبح عدد خلايا مربع كاي خمساً فقط كما في الجدول الأخير ولوجود شرطين هما :

(أ) مجموع القراءات والمشاهدات يساوي ١٥٥

(ب) أن يكون متوسط القراءات $\bar{س} = 2,82$

ومن ذلك يكون عدد درجات الحرية هو $3 = 2 - 5$

وبالتالي تصبح قيمة مربع كاي تحت مستوى ٥٪ من جدول (٤) في نهاية الكتاب هي :

$$٧,٧٢ = (٣)$$

نستنتج من ذلك أن المقدار $\chi^2 = ٤,٥١$ غير معنوية لرفض الفرضية الأولية أو أن المتغير العشوائي المعطى في المثال يتبع توزيع ذي الحدين .

يلاحظ أنه في حالة أن كون قيمة ح معطاة ولا نحتاج إلى تقديرها من قيمة س فإن عدد الشروط المفروضة تصبح واحدًا فقط، وهو أن يكون مجموع المشاهدات ثابتًا .

مثال (٤)

رمى أربع قطع نقدية ٢٠٠ مرة وكان عدد الصورة الظاهر كما في الجدول

التالي :

تكرارات الصور عند رمي أربع قطع نقدية ٢٠٠ مرة

عدد الصور	٠	١	٢	٣	٤
عدد المرات	١٦	٢٥	٧٠	٧٥	١٤

والمطلوب اختبار ائزان الأربع قطع النقدية .

الحل

نفترض في البداية أن الفرضية الأولية هي أن القطع النقدية متزنة ونستخدم توزيع ذي الحدين لتوليد القيم المتوقعة لقراءات مثل هذه التجربة .

واتزان أي قطعة يعني أن $ح = \frac{1}{4}$ وبالتالي نحتاج إيجاد كيفية توزيع ٢٠٠ رمية بين

عدد الصور .

$$ح (س) = ح (س = س) = ح (س) = \left(\frac{1}{4}\right)^س \left(\frac{1}{4}\right)^{٤-س}$$

$$س = ٠, ١, ٢, ٣, ٤$$

حيث

والقيمة المتوقعة في كل مرة هي مت (س) = ن ح ، وتحسب كالتالي

$$\begin{aligned} \text{ح } (0) &= \left(\frac{1}{4}\right) = 0.25 & \text{مت } (0) &= \left(\frac{1}{4}\right) 200 = 50 \\ \text{ح } (1) &= \left(\frac{1}{4}\right) 4 = 1 & \text{مت } (1) &= \left(\frac{1}{4}\right) (4) 200 = 200 \\ \text{ح } (2) &= \left(\frac{1}{4}\right) 6 = 1.5 & \text{مت } (2) &= \left(\frac{1}{4}\right) (6) 200 = 300 \\ \text{ح } (3) &= \left(\frac{1}{4}\right) 4 = 1 & \text{مت } (3) &= \left(\frac{1}{4}\right) (4) 200 = 200 \\ \text{ح } (4) &= \left(\frac{1}{4}\right) = 0.25 & \text{مت } (4) &= \left(\frac{1}{4}\right) 200 = 50 \end{aligned}$$

ومن ذلك نجد أن

عدد الصور	٠	١	٢	٣	٤
مش	١٦	٢٥	٣٠	٧٥	١٤
مت	١٢,٥	٥٠	٧٥	٥٠	١٢,٥

ومن ذلك نجد

$$\text{كا}^2 = \frac{\sum (\text{مش} - \text{مت})^2}{\text{مت}}$$

$$\therefore \text{كا}^2 = \frac{(12,5-14)^2}{12,5} + \frac{(50-75)^2}{50} + \frac{(75-30)^2}{75} + \frac{(50-25)^2}{50} + \frac{(12,5-16)^2}{12,5}$$

$$\therefore \text{كا}^2 = 53,16$$

وبمقارنة هذه القيمة مع قيمة مربع كا^٢ في مستوى ٥٪ وبدرجات حرية عددها ٥ (عدد الخلايا) - ١ (عدد الشروط أو مجموع الرميات) يساوي ٤ في جدول مربع كاي في نهاية الكتاب نجد أن:

$$\text{كا}^2_{0,05} = (4) = 9,49$$

أي أن توزيع ذي الحدين لا يعتبر تطابقه حسناً لتوزيع العينة المعطى أي أن القطع غير متزنة كما سبق أن فرضنا في البداية .

(١٢ - ٣) اختبار حسن المطابقة لتوزيع بواسون

يعتبر توزيع بواسون من التوزيعات المهمة في دراسة العديد من الظواهر العشوائية كما سبق أن أشرنا عند دراسة بعض التوزيعات الإحصائية وكثيراً ما تواجهنا معلومات أو بيانات، ونود التأكد فيما إذا كانت تتبع توزيع بواسون أم لا .

عادة تكون المشاهدات أو التوزيعات الفعلية المشاهدة معطاة سواءً من التجارب أو من أي ظاهرة طبيعية مثلاً . وكل ما نفعله هو إيجاد توزيع بواسون المناظر ومن ثم إيجاد القيم المتوقعة أو المقدرة للتوزيع التكراري نظرياً، ومن ثم نستخدم علاقة مربع كاي المعتادة في الصيغة $\chi^2 = \sum \frac{(مش - مت)^2}{مت}$. وعلى خلاف ما درسنا في توزيع ذي الحدين فإن لتوزيع بواسون حالة واحدة (كانت ح أحياناً مجهولة أو معلومة في توزيع ذي الحدين). ويمكن توليد توزيع بواسون النظري إذا علم متوسط التوزيع وتو مجموع التكرارات الكلي، وبالتالي يوجد شرطان في كل استخدامات توزيع بواسون لحسن المطابقة . لتوضيح ذلك نورد المثال التالي .

مثال (٥)

اختبر حسن مطابقة توزيع بواسون للتوزيع التكراري المعطى بالجدول التالي :

تكرارات متغير بواسون العشوائي

س	٠	١	٢	٣	٤	٥	٦ أو أكثر
ك	١٩	٢٦	٢٧	١٣	١١	٢	٠

يجب أن نحسب أولاً قيمة χ^2 أو χ (معلم توزيع بواسون) كما يلي

$$\chi^2 = ٩٨ ، \quad \chi^2_{س ك} = ١٧٣$$

فإن:

$$1,765 = \frac{173}{98} = \bar{س}$$

وبالتعويض في صيغة بواسون نجد أن:

$$ح(س) = \frac{س^{-1} (1,765)^س e^{-1,765}}{س!}$$

وبالتعويض عن قيم $س = 0, 1, 2, 3, 4, 5, 6$ والضرب في مجموع التكرارات ٩٨ نحصل على القيم النظرية المتوقعة.

حيث إن مت (س) = ٩٨ ح (س)

ومن ذلك نجد أن:

$$\begin{array}{l} \text{مت (0)} = 16,78 \\ \text{مت (1)} = 29,61 \\ \text{مت (2)} = 26,13 \\ \text{مت (3)} = 15,37 \\ \text{مت (4 أو أكثر)} = 3,33 \end{array}$$

في الواقع أمكن إيجاد مت (٥ أو أكثر) كما يلي:

$$\text{مت (5 أو أكثر)} = 98 - (\text{مت (0)} + \text{مت (1)} + \text{مت (2)} + \text{مت (3)} + \text{مت (4)})$$

ومن الواضح أن القيمة المتوقعة الأخيرة أقل من ٥ وبالتالي لا بد من إضافتها إلى القيمة السابقة لها فيكون لدينا مايلي:

س	٠	١	٢	٣	٤ أو أكثر
مت	١٦,٨	٢٩,٦١	٢٦,١	١٥,٣٧	١٠,١
مش	١٩	٢٦	٢٧	١٣	١٣

وبحساب قيمة مربع كاي نجد أن

$$\begin{aligned} \chi^2_{\text{كأ}} &= \frac{\sum (f_o - f_e)^2}{f_e} = \frac{(26,1 - 27)^2}{26,1} + \frac{(29,61 - 26)^2}{29,61} + \frac{(16,8 - 19)^2}{16,8} \\ &+ \frac{(10,1 - 13)^2}{10,1} + \frac{(15,37 - 13)^2}{15,37} + \\ &= 1,964 \end{aligned}$$

حيث إن عدد درجات الحرية تساوي عدد الخلايا ٥ مطروحًا منه عدد الشروط ٢ أي تساوي ٣ ومن جدول مربع كاي (رقم ٤) في نهاية الكتاب وتحت مستوى ٥٪ كا = ٧,٨١ (٣) . . . أي أن القيمة الناتجة ١,٩٦٤ أقل من ٧,٨١ وبالتالي ليست معنوية لرفض الفرضية الأولى بأن القراءات تتبع توزيع بواسون.

(١٢ - ٤) اختبار حسن المطابقة للتوزيع الطبيعي

بنفس الطريقة التي اتبعناها في حساب حسن المطابقة في توزيعي ذي الحدين وبواسون يمكن اختبار حسن مطابقة التوزيع الطبيعي لبعض القراءات أو البيانات التي تواجهنا. ويختلف حساب درجات الحرية عن التوزيعين السابقين لأنه لا بد من تقدير كل من \bar{x} و s^2 أي الوسط والتباين في كل مرة وكذلك تحديد مجموع التكرارات أي أنه توجد ثلاثة شروط في حالة التوزيع الطبيعي ولتوضيح كيفية اختبار حسن المطابقة نورد المثال التالي.

مثال (٦)

بين مدى مطابقة التوزيع الطبيعي لبيانات الجدول التالي :

تكرارات المتغير العشوائي للتوزيع الطبيعي

فئة س	أقل من ١٥	١٥ - ٢٠	٢٠ - ٢٥	٢٥ - ٣٠	٣٠ - ٣٥	أكثر من ٣٥
التكرار	٣	٧	١٥	٢٠	٩	٤

الحل

لحساب ذلك نوجد أولاً الوسط $\bar{س}$ والانحراف المعياري σ بالطرق التي درسناها سابقاً عند دراسة التوزيعات التكرارية.

وجدنا أن $\bar{س} = ٢٥,٧$ و $\sigma = ٦,١٤$ ، حيث اعتبرنا أن الحد الأدنى للفتنة الأولى ١٠ والحد الأعلى للفتنة الأخيرة ٤٠ ($\bar{س} = \frac{\text{مجمك س}}{\text{مجمك}}$ حيث س هي مركز الفتات).

نعين الحدود العليا للفتات ومن ثم نوجد القيم المعيارية لها ولتكن ص ومن ثم نجد الاحتمالات المناظرة لها أوح ($ص \geq ص$) ونحسب من ذلك الاحتمال والتكرار المتوقع لتلك الفتة كما في الجدول.

الفتنة	أقل من ١٥	١٥ - ٢٠	٢٠ - ٢٥	٢٥ - ٣٠	٣٠ - ٣٥	أكثر من ٣٥
الحد الأعلى للفتنة	١٥	٢٠	٢٥	٣٠	٣٥	
القيمة المعيارية للحد الأعلى للفتنة	-١,٧٤	-٠,٩٣	٠,١١	٠,٧٠	١,٥١	-
ح ($ص \geq ص$)	٠,٠٤٠٩	٠,١٧٦٢	٠,٤٥٦٢	٠,٧٥٨٠	٠,٩٣٤٥	١
الاحتمال ح	٠,٠٤٠٩	٠,١٣٥٣	٠,٢٨٠٠	٠,٣٠١٨	٠,١٧٦٥	٠,٠٦٥٥
التكرار المتوقع ح × مجك	٢,٤	٧,٩	١٦,٢	١٧,٥	١٠,٢	٣,٨

يلاحظ أنه لا بد من دمج الفتتين الأولى والثانية وكذلك الفتتين الخامسة والسادسة لأن القيمة المتوقعة في الفتة الأولى والأخيرة أقل من ٥.

لتوضيح طريقة الحساب فإن القيم المعيارية للحدود العليا للفتات نحصل عليها كما يلي:

$$ص = \frac{\bar{س} - س}{\sigma}$$

فمثلاً:

$$\text{ص للفتة الأولى} = \frac{25,7 - 15}{6,14} = 1,74-$$

وهكذا

$$\text{ص للفتة الثانية} = \frac{25,7 - 20}{6,14} = 0,93-$$

أما الصف الرابع وهو الاحتمالات فنجدها من جدول التوزيع الطبيعي المعتاد أما احتمالات الصف الخامس فهي كما يلي:

$$\text{الاحتمال الأول} = 0,0409$$

$$\text{الاحتمال الثاني} = 0,1762 - 0,0409 = 0,1353$$

$$\text{الاحتمال للفتة الثالثة} = 0,4562 - 0,1762 = 0,2800 \text{ . وهكذا}$$

والصف الأخير هو حاصل ضرب كل احتمال في مجموع التكرارات ٥٨ .

وبعد دمج الفئات المشار إليها نحصل على الجدول التالي:

مش	١٠	١٥	٢٠	١٣
مت	١٠,٣	١٦,٢	١٧,٥	١٤

ومن ذلك نحسب قيمة مربع كاي وهي:

$$\text{كا} = \frac{1(14 - 13)^2}{14} + \frac{1(17,5 - 20)^2}{17,5} + \frac{1(16,2 - 15)^2}{16,2} + \frac{1(10,3 - 10)^2}{1,3} = 0,53 =$$

ولوجود أربع خلايا وكذلك لتقدير قيمة كل من المتوسط والانحراف المعياري وكذلك

الالتزام بشرط مجموع التكرارات فإن عدد درجات الحرية = $4 - 3 = 1$ وستكون قيمة مربع كاي تحت مستوى ٥٪ هي :

$$3,84 = (1) \dots \text{كاي}$$

أي أننا نقبل الفرضية الأولى بأن المشاهدات المعطاة في بداية المثال تتبع التوزيع الطبيعي .

(١٢ - ٥) جداول التجانس

في كثير من المسائل العملية والدراسات الإحصائية نحتاج إلى تصنيف البيانات حسب عوامل معينة كما نحدد مقدار العامل في كل مرة مثلاً قد ندرس مستوى النجاح (راسب، جيد، جيد جداً، ممتاز) وعلاقته بجنس الطالب (ذكر، أنثى)، أو علاقة جنس المريض بدرجة حساسيته لمرض معين، أو نوع القمح وعلاقته لنمو أنواع معينة من الفطر فيه . كذلك من الأشياء التي نوردتها كمثال على دراسة الترابط بين العوامل في المجالات الأخرى أيضاً، العلاقة بين تعدد الزوجات والمستوى الاقتصادي للزوج، أو العلاقة بين التدخين للابن والتدخين في حالة أن يكون أحد الوالدين أو كلاهما من المدخنين . الخ .

ولتوضيح فكرة دراسة العلاقة بين عاملين أو أكثر، أو ما يشار إليه أحيانا بموضوع ارتباط العوامل، أو قياس الاستقلال بين العوامل نورد المثالين التاليين .
مثال (٧)

إذا كان في عينة عشوائية مكونة من ١٠٠ طالب وطالبة كانت أعداد الناجحين والراسبين في امتحان الإحصاء التطبيقي كما يلي :

التكرارات المشتركة للطلاب حسب الجنس ونتيجة الإمتحان

الجنس \ النتيجة	ناجح	راسب
طالب	٤٠	١٥
طالبة	٣٥	١٠

ولدراسة العلاقة بين جنس الطلبة ونتائجهم في الامتحان أو أن هذين العاملين مستقلان عن بعضهما نوجد أولاً مجموع الصفوف والأعمدة حيث أن جدول التجانس في هذه الحالة هو 2×2 أي أن له صفان وعمودان .

الجنس \ النتيجة	ناجح	راسب	المجموع
طالب	٤٠	١٥	٥٥
طالبة	٣٥	١٠	٤٥
المجموع	٧٥	٢٥	١٠٠

ومن الجدول الأخير نلاحظ أن احتمال أن يكون الشخص طالباً هو ح (طالب) = $\frac{55}{100}$ وهو مجموع تكراري الصف الأول على مجموع التكرارات .

أما احتمال أن يكون الشخص ناجحاً فهو:

ح (ناجح) = $\frac{75}{100}$ وهو مجموع تكرار العمود الأول على مجموع التكرارات وبالمثل يمكن حساب الاحتمالات الأخرى كالتالي :

$$\text{ح (طالبة)} = \frac{45}{100} , \text{ح (راسب)} = \frac{25}{100}$$

وفي البداية نجعل فرضيتنا الأولية وهو أن لا توجد علاقة بين الجنس والنتيجة في الامتحان أو أن الجنس مستقل عن النتيجة، ولاختبار ذلك نوجد أولاً التكرارات المتوقعة للجدول السابق .

نلاحظ أنه لو كان جنس الشخص (طالباً) لا يؤثر على نجاحه فإن

$$\text{ح (طالب وناجح)} = \text{ح (طالب)} \times \text{ح (ناجح)}$$

$$\text{ح (طالب وناجح)} = \frac{75}{100} \times \frac{55}{100}$$

والقيمة المتوقعة هي حاصل ضرب مجموع التكرارات والاحتمال وبالتالي:

$$\text{مت (طالب وناجح)} = \left(\frac{٥٥}{١٠٠}\right) ١٠٠ = \left(\frac{٧٥}{١٠٠}\right)$$

$$\frac{٧٥ \times ٥٥}{١٠٠} =$$

$$٤١,٢٥ =$$

وبنفس الطريقة يمكن حساب بقية القيم المتوقعة فيكون:

$$\text{ح (طالب وراسب)} = \text{ح (طالب) ح (راسب)}$$

$$\text{ح (طالب وراسب)} = \frac{٢٥}{١٠٠} \times \frac{٥٥}{١٠٠} =$$

$$\text{مت (طالب وراسب)} = \frac{٢٥}{١٠٠} \times \frac{٥٥}{١٠٠} \times ١٠٠ =$$

$$\frac{٢٥ \times ٥٥}{١٠٠} =$$

$$١٣,٧٥ =$$

$$\text{مت (طالبه وناجحة)} = \frac{٧٥}{١٠٠} \times \frac{٤٥}{١٠٠} \times ١٠٠ =$$

$$٣٣,٧٥ =$$

$$\text{مت (طالبه وراسبة)} = \frac{٢٥}{١٠٠} \times \frac{٤٥}{١٠٠} \times ١٠٠ =$$

$$١١,٢٥ =$$

ويكون جدول البيانات المشاهدة والمتوقعة كالتالي:

الجنس \ النتيجة	ناجح	راسب	المجموع
طالب	٤٠	١٥	٥٥
طالبة	٣٥	١٠	٤٥
المجموع	٧٥	٢٥	١٠٠

حيث إن القيم في الجزء الأعلى من الخلية هي القيم المتوقعة .
وبذلك يمكن حساب قيم كا^٢ كما يلي :

$$\text{كا}^2 = \frac{\text{مجم (مش - مت)}^2}{\text{مت}}$$

$$\text{كا}^2 = \frac{^2(11,25-10)}{11,25} + \frac{^2(13,75-15)}{13,75} + \frac{^2(33,75-35)}{33,75} + \frac{^2(41,25-40)}{41,25} = 0,3367 =$$

وبإيجاد قيمة كا^٢ تحت مستوى ٥٪ حيث إن عدد درجات الحرية تساوي عدد الخلايا ٤ وهي ٢ × ٢ مطروحاً منه عدد الشروط وهي ٣ ، وهي مجموع الصفوف والأعمدة حيث إن إعطاء أي مجموع ٣ من صفوف وأعمدة يمكن استنتاج الصف أو العمود الباقي أي ٤ - ٣ = ١ ، أو لو كان عدد الصفوف م والأعمدة ن فإن درجات الحرية تساوي (م - ١) (ن - ١) وفي هذه الحالة

$$\text{درجات الحرية} = (١ - ٢) \times (١ - ٢) = ١$$

ومن ذلك نجد من الجدول أن :

$$\text{كا}^2_{٠,٠٥} (١) = ٣,٨٤$$

ونظراً لأن القراءة المحسوبة لمربع كاي أقل من القراءة الجدولية فإنه لا علاقة بين جنس الشخص ومدى نجاحه في ذلك المقرر.

ملاحظة :

يمكن استخدام نفس الطريقة حتى لو كانت العوامل المدروسة تتوزع بأكثر من صفتين .

(١٢ - ٦) تمارين

- ١ - رميت قطعة نقدية ٢٠٠ مرة، وكانت النتيجة ظهور ١١٥ صورة، و ٨٥ كتابة .
بين ما إذا كانت القطعة متزنة أم لا .

٢ - في خمسين مشاهدة لتغير عشوائي متقطع كانت البيانات المشاهدة والمتوقعة لجميع القيم الممكنة كما يلي:

التكرارات المشاهدة والمتوقعة لتغير عشوائي متقطع

س	صفر	١٠	٢٠	٣٠
مش	١٣	١٠	٢٠	٧
مت	١٠	١٣	١٥	١٢

استخدم اختبار مربع كاي للتأكد من حسن مطابقة توزيع المتغير العشوائي لتلك البيانات إذا علمت بوجود شرطين على البيانات .

٣ - إذا كانت الكتب المستعارة من المكتبة المركزية بجامعة الملك سعود في خمسة أيام في أحد الأسابيع هي:

أعداد الكتب المستعارة حسب أيام العمل الأسبوعية

اليوم	السبت	الأحد	الاثنين	الثلاثاء	الأربعاء
عدد الكتب	٤٥٠	٣٠٠	٤٠٠	٣٥٠	٦٠٠

استخدم اختبار مربع كاي لفحص ما إذا كان هناك زيادة ملموسة للاستعارة في عدد أيام ذلك الأسبوع .

٤ - إذا كانت الأهداف التي سجلها أحد الأندية الرياضية في ٥ مباريات من الدوري العام هي كما يلي:

أعداد الأهداف في خمس مباريات

عدد المباريات	١	٢	٣	٤	٥
عدد الأهداف	٧	٤	٣	١	٠

علق على النتيجة، ومدى مطابقة توزيعي بواسون أو ذي الحدين للبيانات المعطاة، مستخدماً فحص مربع كاي لحسن المطابقة .

- ٥ - إذا كان عدد المكالمات الواردة لأحد المكاتب الحكومية من الساعة الثامنة وحتى الثانية عشرة في ٥٠ يوماً كإيلي:

تكرارات المكالمات في خمسين يوماً

عدد المكالمات	٠	١	٣	٤	٦	٧	أكثر من ٧
عدد الأيام	٣	٢	٩	١٥	١٢	٥	٤

بين مدى مطابقة توزيع بواسون لهذه البيانات .

- ٦ - أطوال ٦٠ طالباً من إحدى المدارس الابتدائية كإيلي:

التوزيع التكراري لأطوال الطلاب

الفئة	أقل من ١١٥	١١٥ - ١٢٠	١٢٠ - ١٣٥	١٣٥ - ١٤٠	أكثر من ١٤٠
التكرار	٥	١٤	٢٠	١٣	٨

بين مدى ملاءمة التوزيع الطبيعي لتمثيل مثل هذه البيانات .

- ٧ - في عينة مكونة من ١٠٠ من الأولاد «أقل من ١٢ سنة» وجد عدد المدخنين منهم الذين أحد والديهم من المدخنين كإيلي:

التكرارات المشتركة لأحد الوالدين والولد حسب عادة التدخين

	الولد	أحد الوالدين
أحد الوالدين	الولد	أحد الوالدين
مدخن	٣٠	١٠
غير مدخن	١٥	٤٥

اختبر العلاقة بين عادة التدخين لدى أحد الوالدين أو كليهما وعادة التدخين عند الأبناء .

- ٨ - في عينة مكونة من ١٥٠ مريضاً بثلاثة أنواع من مرض السرطان كانت البيانات كالتالي:

التكرارات المشتركة لمرض السرطان حسب مكان الإصابة والجنس وعادة التدخين

الجنس والتدخين		ذكر		مكان السرطان	
		غير مدخن	مدخن		
السرطان	غير مدخن	٦	٢٠	١٥	١٩
	مدخن	٧	١١	١٠	٢٢
المعدة	غير مدخن	١٠	٧	١٨	١٥
	مدخن				
الأطراف	غير مدخن				
	مدخن				

ادرس علاقة أنواع السرطان بجنس المريض وبعادة التدخين.

٩ - إذا كانت نسب الأشخاص المتمين لقبيلة ما حسب فصائل الدم الأربع هي ١٥، ٠، ٤٥، ٠، ٢٥، ٠، ١٥، ٠، فإذا كانت التكرارات المشاهدة لعينة من الأشخاص من قبيلة أخرى هي ١٠٠، ١٥٠، ١٥٠، ٣٥٠، ٢٠٠ فناقش فيما إذا كان لأفراد القبيلتين نفس توزيع نسب فصائل الدم.

١٠ - عدد حوادث السيارات المرورية في ١٢ شهراً في إحدى الدول هي كمايلي: ٥٥٠، ٩٠٠، ٦٠٧، ١٣٥٠، ٨٥١، ١٢١٠، ٦١٣، ١٥٢٠، ٨١٢، ٧٢٤، ١١٠٠، ٦٩٠.

بين فيما إذا كانت هذه التكرارات تنسجم مع الافتراض القائل: إن عدد الحوادث الشهرية ثابتة في تلك السنة.

١١ - في تجربة لفحص نظرية مندل للوراثة لأربعة أنواع من سلالات نبات البازلاء ذات البذور الدائرية الصفراء والخضراء، وغير الدائرية الصفراء والخضراء كانت التكرارات المشاهدة للنتيجة كمايلي:

دائرية صفراء = ٣٢٠، دائرية خضراء = ١١٠، غير دائرية صفراء = ١٢٠، غير دائرية خضراء = ٤٠.

بين ما إذا كانت هذه البيانات تنطبق مع نظرية مندل للوراثة، والتي توضح أن هذه النسب من هذه الأنواع من البازلاء هي ٩:٣:٣:١ على الترتيب وذلك باحتمال ٩٥٪.

١٢- الجدول التالي يمثل كميات انتاج أحد آبار البترول بآلاف البراميل في ٢٦٠ يومًا:

توزيع كميات البترول المنتجة بآلاف البراميل في ٢٦٠ يومًا

٥	٢٠	٦٥	٨٠	٦٠	٢٣	٧	عدد الأيام
٦٢	٦١	٦٠	٥٩	٥٨	٥٧	٥٦	كمية البترول المنتجة

بين فيما إذا كانت هذه البيانات تتبع التوزيع الطبيعي .

١٣- في دراسة لفحص تأثير عقارين أ، ب على أحد أنواع الصداع كانت النتائج كما في الجدول التالي:

التكرارات المشتركة لعقارين حسب درجة الشفاء

الحالة / العقار	شفي تمامًا	زاد الصداع	لم يؤثر كليًا
العقار أ	٧٠	١٥	٣٥
العقار ب	٥٥	١٠	١٥

اختبر مدى صحة القول: إن للعقارين نفس التأثير في معالجة ذلك الصداع .