

واجب منزلي حول التحليل العنقودي والإنحدار اللوجستي

التمرين الأول:

يعطى في الجدول التالي الإنفاق الشهري على الغذاء X_1 ، والإنفاق على الملابس X_2 لعينة من 5 طلبة (الوحدة: ألف دج):

الطالب	a	b	c	d	e
الإنفاق على الغذاء X_1	2	8	9	1	8.5
الإنفاق على اللباس X_2	4	2	3	5	1

المطلوب:

1. تحديد الهدف من التحليل العنقودي. وفوائد استخدامه في التسويق.
2. التمثيل البياني للبيانات على معلم، حيث على محور الفواصل X_1 ومحور الترتيب X_2 .
3. تحديد العناقيد الممكنة من ملاحظة الرسم البياني.
4. إعداد مصفوفة التباعد بطريقة المسافة الإقليدية مع تفسيرها وتوضيح خصائصها.
5. إجراء التحليل العنقودي الهرمي لعينة الطلبة بطريقة الربط المنفرد.
6. مثل بيانيا المخطط الشجري (dendrogram) للتحليل العنقودي وعلق عليه.
7. من المخطط الشجري، استنتج عدد العناقيد المناسبة ومفردات كل عنقود، هل يتوافق ذلك مع الإجابة عن السؤال (3).

التمرين الثاني:

قام طبيب بتجربة دواء Rizatriptan للصداع النصفي على عينة من المرضى (12 مريض)، حيث يعطي لهم جرعات من الدواء بتركيز مختلفة x (ملغ)، ويتم قياس تأثير الدواء على المريض ($y=1$) الشعور بالارتياح خلال ساعتين؛ $y=0$ عدم الشعور بالارتياح خلال ساعتين)، نتائج التجربة معطاة في الجدول التالي:

المريض	a	b	c	d	e	f	g	h	i	j	k	l
تركيز الجرعة x	0 (وهمية)	0	2.5	2.5	2.5	5	5	5	5	10	10	10
النتيجة y	0	0	0	0	1	0	0	1	0	1	1	1

1. وضح أن تحليل العلاقة بين تركيز الدواء x ونتيجة العلاج y ، تتطلب تحليل الإنحدار اللوجستي وليس الإنحدار الخطي.
2. مثل انتشار البيانات المعطاة في الجدول على معلم مستوي وعلق على الرسم: x (على محور الفواصل) و y (على محور الترتيب).
3. أحسب معدل الارتياح من الصداع النصفي ومعدل عدم الارتياح خلال الساعتين انطلاقاً من بيانات الجدول.

علماً أنه تم تقدير معالم نموذج الإنحدار اللوجستي انطلاقاً من برنامج SPSS، وكانت النتائج كما في الجدول التالي:

البيان	قيم المعالم B	الخطأ المعياري S.E	اختبار Wald	Sig	درجة الحرية df	Exp(B)
الثابت	0.165	0.037	19.819	0.000	1	1.180
تركيز الجرعة x	-2.490	0.285	76.456	0.021	1	0.083

4. حلل المخرجات في الجدول السابق، ثم شكل معادلة نموذج الإنحدار اللوجستي: $\logit[\text{odds}(x)] = \alpha + \beta x$ ومعادلة نسبة الأرجحية: $\text{odds}(Y) = P_1(x)/P_0(x)$ ، حيث: $P_1(x)$ ؛ $P_0(x)$ هما احتمال الارتياح وعدم الارتياح على الترتيب من أجل تركيز الجرعة x .
5. استنتج تركيز الجرعة التي يتعادل عندها احتمال الارتياح واحتمال عدم الارتياح.
6. باستخدام نموذج الإنحدار وجواب السؤال (5)، استنتج هل يشعر مريضين بارتياح، وقد تم إعطاء أحدهما جرعة 4 ملغ، والآخر جرعة 6.5 ملغ.
7. انطلاقاً من التجربة، ماهي الاستنتاجات التي يمكن التوصل لها؟ ماهي محاذيرها على المرضى؟

أسئلة نظرية:

المطلوب: اختيار الجواب الصحيح في كل سؤال، علماً أن الإجابة تكون بذكر رقم السؤال وحرف الجواب المختار:

1. تسمى عملية تقسيم عناصر المجتمع المدروس إلى عدة مجموعات متعاقبة ومتشابهة داخليا ومتباينة خارجيا، فيكون التباين بين عناصر المجموعات أصغر ما يمكن، والتباين بين المجموعات أكبر ما يمكن:

أ. التحليل العنقودي ب. التحليل التمييزي ج. التحليل العاملي د. تحليل الانحدار هـ. تحليل الارتباط

2. يعبر عن التحليل العنقودي بمصطلح:

أ. Factor analysis ب. Cluster Analysis ج. Discriminant analysis د. Regression analysis

3. يشترط في..... وجود متغير تابع نوعي ثنائي (يأخذ قيمتان 0 أو 1)، أما المتغيرات المستقلة فلا شروط عليها، لتكملة مكان النقط نضيف:

أ. التحليل العنقودي ب. التحليل العاملي ج. تحليل الانحدار د. تحليل الانحدار اللوجستي

4. في ظاهرة ثنائية النتائج (إما نجاح أو فشل)، يطلق على النسبة بين احتمال النجاح واحتمال الفشل:

أ. الخطر المطلق ب. الخطر النسبي ج. نسبة الأرجحية د. الجذر الكامن هـ. التباين المشترك

5. يتطلب التحليل العنقودي إعداد مصفوفة المسافات بين مفردات العينة، هذه المصفوفة لها عدة خصائص، ليس من بينها:

أ. كل عناصرها $\in [-1, +1]$ ب. عناصر القطر = 0 ج. مصفوفة مربعة $(n \times n)$ د. مصفوفة متناظرة هـ. كل ما سبق

6. يتم استخدام مصفوفة التباعد في التحليل العنقودي في حالة ما إذا كانت المتغيرات المدروسة:

أ. متغيرات كمية ب. متغيرات اسمية ج. متغيرات ترتيبية د. متغيرات فئوية

7. في التسويق، تعتبر المبيعات، الأسعار، عدد العملاء وحصصة السوق من:

أ. متغيرات كمية ب. متغيرات اسمية ج. متغيرات ترتيبية د. متغيرات فئوية

8. يطلق على النسبة بين احتمال تحقق حدث واحتمال عد تحققه مصطلح:

أ. الخطر المطلق ب. الخطر النسبي ج. نسبة الأرجحية د. معامل الارتباط هـ. معامل الانحدار

9. تعتبر طريقة الربط المنفرد إحدى طرق تشكيل العناقيد في التحليل العنقودي، وتأخذ في الاعتبار..... المسافات بين المفردات و/

العناقيد التي يريد الباحث تجميعها في عناقيد أكبر، لكلمة مكان النقط نضيف:

أ. أصغر ب. متوسط ج. أكبر د. ليس ما سبق هـ. كل ما سبق

10. رغم تنوع أهداف التحليل العنقودي إلا أن أهمها على الإطلاق هو:

أ. اكتشاف الأنماط ب. التنبؤ بالسلوكيات ج. تصنيف المفردات د. اختزال البيانات هـ. تحديد الخصائص

11. تتمثل عناصر مصفوفة التباعد في التحليل العنقودي في..... بين مفردات العينة، لتكملة مكان النقط نضيف:

أ. معاملات الارتباط ب. معاملات الانحدار ج. المسافات الإقليدية د. التباينات المشتركة هـ. نسب التقارب

12. في اختبار خضوع العلاقة بين المتغيرين المستقل والتابع لنموذج الانحدار اللوجستي، نستخدم:

أ. تحليل التباين ANOVA ب. اختبار F- fisher ج. اختبار T- student د. اختبار R- square هـ. اختبار Wald