جامعة محمد خيضر (بسكرة)
كلية العلوم الاقتصادية والتجارية والتسيير
قسم العلوم التجارية

السنة الجامعية: 2026/2025
المستوى: أولى ماستر تسويق
مقياس: الأساليب الكمية في التسويق 2

حل سلسلة تمارين رقم (02) حول الانحدار اللوجستي

حل التمرين الأول:

1. حساب معدل الشراء عدم الشراء:

من الجدول (1) في التمرين، نلاحظ أن عدد العملاء الذي قاموا بالشراء من المتجر الإلكتروني: $n_1 = 10$، وعدد العملاء الذي تصفو الموقع خلال الأسبوع من دون شراء: $n_0 = 10$، وذلك من بين عينة العملاء ($n=20$)، أي أن المعدل العام للشراء:

$$p = \frac{n_1}{n} = \frac{n_1}{n_1+n_2} = \frac{10}{20} = 0,50$$

وبالمقابل فإن المعدل العام لعدم الشراء:

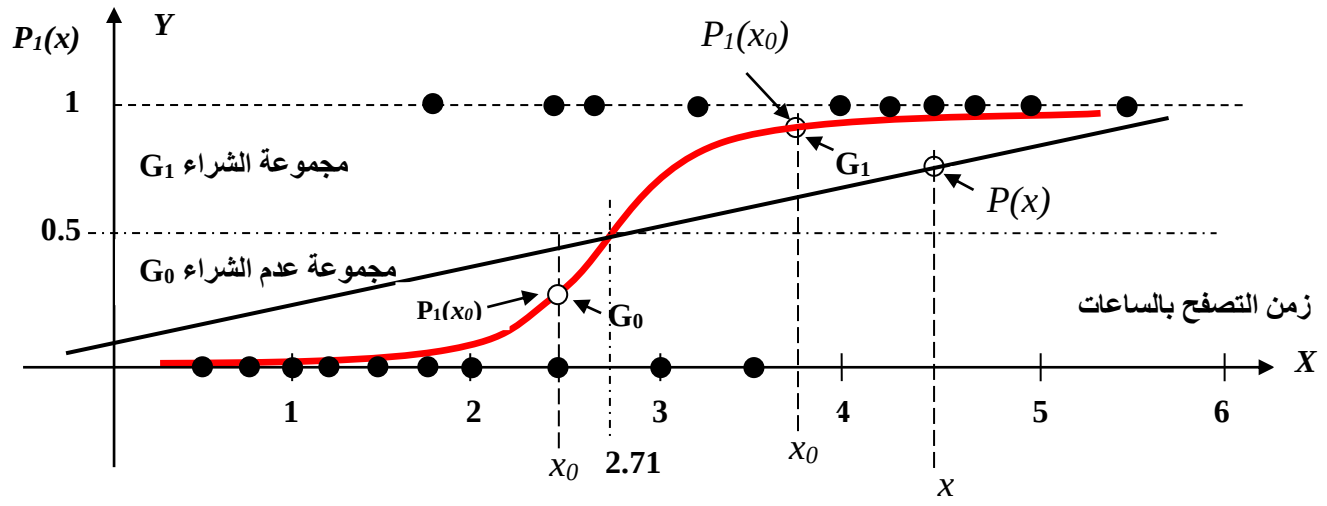
$$q = \frac{n_2}{n} = \frac{n_2}{n_1+n_2} = \frac{10}{20} = 0,50 \qquad q = 1- p = 1- 0,50 = 0,50$$

أما الأرجحية (احتمال الشراء إلى احتمال عدم الشراء) فتحسب كما يلي:

$$odds = \frac{p}{q} = \frac{p}{1- p} = \frac{0,50}{0,5} = 1,00 \qquad odds = \frac{n_1}{n_2} = \frac{n_1/(n_1+n_2)}{n_2/(n_1+n_2)} = \frac{0,50}{0,5} = 1,00$$

نستنتج من العينة المدروسة أن إمكانية أو احتمالية الشراء تساوي إمكانية أو احتمالية عدم الشراء في عينة العملاء المتصفحين المدروسة، وتكتب الأرجحية بالشكل (1: 1).

2. رسم شكل انتشار البيانات:



من رسم شكل الانتشار لهذه البيانات، فإننا نلاحظ أن:

- المتغير المستقل X هو متغير كمي مستمر (عدد الساعات)، حيث يأخذ قيمه في المجال [6 , 0]؛
- المتغير التابع Y فهو متغير منفصل (متقطع) وثنائي القيمة، فهو يأخذ القيمة (Y=1) في حالة الشراء، ويأخذ القيمة (Y=0) في حالة عدم الشراء؛
- إذا قمنا برسم النقاط (x_i , y_i)، حيث: $i = 1 \dots 20$ (20 عميل)، نجد أن قيم Y الصفرية (حالات عدم الشراء) تتموضع على محور الفواصل (OX)، وتقع أغلبها نحو اليسار (بالقرب من نقطة المبدأ O)، أما قيم Y المساوية للـ 1(حالات الشراء)، فتقع على المستقيم ذو المعادلة $Y = 1$، وأغلبها تتموضع على يمين هذا الخط، كما توجد بعض النقط المتقابلة في المنطقة الوسطى. وهذا ما يجعل المنحنى يشبه حرف الفرنسية S، ولذا يسمى المنحنى السيني Sigmoid curve، ودالته الدالة السينية أو دالة سيجموند (Sigmoid function).

3. تحليل جدول المخرجات:

أ. اختبار Wald:

هو وسيلة لاختبار أهمية متغيرات تفسيرية معينة في نموذج إحصائي. وفي الانحدار اللوجستي لدينا متغير تابع ثنائي ومتغير مستقل مفسر واحد أو أكثر، ولكل متغير تفسيري في النموذج سيكون هناك معلمة مرتبطة به، وهنا يعتبر اختبار Wald أحد الطرق لاختبار ما إذا كانت المعلمات المرتبطة بمجموعة من المتغيرات المستقلة التفسيرية تساوي أصفار أم لا، فإذا كان اختبار Wald ذا دلالة معنوية (Sig < 0.05) بالنسبة لمتغير تفسيري معين أو مجموعة من المتغيرات التفسيرية، فإننا نستنتج أن المعلمات المرتبطة بهذه المتغيرات ليست صفرًا، وبالتالي يجب تضمين المتغيرات في النموذج، وإذا لم يكن اختبار والد ذا دلالة، فيمكن حذف هذه المتغيرات التفسيرية من النموذج.

لبيان أهمية معامل الانحدار في نموذج الانحدار اللوجستي نستخدم *Wald*، والذي له توزيع مربع كاي، نقارن مستوى الدلالة المعنوي لـ test wald (في الجدول) مع معيار المعنوية المعتمدة 0.05، ونلاحظ من الجدول الثاني للمعطيات، أن عدد ساعات التصفح X ترتبط معنويا مع احتمال الشراء من المتجر الإلكتروني *P(x)*، وذلك لأن مستوى الدلالة المعنوية Sig لمعامل الانحدار حسب اختبار Wald تساوي 0.0167، وهي أصغر من قيمة مستوى الدلالة المعتمد 0.05.

ب. معادلة نموذج الانحدار اللوجستي:

تمثل معادلة نموذج الانحدار اللوجستي أو معادلة لوجيت بالعلاقة بين عدد ساعات التصفح الإلكتروني x ولوغاريتم احتمال الشراء $\text{logit}[P_1(x)]$، وهي:

$$\text{logit}[P_1(x)]= \left[\frac{P_1(x)}{1- P_1(x)} \right] = \alpha + \beta\, x\ ; \quad x \in [0\ ,\ +\infty[\text{ أو } [0\ ,\ 6]$$

استخرجنا $\text{logit}[P_1(x)]= \left[\frac{P_1(x)}{1- P_1(x)} \right] = 1.5046\, x - 4.0777 = 1.5046[x - 2.71]$ هو ما يفيدنا في الإجابة على السؤال الموالي:

ج. صيغة الأرجحية:

ومنه نجد معادلة الأرجحية (النسبة بين احتمال الشراء واحتمال عدم الشراء) بدلالة x، وهي:

$$odds(Y)= \frac{P_1(x)}{1- P_1(x)} = e^{1.546(x - 2.71)}$$

4. استنتاج فترة التصفح التي يكون يتعادل عندها احتمال الشراء وعدم الشراء: من الصيغة:

$$odds(Y)= \frac{P_1(x)}{1- P_1(x)} = e^{1.546(x - 2.71)}$$

ومنها نجد أنه إذا كان عدد الساعات $x = 2.71$ فإن: $odds(Y) = e^{0} = 1$، وفي هذه الحالة نستنتج أن إمكانية الشراء (odds) تساوي إمكانية عدم الشراء، وان احتمال كل منهما يساوي 0.50، وفي هذه الحالة (عندما $x=2.71$)، القول: أرجحية الشراء مقابل عدم الشراء هي كما يلي: 1 مقابل 1، وكتب كما يلي (1: 1).

5. مناقشة قرارات الشراء من أجل زمن تصفح $x=1$ و$x=6$ بالساعات:

لحساب احتمال الشراء من أجل فترة تصفح x معينة: $P_1(x)$ من النموذج، نجد أن:

$$P_1(x)= \frac{1}{1 + e^{-(-4.0777 + 1.5046\, x)}} = \frac{1}{1 + e^{4.0777 - 1.5046\, x}}$$

أ. عندما يكون $x = 1$ نجد أن احتمال الشراء وعدم الشراء يساويان:

$$P_1(1)= \frac{1}{1 + e^{4.0777 - 1.5046\,(1)}} = 0.0709 \qquad P_0(1)= 1 - 0.071= 0.93$$

وهذا يعني أن احتمال الشراء لمن يتصفح الموقع لمدة ساعة واحدة أسبوعيا ($x=1$)، فإن احتمال الشراء ضعيف جدا (حوالي 7% فقط)، وهو أقل بكثير من احتمال عدم الشراء (حوالي 93%)، لذلك نصنف هذا العميل في المجموعة G_0 (فئة غير المشترين أو العملاء ضعيفي المردودية للمؤسسة).

وتكون نسبة الأرجحية لهذا العميل (الذي يتصفح الموقع ساعة واحدة):

Odds (x=1)= $P_1(x=1)/P_0$ (x=1) =0.0709/ 0.93= 0.076

ويمكن كتابتها بالشكل التالي: (13.11 : 1) ≈ 0.076؛ احتمال الشراء تقريبا 13 مرة احتمال عدم الشراء.

ب. عندما يكون $x = 6$ نجد أن احتمال الشراء $P_1(6)$ وعدم الشراء$P_0(6)$ يساويان:

$$P_1(x=6)= \frac{1}{1 + e^{4.0777 - 1.5046\,(6)}} \approx 0.99 \qquad P_0(x=6) \approx 1 - 0.990.01$$

وهذا يعني أن احتمال شراء العميل الذي يتصفح الموقع مدة (x= 4) ساعات يكاد يصل إلى 100%، ولا مجال لمقارنته باحتمال عدم الشراء (0.01)، لذلك نصنف ذلك هذا العميل في المجموعة G_1 (مجموعة المشترين).

وتكون نسبة الأرجحية لهذا العميل (الذي يتصفح الموقع لمدة 6 سا):

Odds (6)= $P_1(6)/P_0(6)$ =0.99/ 0.01= 99

ويمكن كتابتها بالشكل التالي: (1 : 99) ≈ 99.00؛ احتمال الشراء تقريبا 99 مرة احتمال عدم الشراء. أي أن العميل بعد تصفحه للموقع لمدة 6 ساعات في الأسبوع فإن أكيد تقريبا سيشتري.

ملاحظة: بنفس الطريقة يمكن حساب احتمال الشراء، الأرجحية ودالة لوجيت بالنسبة لقيم x (مدة تصفح الموقع الإلكتروني من طرف عميل ما): 2؛ 3؛ 4، 5 و 6 ساعة كما في الجدول التالي:

الجدول رقم (3): قيم التحليل اللوجستي

عدد ساعات التصفح X	مؤشرات الشراء من الموقع الإلكتروني بعد التصفح		
	قيمة ln (odds)	قيمة odds	احتمال الراء $P_1(x)$
1	**-2.57**	**0.076 ≈ 1 :13.3**	**0.07**
2	**-1.07**	**0.35 ≈ 1 :12.91**	**0.26**
2.71	**0.00**	**1 ≈ 1 :1**	**0.50**
3	**0.44**	**1.56 : 1**	**0.61**
4	**1.94**	**6.69 : 1**	**0.87**
5	**3.45**	**32.33 :1**	**0.97**
6	**4.95**	**99.00 : 1**	**0.99**

6. إنشاء جدول التصنيف Classification Table:

اللوجستي (Predicted) الفعلي (Observed)	الشراء (G_1 , P(x)> 0.5 , x> 2.71)	عدم الشراء (G_0 , P(x)< 0.5 , x< 2.71)	المجموع
الشراء (G_1 , Y=1)	n_{11}= 8	n_{12}= 2	n_1= 10
عدم الشراء (G_0 , Y=0)	n_{21}=2	n_{22}= 8	n_2= 10
المجموع	n_1'= 10	n_1'= 10	n= 20

حساب معدل التصنيف الصحيح R ومعدل التصنيف الخاطئ MR:

يتحقق التصنيف الصحيح عندما يتوافق توقع نموذج الانحدار اللوجستي مع النتائج الفعلية التجريبية، وهو يشير إلى جودة نموذج الانحدار اللوجستي في التنبؤ، ويتم حساب معدل التصنيف الصحيح Correct Classification Rate عن طريق نسبة مجموع عناصر القطر الرئيسي على المجموع الكلي n:

$$R= \frac{n_{11} + n_{22}}{n} = \frac{8 + 8}{20} = \frac{16}{20} = 0{,}80 = 80\%$$

يتحقق التصنيف الخاطئ عندما يفشل نموذج الانحدار في توقع قراء العميل المتصفح (شراء، عدم شراء)، من خلال عدم توافق توقع النموذج مع النتائج الفعلية، وارتفاعه يشير إلى عدم كفاءة نموذج الانحدار اللوجستي في التنبؤ،

ويتم حساب معدل التصنيف الخاطئ Misclassification rate عن طريق نسبة مجموع عناصر القطر غير الرئيسي على المجموع الكلي n:

$$MR= \frac{n_{12} + n_{21}}{n} = \frac{2 + 2}{20} = \frac{4}{20} = 0,20 = 80\% \qquad MR= R= 1- 0,80 = 0,20= 20\%$$

الاستنتاجات:

يتميز نموذج الانحدار اللوجستي عالي الجودة عمومًا بدقة تصنيف إجمالية تتراوح بين 70% و90%، وذلك اعتمادًا على مدى تعقيد مجموعة البيانات ومعيار المجال الصناعي أو الخدمي، ويتم التحقق من صحة النماذج عالية الجودة باستخدام مقاييس تأخذ في الاعتبار عدم توازن البيانات. وبما أن معدل التصنيف الصحيح 80%، فنستنتج أن ننموج الانحدار اللوجستي المستخرج في هذا التمرين ذو جودة وملاءمة عالية.

7. اقتراحات تحسين مبيعات الموقع الإلكتروني للمؤسسة:

بما أن مشتريات العملاء تتزايد بتزايد فترة التصفح للموقع الإلكتروني (علاقة طردية اثبتها تحليل الانحدار اللوجستي)، فإن زيادة جاذبية الموقع، تؤدي إلى بقاء العملاء مدة أطول في تصفح منتجات وخدمات المؤسسة، وبالتالي ارتفاع احتمالات الشراء، ويمكن زيادة جاذبية الموقع اللكتروني من خلال عدة أساليب مثل: التصميم المنظم، المزج بين المرئي والمقروء، سهولة التنقل بين صفحات الموقع، إمكانية الوصول للموقع عبر الأجهزة المختلفة (هاتف محمول، جهاز لوحي، كمبيوتر ...)، معلومات وافية ومفيدة عن منتجات وخدمات الموقع الجديدة، الإعلانات والتخفيضات، التواصل مع العملاء، من خلال نافذة الاقتراحات والتعليقات، الدردشات الحية ... الخ.