

الانحدار: *Regression*

يُعتبر الانحدار أحد أهم الأساليب الإحصائية التي تُستخدم في قياس العلاقات الاقتصادية، حيث يختص بقياس العلاقة بين متغير ما يسمى بالمتغير التابع، ومتغير آخر أو مجموعة من المتغيرات تُسمى بالمتغيرات المستقلة.

الانحدار الخطي البسيط: Simple Linear Regression

يُعتبر الانحدار الخطي البسيط أبسط أنواع نماذج الانحدار، ويوجد هناك نماذج عديدة للعلاقات الاقتصادية البسيطة التي يُمكن قياسها باستخدام أسلوب الانحدار البسيط، مثال ذلك العلاقة بين الاستهلاك كمتغير تابع والدخل المتاح كمتغير مستقل وهو ما يُعرف بدالة الاستهلاك، والعلاقة بين الكمية المطلوبة من سلعة ما وسعرها وهو ما يُعرف بدالة الطلب وغيرها من العلاقات الأخرى.

وتتكون معادلة الانحدار من متغيرات ومعلمات:

$$Y = \alpha + \beta X$$

حيث:

Y: المتغير التابع

X: المتغير المستقل

α : الحد الثابت أو الحد المقطوع من محور المتغير التابع، وتُسمى بالمعلمة التقاطعية أو المعلمة الناقلة نظراً لأن تغييرها يؤدي لانتقال الخط الممثل للعلاقة بالكامل من وضع لآخر

β : تُسمى المعلمة الانحدارية وتمثل ميل الخط المستقيم الممثل للعلاقة، وهي تشير إلى مقدار التغير في المتغير التابع نتيجة لتغير المتغير المستقل بوحدة واحدة وتُحسب رياضياً من العلاقة:

$$\beta = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

تُميز ضمن هذه المعادلة عدة مفاهيم:

β موجب هذا يعني أن المتغير X يؤثر إيجاباً على المتغير Y .

β سالب هذا يعني أن المتغير X يؤثر سلباً على المتغير Y .

عندما تتغير قيمة X بمقدار وحدة واحدة فإن قيمة Y تتغير بمقدار β والذي هو معامل الانحدار.

علماً أن:

$$\alpha \in -\infty, +\infty$$

$$\beta \in -\infty, +\infty$$

مثال: إذا توفرت لديك البيانات الآتية شكّل معادلة الانحدار التي تعين العلاقة بين كمية الإنتاج وكمية السماد المستخدم.

Y الإنتاج (كغ)	X كمية السماد (كغ)
500	30
450	20
700	35
650	40
540	36
440	27
670	50
290	10

$$Y = \alpha + \beta X$$

$$\beta = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

$$\bar{X} = \frac{30+20+35+40+36+27+50+10}{8} = 31$$

$$\bar{Y} = \frac{500+450+700+650+540+440+670+290}{8} = 530$$

Y	X	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$(X_i - \bar{X})^2$
500	30	-1	-30	30	1
450	20	-11	-80	880	121
700	35	4	170	680	16
650	40	9	120	1080	81
540	36	5	10	50	25
440	27	-4	-90	360	16
670	50	19	140	2660	361
290	10	-21	-240	5040	441
				$\Sigma = 10780$	$\Sigma = 1062$

$$\beta = \frac{10780}{1062} = 10.15$$

$$\alpha = \bar{Y} - \beta \bar{X} = 530 - (10.15 \times 31) = 215.35$$

$$Y = 215.35 + 10.15 X$$

أي كلما زادت كمية السماد بمقدار كغ واحد تزداد كمية الإنتاج بمقدار 10.15 كغ (قيمة β)

مثال 2:

الإنتاج	السماذ	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$(X_i - \bar{X})^2$
15	3	-0.86	-6.43	5.53	0.74
20	4	0.14	-1.43	-0.2	0.02
25	6	2.14	3.57	7.64	4.58
18	2	-1.86	-3.43	6.38	3.46
27	5	1.14	5.57	6.35	1.3
22	4	0.14	0.57	0.08	0.02
23	3	-0.86	1.57	-1.35	0.74
$\bar{Y}=21.43$	$\bar{X}= 3.86$			$\Sigma= 24.43$	$\Sigma=10.86$

$$Y = \alpha + \beta X$$

$$\beta = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

$$\beta = \frac{24.43}{10.86} = 2.25$$

$$\alpha = \bar{Y} - \beta \bar{X} = 21.43 - (2.25 \times 3.86) = 12.75$$

$$Y = 12.75 + 2.25 X$$

أي كلما زادت كمية السماذ بمقدار كغ واحد تزداد كمية الإنتاج بمقدار 2.25 كغ (قيمة β)

اختبارات المعنوية لمعاملات الارتباط والانحدار: تُستخدم اختبارات المعنوية للانحدار لقياس درجة الثقة في المعلمات المقدرة من العينة كأساس جيد للوصول لمعاملات المجتمع، وفي الارتباط لمعرفة إن كان الارتباط بين المتغيرين غير ناتج عن الصدفة ومؤكد إحصائياً أم لا.

1- اختبار معنوية معامل الارتباط:

- للقيام بهذا الاختبار نقارن T المحسوبة مع T الجدولية عند درجة حرية n-k ودرجة الخطأ أو مستوى المعنوية 1% و 5% حيث:

$$T = R \sqrt{\frac{n-2}{1-R^2}}$$

حيث: R معامل الارتباط بيرسون، R^2 معامل التحديد، n-2 درجة الحرية.

- فإذا كانت T المحسوبة < T الجدولية يعني هذا أن الارتباط بين المتغيرين المدروسين معنوي ومؤكد إحصائياً أي غير ناتج عن الصدفة ولا يُمكن إهماله.
- T المحسوبة > T الجدولية يعني أن الارتباط بين المتغيرين المدروسين غير معنوي وغير مؤكد إحصائياً أي ناتج عن الصدفة ويُمكن إهماله.

2- اختبار معنوية معامل الانحدار: للقيام بهذا الاختبار نقارن T المحسوبة مع T الجدولية عند درجة حرية n-k ودرجة الخطأ أو مستوى المعنوية 1% و 5% حيث:

$$T = \frac{\hat{b}}{S_b}$$

معلمة الانحدار المقدرة

$$S_b = \frac{\text{Error}}{\sqrt{\sum(x_i - \bar{x})^2}}$$

الانحراف المعياري

$$\text{Error} = \sqrt{\frac{\sum(y_i - \hat{y})^2}{n-2}}$$

- فإذا كانت T المحسوبة < T الجدولية يعني هذا أن المعلمة المقدرة معنوية إحصائياً.
- T المحسوبة > T الجدولية يعني هذا أن المعلمة المقدرة غير معنوية إحصائياً.

مثال: اختبر معنوية معامل الانحدار ومعامل الارتباط في المثال الأول إذا علمت أن T الجدولية عند 0.05 تساوي 1.943 وعند 0.01 تساوي 3.14

الحل:

Y الإنتاج (كغ)	X كمية السماد (كغ)
500	30
450	20
700	35
650	40
540	36
440	27
670	50
290	10

نقوم بحساب \hat{Y} ومن ثم حساب e^2 :

Y	X	\hat{Y}	$e_i = Y - \hat{Y}$	e_i^2
500	30	519.85	-19.85	394.02
450	20	418.35	31.65	1001.72
700	35	570.6	129.4	16744.36
650	40	621.35	28.65	820.82
540	36	580.75	-40.75	1660.56
440	27	489.4	-49.4	2440.36
670	50	722.85	-52.85	2793.12
290	10	316.85	-26.85	720.92
			$\Sigma = 0$	$\Sigma = 26575.88$

$$\text{Error} = \sqrt{\frac{26575.88}{6}} = 66.55$$

$$S_b = \frac{66.55}{\sqrt{1062}} = 2.04$$

$$T = \frac{10.15}{2.04} = 4.98$$

T المحسوبة < T الجدولية عند كل من 0.05 و 0.01 يعني هذا أن المعلمة المقدرة معنوية جداً إحصائياً¹.

$$T = R \sqrt{\frac{n-2}{1-R^2}}$$

$$T = 0.90 \sqrt{\frac{6}{1-0.81}} = 5.06$$

T المحسوبة < T الجدولية عند كل من 0.05 و 0.01 يعني هذا أن الارتباط بين المتغيرين المدروسين معنوي جداً ومؤكد إحصائياً أي غير ناتج عن الصدفة ولا يُمكن إهماله.

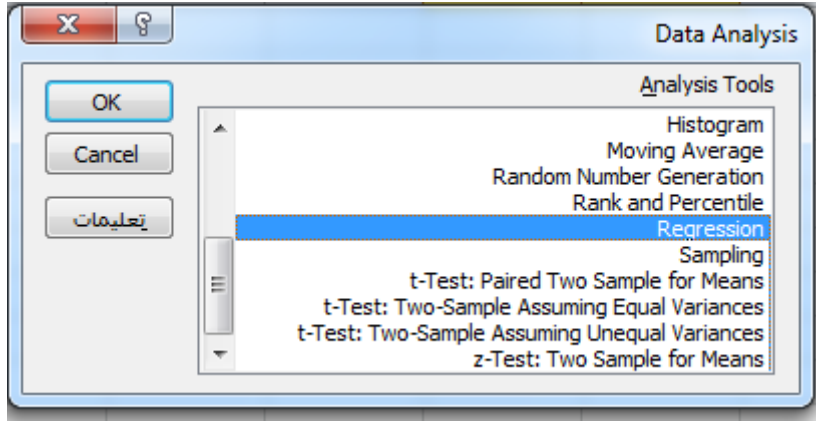
التطبيق على برنامج Excel:

- بعد إدخال البيانات في البرنامج كالتالي:

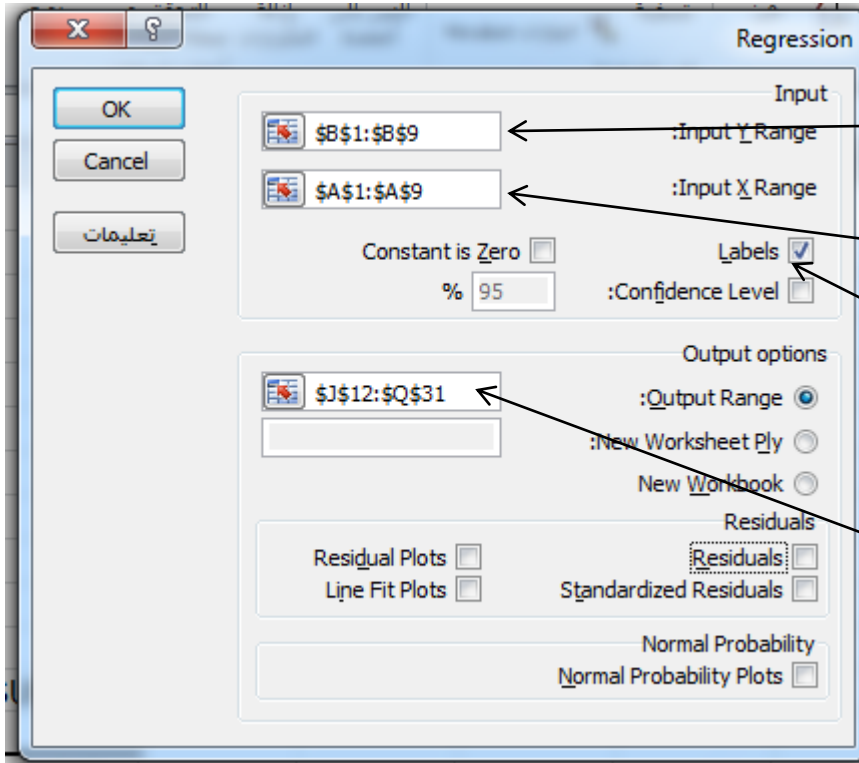
B	A	
Y	X	1
500	30	2
450	20	3
700	35	4
650	40	5
540	36	6
440	27	7
670	50	8
290	10	9

- من علامة التبويب بيانات نضغط على Data Analysis، فتظهر نافذة باسم Data Analysis تحتوي على عدد من طرق التحليل الإحصائي، نختار منها Regression:

¹ إذا كانت t المحسوبة < الجدولية فقط عند مستوى المعنوية 0.05 نقول أن المعلمة المقدرة معنوية (وليس معنوية جداً).



- فيظهر مربع الحوار التالي:



نقوم بتحديد نطاق الخلايا الذي يحوي بيانات المتغير التابع

نقوم بتحديد نطاق الخلايا الذي يحوي بيانات المتغير المستقل

إذا قمنا بتحديد أسماء الأعمدة مع البيانات نختار label

نقوم بتحديد نطاق خلايا فارغ لإظهار النتيجة

- نضغط ok فتظهر النتيجة الآتية:

قيمة المعامل α
 قيمة المعامل β
 T المحسوبة

Regression Statistics	
0.896989	Multiple R
0.804589	R Square
0.772021	Adjusted R Square
66.5531	Standard Error
8	Observations

ANOVA					
Significance	F	MS	SS	df	
0.002526	24.70452	109424.1	109424.1	1	Regression
		4429.316	26575.89	6	Residual
			136000	7	Total

Upper 95.0%	Lower 95.0%	Upper 95%	Lower 95%	P-value	t Stat	Standard Error	Coefficients
380.5955	50.0636	380.5955	50.0636	0.01888	3.188149	67.54063	215.3296 Intercept
15.14783	5.153486	15.14783	5.153486	0.002526	4.970364	2.042237	10.15066 X