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

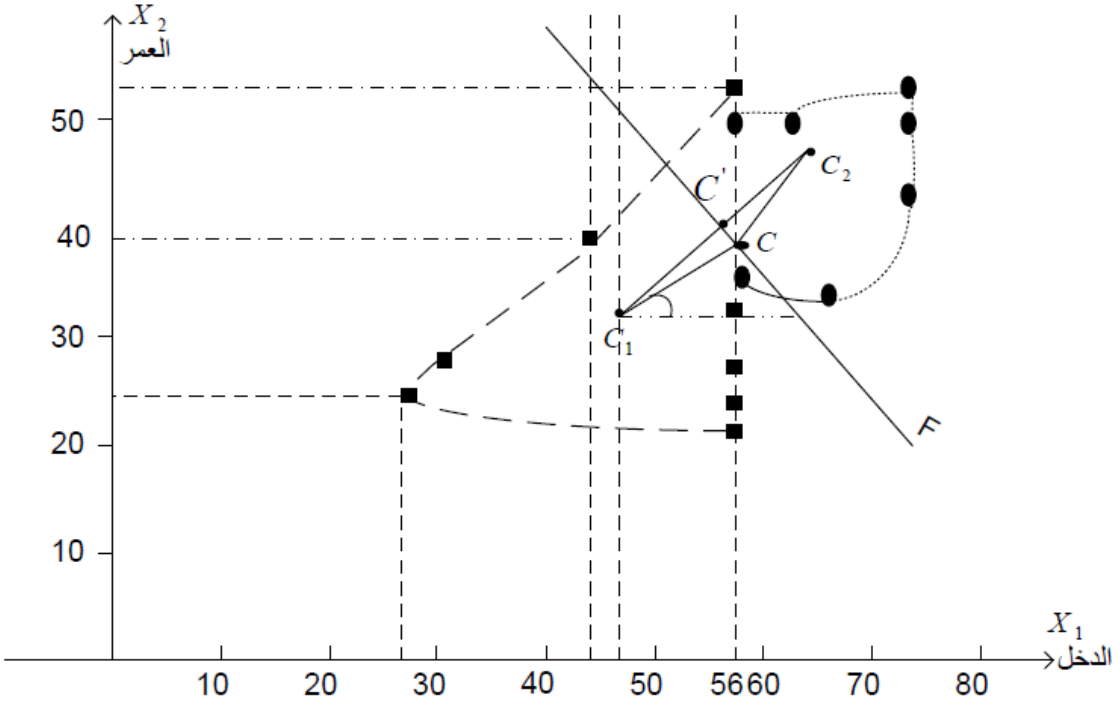
السنة الجامعية: 2025/2024
المستوى: أولى ماستر تسويق
مقياس: الأساليب الكمية في التسويق 2

حل سلسلة تمارين رقم (03) حول التحليل التمييزي

حل التمرين الأول:

1. التمثيل البياني لمجتمع القروض الكلي على معلم مستوى:

عند رسم القروض كنقاط هندسية إحداثياتها (x_1, x_2) في المستوي نجعل المحور الأفقي للدخل x_1، والمحور الرأسي للعمر x_2، نحصلا على الشكل البياني التالي:



2. التحديد والتمثيل البياني للنقاط المتوسطة للمجموعة G_1 وG_2 والمجتمع ككل G_1UG_2 (اتحاد المجموعتين):

ليكن C_1 مركز المجموعة الأولى G_1، واحداثياته $C_1(48\ ,\ 32)$، وC_2 مركز المجموعة الثانية G_2 واحداثياته $C_2(65.14\ ,\ 44.86)$.

كما يمكننا حساب ورسم المركز العام لهذه القروض C، ونحسب إحداثياته من متوسطي الدخل والعمر في العينتين معا، فنحصل على $C(56\ ,\ 38)$، كما يمكن حساب القيمة الوسطى لمركزي المجموعتين من متوسطي إحداثيات المركزين السابقين، فنحصل على القيمة الوسطى للمركزين، ونرمز لو بـ ’C، وتكون إحداثياته:

$$C'[(48+65.14)/2\ ,\ (32+44.86)/2] = C'(56.57\ ,\ 38.43)$$

ملاحظة: إن القيمة الوسطى ’C تقع على القطعة المستقيمة $[C_1C_2]$، وتقع في منتصفها، ولكن المركز العام C يقع خارج القطعة المستقيمة $[C_1C_2]$، كما موضح على الشكل.

السؤال: كيف يمكن صياغة علاقة رياضية للتمييز بين هاتين المجموعتين؟

الجواب: يمكن أن يكون بعدة طرق رياضية نستخدم بعضها:

أولا. طريقة دالة التمييز القانوني Canonical Discriminant Function

تسمى طريقة المستقيم الفاصل بين عناصر المجموعتين، وتتلخص هذه الطريقة بإنشاء مستقيم يمر من نقطة القيمة الوسطى للمركزين $C'(56.57\ ,\ 38.43)$ (أو من أي نقطة مناسبة أخرى)، ويتعامد (أو يتقاطع) مع القطعة المستقيمة الواصلة بين مركزي المجموعتين $C_1(48\ ,\ 32)$، و $C_2(65.14\ ,\ 44.86)$، فنحصل على مستقيم يفرز عناصر العينة إلى مجموعتين جديدتين منفصلتين: المجموعة الأولى تقع تحته، والمجموعة الثانية تقع فوقه.

3. تحديد معادلة المستقيم الفاصل ورسمه:

لإيجاد معادلة هذا المستقيم الفاصل بين هاتين المجموعتين، نقوم بإيجاد ميل القطعة المستقيمة $[C_1C_2]$ الواصلة بين مركزي المجموعتين من العلاقة:

$$m= (44.86 - 32)/(65.14 - 48) = 12.86/17.14 = 0.75029$$

ثم نقوم بحساب ميل المستقيم المتعامد معه من العلاقة:

$$m'= -1/m = -1/0.75029 = -1.3328$$

ثم نقوم بحساب معادلة المستقيم المعامد، والذي يمر من نقطة القيمة الوسطى للمركزين (38.43 , 56.57)C' من العلاقة:

$$X_2- 38.43 = - 1.3328(X_1 – 56.57)$$

وبعد الإصلاح نحصل على معادلة المستقيم التالي:

$$X_2 +1.3328X_1 – 113.8265=0$$

وهنا نلاحظ أن هذا المستقيم يقسم نقاط المستوى إلى نصفين: نصف موجب يقع فوقه ونصف سالب يقع تحته، وجزء معدوم يقع عليه، كما أن هذا المستقيم يفرز عناصر العينة إلى مجموعتين منفصلتين:

المجموعة الأولى: هي النقاط التي تقع تحته (في النص السالب)، وتضم معظم نقاط القروض المتعثرة في G_1 [ماعدا القرض الأول (56,52)].

المجموعة الثانية: هي النقاط التي تقع فوقه (في النص الموجب)، ويضم معظم نقاط القروض غير المتعثرة في G_2 (ماعدا القرض الأخير (36 , 56)].

أي أنه يمكننا استخدام هذا المستقيم (أو أي مستقيم أفضل منه)، كأداة لفرز القروض إلى متعثرة وغير متعثرة باحتمال خطأ معقول.

4. تحديد معادلة دالة التمييز F(x) وكيفية الاستفادة منها في التحليل التمييزي:

ولذلك نفترض أن معادلة المستقيم تساوي دالة F ونكتبها كما يلي: $F= x_2+ 1.3328x_1 – 113.8265$

ولاستخدام هذه المعادلة في تصنيف عناصر العينة أو في تصنيف أي عنصر جديد من المجتمع، نقوم بتعويض إحداثيات صاحب القرض (x_1,x_2) في المعادلة السابقة ونحسب قيمة F، ثم نتخذ القرار (في هذا المثال) حسب إشارة قيمة F كما يلي:

إذا كانت $F\leq 0$، فإن القرض ينتمي إلى المجموعة المتعثرة G_1. أما إذا كانت $F>0$، فإن القرض ينتمي إلى المجموعة غير المتعثرة.

واذا طبقنا ذلك على عناصر العينتين المسحوبتين من G_1 وG_2، نحصل على السطر F في الجدول، اوذا قارنا التصنيف الجديد مع التصنيف الفعلي أو الإداري السابق، نحصل على الجدول التالي:

	القروض المتعثرة (العينة n_1 من المجموعة G_1)									القروض غير المتعثرة (العينة n_2 من المجموعة G_2)							
رقم	1	2	3	4	5	6	7	8	$\bar{x}_1$	1	2	3	4	5	6	7	$\bar{x}_2$
الدخل x_1	56	56	45	56	29	56	56	30	48	72	72	56	72	64	64	56	56.14
العمر x_2	52	23	40	34	26	25	27	29	32	54	50	48	44	47	35	36	44.84
إشارة F	+	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-	+

5. استخدام معادلة التمييز في تصنيف جديد لمجتمع القروض + المقارنة مع التصنيف الأصلي:

تكون نتائج التصنيف الجديد نطلاقا من تطبيق الدالة التمييزية كما يلي: (يتم إضافة التصنيف الأصلي الفعلي الذي انطلقنا منه في البداية، أي المجموعتين G_1 وG_2) :

	التصنيف الجديد		
التصنيف الفعلي الأصلي	المجموعات	G_1	G_2	المجموع
	G_1	7	1	8
	G_2	1	6	7

نلاحظ أن نسبة التصنيف الصحيح تساوي نسبة مجموع عدد العناصر القطرية على حجم العينة الكلية n، أي:

$$P= (7+8)/15=0.8667= 86.67\%$$

وأن نسبة التصنيف الخاطئ تساوي نسبة مجموع العناصر غير القطرية على حجم العينة الكلية n، أي:

$$P= (1+1)/15=0.1333= 13.33\%$$

وهذا يعني أنه يمكننا الاعتماد على المستقيم السابق لتصنيف عناصر المجتمع مع خطأ معقول.

6. تصنيف عميل دخله 55 ألف دج، وعمره 40 سنة: $x_1= 55, x_2= 40$

بتعويض دخل عمر العميل في الدالة التمييزية F نجد: $F= 40+ 1.3328(55) – 113.8265= - 0.5225<0$

بما أن $F\leq 0$، فإن هذا العميل الجديد ينتمي إلى مجموعة القروض المتعثرة. وبالتالي لا انصح البنك بمنح القرض لهاذا العميل.

7. العيوب المتوقعة لطريقة دالة التمييز القانوني (طريقة المستقيم الفاصل):

- المستقيم الفاصل بين المجموعتين ليس هو الفاصل المثالي بينهما، بل هو أحد المستقيمات الفاصلة بينهما، لأنه نقله نحو الأعلى أو الأسفل، نحو اليمين أم اليسار، كما يمكن تغيير ميله m' من خلال تدويره حول النقطة الوسطى بزاوية θ، فنحصل على ما لا نهاية من المستقيمات الفاصلة بين المجموعتين.
- عندما تكون المجموعات متقاطعة ومتشابكة (أغلب عناصر كل مجموعة لا تقع في جهة واحدة من المستقيم)، فإنها تكون غير قابلة للفصل بواسطة مستقيم، وتسمى بالمجموعات غير القابلة للفصل خطيا، وفي هذه الحالة نلجأ إلى استخدام منحنيات مناسبة للفصل بينها، وهو ما يسمى التحليل التمييزي غير الخطي.
- لا تصلح هذه الطريقة إلا في حالة وجود متغيرين ومجموعتين، وهي حالة خاصة فقط في التحليل التمييزي، حيث غالبا تتواجد عدة متغيرات وعدة مجموعات، كما قد تكون المتغيرات كمية أو نوعية، فئوية أو ترتيبية، وهو ما يجعل التمثيل على مستوي غير ممكن.

ثانيا. 8. طريقة الجوار الأقرب (أصغر المسافات عن مراكز المجموعات):

تعتمد هذه الطريقة على حساب مربع المسافة بين كل عنصر x_0 ومركزي المجموعتين (G_1,G_2)، واللذين يساويان حسب الجدول ما يلي:

إحداثيات مركزي المجموعتين بالرموز والأعداد:

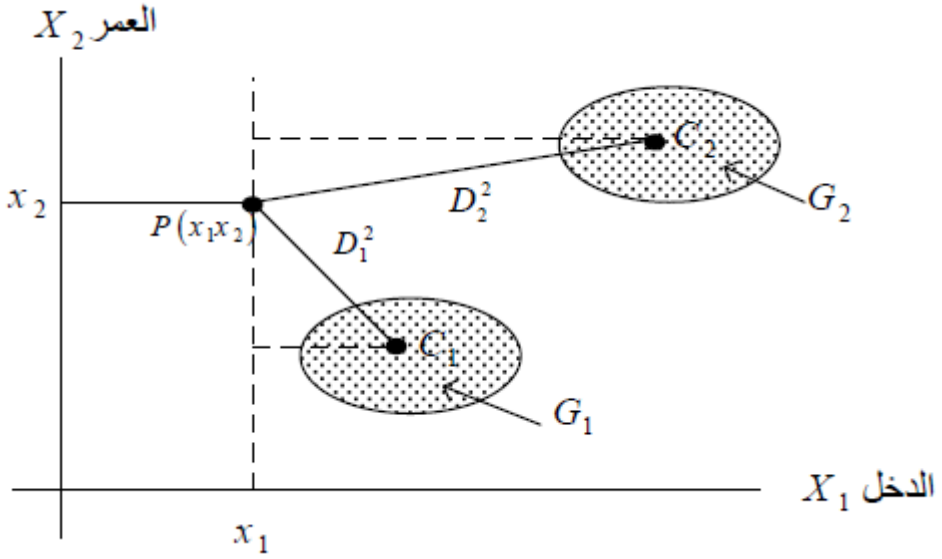
	G_1	G_2
X_1	$\bar{x}_{11}$	$\bar{x}_{12}$
X_2	$\bar{x}_{21}$	$\bar{x}_{22}$

⇨

	G_1	G_2
X_1 الدخل	48	65.14
X_2 العمر	32	44.86

لنأخذ نقطة ممثلة لأحد القروض الجديدة مثل النقطة P، والتي إحداثياتيا يساويان $P(x_1,x_2)$، ولنرمز لمربعي المسافة الهندسية منها حتى مركزي المجموعتين C_1 وC_2 بالرمزين D_1^2 وD_2^2 على الترتيب ونرسمها على الشكل البياني كما يلي:

الشكل (02): المسافات عن مركزي المجموعتين



ولتحديد المجموعة التي ينتمي إليها ذلك القرض الجديد نحسب مربعي المسافتين D_1^2 وD_2^2 ثم نقارن بينهما، ونتخذ القرار حول الانتماء كما يلي:

إذا كانت: $D_1^2 \leq D_2^2$ فإننا ننسب النقطة P إلى المجموعة G_1. أما إذا كانت: $D_1^2 > D_2^2$ فإننا ننسب النقطة P إلى المجموعة G_2.

ولكن هذه المقارنة لا تجوز إلا إذا كانت القيم العددية لهاتين المسافتين مقاسة بوحدات قياس موحدة، وبما أن X_1 يقاس بألف دج، وX_2 يقاس بالسنوات، فلا يجوز مقارنة هذين المتغيرين بطريقة مباشرة. وللتخلص من هذه المشكلة، نقوم بتحويل هذين المتغيرين إلى متغيرين معياريين، وذلك بإجراء التحويل المعياري: $Z= (X- \bar{x})/\sigma$ على كل متغير وفي كل مجموعة على حدة، فنحصل على قيم مجردة من الوحدات، ولتوضيح ذلك نحسب الانحرافات المعيارية لكل متغير وفي كل مجموعة على حدة، فنحصل على الانحرافات المعيارية التالية:

الانحرافات المعيارية لعناصر العينيتن:

	G_1	G_2
X_1	S_{11}	S_{11}
X_2	S_{11}	S_{11}

⇨

	G_1	G_2
X_1	12.036	7.198
X_2	9.769	7.081

وبذلك نجد أن مربع المسافة المعيارية بين النقطة $P(X_1, X_2)$ ومركز المجموعة الأولى $C_1(48 , 32)$

$$D_1^2= \frac{(X_1-48)^2}{12.036} + \frac{(X_2-32)^2}{9.769}$$

وبذلك نجد أن مربع المسافة المعيارية بين النقطة $P(X_1, X_2)$ ومركز المجموعة الثانية $C_2(65.14 , 44.86)$

$$D_2^2= \frac{(X_1-65.14)^2}{7.198} + \frac{(X_2-44.86)^2}{7.081}$$

ومنهما نحصل على قيم D_1^2 وD_2^2 كقيم عددية مجردة من وحدات القياس، لذلك يمكن مقارنتهما وتحديد المجموعة التي تنتمي إليها النقطة P، وهي المجموعة التي تقابل أصغر المربعين D_1^2 وD_2^2.

فمثلا نجد أن مربعي المسافة للقرض الأول من G_1، وهو الذي يقابل النقطة $P(56 , 52)$:

$$D_1^2= \frac{(56-48)^2}{12.036} + \frac{(52-32)^2}{9.769} = 0.44418 + 4.1914= 4.6332$$

$$D_2^2= \frac{(56- 65.14)^2}{7.198} + \frac{(52- 44.86)^2}{7.081} = 1.6124 + 1.0167 = 3.6291$$

وعند المقارنة نجد أن: $D_2^2 < D_1^2$، وهذا يعني أن النقطة P هي اقرب إلى مركز المجموعة G_2، لذلك سنعتبر أن القرض قرضا غير متعثر وينتمي إلى المجموعة G_2، مع أنه ينتمي إلى المجموعة G_1.

وهكذا قمنا بحساب مربعات المسافات المقابلة لبعض عناصر العينتين (عينة القروض المتعثرة وعينة القروض غير المتعثرة)، ووضعناها في سطرين ملحقين بالجدول، وبناء على مقارنة D_1^2 وD_2^2 في هذين السطرين تم تحديد انتماء كل قرض من عناصر عينة البحث، ووضعناها في السطر الأخير، فحصلنا على نفس النتائج التي حصلنا عليها بطريقة المستقيم الفاصل المبينة في الجدول التالي:

	القروض المتعثرة (العينة n_1 من المجموعة G_1)									القروض غير المتعثرة (العينة n_2 من المجموعة G_2)							
رقم	1	2	3	4	5	6	7	8	$\bar{x}_1$	1	2	3	4	5	6	7	$\bar{x}_2$
الدخل x_1	56	56	45	56	29	56	56	30	48	72	72	56	72	64	64	56	56.14
العمر x_2	52	23	40	34	26	25	27	29	32	54	50	48	44	47	35	36	44.84
D_1^2	4.6	1.29	6.88	…	…	…	…	2.33	8.95	..	..	..	..	..	..	0.51	..
D_2^2	2.6	10.14	8.30	...	...	...	...	28.84	1.57	..	..	..	..	..	..	31.8	..
التصنيف	2	1	1	1	1	1	1	1	1	2	2	2	2	2	2	1	2

ملاحظة: يمكن اكمال حساب المسافات المربعة D_1^2 وD_2^2 بنفس الطريقة التي حسبنا فيها المسافتين للخلية رقم (1) للمجموعة الأولى ذات القروض المتعثرة، ثم القيام بعملية التصنيف، سواء للمجموعة (1) أو المجموعة (2)، كما هو واضح في الجدول أعلاه.

ويمكن الاستفادة من هذه الطريقة في تحديد انتماء أي قرض جديد والتنبؤ بمصيره قبل اتخاذ قرار الموافقة على منحه.

ملاحظة: لنأخذ المجموعتين الجديدتين G_1' وG_2'، اللتين أصبحتا بعد التصنيف الجديد يضمان العناصر التالية:

G_1'= {$7_{(1)}$, 1, 2, 3, 4, 5, 6, 8} G_2'={1, 2, 3, 4, 5, 6, $1_{(2)}$}

$7_{(1)}$ هو المؤسسة رقم (7) من المجموعة G_2 التي صارت تنتمي إلى المجموعة الأولى في التصنيف الجديد بطريقة الجوار الأقرب. أي أنه حدث تغير في التصنيف الأولي للمجموعتين G_1 وG_2 الذي انطلقنا منه.

$1_{(2)}$ هو المؤسسة رقم (1) من المجموعة G_1 التي صارت تنتمي إلى المجموعة الثانية في التصنيف الجديد بطريقة الجوار الأقرب. أي أنه حدث تغير في التصنيف الأولي للمجموعتين G_1 وG_2 الذي انطلقنا منه.

واذا حسبنا متوسطات X_1 وX_2 في هاتين المجموعتين الجديدتين نحصل على المتوسطات والتباينات التالية:

الانحرافات المعيارية الجديدة

	G_1	G_2
X_1	12.036	7.198
X_2	6	6.283

المتوسطات الجديدة للمجموعات

	G_1	G_2
X_1	48	56
X_2	30	47.14

أي أن إحداثيات مركز المجموعة الجديدة G_1' صار C_1(48 , 30)، وأن إحداثيات مركز المجموعة الجديدة G_2' صار C_2(56 , 47.14)

وبالتالي مربع المسافة الهندسية بينهما يساوي: $D_1'^2 = (48-56)^2 + (30-47.14)^2 = 357.78$

بينما كان مربع المسافة بينهما حسب التصنيف الفعلي (الإداري) يساوي مربع البعدين المركزين الفعليين اللذين إحداثياتها C_1(48 , 32) وC(56 , 44.86) يساوي: $D_1^2 = (48-56)^2 + (32-44.86)^2 = 211.38$

وهكذا نلاحظ أن التصنيف الجديد جعل مربع المسافة بين مركزي المجموعتين أكبر مما كان عليه قبل ذلك، وصارت 357.78 بدلا من 211.38، وقد استفاد (فيشر) من هذه الملاحظة، واستنبط معيارا خاصاً للتمييز بين عناصر المجموعتين G_1 وG_2.

كما نلاحظ أن التصنيف السابق جعل مجموع مربعات الانحرافات داخل كل مجموعة Sum of Squares Within SSW أصغر أو يساوي مما كان عليه، وبالمقابل جعل مجموع مربعات الانحرافات بين هاتين المجموعتين Sum of squares between (SSB) أكبر مما كان عليه. وهذا يعني أن المجموعتين صارتا أكثر تجانساً داخلياً وأكثر اختلافاً خارجياً، وهو ما نسعى إليه من خلال التحليل التمييزي بين المجموعات. وهذا ما يجعل طريقة الجوار الأقرب أفضل من طريقة المستقيم الفاصل.