احصاء حيوي

الكورس الثاني

(موضوع المحاضرة)

المقارنة بين مجموعتين من البيانات
البقاء على قيد الحياة

Dr.Safwan Nathem Rashed

المقارنة بين مجموعتين من البيانات البقاء على قيد الحياة

من اجل المقارنة بين مجموعتين من البيانات البقاء على قيد الحياة سوف نستخدم اختبار χ^2 والموضح خطوات الاختبار على النحو الاتي:

١. وضع فرضية العدم والبديلة:

H_0 : معدل الخطورة للمجموعة الاولى $=$ معدل الخطورة للمجموعة الثانية

H_1 : معدل الخطورة للمجموعة الاولى $(\neq, <, >)$ معدل الخطورة للمجموعة الثانية

٢. حساب قيمة المختبر الاحصائي بالاعتماد على الصيغة الرياضية الاتية:

$$\chi^2 = \sum_{j=1}^{k} \frac{\left(O_j - E_j\right)^2}{E_j}$$

حيث ان :

Oj : يمثل تكرار المشاهدة Observed frequency

Ej : يمثل تكرار المتوقع Expected frequency

k : يمثل عدد المجموعات حيث ان j=1,2,…,k وهنا لابد من ان تكون اقل قيمة لـ k=2 وذلك لوجود مجموعتين لغرض المقارنة

٣. نقارن قيمة χ^2 المحسوبة مع القيمة الجدولية لها عند مستوى معنوية (α) ودرجة حرية (k−1) وبوجود مجموعتين للمقارنة فان درجة الحرية تساوي واحد، وتعتمد على هذه الصيغة في ايجاد القيمة الحرجة والتي تختلف في البحث عنها في موضوع اختبارات تتعلق بالانحراف المعياري والتباين ، فاذا كانت القيمة المحسوبة اكبر من الجدولية نرفض فرضية العدم ونقبل بالفرضية البديلة والعكس صحيح.

ويطلق على هذا الاختبار ما بين مجموعتين (حسن المطابقة).

لنفرض ان الخطورة النسبية لأي مرض تعتمد على عدد المتوفين بذلك المرض ولفترات معينة وعليه تكون (Oj) تعتمد او تمثل عدد المتوفين (di). وللمقارنة بين مجموعتين من البيانات (اي مرضين) سوف يعتمد على عدد المتوفين بهذا المرض لكل مجموعة ولتكن (di_1, di_2) وعلى التوالي:

$$O_j = \sum_{i=1}^{t} d_{ij} \quad ; t = 1, 2, 3, \ldots \; ; j = 1, 2$$

حيث ان: t عدد المشاهدات

$$O_1 = \sum_{i=1}^{t} d_{i1} \quad ; \quad O_2 = \sum_{i=1}^{t} d_{i2}$$

j=1,2 : تمثل المجموعتين

وعليه فان تكرار كل مجموعة ستكون:

$$O_1 = \sum_{i=1}^{t} d_{i1} \quad ; \quad O_2 = \sum_{i=1}^{t} d_{i2}$$

اما بالنسبة لعدد المتوفين المتوقع للمجموعتين يتم ايجادها على النحو الاتي:

$$E_j = \sum_{i=1}^{t} Ed_{ij} \quad ; \; t = 1, 2, 3, \ldots \; ; \; j = 1, 2$$

اي ان:

$$E_1 = \sum_{i=1}^{t} Ed_{i1} \quad ; \quad E_2 = \sum_{i=1}^{t} Ed_{i2}$$

علماً ان:

$$Ed_j = \frac{d_i}{L_i^*} \times L_{ij}^*$$

$$Ed_1 = \frac{d_i}{L_i^*} \times L_{i1}^* \quad ; \quad Ed_2 = \frac{d_i}{L_i^*} \times L_{i2}^*$$

وان

$$d_i = d_{i1} + d_{i2}$$

$$L_i^* = L_{i1}^* + L_{i2}^*$$

$$L_{ij}^* = L_{ij} + \frac{1}{2}\left(W_{ij} + U_{ij}\right)$$

مثال: البيانات في الجدول الاتي تمثل عدد المتوفين والباقين على قيد الحياة والمعرضين للخطر خلال فترة الدراسة حول مرضين الاول مرض سرطان الكبد والثاني مرض سرطان القولون.

المطلوب/ هل خطورة سرطان الكبد اكبر من خطورة مرض سرطان القولون تحت عند مستوى معنوية 5%.

ti	di1	di2	Li1*	Li2*	di=di1+di2	Li=Li1*+Li2*
0-1	180	20	631	720	200	1351
1-2	160	60	430	700	220	1150
2-3	90	80	290	639	170	929
3-4	70	90	197	556	160	753
4-5	60	120	123	463	180	586
5-6	40	160	62	341	200	403
6-7	20	180	21	180	200	201
total	620	710				

وعليه فان:

$$O_j = \sum_{i=1}^{7} d_{ij} \quad ; t = 1,2,3,4,5,6,7 \quad ; j = 1,2$$

$$O_1 = \sum_{i=1}^{7} d_{i1} = 620 \quad ; \quad O_2 = \sum_{i=1}^{7} d_{i2} = 710$$

ti	Edi1=(di/Li*)*Li1*	Edi2=(di/Li*)*Li2*
0-1	93.400	106.588
1-2	86.087	133.910
2-3	53.068	116.932
3-4	41.859	118.141
4-5	37.782	142.218
5-6	30.770	169.231
6-7	20.896	179.104
total	363.872	966.124

وان:

$$E_j = \sum_{i=1}^{7} Ed_{ij} \quad ; t = 1,2,3,4,5,6,7 \quad ; j = 1,2$$

$$E_1 = \sum_{i=1}^{7} Ed_{i1} = 363.872 \quad ; \quad E_2 = \sum_{i=1}^{7} Ed_{i2} = 966.124$$

H_0 : خطورة مرض سرطان القولون = خطورة مرض سرطان الكبد

H_1 : خطورة مرض سرطان القولون > خطورة مرض سرطان الكبد

$$\chi^2 = \sum_{j=1}^{k} \frac{\left(O_j - E_j\right)^2}{E_j} = \frac{\left(620 - 363.872\right)^2}{363.872} + \frac{\left(710 - 966.124\right)^2}{966.124}$$

$$= 180.287 + 67.400 = 248.187$$

فان القيمة الجدولية لـ χ^2 :

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

القرار: فان حسن المطابقة بين المرضين سيكون من خلال مقارنة القيمة $\chi^2_{cal} > \chi^2_{table}$ المحسوبة لمربع كاي مع القيمة الجدولية التي اوضحت ان

وهذا يدل على رفض فرضية العدم والقبول بالفرضية البديلة، على ان خطورة سرطان الكبد فعلاً اكبر من خطورة سرطان القولون عند مستوى معنوية 5%.

مثال: البيانات الاتية تمثل اوقات البقاء الى حالة الوفات لمجموعتين من المرضى المجموعة الاولى مصابين بمرض سرطان الكلية والمجموعة الثانية مصابين بمرض الفشل الكلوي، قارن بين خطورة المرضين عند مستوى معنوية 5%.

اشهر المتابعة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢	١٣	١٤
مرضى سرطان الكلية	6	6	9*	10	10*	11	14	15*	15	17	19	20	20	22
مرضى الفشل الكلوي	6	8	8	10*	11*	15	17	18	20	21*	22	23	24*	25*

حيث ان: * : تمثل احياء المنسحبين اثناء المتابعة.

١- كون جدول الحياة.

٢- هل ان خطورة الاصابة بمرض سرطان الكلية اكبر من مرض الفشل الكلوي.

Sol/

$$L_1 = \sum d_1 + \sum W_1$$

$$O_j = \sum_{i=1}^{t} d_{ij} \quad ; t = 1,2,3,\ldots \; ; j = 1,2$$

$$E_j = \sum_{i=1}^{t} Ed_{ij} \quad ; \quad Ed_{ij} = \frac{d_i}{L_i^*} \times L_{i\,j}^*$$

$$L_i^* = L_{i1}^* + L_{i2}^* \quad ; \quad d_i = d_{i1} + d_{i2}$$

$$L_{i\,j}^* = L_{i\,j} + \frac{1}{2}\left(W_{ij} + U_{ij}\right) \quad ; \quad L_{i\,j} = \sum d_{i\,j} + \sum W_{i\,j}$$

اشهر المتابعة	di1	di2	Wi1	Wi2	Li1	Li2	Li1*	Li2*	di	Li*	Edi1	Edi2
6	2	1	0	0	14	14	14	14	3	28	1.5	1.5
8	0	2	0	0	12	13	12	13	2	25	0.96	1.04
9	0	0	1	0	12	11	12	11	0	23	0	0
10	1	0	1	1	11	11	11	11	1	22	0.5	0.5
11	1	0	0	1	9	10	9	10	1	19	0.47	0.53
14	1	0	0	0	7	9	8	9	1	17	0.47	0.53
15	1	1	1	0	8	9	7	9	2	16	0.88	1.13
17	1	1	0	0	5	8	5	8	2	13	0.77	1.23
18	0	1	0	0	4	7	4	7	1	11	0.36	0.63
19	1	0	0	0	4	6	4	6	1	10	0.4	0.6
20	2	1	0	0	3	6	3	6	3	9	1	2
21	0	0	0	1	1	5	1	5	0	6	0	0
22	1	1	0	0	1	4	1	4	2	5	0.4	1.6
23	0	1	0	0	0	3	0	3	1	3	0	1
24	0	0	0	1	0	2	0	2	0	2	0	0
25	0	0	0	1	0	1	0	1	0	1	0	0
	11	9	3	5							7.71	12.29

مثال/ واجب: البيانات في تمثل عدد المتوفين والمنسحبين خلال الدراسة لمرضى سرطان الرئة لأشخاص مدخنين وغير مدخنين.

المطلوب/ ١- كون جدول الحياة للمرضى المدخنين وغير المدخنين.

٢- هل ان المرضى المدخنين يتعرضون لخطورة اكبر اي ذات خطورة كبيرة من المرضى غير المدخنين عند مستوى معنوية 5% و 1% .

سنوات المتابعة	di1	di2	Wi1	Wi2
0-1	95	40	0	3
1-2	100	30	1	2
2-3	140	25	0	0
3-4	180	20	2	1
4-5	200	10	1	1