الانحدار الخطي البسيط

أ.النموذج الخطي لمتغيرين: أو ما يعرف بتحليل الانحدار البسيط، ويستخدم لاختبار الفروض حول العلاقة بين متغير تابع (Y) ومتغير مستقل أو مفسر(X). طبعا بعد رسم شكل الانتشار والتأكد من وجود علاقة خطية تقريبية والتي تأخذ الشكل التالي:

$$Y_i = b_0 + b_1 X_i \ldots\ldots\ldots\ldots\ldots\ldots 1$$

طبعا أثناء رسم شكل الانتشار من غير الممكن أت تقع جميع النقاط على نفس الخط تماما، وبالتالي هناك تشويش عشوائي أو حد الخطأ (white noise) وبالتالي نكتب المعادلة أعلاه رقم 1 كما يلي:

$$Y_i = b_0 + b_1 X_i + e_i \ldots\ldots\ldots\ldots\ldots\ldots 2$$

ويفترض في حد الخطأ (e_i) أن يكون:

- ✓ موزع توزيع طبيعي.
- ✓ وسطه أو توقعه الرياضي يساوي صفر.
- ✓ تباينه ثابت.
- ✓ حدود الخطأ غير مترابطة فيما بينها.

ب.كتابة معادلة الانحدار: تعتبر طريقة المربعات الصغرى (OLS) هي أسلوب لتوفيق أفضل خط مستقيم لعينة مشاهدات أو قيم (XY) وهو يتضمن تصغير مجموع المربعات الانحرافات النقاط عن الخط إلى أدنى حد ممكن:

$$Min \sum_{i=1}^{n} (Y_i - \hat{y}_i)^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots 3$$

حيث تشير (Y_i) إلى المشاهدات الفعلية وتشير ($\hat{y}_i$) إلى القيم المناظرة، حيث $e_i = Y_i - \hat{y}_i$

بالعودة للمعادلة رقم 1 نقوم بإيجاد كل من (a) و(b) كما يلي:

$$\hat{b}_1 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} \ldots\ldots\ldots\ldots\ldots\ldots 4$$

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 5$$

حيث: $\hat{b}_0$ و $\hat{b}_1$ **هي مقدرتان للمعلمتين الحقيقيتين** b_0 و b_1

مثال: ليكن لديك المتغيرين X,Y

Y	40	44	46	48	52	58	60	68	74	80
X	6	10	12	14	16	18	22	24	26	32

المطلوب:

- ✓ -أرسم شكل الانتشار.

✓ أوجد المعادلة الخطية بين X و Y

✓ المعادلة: نقوم بالحسابات التالية:

n	Y_i	X_i	$(X_i-\bar{X})$	$(Y_i-\bar{Y})$	$(X_i-\bar{X})\,(Y_i-\bar{Y})$	$(X_i-\bar{X})^2$
1	40	6	-12	-17	204	144
2	44	10	-8	-13	104	64
3	46	12	-6	-11	66	36
4	48	14	-4	-9	36	16
5	52	16	-2	-5	10	4
6	58	18	0	1	0	0
7	60	22	4	3	12	16
8	68	24	6	11	66	36
9	74	26	8	17	136	64
10	80	32	14	23	322	196
Σ	570	180	0	0	**956**	**576**

$$\bar{X}=\frac{180}{10}=18$$

$$\bar{Y}=\frac{570}{10}=57$$

من أجل تقدير المعالم (a ;b) للمعادلة التالية $Y_i=b_1X_i+b_0$ نقوم بالحسابات الآتية:

$$b_1=\frac{\sum_{i=1}^{10}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^{10}(X_i-\bar{X})^2}=\frac{956}{576}=1.66$$

$$b_0=\bar{Y}-b_1\bar{X}=57-(1.66\times18)=27.12$$

على هذا الأساس تصبح معادلة الانحدار الخطي البسيط المقدرة بين المتغيرين X و Y كما يلي:

$$Y_i=1.66X_i+27.12$$

ج.معنوية تقديرات المعالم b0 و b1

من أجل اختبار المعنوية الإحصائية لتقديرات المعالم **b0** و**b1** يلزمنا حساب تباين كل منهما.

المعدلات التالية تعطي تقديرات غير متحيزة لتباين **b0** و **b1**.

حيث تباين المعلمة **b0** تحسب كما يلي:

$$\sigma_{b_0}^2=\frac{\sum_{i}^{n}e_i^2}{n-k}\times\frac{\sum_{i=1}^{n}X_i^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}$$

في حين تباين المعلمة **b1** تحسب كما يلي:

$$\sigma_{b_1}^2=\frac{\sum_{i}^{n} e_i^2}{(n-k).\sum_{i=1}^{n}(X_i-\bar{X})^2}$$

على هذا الأساس تكون Var(b0)أو σ_b^2 و σ_a^2 أو Var(b1) هي الأخطاء المعيارية للتقدير، وكما هو معلوم فإن e_i موزع توزيع طبيعي وبالتالي b0 وb1 تكون هي الأخرى موزعة طبيعيا، وهذا ما يمكنا من استخدام توزيع t بدرجة حرية (n-k).

مثال 2: أختبر معنوية كل من المقدرات b0 وb1 المحسوبة في المثال رقم 1.

من أجل عملية الاختبار نقوم بالحسابات الموضحة في الجدول التالي:

أ.نقوم أولا بحساب ŷ

$\hat{y}_1=1.66(6)+27.12$

$\hat{y}_1=37.08$ وهكذا .....

يمثل ŷ القيم المقدرة من خلال معادلة الانحدار.

ب. نحسب الفرق أو حد الخطأ $e_i=(Y_i-\hat{y})$

K يمثل عدد المقدرات أو المعالم ؛ أي K=2

n	X_i	Y_i		$\hat{y}$	$e_i=(Y_i-\hat{y})$	e^2	X_i^2	$(X_i-\bar{X})^2$	$(Y_i-\bar{Y})^2$
1	6	40		37.08	2.92	8.5264	36	144	289
2	10	44		43.72	0.28	0.0784	100	64	169
3	12	46		47.04	-1.04	1.0816	144	36	121
4	14	48		50.36	-2.36	5.5696	196	16	81
5	16	52		53.68	-1.68	2.8224	256	4	25
6	18	58		57	1	1	324	0	1
7	22	60		63.64	-3.64	13.2496	484	16	9
8	24	68		66.96	1.04	1.0816	576	36	121
9	26	74		70.28	3.72	13.8384	676	64	289
10	32	80		80.24	-0.24	0.0576	1024	196	529
Σ	180	570		/	0.00	**47.3056**	**3816**	**576**	1634

$$\sigma_{b_1}^2=\frac{\sum_{i}^{n} e_i^2}{(n-k).\sum_{i=1}^{n}(X_i-\bar{X})^2}=\frac{47.3056}{(10-2).576}=0.01$$

$$\sigma_{b_1}=\sqrt{0.01}=0.1$$

$$t_{b_1}=\frac{b_1}{\sigma_{b_1}}=\frac{1.66}{0.1}=16.6$$

$$\sigma_{b_0}^2=\frac{\sum_{i}^{n} e_i^2}{n-k}\times\frac{\sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n}(X_i-\bar{X})^2}=\frac{47.3056}{10-2}\times\frac{3816}{10.576}=3.92$$

$$\sigma_{b_0}=\sqrt{3.92}=1.98$$

$$t_{b_0}=\frac{b_0}{\sigma_{b_0}}=\frac{27.12-0}{1.98}=13.7$$

نلاحظ أن كل من t_{b1} و t_{b0} أكبر من $t_{n-k=8}^{0.05}=2.306$ المأخوذة من الجدول بدرجة حرية 8 ومستوى معنوية 5%، وبالتالي b0 و b1 معنوية إحصائية عند 5%.

د.اختبار جودة النموذج

كلما كانت المشاهدات أقرب إلى خط الانحدار كلما زاد التغير في (Y) الذي تفسره معادلة الانحدار المقدرة، فالتغير الإجمالي في (Y) يساوي التغير المفسر زائد تغير البواقي.

<table>
<tr><td colspan="3">$\sum_{i=1}^{n}(Y_i-\bar{Y})^2=\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2+\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2$</td></tr>
<tr><td>مجموع مربع انحرافات القيم في Y عن وسطها</td><td>مجموع مربع انحرافات القيم المقدرة $\hat{Y}$</td><td>مجموع مربع الأخطاء</td></tr>
<tr><td>التغير الإجمالي في (Y)</td><td>التغير المفسر في (Y)</td><td>تغير البواقي في (Y)</td></tr>
<tr><td>SCT=Somme des Carrés Totale</td><td>SCE=Somme des Carrés Expliquée</td><td>SCR= Somme des Carrés Résiduels</td></tr>
<tr><td colspan="3">$\frac{SCT}{SCT}=\frac{SCE}{SCT}+\frac{SCR}{SCT}\Rightarrow 1=\frac{SCE}{SCT}+\frac{SCR}{SCT}$</td></tr>
<tr><td colspan="3">تسمى النسبة بين مجموع انحرافات القيم المقدرة ومجموع انحرافات القيم بمعامل التحديد؛ أي:
$R^2=\frac{SCE}{SCT}=\frac{\sum(\hat{Y}_i-\bar{Y})^2}{\sum(Y_i-\bar{Y})^2}$</td></tr>
<tr><td colspan="3">$R^2=\frac{SCE}{SCT}=1-\frac{SCR}{SCT}\Rightarrow$
$R^2=1-\frac{\sum(Y_i-\hat{Y}_i)^2}{\sum(Y_i-\bar{Y})^2}$</td></tr>
</table>

تتراوح قيمة معامل التحديد $0\leq R^2\leq 1$

مثال: أحسب معامل التحديد للمثال السابق.

n	$(Y_i-\bar{Y})^2$	$(\hat{Y}_i-\bar{Y})^2$	$(Y_i-\hat{Y}_i)^2$
1	289	396.8064	8.5264
2	169	176.3584	0.0784
3	121	99.2016	1.0816
4	81	44.0896	5.5696
5	25	11.0224	2.8224
6	1	0	1
7	9	44.0896	13.2496
8	121	99.2016	1.0816
9	289	176.3584	13.8384
10	529	540.0976	0.0576
Σ	**1634**	**1587.2256**	**47.3056**
	SCT	**SCE**	**SCR**

$$R^2=\frac{SCE}{SCT}=\frac{\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}=\frac{1587.22}{1634}=0.9714$$

$$R^2=1-\frac{SCR}{SCT}=1-\frac{\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}=1-\frac{47.3056}{1634}=1-0.029\simeq 0.9710$$

ملحق Eviews:

Dependent Variable: Y
Method: Least Squares
Date: 10/07/25 Time: 21:45
Sample: 1 10
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	1.659722	0.101321	16.38082	0.0000
C	27.12500	1.979265	13.70458	0.0000

R-squared	0.971049	Mean dependent var	57.00000
Adjusted R-squared	0.967430	S.D. dependent var	13.47426
S.E. of regression	2.431706	Akaike info criterion	4.791920
Sum squared resid	47.30556	Schwarz criterion	4.852437
Log likelihood	-21.95960	Hannan-Quinn criter.	4.725533
F-statistic	268.3312	Durbin-Watson stat	1.783613
Prob(F-statistic)	0.000000		