المحاضرة التاسعة : اختبار كاي تربيع

الأهداف :

* تمكين الطالب من فهم أهمية اختبار كاي تربيع بنوعيه .

*اكتساب مهارة إيجاد قيمة كاي تربيع حسابيا و جدوليا .

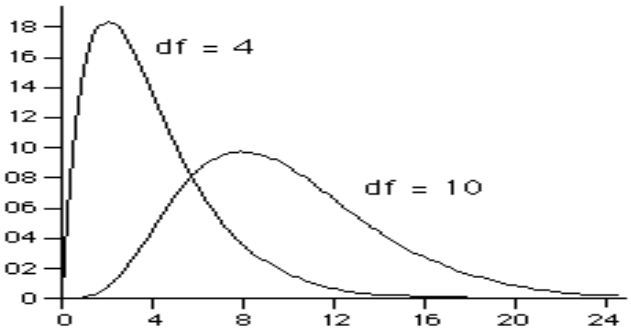
*استيعاب كيفية إجراء اختبار المطابقة و كذلك الاستقلالية .

تمهيد

إن اختبار كاي تربيع هو متغير عشوائي يستخدم في اختبار الفرضيات حيث أنه يصنف إلى اختبارحسن المطابقة (جودة التوفيق) و اختبار الاستقلالية .

يعتبر توزيع كاي تربيع χ^2 من التوزيعات الإحتمالية الشائعة الاستخدام حيث توجد له تطبيقات عديدة بدرجة يمكن معها القول أنه يأتي في المرتبة الثانية للتوزيع المعتدل من حيث كثرة تطبيقاته. و هو يقترب من التوزيع المعتدل كلما ارتفعت درجة الحرية .

χ^2



خصائص اختبار كاي مربع

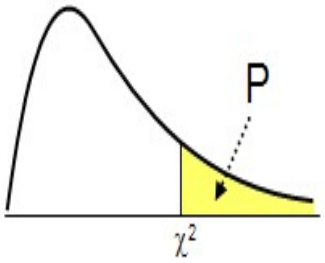
يتميز هذا الاختبار بجملة من المميزات نوجزها فيما يلي :

انه توزيع غير متماثل. -

- انه توزيع غير معرف في الجزء السالب من المستوى-

انه توزيع يبدأ من الصفر ويستمر إلى ما لا نهاية .

أنه توزيع موجب الالتواء (من اليمين).



	P										
DF	**0.995**	**0.975**	**0.20**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
1	0.0000393	0.000982	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.550	10.828
2	0.0100	0.0506	3.219	4.605	5.991	7.378	7.824	9.210	10.597	12.429	13.816
3	0.0717	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.860	16.924	18.467
5	0.412	0.831	7.289	9.236	11.070	12.833	13.388	15.086	16.750	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458
7	0.989	1.690	9.803	12.017	14.067	16.013	16.622	18.475	20.278	22.601	24.322
8	1.344	2.180	11.030	13.362	15.507	17.535	18.168	20.090	21.955	24.352	26.124
9	1.735	2.700	12.242	14.684	16.919	19.023	19.679	21.666	23.589	26.056	27.877
10	2.156	3.247	13.442	15.987	18.307	20.483	21.161	23.209	25.188	27.722	29.588
11	2.603	3.816	14.631	17.275	19.675	21.920	22.618	24.725	26.757	29.354	31.264
12	3.074	4.404	15.812	18.549	21.026	23.337	24.054	26.217	28.300	30.957	32.909
13	3.565	5.009	16.985	19.812	22.362	24.736	25.472	27.688	29.819	32.535	34.528
14	4.075	5.629	18.151	21.064	23.685	26.119	26.873	29.141	31.319	34.091	36.123
15	4.601	6.262	19.311	22.307	24.996	27.488	28.259	30.578	32.801	35.628	37.697
16	5.142	6.908	20.465	23.542	26.296	28.845	29.633	32.000	34.267	37.146	39.252
17	5.697	7.564	21.615	24.769	27.587	30.191	30.995	33.409	35.718	38.648	40.790
18	6.265	8.231	22.760	25.989	28.869	31.526	32.346	34.805	37.156	40.136	42.312
19	6.844	8.907	23.900	27.204	30.144	32.852	33.687	36.191	38.582	41.610	43.820
20	7.434	9.591	25.038	28.412	31.410	34.170	35.020	37.566	39.997	43.072	45.315
21	8.034	10.283	26.171	29.615	32.671	35.479	36.343	38.932	41.401	44.522	46.797
22	8.643	10.982	27.301	30.813	33.924	36.781	37.659	40.289	42.796	45.962	48.268
23	9.260	11.689	28.429	32.007	35.172	38.076	38.968	41.638	44.181	47.391	49.728
24	9.886	12.401	29.553	33.196	36.415	39.364	40.270	42.980	45.559	48.812	51.179
25	10.520	13.120	30.675	34.382	37.652	40.646	41.566	44.314	46.928	50.223	52.620
26	11.160	13.844	31.795	35.563	38.885	41.923	42.856	45.642	48.290	51.627	54.052
27	11.808	14.573	32.912	36.741	40.113	43.195	44.140	46.963	49.645	53.023	55.476
28	12.461	15.308	34.027	37.916	41.337	44.461	45.419	48.278	50.993	54.411	56.892
29	13.121	16.047	35.139	39.087	42.557	45.722	46.693	49.588	52.336	55.792	58.301
30	13.787	16.791	36.250	40.256	43.773	46.979	47.962	50.892	53.672	57.167	59.703
31	14.458	17.539	37.359	41.422	44.985	48.232	49.226	52.191	55.003	58.536	61.098

أولا : اختبار حسن المطابقة

ويهتم هذا النوع من الاختبارات الإحصائية باختبار ما إذا كانت مشاهدات عينة تم اختيارها من مجتمع له توزيع احتمالي معين أو نظرية معينة.

ويستخدم هذا الاختبار عندما تكون البيانات اسمية أو على شكل تكرارات ويقصد بجودة التوفيق هنا مع التكرارات Observedدراسة مدى تشابه تكرارات العينة والتي تسمى عادة بالتكرارات الملاحظة للمتغير موضوع الدراسة فى المجتمع الأصلى.Expectedالمتوقعة

ويستخدم اختبار كا^{2} كطريقة إحصائية للمقارنة بين التكرارين الملاحظ والمتوقع. فإذا كانت العينة ممثلة للمجتمع فى تكراراتها ومتطابقة معه فإن قيمة كا^{2} تكون عادة صفراً وتزداد هذه القيمة لتصبح أكثر من صفر كلما كان هناك فرق بين تكرارات العينة (الملاحظة) وبين تكرارات التوزيع النظرى للمجتمع (المتوقعة).

الغرض منه المقارنة بين التكرارات المتوقعة و التكرارات المشاهدة ، حيث يضع الباحث احتمالا متوقعا لنتائج التجربة في صورة تكرارات متوقعة ثم يقوم بمطابقة النتائج المتوقعة مع النتائج المشاهدة فدور اختبار كاي تربيع هو تحديد ما إذا كان هناك فروق ذات دلالة إحصائية بين ماهو متوقع و ما هو واقع .

في البداية نقوم بحساب قيمة x^2 وفق العلاقة التالية :

$$x^2 = \sum [(o - e)^2 / e]$$

H0 : مجموعة المشاهدات التى تم اختيارها تتبع توزيع احتمالى معين أو نظرية معينة.

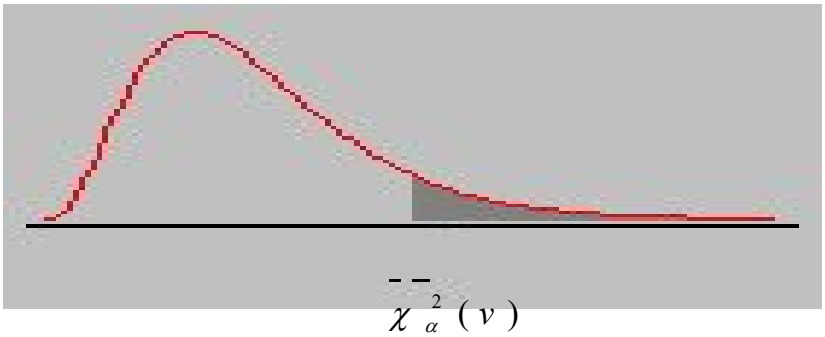
HA : مجموعة المشاهدات التى تم اختيارها لاتتفق مع هذا التوزيع أو نظرية معينة.

احصاء الاختبار:

ويجب أن يكون التكرار P_i المتوقع في أية خلية لا يقل عن 5 حتى يتم حساب إحصائي الاختبار بشكل صحيح.

مناطق الرفض والقبول:

سوف نستخدم جدول مربع كاي لتعيين القيمة المجـــــــدولة (الحرجة) حيث ، و هي عدد المشاهدات النهائية بعد عملية الدمج إن وجدت.



$\chi^2_{\alpha}(\nu)$

القرار:

نقبل فرض العدم إذا كانت (قيمة كاي تربيع المحسوبة أصغر من القيمة المجدولة) أي أن :

$$\chi_0^2 < \chi_\alpha^2(\nu)$$

ونرفض فرض العدم إذا كانت (قيمة كاي تربيع المحسوبة أكبر من القيمة المجدولة) أي أن :

$$\chi_0^2 > \chi_\alpha^2(\nu)$$

مثال:

اختار أحد الباحثين عينة حجمها n=800 شخصا من أحد المدن، وكان توزيعهم حسب فصيلة الدم كالتالي:

فصيلة الدم	A	B	AB	O
عدد الأشخاص (التكرار المشاهد)	200	150	100	350

هل يتفق هذا التوزيع مع توزيع أفراد مدينةأخرى كان توزيع فصيلة دمهم حسب النسب التالية:

فصيلة الدم	A	B	AB	O
النسب المئوية للأشخاص	25%	15%	15%	45%

استخدم مستوى معنوية $\alpha = 0.05$

الحل:

الفروض الإحصائية:

H_0 : توزيع فصيلة الدم في العينة يتفق مع التوزيع المناظر للمدينة الأخرى.

H_A : توزيع فصيلة الدم في العينة لا يتفق مع التوزيع المناظر للمدينة الأخرى.

مستوى المعنوية:

مستوى معنوية α = 0.05

احصاء الأختبار:

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

حيث تمثل O_i التكرار المشاهد للنتيجة رقم i ،

تمثل E_i التكرار المتوقع المناظر للنتيجة رقم i حيث $E_i = np_i$

$$E_1 = np_1 = 800(0.25) = 200$$
$$E_2 = np_2 = 800(0.15) = 120$$
$$E_3 = np_3 = 800(0.15) = 120$$
$$E_4 = np_4 = 800(0.45) = 360$$

ونلاحظ أن جميع المشاهدات المتوقعة أكبر من 5 وأيضا حجم العينة اكبر من 50 ، لذا يمكن تعيين احصاء الاختبار كاي تربيع لاختبار هذه البيانات، وبكتابة كلا من المشاهدات والقيم المتوقعة معا في جدول واحد كالتالي:

فصيلة الدم	A	B	AB	O
عدد الأشخاص (التكرار المشاهد)	200	150	100	350
التكرار المتوقع	200	120	120	360

وبتطبيق معادلة كاي تربيع للحصول على قيمة كاي المحسوبة ويتم ذلك كالتالي:

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \left[\frac{(200-200)^2}{200} + \frac{(150-120)^2}{120} + \frac{(100-120)^2}{120} + \frac{(350-360)^2}{360} \right] = 11.11$$

مناطق الرفض والقبول:

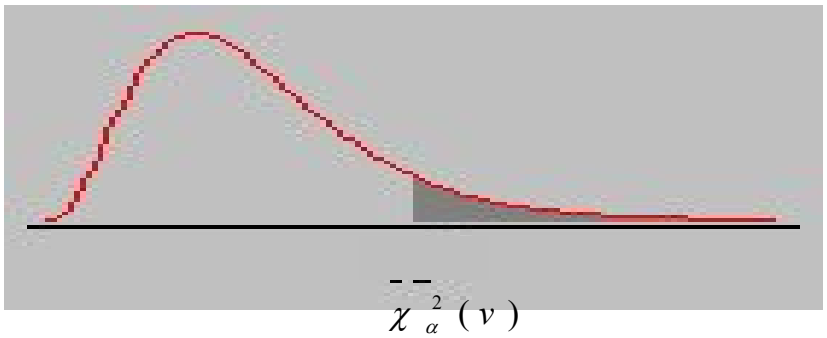
عند استخدام جدول مربع كاي لتعيين القيمة المجدولة $\chi_\alpha^2(v)$ حيث v=4-1=3 ،

وبالتالي فإن قيمة كاي تربيع المجدولة هي : 7.815

القرار:

$$\chi_0^2 > \chi_\alpha^2(v)$$

ونلاحظ أن لذا سوف نرفض فرضية العدم وبالتالي فإن توزيع فصيلة الدم في المدينتين مختلف



$$\chi^2_{0.05}(3) = 7.815$$

ثانيا : اختبار مربع كاي للاستقلالية (الإعتمادية)

Testing of Independence

كاي تربيع للاستقلالية (Chi-Square test of independency) هو اختبار بسيط يقوم به الباحث لمعرفة ما إذا كان هناك علاقة بين شيئين أو متغيرين. يجرى هذا الاختبار عن طريقة مقارنة قيمة يحددها الباحث مسبقا تعرف بمستوى المعنوية (الفا) بالقيمة المسماة p-Value تحسب من البيانات التوفرة، حيث سيتضح عن طريق المقارنة بين القيمتين ما إذا كانت هنالك علاقة بين الاثنين أم لا

فرضية العدم (Null hypothesis):لا توجد أي علاقة بين المتغيرين ويرمز لهذه الفرضية H0 والذي يتم افتراض صحته عند القيام بالاختبار.

عند القيام بالاختبار لمتغيرين، تكتب هذه الفرضية بهذه الطريقة: V_1مستقل عن V_2 ، حيث V_1 و V_2 تمثل المتغيرين تحت الدراسة. ويمكن كتابة فرض العدم الإحصائي بالشكل التالي:

H_0: V_1 is independent of V_2

الفرض البديل (Alternative hypothesis): توجد علاقة بين المتغيرين تحت الدراسة ويرمز لهذه الفرضيةH_A وتكتب الطريقة التالية: V_1غير مستقل أو يتبع لـ V_2 ، حيث V_1 و V_2 المتغيرين تحت الدراسة. ويمكن كتابة الفرض البديل بالشكل التالي:

H_A: V_1 is dependent on V_2

مستوى المعنوية (Level of Significance) الفـا:

عند إجراء إختبار كاي تربيع فإن على الباحث اختيار قيمة تسمى Level of Significance أو مستوى المعنوية (الفا) وهذه القيمة يمكن القول بأنها تمثل احتمال الوقوع في خطأ في الاختبار يسمى الخطأ من النوع الأول وهو رفض فرض العدم H_0 مع أنه صحيح. بمعنى أن يستنتج الباحث بناء على البيانات المتوفرة أن هنالك علاقة بين المتغيرين مع أنه لا توجد علاقة وهو استنتاج خاطئ.

هذه القيمة التي يحددها الباحث يقوم بمقارنتها بقيمة تسمى p-value والتي يمكن حسابها يدويا أو باستخدام أحد البرامج الإحصائية وذلك من البيانات التي جمعها الباحث.

غالبا في الأبحاث ما يتم استخدام قيمة الفا أو Level of Significance على أنها 0،01 أو 0،05، و الاختيار يرجع للباحث ومدى مجال الخطأ الذي يود أن يسمح به، حيث في حالة إختيار الفا = 0،01 فإن نتيجة الاختبار تكون أدق.

المختبر الإحصائي:

التكرار المتوقع:

$$\frac{\text{مجموع الصف} \times \text{مجموع العمود}}{\text{حجم العينة}}$$

نكرر تطبيق هذه المعادلة لجميع الصفوف والأعمدة لكلا المتغيرين

تحديد درجات الحرية:

درجات الحرية = (عدد الصفوف – 1) × (عدد الأعمدة – 1)

تحديد قيمة كا^2 المجدولة:

يتم بعد ذلك تحديد قيمة كا^2 المجدولة من خلال الرجوع إلى جدول كا^2 عند درجة حرية محددة وفقا لمعطيات الدراسة

القرار:

نقارن كا^2 المحسوبة بالجدولية، فعندما تكون قيمة كا^2 المحسوبة أكبر من قيمة كا^2 المجدولة فإننا نرفض الفرضية الصفرية أو فرض العدم والتي تنص على أنه لا توجد أي علاقة بين المتغيرين ونقبل الفرض البديل والتي تثبت وجود علاقة بين المتغيرين تحت الدراسة.

أما إذا كانت قيمة كا^2 المحسوبة أقل من قيمة كا^2 المجدولة فإننا نقبل الفرضية الصفرية أو فرض العدم

مثال:

في دراسة للعلاقة بين التقدير الذي يحصل عليه الطالب في الجامعة وجنسه أخذت عينة من نتائج الطلاب الذكور و الاناث وكانت كما يلي:

أولا: الاناث

راسب	راسب	راسب	راسب	جيد جدا	ممتاز	مقبول	ممتاز
جيد	جيد جدا	جيد جدا	جيد	مقبول	مقبول	مقبول	راسب
مقبول	راسب	مقبول	مقبول	مقبول	راسب	جيد جدا	جيد جدا
جيد	جيد	ممتاز	جيد جدا	ممتاز	جيد	جيد	جيد
					جيد جدا	ممتاز	جيد

ثانيا: الذكور

راسب	جيد	جيد	جيد	راسب	جيد جدا	راسب	جيد جدا
جيد جدا	جيد	راسب	راسب	راسب	راسب	راسب	مقبول
مقبول	ممتاز	ممتاز	راسب	راسب	مقبول	مقبول	ممتاز
ممتاز	جيد	جيد	مقبول	راسب	راسب	جيد	جيد
			جيد جدا	ممتاز	جيد	جيد جدا	ممتاز

والمطلوب:

هل توجد علاقة بين تقدير الطالب وجنسه عند مستوى الدلالة 0.05 ؟ = α

النتيجة / الجنس	ذكـــــــــــور	إنـــــــــاث	المجـــمـــــــوع
راســـــــــــب	12 (**9.76**)	7 (**9.24**)	19
مقبـــــــــــول	5 (**6.68**)	8 (**6.32**)	13
جـــــيـــــــــد	9 (**8.74**)	8 (**8.26**)	17
جيـــــد جـــــدا	5 (**6.17**)	7 (**5.83**)	12

ممـــتـــــاز	6 (5.65)	5 (5.35)	11
المـــجـــمـوع	37	35	72

يبين **الجدول السابق (التكرارات المتوقعة بين قوسين)** أن عدد البيانات المدخلة 72 ، عدد الذكور 37 (منهم 12 راسب وقيمتها المتوقعة 9.76 ، 5 مقبول وقيمتها المتوقعة 6.68 ، 9 جيد وقيمتها المتوقعة 8.74 ، 5 جيدجدا وقيمتها المتوقعة 6.17 ، و 6 ممتاز وقيمتها المتوقعة 5.65)

والاناث 35 (منهم 7 راسب وقيمتها المتوقعة 9.24 ، 8 مقبول وقيمتها المتوقعة 6.32 ، 8 جيد وقيمتها المتوقعة 8.26 ، 7 جيدجدا وقيمتها المتوقعة 5.83 ، و 5 ممتاز وقيمتها المتوقعة 5.35)

القيمة المحسوبة :

$$x^2 = \sum [(o - e)^2 / e] = [(12-9.76)^2/9.76] + [(5-6.68)^2/6.68] + [(9-8.76)^2/8.76] + [(5-6.17)^2/6.17] + [(6-5.65)^2/5.65] + [(7-9.24)^2/9.24] + [(8-6.32)^2/6.32] + [(8-8.26)^2/8.26] + [(7-5.83)^2/5.83] + [(5-5.35)^2/5.35]$$

$$x^2 = \sum [(o - e)^2 / e] = 0.514 + 0.4225 + 0.0065 + 0.2218 + 0.0217 + 0.5430 + 0.4466 + 0.0082 + 0.2348 + 0.0229 = \boxed{2.442}$$

القيمة المجدولة :

عند مستوى المعنوية 0.05 و درجات الحرية 4 فإن قيمة كاي ترلبيع المجدولة هي : **9.488**

اتخاذ القرار : بما أن القيمة المحسوبة أقل من المجدولة فإننا نقبل الفرضية الصفرية : لا توجد علاقة بين تقدير الطالب و جنسه .

التقويم التشخيصي :

التمرين الأول : (10 نقاط)

إذا كانت علامات الطلبة في جامعة ما تخضع لتوزيع طبيعي متوسطه الحسابي قدره 72 و انحراف معياري قدره 7 . أوجد مايلي :

(1) إذا كانت هذه الجامعة تمنح جائزة تقديرية لأعلى 5 ٪ من الطلبة فما هي أقل علامة تحصل على جائزة تقديرية ؟ (4 ن)

(2) إذا كان عدد الطلبة في هذه الجامعة 20 000 طالب و كان عدد الناجحين 14000 فماهي علامة النجاح ؟ (4 ن)

(3) المئين 75 . (2 ن)

التمـرين الثاني : (10 نقاط)

" هل شاركت عي الانتخابات السابقة ؟ " هذا السؤال تم طرحه على عينة حجمها 200 شخص فتوزعت الإجابات على النحو التالي :

الانتخاب / فئة السن	نعم	لا	المجموع
شباب	60	20	80
كهول	40	30	70
مسنون	30	20	50
المجموع	130	70	200

المطلوب : هل توجد علاقة بين المشاركة في الانتخابات و الفئة العمرية ؟ و ذلك عند مستوى معنوية 5 ٪ .

المراجع

1 ــ سعدي شاكر حمودي، مبادء علم الإحصاء وتطبيقاته، دار الثقافة للنشر والتوزيع، عمان،2009.

2 ــ محمد حسين محمد رشيد ، الاحصاء في التربية ، ط .1 ،دار صفاء للنشر و التوزيع ، عمان ، 2009

3 ـ.نافذ محمد بركات ، تطبيقات حاسوبية في الاقتصاد