الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة الشهيد حمة لخضر الوادي

قسم العلوم الاجتماعية

محاضرات في مقياس المعالجة الإحصائية للبيانات

محاضرة رقم 04 نماذج من المعالجة الاحصائية اللابارامترية في المستوى الاسمي

موجهة لطلبة السنة أولى ماستر إرشاد وتوجيه

إعداد الأستـــــــاذ:

د. محمد سبــع

السنة الجامعية: 2020 / 2021

محتوى المادة:

مقدمة

- 1. مدخل إلى علم الإحصاء : مفاهيم أساسية
- 2. مراجعة عامة حول مبادئ الإحصاء الوصفي
  - 1.2. مقاييس النزعة المركزية
  - 2.2. مقاييس التشتت
- 3. مدخل إلى الإحصائ المعمق (الاستدلالي)
  - 1.3. شرطيات الإحصاء البارامتري (المعلمي) والإحصاء اللابارامتري (اللامعلمي)
  - 2.3. معايير اختيار الإحصاء البارامتري (المعلمي) والإحصاء اللابارامتري (اللامعلمي)
- 4. نماذج من المعالجة الإحصائية اللابارمترية (اللامعلمية) في المستوى الاسمي
  - 1.4. إختبار كا^2 chi square لحسن المطابقة
  - 2.4. اختبار ماك نيمر
  - 3.4. معامل ارتباط فاي
  - 4.4. معامل التوافق
- 5. نماذج من المعالجة الإحصائية اللابارمترية (اللامعلمية) في المستوى الرتبي
  - 1.5. اختبار ويلكوكسن للفرق بين الرتب
  - 2.5. اختبار فريدمان لتحليل التباين
  - 3.5. اختبار فروق الرتب سبيرمان
- 6. نماذج من المعالجة الإحصائية البارمترية (المعلمي) في المستوين النسبي / الفتري
  - 1.6. معامل ارتباط بيرسون
  - 2.6. اختبار ت لدلالة الفروق بين المتوسطات
  - 3.6. تحليل التباين الأحادي
- قائمة المراجع
- قائمة الملاحق:
- جداول الدلالة الاحصائية

- . نماذج من المعالجة الإحصائية اللابارمتربة (اللامعلمية) في المستوى الاسمي
  - 1.4. إختبار كا^2 chi square
  - 2.4. معامل ارتباط فاي

• 1.4. إختبار كا² chi square test

ترجع النشأة الأولى لاختبار كا² إلى البحث الذي نشره كارل بيرسون فى أوائل القرن العشرين وهى تعد من أهم اختبارات الدلالة الإحصائية وأكثرها شيوعاً لأنها لا تعتمد على شكل التوزيع ولذا فهى تعد من المقاييس اللابارامترية أى مقاييس التوزيعات الحرة ولأنها تحسب لكل خلية من خلايا أى جدول تكرارى ثم تجميع القيم الجزئية للحصول على القيمة الكارلية لـ كا². (01) ورغم شيوع استخدام كا² لحساب دلالة فروق التكرار أو البيانات العددية التي يمكن تحويلها إلى تكرار مثل النسب والاحتمال إلا أن استخدامات اختبار متعددة منها(02) :

1. تقدير فترة الثقة
2. أختبار تساوي التباين والاستقلالية
3. اختبار حسن المطابقة
4. اختبار عدد النسب

وسيرتكز الحديث ههنا على الاستخدامات الأوسع لاختبار كا² chi square test χ^2 وهي

أولا اختبار مربع كاي لحسن المطابقة chi square test Testing of Goodness of Fit

يتم استخدام χ^2 في البيانات التي تقع في تصنيفات متعددة والتي يبلغ عددها اثنين أو أكثر ، مثل الإجابة عن أسئلة الاستبيان (نعم ، لا) (موافق بشدة، موافق، محايد معارض، معارض بشدة) والتي تتطلب الإجابة عنها اختيار بديل من عدة بدائل، كنوع التخصص الذي يرغب الطالب في الالتحاق به، أي أن كا² تستخدم في حالة البيانات الإسمية، ويطلق عليه في هذه الحالة Testing of Goodness of Fit نظرا لأنه يستخدم في حالة الكشف عن دلالة الفروق بين الأعداد المشاهدة، أو التكرارات الملاحظة، أو الاستجابات الواقعة في كل تصنيف وتسمى observed frequency وبين العدد المتوقع المعتمد على الفرض الصفري، أو التكرارات المتوقعة وهي التكرارات النظرية للمتغير موضوع الدراسة في المجتمع الأصلي وتسمى expected frequency .

فإذا كانت قيمة كا² = 0 فهذا يدل على أن عينة البحث ممثلة للمجتمع في تكرارها ومتطابقة معه، أما إذا كانت كا² > 0 فهذا يدل على وجود فروق بين تكرارات العينة الملاحظة وبين تكرارات التوزيع النظري للمجتمع (التكرارات المتوقعة) ويكون الفرض الصفري هنا حول المجتمع الأصلي الذي تسحب منه العينة، فهو يفترض عدم وجود فروق دالة إحصائيا بين تكرارات العينة الملاحظة أو المشاهدة والتكرارات المتوقعة فإذا ما تم رفض الفرض الصفري (تطابق العينة مع المجتمع) فيتم قبول الفرض البديل للبحث والذي عادة ما يكون عكس الفرض الصفري ويتم حساب قيمة كا² إجمالا انطلاقا من صيغة المعادلة التالية: (3)

$$\boxed{\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}}$$

[16]

حيث أن f_0 يمثل التكرارات المشاهدة و f_e يمثل التكرارات المتوقعة

فى جميع الحالات نخرج من الحسابات بقيمة كا² المحسوبة نقارنها بقيمة كا² الجدولية كالتالى :

- إذا كانت كا² المحسوبة > كا² الجدولية فان كا² تكون دالة إحصائية .
- إذا كانت كا² المحسوبة < كا² الجدولية فان كا² ليست دالة إحصائية .

مثال (11): الجدول التالي يوضح آراء 80 طالبا سنة أولى ماستر حول إمكانية تأجيل الامتحانات فكانت استجاباتهم على النحو التالي :

الرأي	موافق	غير موافق	المجموع
التكرار	60	40	100

السؤال: هلة هناك فروق دالة احصائيا عند مستوى 0.05 بين تكرارات استجابات الطلبة

الفرضية

- ليس هناك فرق بين التكرارات $f_E = f_O$ $H_0: \chi^2 = 0$
- هناك فروق بين التكرارات $f_E \neq f_O$ $H_1: \chi^2 > 0$

f_e = حاصل قسمة مجموع التكرارات على عدد البدائل

$$f_e = \frac{\sum f}{C}$$

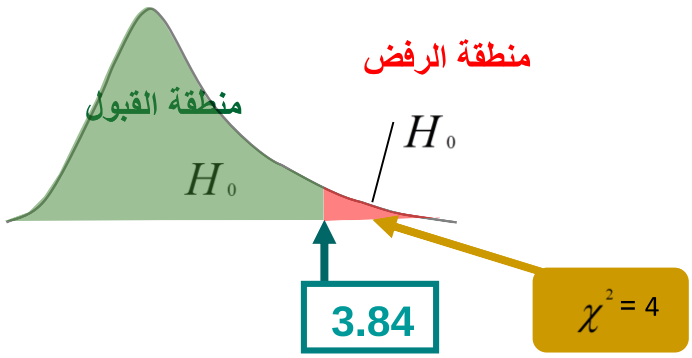
لحساب قيمة كا² نتبع الخطوات الموضحة في الجدول

البدائل	التكرارات المشاهدة f_o	التكرارات المتوقعة f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
موافق	60	50	10	100	2=50/100
غير موافق	40	50	10	100	2=50/100
المجموع	100	100	/	/	4

حساب درجة الحرية : $df = c-1$ حيث أن c هو عدد البدائل (موافق/ غير موافق) أو الاعمدة

بالبحث فى جداول كا² عند درجة حرية = 1 ومستوى دلالة 0.05 نجد قيمة كا² الجدولية = 3.84

$$\chi^2_{(df\ 1, \alpha 0.05)} = 3.84 \cdot$$



اتخاذ القرار : بما أن القيمة المحسوبة $\chi^2 = 4$ أكبر تماما من القيمة الجدولية (3.84) نرفض الفرض الصفري ونقبل الفرض البديل الذي يقول هناك فروق دالة احصائيا بين تكرارات استجابات الطلبة

ثانيا اختبار مربع كاي للاستقلالية chi square test Teting of Independences

كاي تربيع للاستقلالية (Chi-Square test of independency) هو اختبار بسيط يقوم به الباحث لمعرفة ما إذا كان هناك علاقة بين متغيرين. اسميين في جدول تقاطعي يقرن بين المتغيرين فإذا كان المتغير الأول مثل الجنس ينقسم إلى قيمسن (ذكور/ إناث) والمتغير الثاني التحصيل ينقسم بدوره إلى قسمين (ناجح/ راسب)نسمي هذا الجدول التقاطعي بجدول 2x2 وهكذا ...الخ.

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: V1مستقل عن V2 ، حيث V1 و V2 تمثل المتغيرين تحت الدراسة. ويمكن كتابة فرض العدم الإحصائي بالشكل التالي:

H0: V1 is independent of V2

الفرض البديل (Alternative hypothesis): توجد علاقة بين المتغيرين تحت الدراسة ويرمز لهذه الفرضية HA وتكتب الطريقة التالية: V1 غير مستقل أو يتبع لـ V2 ، حيث V1 و V2 المتغيرين تحت الدراسة. ويمكن كتابة الفرض البديل بالشكل التالي:

HA: V1 is dependent on V2

مستوى المعنوية (Level of Significance) الفا:

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 H0 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.

اعتمادا على القانون العام لاختبار كا^2 رقم [16]

التكرار المتوقع للخلية= $\dfrac{\text{مجموع الصف} \times \text{مجموع العمود}}{\text{حجم العينة}}$

$$f_e = \frac{(\sum r)\times(\sum c)}{n}$$

نكرر تطبيق هذه المعادلة لجميع على كل الخلايا في 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 رقم (11) أراد الفريق العامل باحد المستشفيات معرفة العلاقة بين التدخين والإصابة بسرطان الرئة لدى عينة بلغ عدد أفرادها 200 فرد فتحصلوا على الجدول التالي:

	مصاب	غير مصاب	المجموع
مدخن	78	32	110
غير مدخن	30	60	90
المجموع	108	92	200

السؤال: هل هناك علاقة بين التدخين والإصابة بمرض السرطان

ملاحظة: التساؤل هنا عن العلاقة بمعنى اللااستقلالية ويمكن معرفته بالتأكد من وجود اختلاف معنوي بين المدخنين وغير المدخنين، أومن خلال البحث عن العلاقة بين المتغيرين وفي كلا الحالتين نتبع الخطوات التالية باستخدام كاي تربيع. (أنظر 04)

الفرضية الصفرية H0 – ليس هناك علاقة بين التدخين والإصابة بالسرطان

أو يمكن القول H0 – لا توجد فروق دالة بين تكرارات المصابين وغير المصابين تبعا لمتغير التدخين

	مصاب		غير مصاب			مصاب	غير مصاب	مصاب	غير مصاب	مصاب	غير مصاب
	f_o	f_e	f_o	f_e	مج	$f_o - f_e$	$f_o - f_e$	$(f_o - f_e)^2$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$	$\frac{(f_o - f_e)^2}{f_e}$
مدخن	**78**	59.4	**32**	50.6	**110**	18.6	-18.6	345.96	345.96	5.82	6.83
غير مدخن	**30**	48.6	**60**	41.4	**90**	-18.6	18.6	345.96	345.96	7.11	8.35
مج	**108**		**92**		**200**		/	/		المجموع= 28.11=8.35+7.11+6.83+5.82	

القيمة المحسوبة : كا^2 = 28.11

حساب درجة الحرية: df = (عدد الاعمدة –1) x (عدد الصفوف –1)

بالبحث فى جداول كا^2 عند درجة حرية = 1 ومستوى دلالة 0.05 نجد قيمة كا^2 الجدولية = 3.84

$$\chi^2_{(df\ 1, \alpha 0.05)} = 3.84 \cdot$$

اتخاذ القرار : بما أن القيمة المحســـوبة χ^2 =28.11 أكبر تماما من القيمة الجدولية (3.84) نرفض الفرض الصفري ونقبل الفرض البديل الذي يقول هناك فروق دالة احصائيا بين تكرارات استجابات الطعلاقة بين التدخين والإصابة بالسرطان.

ثانيا معامل ارتباط فاي (Phi Correlation Coefficient)

يستخدم معامل ارتباط فاي في حساب العلاقة بين متغيرين منفصلين اسميين وتحديدا في الحالات التي ينقسم فيها كل متغير إلى نوعين مختلفين مثل الصفات ونقيضها ذكور إناث ، علمي أدبي، لذا فهو يصلح لتحليل مفردات أسئلة الاختبارات النفسية التي تنقسم إلى قسمين مثل (نعم، لا) .(05)

ويتم حساب قيمة معامل فاي انطلاقا من المعادلة رقم [17]

$$R_{\emptyset} = \frac{AD - BC}{\sqrt{(A+B)(C+D)(A+C)(B+D)}} \quad \textbf{[17]}$$

مثال رقم (12)

لمعرفة العلاقة الارتباطية بين نوع الجنس والتحصيل الدراسي جرى تقسيم عينة من الطلبة وفق ما هو موضح في الجدول

المتغير	ناجح	راسب	المجموع
ذكور	A 35	B 37	A+B 72
إناث	C 14	D 34	C+D 48
المجموع	A+C 49	B+D 71	120

الفرضية الصفرية H0 . لا توجد علاقة ارتباطية دالة إحصائيا بين الجنس والتحصيل الدراسي
الفرضية البديلة: توجد علاقة ارتباطية دالة احصائيا بين الجنس والتحصيل الدراسي إذا كانت القيمة المحسوبة كا^2 أكبر من كا^2 الجدولية
ملاحظة: يتم تحويل قيمة فاي إلى كا^2 من خلال القانون التالي:

$$\chi^2 = (R\varnothing)^2 \times N$$

حساب قيمة فاي: $R\varnothing = \frac{(35x34)-(37x14)}{\sqrt{(49)(71)(48)(72)}}$

$= 0.19$

هناك علاقة ارتباطية طردية ضعيفة

للتحقق من دلالة قيمة فاي تحولها إلى كا^2 حيث أن

كا2 = فاى2 × ن

كا2 = (0.19)2 × 120 = 4.33

حساب كا2 الجدولية :

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة :

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

= (2 – 1) × (2 – 1) = 1 × 1 = 1

مستوى الدلالة = 0.05 .

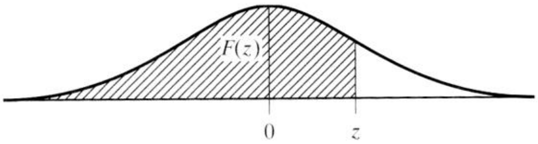
بالبحث فى جداول كا2 عند درجة حرية = 1 ومستوى دلالة 0.05 نجد قيمة كا2 الجدولية = 3.841

تحديد مدى دلالة كا2 :

نقارن قيمة كا2 المحسوبة بقيمة كا2 الجدولية نجد أن :

قيمة كا2 المحسوبة = 4.33 > قيمة كا2 الجدولية = 3.841

لذا فان كا2 دالة إحصــائية عند مســتوى دلالة 0.05 .وهو ما يشــير إلى رفض الفرض الصــفري وقبول الفرض البديل الذي يشير إلى وجود علاقة ارتباطية دالة احصائيا بين الجنس والتحصيل الدراسي.



$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6661	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6913	.6950	.6985	.7019	.7054	.7083	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9432	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9658	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9716	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9812	.9826	.9830	.9834	.9838	.9842	.9846	.9854	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9937	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

(Used with permission of Dover Publications, from E. L. Crow; F. A. Davis; and M. W. Maxfield, *Statistics Manual*. © 1960, Dover Publications.)