## 2-تقدير معالم نموذج الانحدار الخطي البسيط:

إن الهدف الأساسي في تحليل الانحدار هو أن نحدد معادلة انحدار لها معنى وتوفق بيانات عينة الدراسة بحيث يكون تباين الخطأ أصغر ما يكون.

ولهذا فمعادلة خط الانحدار تكتب من الشكل: $\hat{y}=b_0+b_1x$

حيث أن:

$b_0$: هي نقطة تقاطع خط الانحدار مع محور التراتيب (وهي قيمة y لما $x=0$).

$b_1$: تمثل ميل خط الانحدار، كما أنها تبين التغير الذي يحدث في y نتيجة التغير الذي يحدث في x.

$\hat{y}$: هي القيمة المتوقعة للمتغير التابع.

يتم إيجاد قيمتي $b_0$، $b_1$ بالاعتماد على طريقة المربعات الصغرى.

- **طريقة المربعات الصغرى:**

تقوم هذه الطريقة على مبدأ أساسي وهو جعل مجموع مربعات الأخطاء أقل ما يمكن بغية الحصول على أفضل خط مستقيم يمثل العلاقة بين المتغيرين x،y والذي يسمى بخط المربعات الصغرى.

إن الإحصاءات $b_0$، $b_1$ تعرف على أنها تقديرات المربعات الصغرى للمعالم $B_0$، $B_1$ على التوالي، ويتم تحديدهما باستخدام طريقة المربعات الصغرى والتي تقتضي أن تكون الدالة

SSE أصغر ما يمكن، ولهذا يلزمنا حساب المشتقين الجزئيين للدالة SSE بالنسبة لـ $b_0$،

$b_1$ وجعل الناتج يساوي الصفر، فنحصل على :

$$SSE = \sum_{i=1}^{n} e_i^2$$

$$SSE = \sum_{i=1}^{n} \ [y_i - (b_0 + b_1\ x_i\ )]^2$$

عندما نشتق الدالة SSE بالنسبة لـ $b_0$ ونجعل الناتج يساوي الصفر

$$\frac{\delta SSE}{\delta b_0} = 0 \rightarrow -2(\sum \left(y_i - (b_0 + b_1 x_i)\right) = 0$$

ثم نشتق الدالة SSE بالنسبة لـ $b_1$ ونجعل الناتج يساوي الصفر

$$\frac{\delta SSE}{\delta b_1} = 0 \rightarrow -2 \sum x_i \left(y_i - (b_0 + b_1 x_i)\right) = 0$$

بقسمة طرفي هاتين المعادلتين على( $2-$) نجد:

$$\sum \left(y_i - (b_0 + b_1 x_i)\right) = 0 \ldots \ldots \ldots \ldots \ldots .. (1)$$

$$\sum x_i \left(y_i - (b_0 + b_1 x_i)\right) = 0 \ldots \ldots \ldots \ldots (2)$$

من المعادلتين (1) و(2) نجد:

$$\sum y_i = nb_0 + b_1 \sum x_i \ldots \ldots \ldots \ldots . \ldots \ldots (3)$$

$$\sum x_i y_i = b_0 \sum x_i + b_1 \sum x_i^2 \ldots \ldots \ldots \ldots . (4)$$

ومنه يكون لدينا:

$$\rightarrow \begin{cases} b_0 = \dfrac{\sum y_i - \mathrm{b}_1 \sum \mathrm{x_i}}{n} \\ \sum \mathrm{x_i} y_{\mathrm{i}} = \mathrm{b}_0 \sum \mathrm{x_i} + \mathrm{b}_1 \sum \mathrm{x_i^2} \end{cases}$$

وبحل هاتين المعادلتين حلا مشتركا نجد:

$$b_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

ملاحظة:

عند استخدام الرموز SSx, SSy, SSxy فإن $b_1$ تكون بالشكل التالي:

$$b_1 = \frac{SSxy}{SSx}$$

**مثال01:**

اعتبر بيانات العينة التالية :

| x | y |
|---|---|
| 1 | 6 |
| 10 | 12 |
| 2 | 10 |
| 2.5 | 18 |
| 3 | 18 |

المطلوب:

1- كون شكل انتشار لهذه البيانات، وكيف تبدو العلاقة الخطية؟

2- حدد خط المربعات الصغرى؟

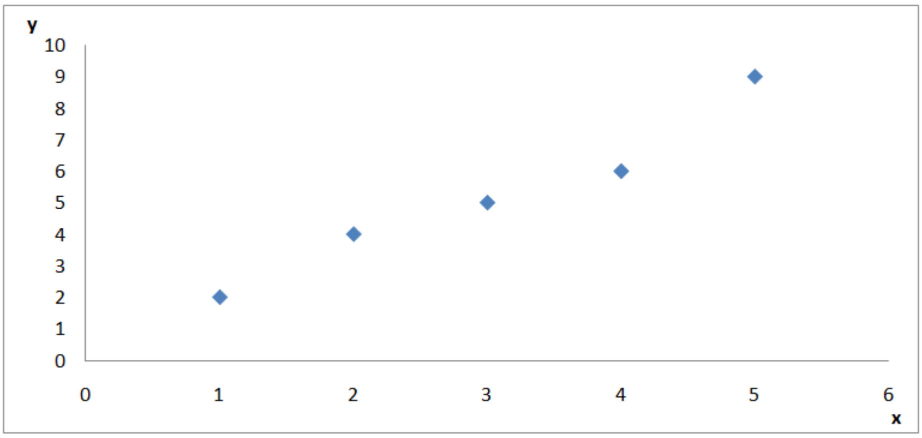
3- أحسب مجموع مربعات الأخطاء ؟

4- أحسب نسبة انخفاض الخطأ في التقدير نتيجة استخدام خط المربعات الصغرى؟

الحل:

1- تكوين شكل انتشار البيانات:

الشكل رقم(01):الشكل الانتشاري للبيانات




انطلاقا من الشكل الانتشاري للبيانات نلاحظ وجود علاقة خطية مقبولة بين المتغيرين $x$،$y$ لأن معظم النقاط تتموقع بالقرب من خط المستقيم.

2- تحديد خط المربعات الصغرى:

سنحاول إيجاد معادلة خط الانحدار بتحديد قيمتي $b_0$، $b_1$ كما يلي:

$$b_1 = \frac{SSxy}{SSx}$$

$$b_0 = \bar{y} - b_1\bar{x}$$

| $X^2$ | $xy$ | $y$ | $x$ |
|---|---|---|---|
| 1 | 6 | 6 | 1 |
| 2.25 | 18 | 12 | 1.5 |
| 4 | 20 | 10 | 2 |
| 6.25 | 45 | 18 | 2.5 |
| 9 | 54 | 18 | 3 |
| 22.5 | 143 | 64 | 10 |

$$SS_x = 2.5$$

$$SS_y = 108.8$$

$$SSxy = 15$$

$$\bar{y} = \frac{\sum y_i}{n} = 12.8 \qquad , \qquad \bar{x} = \frac{\sum x_i}{n} = 2$$

$$b_1 = \frac{15}{2.5} = 6$$

$b_0$=0.8

إذن معادلة الانحدار هي: $\hat{y} = 0.8 + 6x_i$

3-حساب مجموع الأخطاء ومجموع مربعاتها:

| $e_i^2$ | $y_i^2$ | $y$ | $x$ |
|---|---|---|---|
| 0.64 | 36 | 6 | 1 |
| 4.84 | 144 | 12 | 1.5 |
| 7.84 | 100 | 10 | 2 |
| 4.84 | 324 | 18 | 2.5 |
| 0.64 | 324 | 18 | 3 |
| 18.8 | 928 | 64 | 10 |

*4-حساب نسبة انخفاض في الأخطاء نتيجة استخدام المربعات الصغرى:*

$$\frac{SS_y - SSE}{\mathrm{SS_y}}$$

*لدينا:*

$$SSE = 18.8$$

$$SSy = 108.8$$

$$\frac{108.8 - 18.8}{108.8} = 0.8272$$

إذن فنتيجة استخدام خط الانحدار قللت الأخطاء المربعة بمقدار 82.72%.

**-خصائص مقدرات المربعات الصغرى:**

تتميز خصائص مقدرات المربعات الصغرى بالخصائص التالية:

**أ-خاصية الخطية:**

تعتبر مقدرات المربعات الصغرى دوال خطية للعنصر العشوائي التابع Y أي أن:

$$b_1 = \frac{\sum(x_i - \bar{x})}{SSx} y_i$$

$$b_0 = \sum(\frac{1}{n} - \frac{\bar{x}(x_i - \bar{x})}{SSx})y_i$$

**ب- خاصية عدم التحيز:**

يسمى المقدر مقدر غير متحيز إذا كانت القيمة المتوقعة للمقدر تساوي القيمة الحقيقية للمعلمة بمعنى أخر أن:

$E(b_1)= B_1$

$E(b_0)= B_0$

**ج- خاصية الكفاءة:**

إذا كان هناك أكثر من مقدر غير متحيز لمعلمة ما وكان تباين احدهم أقل من تباينات المقدرات الأخرى فان التباين الأقل أكفأ من بقية المقدرات الأخرى.

ومنه فتباين المقدر $b_1$ :

$$\sigma^2 b_1 = \frac{\sigma^2\zeta}{SSx}$$

ولان تباين الخطأ $\sigma^2\zeta$ مجهول فإننا نقدره بـ $S_e^2$

ومنه يصبح :

$$\sigma^2 b_1 = \frac{S_e^2}{SSx}$$

أما تباين المقدر $b_0$ :

$$b_0 = \sigma^2 \zeta \left(\frac{1}{n} + \frac{\bar{x}}{SSx}\right)$$

أو: $b_0 = S_e^2 \left(\frac{1}{n} + \frac{\bar{x}}{SSx}\right)$