المحور01: الانحدار الخطي البسيط

المحاضرة 01 :

أولا: تقديم نموذج الانحدار الخطي البسيط

1- الشكل العام: النموذج الخطي البسيط يأخذ الشكل التالي:

$$Y_t = \alpha + \beta \cdot X_t + \varepsilon_t$$

سمي النموذج خطيا لأن العلاقة بين المتغير التابع والمستقل علاقة خطية، وسمي البسيط لأن عدد المتغيرات المستقلة متغير واحد فقط، و $\alpha \& \beta$ معلمات أو معاملات النموذج.

حيث: Y : المتغير التابع، أو المتغير الداخلي.

X : المتغير المفسر، أو المتغير المستقل.

ε : المتغير العشوائي.

α و β : معلمات للتقدير.

t : مؤشر الزمن.

تمثل المعلمة α الجزء الثابت، وهو الجزء المقطوع من المحور الرأسي، وهو عبارة عن قيمة متوسط المتغير التابع لما تنعدم قيمة المتغير المستقل، بينما تمثل المعلمة β معامل الانحدار أو ميل الخط المستقيم، وتعبر عن مقدار التغير في المتغير التابع نتيجة لتغير المتغير المستقل بوحدة واحدة، وتبين اشارتها إذا ما كانت العلاقة بين المتغير التابع والمستقل علاقة طردية أو عكسية.

إن إدخال المتغير العشوائي ε_t في النموذج القياسي له عدة مسوغات أهمها أنه عبارة عن مجموعة شاملة تتضمن كل تلك المتغيرات التي لا يمكن قياسها بسهولة، قد يمثل هذا الحد المتغيرات التي لا يمكن إدراجها في النموذج لعدم توفر البيانات، أو أخطاء في القياس في البيانات، أو العشوائية الموجودة في السلوك البشري.

2- الفرضيات الاحتمالية التي يقوم عليها النموذج الخطي البسيط:

إن الطريقة المستعملة في تقدير معاملات نموذج الانحدار الخطي البسيط هي طريقة المربعات الصغرى العادية "OLS"، التي تم وضعها من طرف : Carl Friedrich Gauss، بناءا على بعض الفرضيات التي تجعل منها الطريقة الأكثر استعمالا، و تدور هذه الفرضيات حول طبيعة وشكل المتغير العشوائي، وهي:

- $\forall t \quad E(\varepsilon_t) = 0$: وتنص هذه الفرضية على أن الأخطاء لا تدخل في تفسير Y ، حيث تعبر عن قيم عشوائية تأخذ قيما سالبة، موجبة أو معدومة، لا يمكن قياسها وتحديدها بدقة، تخضع للقوانين الاحتمالية، حيث أن أملها الرياضي أو متوسطها يكون معدوما.
- $\forall t \quad V(\varepsilon_t) = E(\varepsilon_t^2) = \delta_\varepsilon^2$: ثبات أو تجانس التباين (HOMOSCEDASTICITY). أي أن تشتت الأخطاء حول متوسطها معدوم .
- $\forall i \neq j \quad Cov(\varepsilon_i \varepsilon_j) = E(\varepsilon_i \varepsilon_j) = 0$: عدم وجود ارتباط ذاتي بين الأخطاء، أي أن التباينات المشتركة بين الأخطاء تكون معدومة.
- $Cov(x_t . \varepsilon_t) = 0$: عدم وجود ارتباط بين المتغير المستقل والمتغير العشوائي.
- $\varepsilon_t \rightarrow N(0 . \delta_\varepsilon^2)$: التوزيع الاحتمالي للمتغير العشوائي هو التوزيع الطبيعي.

فرضيات أخرى:

- المتغيرات Y و X محددة بدون خطأ.
- قيم المتغير X غير عشوائية.

ثانيا: مقدرات معلمات النموذج الخطي البسيط:

يتم تقدير نموذج الانحدار الخطي البسيط بطريقة المربعات الصغرى العادية (ORDINARY LEAST SQUARES)، التي تهدف إلى الحصول على مقدرات $\hat{\alpha}$. $\hat{\beta}$ تعطي مجموع مربعات انحراف القيم المقدرة عن القيم الحقيقية في أدنى قيمة لها.

مقدرات المعلمات بطريقة التقدير OLS (بدون برهان رياضي) مساوية دائما لـ:

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \cdot \bar{X}$$

$$\hat{\beta} = \frac{\sum X_t Y_t - n\bar{Y}\bar{X}}{\sum X_t^2 - n\bar{X}^2} = \frac{\sum(Y_t - \bar{Y}) \cdot (X_t - \bar{X})}{\sum(X_t - \bar{X})^2} = \frac{Cov(X_t . Y_t)}{V(X_t)}$$

مثال:

ترغب إحدى الشركات في تحديد العلاقة بين إنفاقها على الدعاية والاعلانات وعوائد المبيعات، كلاهما بالمليون دينار جزائري، فإذا كانت لدينا البيانات التالية عن تطور هاذين المتغيرين من 2014 إلى 2023 كمايلي:

2023	2022	2021	2020	2019	2018	2017	2016	2015	2014	السنة

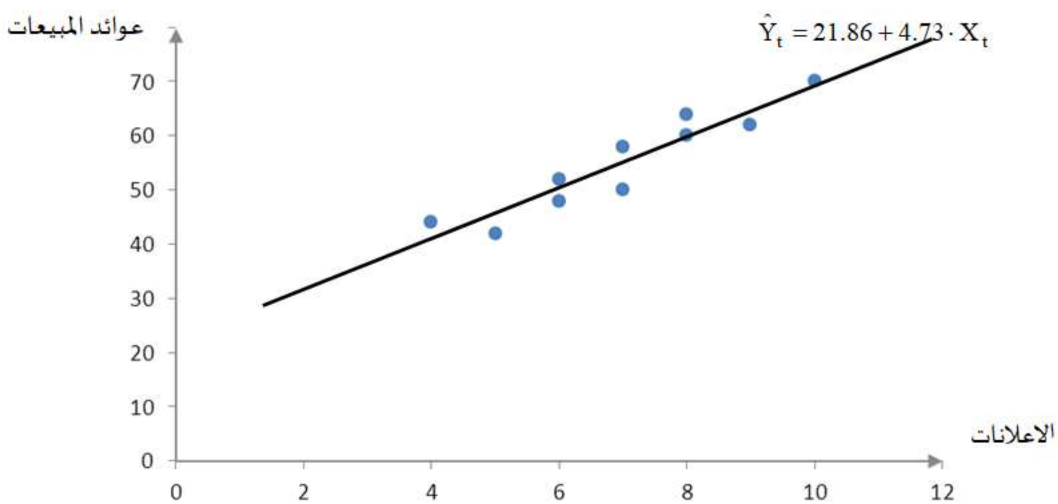
الاعلانات	4	5	6	6	7	8	7	9	8	10
المبيعات	44	42	52	48	50	60	58	62	64	70

المطلوب:

- مثل بيانيا بيانات الجدول بسحابة النقاط، ماذا تستنتج؟
- قدر النموذج الخطي البسيط الذي يقيس أثر الانفاق على الاعلانات على عوائد المبيعات في هذه الشركة، وفسر النتائج.
- حساب القيم المقدرة $\hat{Y}_t$ واستنتاج بواقي التقدير e_t.

الحل:

1- سحابة النقاط:



من خلال سحابة النقاط نجد أن العلاقة بين الانفاق على الاعلانات وعوائد مبيعات الشركة علاقة خطية، أي أنه يمكننا التعبير عنها بخط انحدار يمر بين مختلف الثنائيات المشكلة من قيم الانفاق على الاعلانات وقيم عوائد المبيعات، هذا من جهة، ومن جهة أخرى نجد أن هذه العلاقة هي علاقة طردية، بحكم أن زيادة الانفاق على الاعلانات تؤدي إلى زيادة عوائد الأرباح في هذه الشركة. وهو ما سنتأكد منه من خلال إشارة معامل الانحدار الذي سنحسبه لاحقا.

2- تقدير النموذج الخطي البسيط وتفسير النتائج:

نضع:

X_t : الانفاق على الاعلانات، Y_t : عوائد المبيعات

وبالتالي يكون النموذج المراد تقديره كمايلي:

$$Y_t = \alpha + \beta \cdot X_t + \varepsilon_t$$

أما مقدرات طريقة OLS لهذا النموذج فتعطى كما يلي:

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X}$$
$$\hat{\beta} = \frac{\sum(Y_t - \overline{Y}) \cdot (X_t - \overline{X})}{\sum(X_t - \overline{X})^2}$$

العمليات الحسابية موضحة بالجدول الموالي:

السنة	X_t	Y_t	$X_t - \overline{X}$	$Y_t - \overline{Y}$	$(X_t - \overline{X})(Y_t - \overline{Y})$	$(X_t - \overline{X})^2$	$\hat{Y}_t$	e_t	$(Y_t - \overline{Y})^2$
2009	4	44	− 3	− 11	33	9	40.80	3.19	121
2010	5	42	− 2	− 13	26	4	45.53	-3.53	169
2011	6	52	− 1	− 3	3	1	50.26	1.73	9
2012	6	48	− 1	−7	7	1	50.26	-2.26	49
2013	7	50	0	− 5	0	0	54.99	-4.99	25
2014	8	60	1	5	5	1	59.73	0.26	25
2015	7	58	0	3	0	0	54.99	3.00	9
2016	9	62	2	7	14	4	64.46	-2.46	49
2017	8	64	1	9	9	1	59.73	4.26	81
2018	10	70	3	15	45	9	69.19	0.80	225
$\sum$	**70**	**550**	**0**	**0**	**142**	**30**	**550**	**0**	**762**

لدينا:

$$\overline{X} = \frac{\sum X_t}{10} = \frac{70}{10} = 7$$
$$\overline{Y} = \frac{\sum Y_t}{10} = \frac{550}{10} = 55$$

ومنه نجد:

$$\hat{\beta}=\frac{\sum(Y_t-\overline{Y})\cdot(X_t-\overline{X})}{\sum(X_t-\overline{X})^2}=\frac{142}{30}=4.73$$
$$\hat{\alpha}=\overline{Y}-\hat{\beta}\cdot\overline{X}=55-4.73(7)=21.86$$

وبالتالي يكون النموذج الخطي البسيط المقدر كمايلي:

$$\hat{Y}_t=21.86+4.73\cdot X_t$$

تفسير نتائج التقدير:

- تشير المعلمة $\hat{\alpha}=21.86$ إلى وجود مقدار ثابت من عوائد المبيعات مقداره 21.86 مليون دج لا يتأثر بتغير الانفاق على الاعلانات.

- تشير المعلمة $\hat{\beta}=+4.73$ إلى وجود علاقة طردية بين الانفاق على الاعلانات وعوائد المبيعات من جهة، وأن كل تغير في الانفاق على الاعلانات بوحدة واحدة يؤدي إلى تغير عوائد المبيعات في نفس الاتجاه بـ: 4.73 وحدة.

- حساب القيم المقدرة $\hat{Y}_t$ واستنتاج بواقي التقدير e_t.

لدينا:

$$\begin{aligned}
&\hat{Y}_t=21.86+4.73\cdot X_t\\
&\hat{Y}_{2009}=21.86+4.73\cdot X_{2009}=21.86+4.73(4)=40.80 \Rightarrow e_{2009}=Y_{2009}-\hat{Y}_{2009}=44-40.80=+03.19\\
&\hat{Y}_{2010}=21.86+4.73\cdot X_{2010}=21.86+4.73(5)=45.53 \Rightarrow e_{2010}=Y_{2010}-\hat{Y}_{2010}=42-45.53=-03.53\\
&\vdots\\
&\vdots\\
&\vdots\\
&\hat{Y}_{2018}=21.86+4.73\cdot X_{2018}=21.86+4.73(10)=69.19 \Rightarrow e_{2018}=Y_{2018}-\hat{Y}_{2018}=70-69.19=+0.80
\end{aligned}$$

ملاحظة:

لدينا:

$$\begin{aligned}
&e_t=Y_t-\hat{Y}_t=Y_t-\hat{\alpha}-\hat{\beta}X_t=Y_t-(\overline{Y}-\hat{\beta}\overline{X})-\hat{\beta}X_t=Y_t-\overline{Y}+\hat{\beta}\overline{X}-\hat{\beta}X_t=Y_t-\overline{Y}-\hat{\beta}(X_t-\overline{X})\\
&\Rightarrow \sum e_t=\sum\left(Y_t-\overline{Y}-\hat{\beta}(X_t-\overline{X})\right)=\sum(Y_t-\overline{Y})-\hat{\beta}\sum(X_t-\overline{X})=0-\hat{\beta}(0)=0
\end{aligned}$$

لدينا أيضا:

$$\sum e_t=0 \quad \Rightarrow \sum e_t=\sum(Y_t-\hat{Y}_t)=0 \Rightarrow \sum Y_t-\sum\hat{Y}_t=0 \Rightarrow \sum Y_t=\sum\hat{Y}_t$$