جامعة محمد خيضر - بسكرة
كلية العلوم الاقتصادية والتجارية وعلوم التسيير
السنة الثانية علوم التسيير

سلسلة التمارين رقم 01 في مقياس إحصاء 4

التمرين(01): إليك البيانات التالية الخاصة بالمتغيرين X و Y.

X	2	4	6	8	10
Y	3	6	6	9	11

المطلوب:
1-كون شكل الانتشار.
2-حدد خط المربعات الصغرى.
3-أرسم خط المربعات الصغرى، ثم عين الأخطاء أو البواقي على شكل الانتشار السابق.
4-أحسب مجموع مربعات الأخطاء، وماذا تلاحظ؟.
5-أحسب نسبة انخفاض الخطأ في التقدير نتيجة استخدام خط المربعات الصغرى.

التمرين(02): يرغب صاحب مطعم الزيبان في الحصول على نموذج يوضح إلى أي مدى يكون إيراد الفترة الصباحية مرتبط بعدد الأشخاص الذين يطلبون الوجبة. البيانات التالية لست فترات صباحية حديثة، مع العلم أن الإيرادات معبرا عنها بمائة دولار:

الفترة الصباحية	عدد الزبائن	الإيراد
1	15	5
2	20	8
3	50	12
4	30	10
5	25	9
6	40	13

المطلوب:
1- أرسم لوحة الانتشار وماذا تلاحظ ؟
2- حدد خط المربعات الصغرى ومثله بيانيا. وفسر تقديرات الميل والجزء المقطوع.
3- إلى أي مدى لقيم X تكون تفسيراتك في السؤال الثاني صحيحة؟ ولماذا تكون محدودة في هذا المدى؟
4- حدد نسبة انخفاض الخطأ في التقدير نتيجة استخدام خط الانحدار أو خط المربعات الصغرى.
5- قدر تباين الخطأ ثم حدد تباين كل من b_1 و b_0.

6-أحسب معامل التحديد، وما هو التفسير المناسب لذلك ؟

التمرين (03): البيانات التالية تمثل متوسط الدخل الشهري (X) ومتوسط الإنفاق الشهري (Y) لخمس عائلات:

متوسط الدخل الشهري	200	320	480	260	620
متوسط الإنفاق الشهري	180	240	310	160	400

المطلوب:

1- قدر معلمتي نموذج الانحدار الخطي البسيط.
2- تنبؤ بمتوسط الإنفاق الشهري لعائلة ما، اذا كان متوسط دخلها الشهري 640 دولار.

التمرين (04): بفرض أنه لدينا بيانات العينة التالية:

X	1	2	3	4	5	6
Y	2	4	4	6	9	10

المطلوب:

1- كون شكل الانتشار لهذه البيانات. وهل تبدو العلاقة الخطية مقبولة ؟
2- بافتراض أن التوفيق الخطي مناسبا، حدد خط المربعات الصغرى وفسر ميله والجزء المقطوع ؟
3- أحسب تباين البواقي S_e^2 ومعامل التحديد r^2 وفسر معنى كل منهم ؟

بالتوفيـق للجميـع

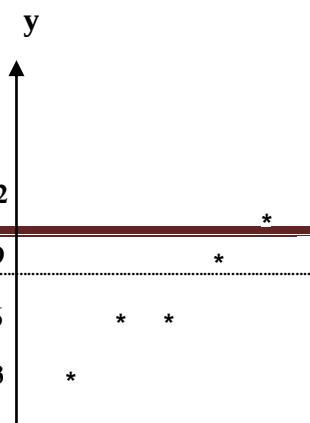
أسرة المقياس

حلول السلسلة الأولى / 2025

التمرين (01):

1- شكل الانتشار

الشكل الانتشاري للبيانات



من خلال شكل الانتشار هذا نلاحظ أن النقاط تتموقع بالقرب من الخط المستقيم، وبالتالي نحكم بوجود علاقة خطية بين المتغيرين X و Y، وهذا الخط يسمى بخط الانحدار المقدر أو خط المربعات الصغرى وتكتب معادلته بالشكل التالي: $\hat{y}_i = b_0 + b_1X_i$

2-تحديد خط المربعات الصغرى أي تحديد ($\hat{y}_i = b_0 + b_1X_i$): أي يجب ايجاد قيمة كل من b_0 و b_1 وذلك وفق الطريقة التالية:

رقم المشاهدة	X	Y	XY	X^2	Y^2
1	2	3	6	4	9
2	4	6	24	16	36
3	6	6	36	36	36
4	8	9	72	64	81
5	10	11	110	100	121
∑	**30**	**35**	**248**	**220**	**283**

من خلال هذا الجدول نجد:

$$\overline{X} = \frac{30}{5} = 6 \quad , \quad \overline{Y} = \frac{35}{5} = 7$$

$$SS_{XY} = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n} = 248 - \frac{(30)(35)}{5} = 38$$

$$SS_X = \sum X_i^2 - \frac{(\sum X_i)^2}{n} = 220 - \frac{(30)^2}{5} = 40$$

$$SS_Y = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} = 283 - \frac{(35)^2}{5} = 38$$

وبالتعويض نجد:

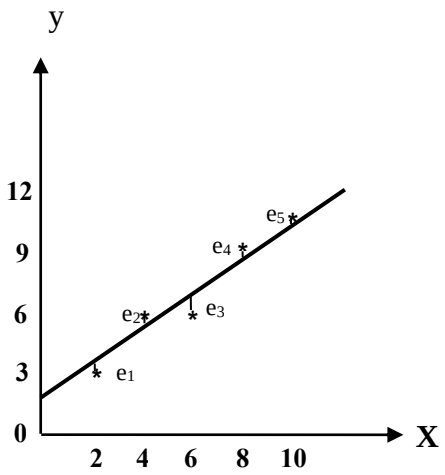
$$b_1 = \frac{SS_{XY}}{SS_X} = \frac{38}{40} = \mathbf{0.95}$$

$$b_0 = \bar{Y} - b_1 \bar{X} = 7 - 0.95(6) = \mathbf{1.3}$$

وبذلك يكون خط الانحدار المقدر أو خط المربعات الصغرى هو:

$$\hat{y}_i = \mathbf{1.3 + 0.95\, X_i}$$

3-رسم خط المربعات الصغرى وتعيين الأخطاء عليه: أنظر الشكل الموالي:



4- حساب مجموع مربعات الأخطاء:

في الجدول الموالي سوف نوضح كيفية حساب الأخطاء أو البواقي ومجموع مربعاتها:

X	Y الفعلية	القيمة المتوقعة أو التقديرية $\hat{y}_i = 1.3 + 0.95\, X_i$	الأخطاء $e_i = Y_i - \hat{y}_i$	مربعات الأخطاء $e_i^2 = (Y_i - \hat{y}_i)^2$
2	3	3.2	-0.2	0.04
4	6	5.1	0.9	0.81
6	6	7	-1	1
8	9	8.9	0.1	0.01
10	11	10.8	0.2	0.04
المجـمـــوع			$\sum(Y_i - \hat{y}_i) = 0$	$\sum(Y_i - \hat{y}_i)^2 = 1.90$

نلاحظ من خلال هذا الجدول أن مجموع مربعات الأخطاء الناتجة عن خط المربعات الصغرى يُقدر بـ : 1.90، حيث مجموع مربعات الأخطاء سيكون أكبر لأي خط آخر لهذه البيانات.

5-حساب نسبة انخفاض الخطأ في التقدير نتيجة استخدام خط المربعات الصغرى:

يتم ذلك بالاعتماد على العلاقة التالية:

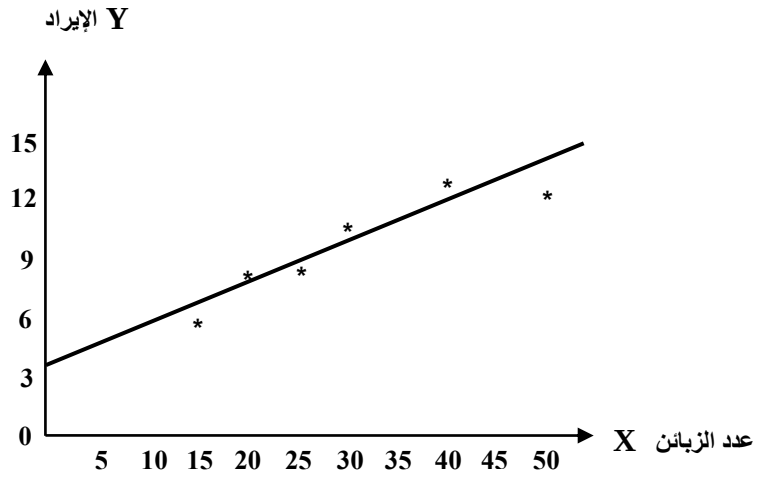
$$\frac{SS_Y - SSE}{SS_Y} = \frac{38 - 1.90}{38} = 0.95$$

وعليه يمكن القول بأن خط المربعات الصغرى قلل الأخطاء المربعة بمقدار 0.95 أو بنسبة 95%.

التمرين (02):

1-رسم لوحة الانتشار

شكل الانتشار وخط المربعات الصغرى للإيراد مقابل عدد الزبائن



من خلال شكل الانتشار السابق نلاحظ أن النقاط تتموقع بالقرب من الخط المستقيم، وعليه نحكم بوجود علاقة خطية بين المتغيرين محل الدراسة (عدد الزبائن) و (الإيراد)، وهذا الخط يسمى بخط المربعات الصغرى وتكتب معادلته بالشكل التالي:

$\hat{y}_i = b_0 + b_1 X_i$

2-تحديد خط المربعات الصغرى أي تحديد ($\hat{y}_i = b_0 + b_1 X_i$): أولا يجب إيجاد قيمة كل من b_0 و b_1 وذلك وفق الطريقة التالية:

رقم المشاهدة	X	Y	XY	X^2	Y^2
1	15	5	75	225	25
2	20	8	160	400	64
3	50	12	600	2500	144
4	30	10	300	900	100
5	25	9	225	625	81
6	40	13	520	1600	169

∑	**180**	**57**	**1880**	**6250**	**583**

من خلال هذا الجدول نجد:

$$\overline{X} = \frac{180}{6} = 30 \quad , \quad \overline{Y} = \frac{57}{6} = 9.5$$

$$SS_{XY} = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n} = 1880 - \frac{(180)(57)}{6} = 170$$

$$SS_X = \sum X_i^2 - \frac{(\sum X_i)^2}{n} = 6250 - \frac{(180)^2}{6} = 850$$

$$SS_Y = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} = 583 - \frac{(57)^2}{6} = 41.5$$

وبالتعويض نجد:

$$b_1 = \frac{SS_{XY}}{SS_X} = \frac{170}{850} = \mathbf{0.2}$$

$$b_0 = \overline{Y} - b_1 \overline{X} = 9.5 - 0.2(30) = \mathbf{3.5}$$

وبذلك يكون خط المربعات الصغرى هو: $\hat{\mathbf{y}}_\mathbf{i} = \mathbf{3.5} + \mathbf{0.2}\,\mathbf{X}_\mathbf{i}$

- **تفسير تقديرات الميل والجزء المقطوع:**

تقدير الميل هو (b_1= 0.2)، وذلك يعني أن لكل زبون إضافي،متوسط الإيراد المقدر يزيد بمقدار 0.2 وحدة أو 20 دولار. وبما أن البيانات المعطاة لا تشتمل على معلومات عندما تكون (X= 0)، فان تفسير تقدير الجزء المقطوع (b_0=3.5)، يكون ليس له معنى؛ لأنه في الواقع نعلم جيدا أنه لن يكون هناك أي إيراد من الوجبات إذا لم يكن هناك أي زبون.

3-انطلاقا مما سبق، فان تفسير النموذج السابق يكون صحيح إحصائيا فقط للأيام التي يكون فيها عدد الزبائن بين 15 و 50. وحيث أنه لا توجد لدينا بيانات خارج هذا المدى، فإننا لا نستطيع أن نستنتج إحصائيا أن نفس العلاقة الخطية السابقة تكون مناسبة عندما يكون عدد الزبائن أقل من 15 أو يزيد عن 50.

4-تحديد نسبة انخفاض الخطأ في التقدير نتيجة استخدام خط الانحدار المقدر أو خط المربعات الصغرى:

يتم ذلك وفق العلاقة التالية:

$$\frac{SS_Y - SSE}{SS_Y}$$

يجب أولا إيجاد قيمة SSE، ويكون ذلك وفق العلاقة التالية:

$$SSE = \frac{SS_x . SS_Y - (SS_{xy})^2}{SS_x} = \frac{(850).(41.5) - (170)^2}{(850)} = 7.5$$

وعليه فان نسبة انخفاض الخطأ في التقدير المطلوبة هي:

$$\frac{SS_Y - SSE}{SS_Y} = \frac{41.5 - 7.5}{41.5} = 0.82$$

وعليه يمكننا القول أن خط المربعات الصغرى قلل الأخطاء المربعة بمقدار 0.82 أو بنسبة 82%.

5-تقدير تباين الخطأ وتحديد تباين كل من b_1 و b_0.

أ- تقدير تباين الخطأ σ_ε^2: يتم تقدير هذا الأخير من خلال تباين البواقي أو تباين الفروق والذي نرمز له بالرمز S_e^2 حيث:

$$S_e^2 = \frac{SSE}{n-2} = \frac{7.5}{6-2} = 1.875$$

ب-تحديد تباين كل من b_1 و b_0.

$$\sigma_{b_1}^2 = \frac{\sigma_\varepsilon^2}{SS_X} \simeq \frac{S_e^2}{SS_X} = \frac{1.875}{850} = 0.0022$$

$$\sigma_{b_0}^2 = \sigma_\varepsilon^2\left[\frac{1}{n} + \frac{(\overline{X})^2}{SS_X}\right] \simeq S_e^2\left[\frac{1}{n} + \frac{(\overline{X})^2}{SS_X}\right] = 1.875\left[\frac{1}{6} + \frac{(30)^2}{850}\right] = 2.297$$

6-حساب معامل التحديد وتفسيره:

$$r^2 = \frac{SSR}{SST} = \frac{SST - SSE}{SST} = 1 - \frac{SSE}{SST} = 1 - \frac{SSE}{SS_y} = 1 - \frac{7.5}{41.5} = 0.8193$$

وهذا يعني أن 81.93% من التغيرات اليومية في إيرادات العينة تُفسر عن طريق الاختلافات اليومية في عدد الزبائن، أما النسبة المتبقية 18.07% من الاختلافات اليومية في إيرادات العينة فهي راجعة للخطأ العشوائي.