العلوم الإقتصادية
العدد الخامس والثلاثون / المجلد التاسع / آذار , 2014

تطبيق أختبارات أحصائية مختلفة لقياس التوافق على بيانات صحية تتعلق بالأمراض المزمنة مأخوذة من مستشفى اليرموك التعليمي

م.م.هديل مهدي كاظم م.م . ميسون حميد فرج
وزارة التربية – ديوان الرقابة المالية

المستخلص :

يتناول هذا البحث تقديم وتطبيق أختبارات أحصائية مختلفة، لأختبار العلاقات بين الأمراض المزمنة الستة مع الجنس والعمر، وقد رمزنا للأمراض بالأحرف (A, B, C, D, E, F)، ويشير (A) الى مرض داء السكري، (B) الى أرتفاع ضغط الدم، (C) الى أمراض القلب، (D) الى أمراض المخ، (E) الى مرض قرحة المعدة، و (F) الى الربو القصبي. جمعت البيانات من مستشفى اليرموك / وزارة الصحة العراقية ولفئات عمرية مختلفة هي (أقل من 15 سنة)، (من 15 سنة الى 44 سنة)، وفئة (45 الى 64 سنة) ثم الفئة الأخيرة (65) فأكثر، وللرجال والنساء وتم ترتيب البيانات بنظام $(Excel)$ في جداول خاصة لتسهيل تطبيق الأختبارات الأحصائية المختلفة ولكي تكون مصدر لكل الباحثين في المجالات الصحية لمساعدتهم في التعرف على تأثير المتغيرات والتفاعل بين المتغيرات والفروق المعنوية، وقد كتبت برامج خاصة بلغة $(Minitab)$، وطبقت أحصاءة (χ^2) وأحصاءة $(Kappa)$، وأختبار الأتجاه في ثنائي الحدين، وأختبار معتمد على الأوزان، وأختبار التوافق، وبعد شرح كل أختبار والأشارة الى البرنامج الخاص به، طبقت البرامج وفسرت النتائج التي تم الحصول عليها.

تاريخ قبول النشر 2013/12/31

1) الهدف من البحث

يهدف البحث الى تطبيق أختبارات أحصائية متنوعة تم أختيارها من قبل الباحثين، وكتبت لها برامج حاسوب بأستخدام ($Minitab$) لبيان العلاقات المختلفة بين مجموعة الأمراض الستة المزمنة والتي رمزنا لها بالأحرف (A, B, C, D, E, F)، وعلاقتها بالعمر والجنس، وقد أخذت البيانات عن الأمراض المزمنة من وزارة الصحة العراقية لعام (2008)، والأختبارات التي ستطبق هي أختبار توافق في جدول صفوف وأعمدة ($R \times C$) ،اختبار هل يوجد أتجاه متزايد في الأصابة بهذه الأمراض وعلاقته بالجنس وكذلك تطبيق أحصاءة ($Kappa$)، وأختبار التطابق، وجميعها أساليب وأختبارات أحصائية بسيطة ومهمة تساعد الأطباء والباحثين في متابعة المرضى وأتخاذ القرارات.

2) المقدمة [2][13]

تتميز الطرق الأحصائية بتطبيقات واسعة لها في المجالات الصحية والفيزيائية والبايولوجية وغيرها، وأن معظم المؤسسات الصحية تعاني من نقص كبير في التطبيقات الأحصائية، وأن وجدت فهي لا تتعدى كونها سجلات وجداول وبيانات تجمعت بمرور الزمن، لذلك أرتأينا البحث في المجال الصحي لبيانات التطبيقات الحديثة المتبعة في المجالات الصحية والمختبرات العالمية وخاصة ما يتعلق بتحليل التجارب الحياتية والعوارض المرضية وأكتشاف الجرعات المؤثرة الفعالة وغيرها، وبعد مراجعتنا المستمرة لمؤسسة اليرموك الطبية التعليمية بأعتبارها من المؤسسات الطبية الكبيرة، تم الحصول على البيانات الخاصة بالأمراض المزمنة والعمر والجنس ورتبت حسب الجدول المرفق بالبحث، ولأن البيانات يمكن وضعها في جداول توافق من نوع (2×2) أو ($2 \times k$) أو ($r \times c$) أو ($r \times c \times l$)، لذلك استخدمت مقاييس أحصائية مختلفة وهي مقاييس بسيطة ولكنها ذات نتائج فعالة في أختبار الفرضيات المختلفة الخاصة بالبيانات عن الأمراض المزمنة، وقد اعتمد أختبار (χ^2) وأختبار ($Kappa$) وأختبار (χ^2) للأتجاه واختبار الأنحدار، وتم شرح صيغة كل أحصاءة وكيفية تطبيقها على البيانات الصحية الخاصة بالأمراض المزمنة، والأختبارات الأربعة التي طبعت

في هذا البحث وأستخرجت نتائجها حسب البرامج الخاصة بها، يمكن أعتمادها لأية مجموعة من البيانات المصنفة والتي من الممكن تفريغها في جداول توافق من نوع (2×2) أو $(2 \times c)$ أو $(r \times c)$ أو $(r \times c \times l)$.

3) الجانب النظري

3–1 أختبار التوافق في جدول توافق $(R \times C)$

$$Tests\ for\ association\ of\ (R \times C)\ contingency\ tables$$

يطبق هذا الأختبار لبيانات مأخوذة من مجتمع ثنائي الحدين وأنها مصنفة في أصناف تتكون من (R) من الصفوف و (C) من الأعمدة وأن التكرار المشاهد لكل خلية هو (O_{ij}) ويكون ترتيب البيانات كما في الجدول التالي:

جدول (1): بيانات جدول توافق مرتبة بشكل صفوف وأعمدة.

Row Variable	Column Variable					
		A_1	A_2	⋯	A_C	$Total$
	B_1	O_{11}	O_{12}	⋯	O_{1c}	R_1
	B_2	O_{21}	O_{22}	⋯	O_{2c}	R_2
	⋮	⋮	⋮	⋯	⋮	⋮
	B_R	O_{R1}	O_{R2}	⋯	O_{Rk}	R_R
	$Total$	C_1	C_2	⋯	C_k	

لأختبار فرضية الأستقلالية بين متغير الصفوف ومتغير الأعمدة (مثلاً العلاقة بين العمر والأصابة بالأمراض المزمنة) أي أن أحتمالات الأصابة متساوية لكلا الجنسين.

$$H_0: p_{11} = p_{12} = \cdots = p_{R(C-1)}$$

$$H_1: p_{11} \neq p_{12} \neq \cdots \neq p_{R(C-1)}$$

فأن أحصاءة الأختبار هي:

$$\chi^2 = \sum_{i=1}^{R}\sum_{j=1}^{C} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_{(R-1)(C-1)} \quad \dots (1)$$

حيث تشير (i) الى عدد الصفوف

وتشير (j) الى عدد الأعمدة وتستخرج التكرارات المتوقعة لكل خلية من تطبيق المعادلة التالية:

$$E_{ij} = \frac{E_i \times E_j}{Total} \quad \dots (2)$$

ويساوي حاصل ضرب مجموع الصف X مجموع العمود الواقعة فيه الخلية مقسوماً على المجموع الكلي للمشاهدات، وهناك شروط ضرورية لتطبيق هذا الأختبار، وهي:

1–أن لا يكون عدد الخلايا التي تكرارها أقل من خمسة هو أكثر من (0.2) من مجموع عدد الخلايا.

2– لا توجد خلية تكرارها المتوقع أقل من واحد، وبعد تطبيق صيغة (χ^2) المحسوبة مع (χ^2) الجدولية لدرجة حرية $(R-1)(C-1)$ ومستوى معنوية مفترض من الباحث (0.05) مثلاً فاذا كانت قيمة (χ^2_{cal}) أكبر من (χ^2_{tab}) نرفض فرضية العدم (H_0) التي تنص على عدم وجود علاقة بين العمر والإصابة بالأمراض المزمنة، والعكس صحيح.

2–3 أختبار وجود أتجاه متزايد في نسبة الأصابة بأعتماد توزيع ثنائي الحدين

Chi – Square test for trend in Binomial Proportions

يطبق هذا الأختبار لمقارنة عدة نسب (k) من مجتمع ثنائي الحدين وممثلة في جداول $(2 \times k)$ مثلاً يشير العدد (2) الى عدد اصناف المتغير (A)، وأن (k) تمثل أصناف المتغير (B)، وأن البيانات تفرغ في جداول كالآتي:

جدول (2): جدول بيانات مصنفة من نوع $(2 \times k)$.

		Disease				
		A_1	A_2	$\cdots$	A_k	*Total*
Male *Female*	B_1	x_1	x_2	$\cdots$	x_k	x
	B_2	$n_1 - x_1$	$n_2 - x_2$	$\cdots$	$n_k - x_k$	$n - x$
	Total	n_1	n_2	$\cdots$	n_k	n

يتم أولاً أيجاد النسب التقديرية:

$$\hat{p}_i = \frac{x_i}{n_i}, \quad \overline{p} = \frac{x}{n} \qquad \ldots(3)$$

وأن الفرضية المطلوب أختبارها تساوي أحتمالات الأصابة لكل الأمراض بدون أتجاه نحو فئة عمرية معينة وهي:

$H_0: p_1 = p_2 = \cdots = p_k$ بدون أتجاه

$H_a: S_1 p_1 \neq S_2 p_2 \neq \cdots \neq S_k p_k$ يوجد أتجاه في أحتمالات الأصابة

وأن (S_i) هو متغير وزن ($Score\ Variable$) يحدد من قبل الباحث، وأن (p_i) تمثل نسبة الأصابة بالمرض وتحسب وفق العلاقة (3) ثم بعد ذلك يمكن حساب النسب التقديرية من المعادلة (4)، وتعتمد ةهذه المعادلة في حالة رفض الفرضية (H_0).

$$p_j = \alpha + \beta S_i \qquad \dots \ (4)$$

يتم أولاً تحديد نسب الأصابة وفق العلاقة (3) ثم نوجد المقدار:

$$A = \sum_{i=1}^{k} n_i(\hat{p}_i - \overline{p})(S_i - \overline{S}) = \sum_{i=1}^{n} x_i S_i - x\overline{S}$$

$$= \sum_{i=1}^{n} x_i S_i - \frac{x\sum_{i=1}^{k} n_i S_i}{n} \qquad \dots . (5)$$

كذلك نوجد المقدار:

$$B = \overline{pq}\left[\sum_{i=1}^{n} n_i S_i^2 - \frac{(\sum_{i=1}^{n} n_i S_i)^2}{n}\right] \qquad \dots . (6)$$

ثم نوجد أحصاءة الأختبار:

$$\chi^2_{(1)cal} = \frac{A^2}{B} \qquad \dots . (7)$$

وتقارن مع القيمة الجدولية $\chi^2_{(1)tab}$ فاذا كانت المستخرجة أكبر من الجدولية نرفض فرضية العدم (H_0) وتقبل بالفرضية البديلة (H_1). وفي حالة وجود أتجاه لتزايد النسب نطبق المعادلة (4)، وهي معادلة أنحدار خطي بسيط تعتمد لتقدير النسب اللاحقة.

تمت كتابة برنامج بلغة ($Minitab$) وأعطيت أوزان هي ($S_i = 1,2,3,4,5,6$) حسب عدد الأمراض الواردة في البحث، ثم أختبرت البيانات التي تمثل العلاقة بين الجنس والأصابة بالأمراض المزمنة حسب الأختبار الموضح في المعادلة (7).

وفيما يأتي نتائج أختبار تزايد نسب الأتجاه العام.

لأختبار وجود الأتجاه حول ظاهرة معينة تم تفريغ البانات في جداول من نوع ($2 \times k$)، حيث تشير الصفوف الى الجنس وتشير الأعمدة الى فئة العمر، وقد أعطيت أوزان ($scores$) عددها (k)، حيث تم تطبيق أختبار الأتجاه بواسطة ثنائي الحدين وكما يأتي:

1– داء السكروالجنس

$Class$					$Total$
	1	2	3	4	
$Male$	11	34	77	112	234
$Female$	4	44	82	110	240
$Total$	15	78	159	222	474
p_0	2.7500	0.7727	0.9390	1.0182	
p_e	2.1244	1.6214	1.1185	0.6156	
e_i^2	0.3914	0.7203	0.0322	0.1621	1.3060

لا يوجد أتجاه $H_0: P_1 = P_2 = P_3 = P_4 = \cdots$

يوجد أتجاه $H_1: S_1P_1 = S_2P_2 = S_{13}P_3 = S_4P_4 = \cdots$

Chi – Square for trend in Binomial proportion.

$$\chi^2_{cal.} = 264.9712 \qquad \chi^2_{tab.} = 7.815$$

القرار رفض (H_0).

وعند مقارنة $(\chi^2_{cal.})$ المحسوبة مع $(\chi^2_{tab.})$ الجدولية نرفض فرضية العدم، أي أن هناك أتجاه عام لتزايد الأصابة بالمرض لكلا الجنسين.

2–ضغط الدم والجنس

Class					Total
	1	2	3	4	
Male	0	10	93	85	188
Female	0	20	113	78	211
Total	0	30	206	163	399
p_0	0	0.5000	0.8230	1.0897	
p_e	0.0644	0.4236	0.7828	1.1420	
e_i^2	0.0041	0.0058	0.0016	0.0027	0.0143

$H_0: P_1 = P_2 = P_3 = P_4 = \cdots$ لا يوجد أتجاه

$H_1: S_1P_1 = S_2P_2 = S_{13}P_3 = S_4P_4 = \cdots$ يوجد أتجاه

Chi – Square for trend in Binomial proportion.

$$\chi^2_{cal.} = 543.9022 \qquad \chi^2_{tab.} = 7.815$$

القرار رفض (H_0).

وعند مقارنة المحسوبة مع الجدولية نرفض فرضية العدم، أي أن هناك أتجاه عام لتزايد الأصابة بالمرض لكلا الجنسين، كلما تقدم العمر.

3– داء القلب والجنس

$Class$					$Total$
	1	2	3	4	
$Male$	2	5	40	58	105
$Female$	2	3	46	27	78
$Total$	4	8	86	85	183
p_0	1.0000	1.6667	0.8696	2.1481	
p_e	1.0240	1.2887	1.5535	1.8182	
e_i^2	0.0006	0.1428	0.4677	0.1089	0.7200

$H_0: P_1 = P_2 = P_3 = P_4 = \cdots$ لا يوجد أتجاه

$H_1: S_1P_1 = S_2P_2 = S_{13}P_3 = S_4P_4 = \cdots$ يوجد أتجاه

Chi – Square for trend in Binomial proportion.

$$\chi^2_{cal.} = 944.4667 \qquad \chi^2_{tab.} = 7.815$$

Decision: Reject H_0 $\quad beta = (-0.2949, 0.3592)$

وعند مقارنة المحسوبة مع الجدولية نرفض فرضية العدم، أي أن هناك أتجاه عام لتزايد الأصابة بالمرض لكلا الجنسين.

4-أمراض المخ والجنس

Class					Total
	1	2	3	4	
Male	0	36	76	96	208
Female	0	33	111	86	230
Total	0	69	187	182	438
p_0	0	1.0909	0.6847	1.1163	
p_e	0.2816	0.5758	0.8701	1.1644	
e_i^2	0.0793	0.2653	0.0344	0.0023	0.3813

$H_0: P_1 = P_2 = P_3 = P_4 = \cdots$ لا يوجد أتجاه لتزايد الأصابة

$H_1: S_1P_1 = S_2P_2 = S_{13}P_3 = S_4P_4 = \cdots$ يوجد أتجاه

Chi – Square for trend in Binomial proportion.

$$\chi^2_{cal.} = 2101.7 \qquad \chi^2_{tab.} = 7.815$$

القرار رفض (H_0).

وعند مقارنة المحسوبة مع الجدولية نرفض فرضية العدم، أي أن هناك أتجاه عام لتزايد الأصابة بالمرض لكلا الجنسين.

5– **قرحة المعدة والجنس**

	Class				Total
	1	2	3	4	
Male	0	30	52	31	109
Female	0	37	48	24	113
Total	0	67	100	55	222
p_0	0	0.8108	1.0833	1.2917	
p_e	0.1743	0.5891	1.0038	1.4186	
e_i^2	0.0304	0.0492	0.0063	0.0161	0.1020

لا يوجد أتجاه $H_0: P_1 = P_2 = P_3 = P_4 = \cdots$

يوجد أتجاه $H_1: S_1P_1 = S_2P_2 = S_{13}P_3 = S_4P_4 = \cdots$

Chi – Square for trend in Binomial proportion.

$$\chi^2_{cal.} = 7254.9565 \qquad \chi^2_{tab.} = 7.815$$

القرار رفض (H_0).

وعند مقارنة المحسوبة مع الجدولية نرفض فرضية العدم، أي أن هناك أتجاه عام لتزايد الأصابة بالمرض لكلا الجنسين.

6-الربو القصبي والجنس

	Class				Total
	1	2	3	4	
$Male$	0	17	79	12	108
$Female$	0	63	31	0	94
$Total$	0	80	110	12	202
p_0	0	0.2698	2.5484	0	
p_e	0.3628	0.5906	0.8185	1.0463	
e_i^2	0.1316	0.1029	2.9926	1.0948	4.3219

لا يوجد أتجاه $H_0: P_1 = P_2 = P_3 = P_4 = \cdots$

يوجد أتجاه $H_1: S_1P_1 = S_2P_2 = S_{13}P_3 = S_4P_4 = \cdots$

Chi – Square for trend in Binomial proportion.

$$\chi^2_{cal.} = 228.1094 \qquad \chi^2_{tab.} = 7.815$$

القرار رفض (H_0).

وعند مقارنة المحسوبة مع الجدولية نرفض فرضية العدم، أي أن هناك أتجاه عام لتزايد الأصابة بالمرض لكلا الجنسين.

3–3 أحصاءة $Kappa$

تقيس هذه الأحصاءة درجة التوافق أو التطابق بين متغيرين فرغت نتائجهما في جدول (2×2) والمعرف بالمعادلة التالية:

$$K = \frac{p_0 - p_e}{1 - p_e} \quad \dots (8)$$

حيث تشير (p_0) الى الأحتمال المشاهد للتطابق.

$$p_0 = \frac{1}{n}(f_{11} + f_{22}) \quad \dots (9)$$

و تشير (p_0) الى الأحتمال المتوقع للتطابق.

$$p_e = \frac{1}{n^2}(R_1C_1 + R_2C_2) \quad \dots (10)$$

تكون البيانات متمثلة بالجدول:

	Yes	No	$Total$
Yes	f_{11}	f_{12}	R_1
No	f_{21}	f_{22}	R_2
$Total$	C_1	C_2	n

وعند تطبيق أحصاءة كابا حسب البرنامج الخاص بها، تم تشكيل علاقات ثنائية في جدول (2×2)، مثلاً الأصابة بأحد الأمراض المزمنة مع الجنس، ثم تحسب أحصاءة كابا فأذا كانت:

توافق عالي $K > 0.75$

توافق جيد $K \in [0.4, 0.75]$

توافق ضعيف $K \in [0,0.4]$

وتعتمد قيمة (K) على قيمة (Z)، حيث أن:

$$Z = \frac{K}{Se(K)} \sim N(0,1)$$

$$Var(K) = \frac{1}{n(1-p_e)^2}\left\{p_e + p_e^2 - \sum_{i=1}^{C}\frac{1}{n^3}[R_i C_i (R_i + C_i)]\right\} \quad \ldots (11)$$

... (11)

وتعتمد صيغة (Z) لأختبار الفرضية:

$$H_0: K = 0$$

$$H_1: K \neq 0$$

وتمت كتابة برنامج خاص لأحصاءة $(Kappa)$ وأجريت أختبارات لعلاقات مختلفة.

وحيث أن تطبيق أحصاءة كابا يتطلب أعداد جداول توافق (2×2)، وكذلك أختبار النسب أيضاً يتطلب يتطلب أعداد جداول توافق (2×2)، لذلك تم تكوين جداول (2×2) للعلاقة بين داء السكر والجنس للفئة العمرية (44 – 15) وطبقت أحصاءة كابا، وأختبار النسب على البيانات في الجدول (4) والجداول $(5,6,7,8,9,10,11,12)$، وتم شرح نتائج الأختبارين لكل جدول (2×2) بهدف أختصار النتائج.

4-3 أختبار النسب *Test of Proportions*

حيث أن أحصاءة كابا تتطلب تكوين جداول توافق من نوع (2×2)، لذلك أرتأينا أيضاً شرح أختبار تساوي نسب الأصابة، وهو أختبار يتطلب أيضاً جداول (2×2)، ولكي نختبر الفرضية التي تنص على تساوي نسبة الأصابة بالمرض (A) لكل من النساء والرجال، ويعتمد لأختبار الفرضية:

$$H_0: p_1 = p_2 \Rightarrow \overline{p}_1 - \overline{p}_2 = 0$$

وأن:

$$\overline{p}_1 = \frac{a}{a+b} \quad , \quad \overline{p}_2 = \frac{c}{c+d}$$

حيث أن البيانات مفرغة في جداول (2×2):

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	

وتحت الفرضية $_0$ يفترض أن:

$$H_0: p_1 = p_2 = p$$

$$H_1: p_1 \neq p_2 \neq p$$

$$\overline{p} = \frac{a+c}{a+b+c+d}$$

أما أحصاءة الأختبار (أختبار تساوي النسب) فهي تعتمد على قيمة (Z) المحسوبة من المعادلة:

$$Z = \frac{|\overline{p}_1 - \overline{p}_2| - \frac{1}{2}\left(\frac{1}{a+b} + \frac{1}{c+d}\right)}{\sqrt{\overline{p}\overline{q}\left(\frac{1}{a+b} + \frac{1}{c+d}\right)}} \sim N(0,1) \qquad \ldots(12)$$

وكانت جميع نتائج أختبار $(Two\ samples\ for\ Binomial)$ تشير الى قبول الفرضية (H_0).

وفيما يلي نتائج تطبيق أحصاءة كابا، وأختبار تساوي النسب على بيانات الأمراض المزمنة، حيث تم إعداد جداول ثنائية ومن ثم كتبت برامج خاصة باختبارات $(Kappa, Two\ Sample\ Test)$، وطبق الأختبار للفئات العمرية $(15 - 44)$ بالشكل التالي:

1- داء السكر والجنس، الفئة العمرية$(15 - 44)$ سنة.

جدول (4): العلاقة بين الجنس والأصابة بداء السكر

	Yes	*No*	*Total*
Male	38	3648	3686
Female	44	3820	3864
Total	82	7468	7550

H_0: لا توجد علاقة بين الجنس والأصابة بداء السكر

H_1: توجد علاقة بين الجنس والأصابة بداء السكر

1- *Kappa Test*

$$Z_c = -0.4517 \quad decesion: accept\ H_0$$

2- *Two – Samples Test for Binomial Proportion*

$$Z_c = 0.0706 \quad decesion: accept\ H_0$$

2- ضغط الدم والجنس، الفئة العمرية$(15 - 44)$ سنة.

جدول (5): العلاقة بين الجنس والأصابة بضغط الدم.

	Yes	*No*	*Total*
Male	33	4543	4576
Female	36	4130	4266
Total	96	8673	8772

H_0: لا توجد علاقة بين الجنس وضغط الدم

H_1: توجد علاقة بين الجنس وضغط الدم

1- *Kappa Test*

$$Z_c = -0.7545 \quad decesion: accept\ H_0$$

2- *Two – Samples Test for Binomial Proportion*

$$Z_c = 0.1123 \quad decesion: accept\ H_0$$

3- القلب والجنس، الفئة العمرية(44 – 15) سنة.

جدول (6): العلاقة بين الجنس وأمراض القلب.

	Yes	*No*	*Total*
Male	10	648	658
Female	20	623	643
Total	30	1271	1301

H_0: لا توجد علاقة بين الجنس وأمراض القلب

H_1: توجد علاقة بين الجنس وأمراض القلب

1- *Kappa Test*

$$Z_c = -1.9112 \quad decesion: Reject\ H_0$$

2- *Two – Samples Test for Binomial Proportion*

$$Z_c = 0.5184 \quad decesion: accept\ H_0$$

4–أمراض المخ والجنس، الفئة العمرية$(15 - 44)$ سنة.

جدول (7): العلاقة بين الجنس وأمراض المخ.

	Yes	*No*	*Total*
Male	3	575	578
Female	5	466	471
Total	8	1041	1049

H_0: لا توجد علاقة بين الجنس وأمراض المخ

H_1: توجد علاقة بين الجنس وأمراض المخ

1- *Kappa Test*

$$Z_c = -1.0047 \quad decesion: accept\ H_0$$

2- *Two – Samples Test for Binomial Proportion*

$$Z_c = 0.1134 \quad decesion: accept\ H_0$$

5–قرحة المعدة والجنس، الفئة العمرية$(15 - 44)$ سنة.

جدول (8): العلاقة بين الجنس وقرحة المعدة.

	Yes	*No*	*Total*
Male	36	1669	1705
Female	17	1569	1586
Total	53	3238	3291

H_0: لا توجد علاقة بين الجنس وقرحة المعدة

H_1: توجد علاقة بين الجنس وقرحة المعدة

1- *Kappa Test*

$$Z_c = 2.3673 \quad decesion: Reject\ H_0$$

2- *Two – Samples Test for Binomial Proportion*

$$Z_c = 0.5613 \quad decesion: accept\ H_0$$

6ـ الربو القصبي والجنس، الفئة العمرية$(44 - 15)$ سنة.

جدول (9): العلاقة بين الجنس والربو القصبي.

	Yes	*No*	*Total*
Male	30	3585	3615
Female	37	2869	2906
Total	67	6454	6521

H_0: لا توجد علاقة بين الجنس والربو القصبي

H_1: توجد علاقة بين الجنس والربو القصبي

1- *Kappa Test*

$$Z_c = -1.7646 \quad decesion: Reject\ H_0$$

2- *Two – Samples Test for Binomial Proportion*

$$Z_c = 0.3330 \quad decesion: accept\ H_0$$

أما الأختبار للفئات العمرية $(64 - 45)$ فكان بالشكل التالي:

1- داء السكر والجنس، الفئة العمرية(45 − 64) سنة.

جدول (10): العلاقة بين الجنس والأصابة بداء السكر

	Yes	No	Total
Male	82	3437	3519
Female	77	2796	2873
Total	159	6233	6392

H_0: لا توجد علاقة بين الجنس والأصابة بداء السكر

H_1: توجد علاقة بين الجنس والأصابة بداء السكر

1- *Kappa Test*

$$Z_c = -0.8935 \quad decesion: accept\ H_0$$

2- *Two − Samples Test for Binomial Proportion*

$$Z_c = 0.0706 \quad decesion: accept\ H_0$$

2- ضغط الدم والجنس، الفئة العمرية(45 − 64) سنة.

جدول (11): العلاقة بين الجنس والأصابة بضغط الدم.

	Yes	No	Total
Male	76	4401	4477
Female	111	3552	3663
Total	187	7953	8140

H_0: لا توجد علاقة بين الجنس وضغط الدم

H_1: توجد علاقة بين الجنس وضغط الدم

1- $Kappa\ Test$

$$Z_c = -3.9928 \quad decesion: Reject\ H_0$$

2- $Two - Samples\ Test\ for\ Binomial\ Proportion$

$$Z_c = 1.1811 \quad decesion: accept\ H_0$$

3–أمراض القلب والجنس، الفئة العمرية(45 – 64) سنة.

جدول (12): العلاقة بين الجنس وأمراض القلب.

	Yes	No	Total
Male	93	1183	1276
Female	113	1171	1284
Total	206	2354	2560

H_0: لا توجد علاقة بين الجنس وأمراض القلب

H_1: توجد علاقة بين الجنس وأمراض القلب

1- $Kappa\ Test$

$$Z_c = -1.4064 \quad decesion: Accept\ H_0$$

2- $Two - Samples\ Test\ for\ Binomial\ Proportion$

$$Z_c = 0.7257 \quad decesion: accept\ H_0$$

4-أمراض المخ والجنس، الفئة العمرية(45 − 64) سنة.

جدول (13): العلاقة بين الجنس وأمراض المخ.

	Yes	No	Total
Male	40	386	426
Female	46	319	365
Total	86	705	791

H_0: لا توجد علاقة بين الجنس وأمراض المخ

H_1: توجد علاقة بين الجنس وأمراض المخ

1- $Kappa\ Test$

$$Z_c = -1.4472 \quad decesion: accept\ H_0$$

2- $Two - Samples\ Test\ for\ Binomial\ Proportion$

$$Z_c = 0.8332 \quad decesion: accept\ H_0$$

5-قرحة المعدة والجنس، الفئة العمرية(45 − 64) سنة.

جدول (14): العلاقة بين الجنس وقرحة المعدة.

	Yes	No	Total
Male	74	1561	1635
Female	31	1643	1674
Total	105	3204	3309

H_0: لا توجد علاقة بين الجنس وقرحة المعدة

توجد علاقة بين الجنس وقرحة المعدة: H_1

1- *Kappa Test*

$$Z_c = 4.3876 \quad decesion: Reject\ H_0$$

2- *Two – Samples Test for Binomial Proportion*

$$Z_c = 1.5045 \quad decesion: accept\ H_0$$

6–الربو القصبي والجنس، الفئة العمرية(45 – 64) سنة.

جدول (15): العلاقة بين الجنس والربو القصبي.

	Yes	*No*	*Total*
Male	30	3585	3615
Female	37	2869	2906
Total	67	6454	6521

لا توجد علاقة بين الجنس والربو القصبي: H_0

توجد علاقة بين الجنس والربو القصبي: H_1

1- *Kappa Test*

$$Z_c = 0.1686 \quad decesion: Accept\ H_0$$

2- *Two – Samples Test for Binomial Proportion*

$$Z_c = 0.0193 \quad decesion: accept\ H_0$$

3-5 أختبار التطابق لجدول توافق $(2 \times k)$

يتكون الجدول من صفين و (k) من الأعمدة التي لها ترتيب معين وهنا لدينا جدول فيه فئات العمر مرتبة في الأعمدة ومتغير الصفوف هو الجنس (ذكر، أنثى) سوف نطبق ما يسمى $(Generalized\ odds)$ على البيانات المتمثلة بالجدول الآتي:

	1	2	…	k	$Total$
1	a_1	a_2	…	a_k	A
2	b_1	b_2	…	b_k	B
$Total$	n_1	n_2	…	n_k	N

ويحسب رقم التطابق، $\{C\ (of\ Concordances)\}$:

$$C = a_1(b_2 + b_3 + \cdots + b_k) + a_2(b_3 + \cdots + b_k) + \cdots + a_{k-1}b_k$$

ثم نحسب المقدار:

$$D = b_1(a_2 + a_3 + \cdots + a_k) + b_2(a_3 + \cdots + a_k) + \cdots + b_{k-1}a_k$$

ولأجراء الأختبار نوجد:

$$S = C - D$$

ثم نحسب أحصاءة (Z) من:

$$Z = \frac{S - \mu_s}{\sigma_s}$$

حيث $(\mu_S = 0)$ هو المتوسط تحت فرضية العدم H_O ، وان:

$$\sigma_S = \left[\frac{AB}{3N(N-1)}(N^3 - n_1^3 - n_2^3 - n_3^3 - \cdots - n_k^3)\right]^{\frac{1}{2}}$$

ثم تقارن قيمة (Z) المحسوبة مع (Z) الجدولية، فاذا كانت (Z) المحسوبة اكبر من (Z) الجدولية نرفض H_O .

وعند كتابة برنامج خاص بالأحصاءة (Z) للبيانات التي تمثل العلاقة بين الجنس والأصابة بأحد أنواع الأمراض المزمنة حيث أن:

لا توجد علاقة بين الجنس والأصابة بأحد الأمراض المزمنة الستة: H_0

توجد علاقة بين الجنس والأصابة بأحد الأمراض المزمنة الستة: H_1

ووجد أن (Z) المحسوبة ($Z = 0.94565$) أصغر من الجدولية عند مستوى معنوية (5%)، لذلك نقبل الفرضية (H_0) التي تنص على عدم وجود علاقة بين الجنس والأصابة بأحد الأمراض المزمنة.

الأستنتاجات

1- عند تطبيق أحصاءة كابا، وأختبار تساوي النسب (نسب الأصابة بالمرض) بين الرجال والنساء، تم قبول فرضية العدم (H_0) لكل من أحصاءة كابا وأختبار تساوي نسب ذي الحدين لكل من الفرضيات الخاصة بين العلاقات الثنائية، أي جداول التوافق (2×2) للعلاقة بين الجنسين والأمراض الستة المختلفة، كما هو واضح في الجداول السابقة.

2- أشارت أحصاءة أختبار التطابق لجدول توافق $(2 \times k)$ أيضاً الى قبول الفرضية (H_0) لأن $(Z_{cal.} = 0.94563)$ في حين كانت قيمة (Z) الجدولية عند مستوى معنوية 5% هي (1.96)، وبذلك نقبل (H_0) ونرفض (H_1).

3- عند أختبار العلاقة بين الأصابة بالمرض مع الجنس (ذكر، أنثى)، وللفئات العمرية المختلفة، (جداول (2×4))، كانت النتائج تشير الى رفض (H_0) التي تنص على عدم وجود أتجاه لتزايد الأصابة بأحد الأمراض المزمنة بالنسبة للعمر، حيث كانت :

$$\chi^2_{cal.} = 494.826, \quad \chi^2_{cal.} = 12.8666, \chi^2_{cal.} = 1089.76$$

وجميعها أكبر من $(\chi^2_{cal.})$، لذلك نرفض الفرضية (H_0).

4- يمكن توسيع العمل في هذا المجال وأدخال أختبارات أخرى معتمدة على أستخدام تحليل الأنحدار في حالة كون الفروق معنوية، مما يستدعي ذلك تجزأة مركبات الأنحدار لمعرفة سبب الأنحرافات المعنوية، ويمكن كذلك تقدير معامل الأنحادر الخطي لأعتماد المتغير (y) على العمر (x)، وأستخراج التباين الكلي، وأجراء أختبار الفرضيات.

5- يمكن أعتماد أحصاءة (χ^2) في أيجاد متوسط مربعات الخطأ التجريبي، فكلما كانت قيمة متوسط مربعات الخطأ التجريبي صغيرة دل ذلك على دقة الأختبار.

References

[1] AndrewR. Webb (2002)، "Statistical Pattern Recognition 2nd edition، John Wiley & Sons، LTD،.

[2]Chambers، R. L.، and D. G. Steel. 2001. "Simple methods for ecological inference in 2×2 tables". J. Roy. Statist. Soc. Ser. A **164**: 175-192.

[3] Chapt L. E. "Introduction to Biostatistics" John Wiley & Sons، LTD، 2003.

[4]Daniels، M. J.، and C. Gatsonis. 1999. "Hierarchical generalized linear models in the analysis ofvariations in health care utilization". J. Amer. Statist. Assoc. **94**: 29-42.

[5] Dawant B. M. & Garbay، CX. Eds، "Special topics section on Biomedical Dad a fusion" IEEE Transaction on Biomedical Engineering، 46(10)، 1999.

[6] Everitt، B. S. (1977)، "The analysis of contingency tables"، London، Chapman and Hall.

[7]Fisher، R. A. 1922. "On the interpretation of chi-square from contingency tables"، and thecalculation of P. J. Roy. Statist. Soc. **85**: 87-94.

[8] Finny، D. J. (1964،1971)، "Statistical Methods in Biological Assy"، 2nd edition، London;Griffin.

[9] Kutner، M.H.، Nachtsheim، C.J. and Neter، J.، 2006، " Applied linear Regression model"، 5th edition.

[10]Lipsitz، S. R.، and G. Fitzmaurice. 1996. "The score test for independence in $R \times C$ contingencytables with missing data". Biometrics**52**: 751-762.

[11]Rayner، J. C. W.، and D. J. Best. 2001. "AContingency Table Approach to Nonparametric Testing".London: Chapman & Hall.

[12] Risto Lehtonen & Erkki Pahkinen، " Practical Methods for Designed Analysis of complex surveys" 2nd edt. John Wiley & Sons، LTD، 2004.

[13] Simonoff، J. S. (2003). "Analyzing Categorical Data". NewYork: Springer.

[14]Stokes، M. E.، C. S. Davis، and G. G. Koch (2000). "Categorical Data Analysis Using the SAS System"، 2nd edn. Cary، NC: SAS Institute.

[15] Van der Ark، L.A.، Croon، M.A.، & Sijtsma، K. (editors) (2005). New Developments in CategoricalData Analysis for the Social and Behavioral Sciences. Mahwah، NJ: LawrenceErlbaum.

ملحق

نتيجة أختبار وجود أتجاه متزايد في نسبة الأصابة بأعتماد توزيع ثنائي الحدين حسب برنامج (*Minitab*) وكانت النتائج هي:

```
table 2x6
********
      2 rows read.
Storing in file: bin.mtj
Executing from file: bin.mtj

Data Display


 Row     C1     C2     C3     C4      C5      C6     SUMR

   1   8965   4573   1493   2959   12064   10457   40511
   2   8580   4622   1116   2758   11366   10122   38564


Data Display


 Row      xi   ni-xi      ni         Pi

   1    8965    8580   17545   0.510972
   2    4573    4622    9195   0.497335
   3    1493    1116    2609   0.572250
   4    2959    2758    5717   0.517579
   5   12064   11366   23430   0.514895
   6   10457   10122   20579   0.508139
   7   40511   38564   79075   0.512311


Data Display


 Row    SUMR      xi   ni-xi      ni         Pi   Si

   1   40511    8965    8580   17545   0.510972    1
   2   38564    4573    4622    9195   0.497335    2
   3            1493    1116    2609   0.572250    3
   4            2959    2758    5717   0.517579    4
   5           12064   11366   23430   0.514895    5
   6           10457   10122   20579   0.508139    6
```

```
Data Display


 Row        A         B       Chi-Sq   chi-tab     sig   df

   1   78.375   75455.5   0.0814074    3.8415   0.950    1
   2                                   10.8277   0.999    1
Executing from file: bin24.MTB
Retrieving worksheet from file: bin24.MTW
Worksheet was saved on  9/22/2012
********
table 2x4
********
      2 rows read.
Storing in file: bin.mtj
Executing from file: bin.mtj

Data Display


 Row     C1      C2      C3     C4    SUMR

   1   9151   13790   48289   2871   74101
   2   7408   12204   13755   5197   38564


Data Display


 Row      xi   ni-xi       ni         Pi

   1    9151    7408    16559   0.552630
   2   13790   12204    25994   0.530507
   3   48289   13755    62044   0.778302
   4    2871    5197     8068   0.355850
   5   74101   38564   112665   0.657711


Data Display


 Row      P-bar      xi   ni-xi      ni         Pi   Si

   1   0.657711    9151    7408   16559   0.552630    1
   2             13790   12204   25994   0.530507    2
   3             48289   13755   62044   0.778302    3
   4              2871    5197    8068   0.355850    4
```

Data Display

Row	A	B	Chi-Sq	chi-tab	sig	df
1	4351.2	17373.5	1089.76	3.8415	0.950	1
2				10.8277	0.999	1

Row	C1	C2	C3	C4	SUMR
1	394	417	136	17	964
2	325	431	200	10	966

Data Display

Row	xi	ni-xi	ni	Pi
1	394	325	719	0.547983
2	417	431	848	0.491745
3	136	200	336	0.404762
4	17	10	27	0.629630
5	964	966	1930	0.499482

Data Display

Row	P-bar	xi	ni-xi	ni	Pi	Si
1	0.499482	394	325	719	0.547983	1
2		417	431	848	0.491745	2
3		136	200	336	0.404762	3
4		17	10	27	0.629630	4

Data Display

Row	A	B	Chi-Sq	chi-tab	sig	df
1	-59.6704	276.729	12.8666	3.8415	0.950	1
2				10.8277	0.999	1

Data Display

Row	C1	C2	C3	C4	SUMR
1	10116	14956	15920	5137	46129
2	7378	13379	15425	6417	42599

Data Display

Row	xi	ni-xi	ni	Pi
1	10116	7378	17494	0.578255
2	14956	13379	28335	0.527828
3	15920	15425	31345	0.507896
4	5137	6417	11554	0.444608
5	46129	42599	88728	0.519892

Data Display

Row	P-bar	xi	ni-xi	ni	Pi	Si
1	0.519892	10116	7378	17494	0.578255	1
2		14956	13379	28335	0.527828	2
3		15920	15425	31345	0.507896	3
4		5137	6417	11554	0.444608	4

Data Display

Row	A	B	Chi-Sq	chi-tab	sig	df
1	-3136.70	19883.5	494.826	3.8415	0.950	1
2				10.8277	0.999	1

Data Display

Row	C1	C2	C3	C4	SUMR
1	1107	1159	739	65	3070
2	1112	1231	648	45	3036

Data Display

Row	xi	ni-xi	ni	Pi
1	1107	1112	2219	0.498873
2	1159	1231	2390	0.484937
3	739	648	1387	0.532805
4	65	45	110	0.590909
5	3070	3036	6106	0.502784

Data Display

Row	P-bar	xi	ni-xi	ni	Pi	Si
1	0.502784	1107	1112	2219	0.498873	1
2		1159	1231	2390	0.484937	2
3		739	648	1387	0.532805	3
4		65	45	110	0.590909	4

Data Display

Row	A	B	Chi-Sq	chi-tab	sig	df
1	69.7041	996.134	4.87752	3.8415	0.950	1
2				10.8277	0.999	1