المحور الثالث: الانحدار الخطي البسيط

أولا: تقديم النموذج

يسمى هذا النموذج بالبسيط لأنه يتضمن متغيرين أحدهما مستقل(مؤثر)والآخر تابع(مفسر)، كما يسمى أيضا ب"الخطي "لأنه يعطى بصيغة خطية من الشكل $y=\alpha+\beta x$ وهي من الدوال ،التي تمثل بيانيا بخط مستقيم، لأن ميلها ثابت b

يعتبر الانحدار الخطي البسيط أبسط أنواع نماذج الانحدار، بحيث يوجد الكثير من العلاقات الاقتصادية التي يمكن قياسها باستخدام هذا الأسلوب لتحليل تأثير بعض السياسات التي تتضمن تغير قيم لفرد معين. كأن يتم تحليل تأثير النفقات الاعلانية على كمية المبيعات، وعلاقة الكمية المطلوبة من السلعة وسعرها، وأيضا مستوى الأسعار مع معدل التضخم.

ويمكن التعبير عن العلاقة بين المتغيرين بالعلاقة التالية:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1.....n$$

X_i: المتغير المستقل

Y_i :المتغير التابع

α وβ هما معلمتا أو معاملا النموذج.

ε : فيعبر عن الخطإ أو باقي المتغيرات المفسرة ل x التي لم يتضمنها النموذج.

إن النموذج المشار إليه في العلاقة أعلاه لا يمثل سوى تبسيطًا للواقع، إذ لا يمكن في الواقع تفسير المتغير Yاعتمادًا على المتغير Xفقط، لأن هناك العديد من المتغيرات الأخرى التي قد يكون لها تأثير متفاوت على الظاهرة المدروسة. ولهذا السبب يتم تضمين حد عشوائي نرمز له بـ εi، والذي يعبر عن مجموع المتغيرات التي لم يتضمنها النموذج.

ويُطلق على هذا الحد اسم **المتغير العشوائي**، وهو يجسد مختلف العوامل غير الملاحظة أو المهملة التي قد تؤثر في المتغير التابع.

ويتضمن هذا الحد العشوائي ثلاثة أنواع من الأخطاء، وهي:

1. **خطأ في الصياغة (أو تحديد النموذج)** :ويحدث عندما لا تعكس الصيغة المقترحة للنموذج بشكل كامل العلاقة الحقيقية بين المتغير المستقل والمتغير التابع، مما يؤدي إلى إهمال جزء من تأثير المتغيرات المفسرة.

2. **خطأ في القياس** :ويقع عندما لا تعبر البيانات المستخدمة بدقة تامة عن الظواهر الاقتصادية أو الإحصائية المدروسة.
3. **خطأ في المعاينة** :حيث قد تختلف التقديرات المتحصل عليها من عينة إلى أخرى نتيجة اختلاف العينات المختارة من المجتمع الإحصائي.

ثانيا: فرضيات النموذج

الفرضية الأولى: فرضية خطية العلاقة بين المتغيرات

تفترض هذه الفرضية أن العلاقة بين المتغير المستقل Xوالمتغير التابع Yعلاقة خطية، أي يمكن تمثيلها بمعادلة خط مستقيم. ويُعبر عن هذه العلاقة عادة بالصيغة التالية:

$$Yi=\alpha+\beta Xi+\varepsilon i$$

الفرضية الثانية: الأمل الرياضي للأخطاء معدوم

$E(\varepsilon i)=0 \quad \forall i=1,2,...,n$

الفرضية الثالثة : تجانس (ثبات) تباين الأخطاء Homoscedasticity

$Var(\varepsilon_i)=E(\varepsilon_i^2)=\sigma_\varepsilon^2 \quad \forall i=1,2,...,n$

الفرضية الرابعة: عدم وجود ارتباط ذاتي بين الأخطاء

$Cov(\varepsilon_i,\varepsilon_j)=E(\varepsilon_i\varepsilon_j)=0 \quad \forall i\neq j$

-الفرضية الخامسة: استقلال الأخطاء عن المتغير المستقل

$Cov(X_i,\varepsilon_i)=0 \quad \forall i=1,2,...,n$)

الفرضية السادسة: التوزيع الطبيعي للأخطاء

تفترض هذه الفرضية أن الحد العشوائي يتبع التوزيع الطبيعي بمتوسط يساوي صفر وتباين ثابت، أي:

$\varepsilon i \sim N(0,\sigma_2)$

إذا تبيّن أن إحدى فرضيات نموذج الانحدار الخطي البسيط غير محققة، أي حدث خرق لإحدى هذه الفرضيات، فإنه يتعين معالجة هذا الخلل باستخدام الأساليب الإحصائية المناسبة. وبعد تصحيح المشكلة، يتم إعادة تقدير النموذج من جديد بهدف الحصول على تقديرات أكثر دقة وموثوقية لمعاملات الانحدار، وبالتالي تحسين قدرة النموذج على تفسير العلاقة بين المتغيرات المدروسة.

ثالثا: تقدير معالم نموذج الإنحدار الخطي البسيط

عند تمثيل العلاقة بين المتغير المستقل Xوالمتغير التابع Yبيانياً، نحصل على مجموعة من النقاط تعرف بسحابة النقاط. غير أن التعامل الرياضي المباشر مع هذه السحابة يعد أمراً معقداً، نظراً لصعوبة نمذجتها بدقة والتنبؤ بقيمها في المستقبل.

ولهذا يلجأ الإحصائيون إلى تقريب هذه السحابة بمستقيم يكون أقرب ما يمكن إلى جميع نقاطها، ويُطلق عليه **مستقيم الانحدار** .ويسمح هذا المستقيم للباحث بالاستفادة من خصائصه الرياضية من أجل تفسير العلاقة بين المتغيرين وتحديد اتجاهها، كما يساعد في التنبؤ بالقيم المستقبلية للمتغير التابع.

ويهدف تقدير نموذج الانحدار إلى تحديد معاملات هذا المستقيم، أي الحد الثابت ومعامل الانحدار، اللذين يحددان شكل العلاقة بين المتغيرين.

ويتم تقدير معاملات مستقيم الانحدار بالاعتماد على مفاهيم **نظرية التقدير الإحصائي**، وتعد من أشهر الطرق المستخدمة في ذلك **طريقة المربعات الصغرى العادية** (Ordinary Least Squares – OLS)، والتي تقوم على اختيار المعاملات التي تجعل مجموع مربعات الفروق بين القيم الفعلية والقيم المقدرة للمتغير التابع أصغر ما يمكن.

- **طريقة المربعات الصغرى**

تعتمد **طريقة المربعات الصغرى العادية** على فكرة تقليل أو تدنية مجموع مربعات الانحرافات بين القيم الفعلية للمتغير التابع والقيم المقدرة له من خلال نموذج الانحدار. أي أن الهدف من هذه الطريقة هو إيجاد معاملات الانحدار التي تجعل مجموع مربعات الأخطاء أصغر ما يمكن.

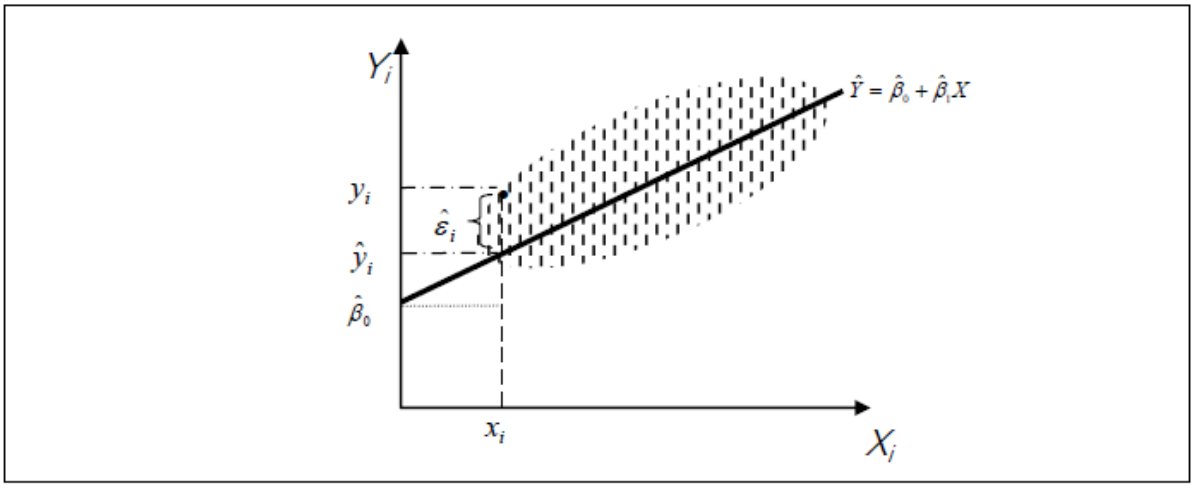
ويُعبر عن ذلك رياضيًا بالعلاقة التالية:

$$\sum_{i=1}^{n} \hat{\varepsilon}_i^2 \text{ ، حيث: } \hat{\varepsilon}_i = Y_i - \hat{Y}_i .$$

^

وتمثل εi **الخطأ العشوائي**، أي الفرق بين القيمة الفعلية للمتغير التابع Yi والقيمة المقدرة له Yi من خلال نموذج الانحدار.

وبالتالي فإن طريقة المربعات الصغرى العادية تقوم باختيار قيم معاملات الانحدار بحيث يكون **مجموع مربعات هذه الأخطاء العشوائية أصغر ما يمكن**، وهو ما يسمح بالحصول على أفضل مستقيم يمر بالقرب من نقاط سحابة البيانات.



يوضح الشكل التالي الانحرافات الموجودة بين نقاط سحابة البيانات التي تمثل القيم الفعلية للمتغير التابع، والنقاط الواقعة على مستقيم الانحدار التي تمثل القيم المقدرة. ووفقًا لطريقة المربعات الصغرى العادية، يجب أن يكون هذا المستقيم هو أفضل مستقيم ممكن، أي الأقرب إلى جميع نقاط السحابة.

وبعبارة أخرى، تقوم فكرة هذه الطريقة على إيجاد المستقيم الذي يقلل إلى أدنى حد ممكن مجموع مربعات الانحرافات بين القيم الفعلية الممثلة بنقاط السحابة والقيم المقدرة الواقعة على مستقيم الانحدار، بحيث يكون هذا المستقيم أفضل تمثيل للعلاقة بين المتغيرين ويعوض سحابة النقاط في عملية التحليل الإحصائي.ويمكن كتابة ذلك رياضيا كالأتي:

$$\min \sum_{i=1}^{n} \hat{\epsilon}_i^2 = \min_{\hat{\alpha},\hat{\beta}} \sum_{i=1}^{n} (Y_i - \hat{\alpha} - \hat{\beta} X_i)^2$$

والشرط اللازم لتدنية هذا المقدار هو أن تكون المشتقات الجزئية بالنسبة للمعلمات معدومة أي :

$$\begin{cases} \frac{\partial}{\partial \alpha}\sum_i (Y_i - \alpha - \beta X_i)^2 = 0 \Leftrightarrow -2\sum_i (Y_i - \alpha - \beta X_i) = 0 \Leftrightarrow \sum_i Y_i - n\alpha - \beta \sum_i X_i = 0 \ldots . (I) \\ \frac{\partial}{\partial \beta}\sum_i (Y_i - \alpha - \beta X_i)^2 = 0 \Leftrightarrow -2\sum_i X_i (Y_i - \alpha - \beta X_i) = 0 \Leftrightarrow \sum_i X_i Y_i - \alpha \sum_i X_i - \beta \sum_i X_i^2 = 0 \ldots (II) \end{cases}$$

بعد حل جملة المعادلتين نتحصل على الأتي:

$$\hat{\beta} = \frac{n\sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{n\sum_i X_i^2 - (\sum_i X_i)^2} = \frac{\sum_i X_i Y_i - n\overline{XY}}{\sum_i X_i^2 - n\overline{X}^2} = \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2}$$

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X}$$

ويكتب النموذج المقدر(المعبر عنه في التمثيل البياني بخط الإنحدار)بالصيغة التالية:

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i$$

مثال تطبيقي

ليكن لديك الجدول التالي الذي يمثل الكمية المطلوبة y_i وسعرها x_i

y_i	18	14	9	7	4	3	1
x_i	1	2	3	4	5	6	7

المطلوب : تقدير العلاقة وإيجاد معادلة انحدار الكمية المطلوبة على السعر x_i باستخدام طريقة OLS .

الحل

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i$$

حيث أن:

$$\hat{\beta} = \frac{n\sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{n\sum_i X_i^2 - \left(\sum_i X_i\right)^2}$$

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X}$$

	x_i	y_i	$x_i\, yi$	x^2
	1	18	18	1
	2	14	28	4
	3	9	27	9
	4	7	28	16
$\overline{y} = \frac{56}{7} = 8$	5	4	20	25
$\overline{x} = \frac{28}{7} = 4$	6	3	18	36
	7	1	7	49
Somme	28	56	146	140

وبتعويض المقادير المختلفة في العلاقة أعلاه نحصل على ما يلي :

$$\tilde{\beta} = \frac{7(146) - (28)(56)}{7(140) - 140^2} = -2.7857$$

$$\tilde{\alpha} = 8 - 2.78(4) = 19.143$$

وبالتالي معادلة الانحدار هي:

$$\hat{Y}_i = 19.143 - 2.7857 X_i$$

مثال 2:

نرغب في تقدير العلاقة الخطية بين الاستهلاك C والدخل R لمجموعة من العائلات، وذلك حسب نموذج الانحدار الخطي البسيط التالي:

$C_i=\alpha+\beta R_i+\mu_i \qquad i= 1,2,...,n$

لهذا الغرض تم اختيار عينة مكونة من 12 عائلة وكانت النتائج كما يلي (ألف وحدة نقدية):

i	1	2	3	4	5	6	7	8	9	10	11	12
Ci	4.8	9.5	9	10.5	10.8	12.2	13.5	16	17.5	18	22	25
Ri	5	10	10	11	12	13	15	18	20	20	25	30

المطلوب:

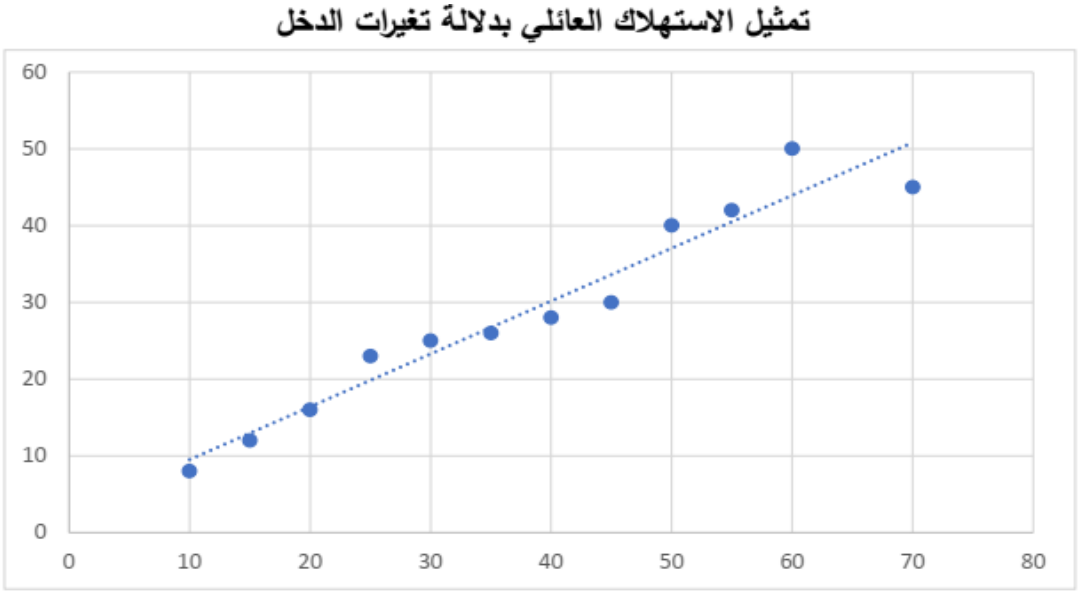
- مثل سحابة النقاط بيانيا.
- قدر معامل معادلة الانحدار
- ماذا تمثل معاملات الانحدار اقتصادياً؟ وهل نتائج النموذج المقدر موافقة للنظرية الاقتصادية؟
- باستعمال معادلة الانحدار، احسب القيم الجديدة المقدرة للاستهلاك.
- احسب البواقي المقدرة

الحل

1. تمثيل سحابة النقاط بيانياً

لتمثيل النقاط، نضع الدخل R_i على المحور الأفقي والاستهلاك C_i على المحور الرأسي.

سنلاحظ أن النقاط تتجه من أسفل اليسار إلى أعلى اليمين، مما يشير إلى علاقة طردية بين الدخل والاستهلاك.



n=12

i	C_i	R_i	R_iC_i	R_i^2
1	4.8	5	24.0	25
2	9.5	10	95.0	100
3	9.0	10	90.0	100
4	10.5	11	115.5	121
5	10.8	12	129.6	144
6	12.2	13	158.6	169
7	13.5	15	202.5	225
8	16.0	18	288.0	324
9	17.5	20	350.0	400
10	18.0	20	360.0	400
11	22.0	25	550.0	625
12	25.0	30	750.0	900
المجموع	**168.8**	**189**	**3113.2**	**3533**

$\sum C_i=168.8,\ \sum R_i=189,\ \sum R_iC_i=3113.2,\ \sum R_i^2=3533$

2. تقدير معالم النموذج

تقدير $\hat{\beta}$

$$\hat{\beta} = \frac{n\sum R_i C_i - \sum R_i \sum C_i}{n\sum R_i^2 - (\sum R_i)^2}$$

$$\hat{\beta} = \frac{12 \times 3113.2 - 189 \times 168.8}{12 \times 3533 - (189)^2}$$

$$\hat{\beta} = \frac{37358.4 - 31903.2}{42396 - 35721}$$

$$\hat{\beta} = \frac{5455.2}{6675} \approx 0.8173$$

تقدير $\hat{\alpha}$

$$\bar{R} = \frac{\sum R_i}{n} = \frac{189}{12} = 15.75$$

$$\bar{C} = \frac{\sum C_i}{n} = \frac{168.8}{12} \approx 14.0667$$

$$\hat{\alpha} = \bar{C} - \hat{\beta}\bar{R}$$

$$\hat{\alpha} = 14.0667 - (0.8173 \times 15.75)$$

$$\hat{\alpha} = 14.0667 - 12.8725 \approx 1.1942$$

معادلة الانحدار المقدرة

$$\boxed{\hat{C}_i = 1.1942 + 0.8173\, R_i}$$

3.التفسير الاقتصادي ومو افقته للنظرية الاقتصادية

- $\hat{\beta}$=**0.8173** يعني أن زيادة الدخل بمقدار ألف وحدة نقدية واحدة تؤدي إلى زيادة الاستهلاك بمقدار **0.8173ألف وحدة نقدية** (أي حوالي 817.3 وحدة نقدية).
- $\hat{\alpha}$=**1.1942** هو مستوى الاستهلاك عندما يكون الدخل صفراً (الاستهلاك الذاتي). قيمته موجبة ومنطقية اقتصادياً.
- **مو افقة النظرية الاقتصادية:**

نعم، لأن العلاقة طردية (الميل موجب)، والميل أقل من 1، وهذا يتوافق مع نظرية كينز للاستهلاك (الميل الحدي للاستهلاك بين 0 و 1).

اما بالنسبة للثابت α فاشارته موجبة وهو متوافق مع النظرية الاقتصادية ، لأن هذه الأخيرة تفترض ضرورة وجود حد أدنى من الاستهلاك (اكبر من الصفر) حتى ولو كان الدخل معدوم

4. حساب القيم المقدرة للاستهلاك

نستخدم المعادلة التالية:

$$\hat{C}_i=1.1942+0.8173R_i$$

i	R_i	$\hat{C}_i$
1	5	1.1942 + 4.0865 = 5.2807
2	10	1.1942 + 8.173 = 9.3672
3	10	9.3672
4	11	1.1942 + 8.9903 = 10.1845
5	12	1.1942 + 9.8076 = 11.0018
6	13	1.1942 + 10.6249 = 11.8191
7	15	1.1942 + 12.2595 = 13.4537
8	18	1.1942 + 14.7114 = 15.9056
9	20	1.1942 + 16.346 = 17.5402
10	20	17.5402
11	25	1.1942 + 20.4325 = 21.6267
12	30	1.1942 + 24.519 = 25.7132

5. حساب البواقي المقدرة

$$\hat{e}_i = C_i - \hat{C}_i$$

i	Ci	$\hat{C}i$	$\hat{e}_i$
1	4.8	5.2807	-0.4807
2	9.5	9.3672	0.1328
3	9.0	9.3672	-0.3672
4	10.5	10.1845	0.3155
5	10.8	11.0018	-0.2018
6	12.2	11.8191	0.3809
7	13.5	13.4537	0.0463
8	16.0	15.9056	0.0944
9	17.5	17.5402	-0.0402
10	18.0	17.5402	0.4598
11	22.0	21.6267	0.3733
12	25.0	25.7132	-0.7132

ملاحظة :مجموع البواقي ≈ 0 (بسبب تقريب الأرقام).

رابعا:خصائص طريقة المربعات الصغرى

ان المقدر يتضمن عدة خصائص ، وهذه الخصائص نجد اغلبها في طريقة المربعات الصغرى العادية *OLS*.

1. عدم التحيز

وإذا عدنا إلى مقدرتي المربعات الصغرى فإن $\hat{\alpha}$مقدرة غير متحيزة للمعلمة α .عدم التحيز يتطلب بأن القيمه المتوقعة لـ $\hat{\alpha}$و التي هي قيمة المعلومة الحقيقية بمعنى آخر متوسط $E(\hat{\alpha}) = \alpha$.و $E(\hat{\beta}) = \beta$

إذا جمعت عينات كثيرة وفي كل عينه نحسب $\hat{\alpha}$او $\hat{\beta}$يتم أخذ المتوسط. ذلك المتوسط نظريا يجب أن يتساوى مع المعلمة الحقيقية αو βعلى التوالي

2. أفضل مقدرات خطية غير متحيزة(BLUE)

وفقًا لنظرية **غاوس-ماركوف** (Gauss-Markov)، التي تنص على ما يلي:

"من بين جميع المقدرات الخطية وغير المتحيزة، تكون مقدرات المربعات الصغرى العادية $\hat{\alpha}$ و βهي أفضل مقدرتين خطيتين وغير متحيزتين، وذلك لأنها تمتلك أصغر تباين ممكن مقارنة مع بقية المقدرات الخطية وغير المتحيزة الأخرى".تتميز هذه النظرية بشكل خاص بخاصية **أقل تباين للمقدرات** (Minimal Variance)،:

3. **خاصية الاتساق :**نقول أن المعلمات المقدرة هي مقدرات متسقة إذا تحقق مل يلي: كلما كبر حجم العينة فان قيمة المعلمة المقدرة $\hat{B}$ تقترب من قيمة المعلمة الحقيقية B، وتباين المعلمة المقدرة يقترب من الصفر، يمن التعبير رياضيا ذلك كما يلي:

$$\lim_{n \to \infty} E(\hat{B}) = B \quad , \qquad \lim_{n \to \infty} E(\hat{\propto}) = \propto$$

$$limv(\hat{B}) = 0 \qquad limv(\hat{\propto}) = 0$$

خامسا: توزيع المعاينة للمقدرات والتقدير بالمجال للمقدرات:

بمعرفة توزيع المقدرات α^ و β^ ، يمكن تكوين فترات ثقة وإجراء الاختبارات الإحصائية للفرضيات المتعلقة بمعالم النموذج الأصلية αو β

1) تقدير تباين الخطأ وتباين المعلمات

نقوم أولا بالبحث عن مقدر تباين الخطأ

- **تقدير تباين الخطأ**

نعلم أن البواقي تحسب بالعلاقة التالية:

$$e_i = Y_i - \hat{Y}_i$$

حيث تمثل Yi القيم الحقيقية المعطاة، بينما تمثل $\hat{Y}_i$ القيم المقدرة بعد التعويض في معادلة الإنحدار المقدرة، ومن خلال مجموعة من العلاقات الرياضية والبراهين 1نصل إلى العبارة التالية لمقدر تباين الخطأ

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_i e_i^2}{n-2} = \frac{\sum_i \left(Y_i - \hat{Y}\right)^2}{n-2}$$

- تقدير تبان المقدرات α وβ

- **تقدير تبان المقدرة β**

$$(\hat{\sigma}^2)_{\hat{\beta}} = \mathrm{Var}(\hat{\beta}) = \frac{\hat{\sigma}^2{}_\varepsilon}{\sum X_i^2 - n\bar{X}^2}$$

- **تقدير تبان المقدرة α**

$$(\hat{\sigma}^2)_{\hat{\alpha}} = \mathrm{Var}(\hat{\alpha}) = \hat{\sigma}_\varepsilon^2 \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum X_i^2 - n\bar{X}^2}\right)$$

$$(\hat{\sigma}^2)_{\hat{\alpha}} = \mathrm{Var}(\hat{\alpha}) = \hat{\sigma}_\varepsilon^2 \cdot \frac{\sum X_i^2}{n\sum (X_i - \bar{X})^2}$$

- **التباين المشترك لـ α و β**

$$\mathrm{Cov}(\hat{\alpha},\hat{\beta}) = -\frac{\hat{\sigma}_u^2 \cdot \bar{X}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

2 مجال الثقة للمعلمات

يعطى مجال ثقة بالنسبة للمعالم αو β كما يلي:

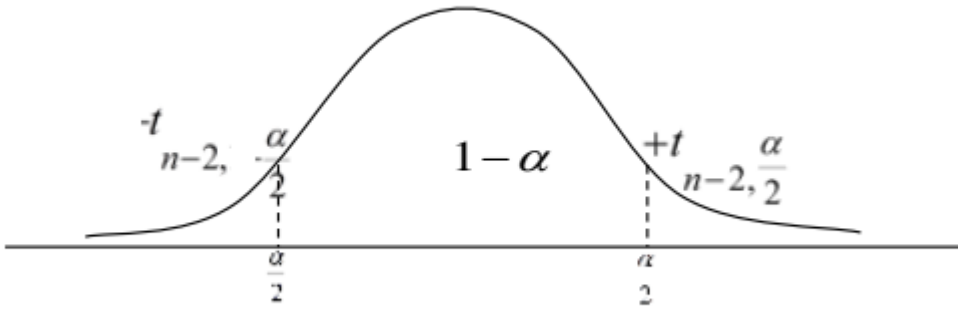
$$\alpha \in \left[\hat{\alpha} \pm \hat{\sigma}_{\hat{\alpha}} \times t_{n-2}^{\frac{\alpha}{2}}\right]$$

$$\beta \in \left[\hat{\beta} \pm \hat{\sigma}_{\hat{\beta}} \times t_{n-2}^{\frac{\alpha}{2}}\right]$$

α/2

t_{n-2} القيمة الحرجة (المقروءة) من توزيع Student عند درجة حرية $n-2$ ومستوى معنوية α كما في الشكل الموالي:

توزيع المعاينة لـ $\hat{\alpha}$ ثنائي الطرف



ملاحظة

كلما كان مجال الثقة ضيقا وصغيرا كان المقدر احسن نظرا لصغر الانحراف المعياري لكل مقدر.

سادسا: اختبار الفرضيات(اختبار معنوية α وβ)

يهدف اختبار الفرضيات المتعلق بمعنوية معاملي الانحدار α و β إلى التحقق من مدى دلالتهما الإحصائية في تفسير العلاقة بين المتغيرات، وذلك من خلال اختبار ما إذا كان كل منهما يختلف معنويًا عن الصفر عبر صياغة فرضية العدم التي تنص على انعدام التأثير مقابل فرضية بديلة تفيد بوجوده.

1. معنوية الثابت α

- نختبر هل للثابت قيمة حقيقية أم لا
- الفرضيات:

$$H_0 : \alpha = 0 \quad \text{(غير معنوي)}$$

$$H_1 : \alpha \neq 0 \quad \text{(معنوي)}$$

2. معنوية الميل β

- وهو الأهم 👇 لأنه يمثل تأثير X على Y
- الفرضيات:

$$H_0 : \beta = 0 \quad \text{(لا يوجد تأثير)}$$

$$H_1 : \beta \neq 0 \quad \text{(يوجد تأثير)}$$

👉 إذا كان **معنويًا**:

- فهذا يعني أن المتغير X يفسر جزءًا من التغير في Y

👉 إذا كان **غير معنوي**:

- فهذا يعني أن X لا يؤثر إحصائيًا على Y

ويتم الاعتماد في ذلك على **اختبار (t)** لستيودنت، حيث تقارن القيمة المحسوبة لكل معامل بالقيمة الجدولية.

فإذا كانت القيمة المحسوبة أكبر من القيمة الجدولية عند مستوى الدلالة المعتمد (غالبًا 0.05) وعند مستوى معنوية n-2 يتم **رفض الفرضية العدمية** وقبول الفرضية البديلة، مما يدل على معنوية المعلمة، أما إذا كانت القيمة المحسوبة أصغر من القيمة الجدولية فتقبل الفرضية العدمية، مما يشير إلى عدم معنوية المعلمة. وتبرز أهمية هذا الاختبار خصوصًا بالنسبة للمعامل β، باعتباره يمثل أثر المتغير المستقل في تفسير التغيرات التي تطرأ على المتغير التابع، حيث إن ثبوت معنويته يدل على وجود علاقة ذات دلالة إحصائية، في حين أن عدم معنويته يشير إلى ضعف أو غياب هذا التأثير.

ويتم هذا الاختبار بحساب الإحصائية التالية:

معنوية المعلمة β

$$T_c = \left| \frac{\hat{\beta} - \beta}{\delta_{\hat{\beta}}} \right|$$

وبما أن الفرضية H_0 تنص على انعدام β وبذلك تصبح قيمة T كالتالي :

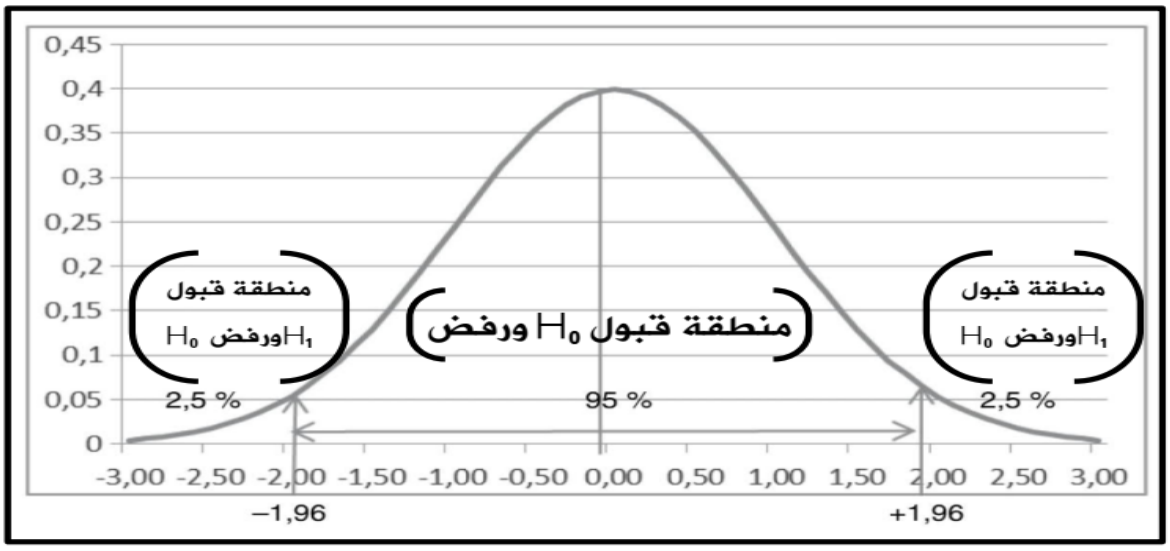
$$T_c = \left| \frac{\hat{\beta}}{\delta_{\hat{\beta}}} \right|$$

حيث $\hat{\beta}$ القيمة المقدرة ل β

$\delta_{\hat{\beta}}$:الانحراف المعياري لـ $\hat{\beta}$

- يتم رفض الفرضية H_0 اذا كان $Tt < Tc$ اي ان $\beta \neq 0$ معناه ان المتغير له تأثير في النموذج وذو معنوية
- ويتم قبول الفرضية H_0 اذا كان $Tt > Tc$ اي ان $\beta = 0$ معناه ان المتغير ليس له تأثير في النموذج .

ونتبع نفس الطريقة بالنسبة للمعلمة α لدراسة معنويتها



يوضح الشكل أعلاه أن منطقة قبول H_1 ورفض H_0 تتواجد في طرفي المنحنى، أي أن قيم Student المحسوبة **مثلا** التي تكون أكبر من 1.96 أو أقل من 1.96-، أي $\left[|t_{Cal}| > t_{n-2}^{0.05}\right]$ تعبر عن قبول الفرضية H_1 ورفض H_0 أي عندها نقول أن المعلمة معنوية أو دالة إحصائيا عند مستوى المعنوية 5 %.

سابعا: تحليل جدول التباين ANOVA و إيجاد معامل التحديد

1. **تحليل جدول التباين ANOVA** يمكن الحصول على معادلة تحليل التباين كالتالي:

نعلم أن:

$$Y_i = \hat{Y}_i + \hat{\varepsilon}_i$$

$$Y_i - \overline{Y} = \hat{Y}_i - \overline{Y} + \hat{\varepsilon}_i$$

وبتربيع طرفي المعادلة أعلاه وجمعها بالنسبة لكل i نجد:

$$\sum_i (Y_i - \bar{Y})^2 = \sum_i (\hat{Y}_i - \bar{Y})^2 + \sum_i \hat{\varepsilon}_i^2$$

وتعد هذه المعادلة مفيدة جدا لخدمة أغراضنا فيما يتعلق بقياس القدرة التفسيرية(ايجاد معامل التحديد)، ولذا من المهم أن نفحص بعناية معنى كل حد من حدودها:

- Total Sum of Squares (TSS) : $\sum_i (Y_i - \bar{Y})^2$ هو مجموع مربعات الانحرافات الكلية في المتغير Y
- Explained Sum of Squares (ESS) : $\sum_i (\hat{Y}_i - \bar{Y})^2$ فهو مجموع مربعات الانحرافات المشروحة
- Residual Sum of Squares (RSS) : ويبقى الحد الأخير $\sum_i \hat{\varepsilon}_i^2$ الذي هو مجموع مربعات البواقي

نعيد صياغة المعادلة السابقة على الشكل : $TSS = ESS + RSS$

أي أن: التباين الإجمالي= التباين المفسر + التباين ير المفسر

وتسمى هذه العبارة عبارة تحليل التباين.

ويوضح الجدول التالي ما يسمى بجدول تحليل التباين في حالة نموذج الإنحدار الخطي البسيط

جدول تحليل التباين لنموذج انحدار خطي بسيط

منشأ التباين	مجموع المربعات	درجة الحرية	متوسط المربعات
X	$ESS = \sum \left(\hat{Y}_i - \bar{\hat{Y}}\right)^2$	1	$ESS/1$
البواقي	$RSS = \sum \left(Y_i - \hat{Y}_i\right)^2$	$n-2$	$RSS/n-2$
الكلي	$TSS = \sum \left(Y_i - \bar{Y}\right)^2$	$n-1$	

2. اختبار القوة التفسيرية للنموذج (جودة التوفيق) (معامل التحديد termination (Coefficient) (R^2)

معامل التحديد R^2 يقيس ويشرح نسبة التغيرات التي تحدث في المتغير التابع Y_i والمشروحة بواسطة تغيرات المتغير المستقل Xi فهي نسبة تأثير المتغير المستقل على المتغير التابع، فهو إذن مقياس للقدرة التفسيرية للنموذج أي يختبر جودة التوفيق و الارتباط.

بمعنى أخر يعبر معامل التحديد عن نسبة تفسير النموذج للظاهرة المدروسة، فإذا وجدنا مثلا أن معامل التحديد يساوي 0.8 فنفسر ذلك بأن المتغير المستقل أو الخارجي يفسر المتغير التابع أو الداخلي بنسبة 80%وأن الباقي % 20ترجع إلى متغيرات أخرى لم تدرج في النموذج.

◄ ملاحظة: يمكن إيجاد قيمة معامل التحديد بتربيع قيمة معامل الإرتباط

من خلال معادلة التباين الكلي نستخرج معامل التحديد

$$TSS = ESS + RSS$$

وبتقسيم كل الأطراف على الانحرافات الكلية TSS نجد :

$$1 = \frac{ESS}{TSS} + \frac{RSS}{TSS}$$

وعليه نعرف معامل التحديد $R^2 = r^2$ كما يلي:

$$R^2 = r^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} = 1 - \frac{\sum \hat{\varepsilon}_i^2}{\sum (Y_i - \bar{Y})^2}$$

وتكون قيمة معامل التحديد محصورة بين 0 و1

$$0 \leq R^2 \leq 1$$

3. معامل الارتباط r

الفرق الجوهري بين معامل الارتباط ومعامل التحديد يتمثل في طبيعة الدور الذي يعكسه كل منهما؛ فمعامل الارتباط يقيس قوة واتجاه العلاقة بين متغيرين دون التمييز بين المتغير المستقل والتابع، أي دون اعتبار للسببية. في المقابل، لا يكتفي معامل التحديد بقياس درجة الارتباط، بل يعبّر عن مدى مساهمة المتغير المستقل X_i في تفسير التغيرات الحاصلة في المتغير التابع، أي أنه يعكس جانبًا تفسيريًا يرتبط باتجاه العلاقة داخل نموذج الانحدار.

$$r = \sqrt{R^2}$$

وجديرٌ بالإشارة إلى أن معامل الارتباط يمكن أن يكون موجبًا أو سالبًا، وتتراوح قيمته بين:

$$-1 \leq r \leq 1$$

دراسة معنوية معامل الارتباط-

ليكن r معامل الارتباط لعينة أخذت من المجتمع .نريد أن نختبر إذا كان معامل الارتباط يختلف معنويا عن الصفر(عند مستوى معنوية α ودرجة حرية n-2)، وهذا يعني اختبار الفرضيات التالية:

H_0: r=0

H_1: r≠0

لإجراء هذا الاختبار فإننا نستخدم توزيع t (توزيع ستيودنت) فتكون قيمة t بالعلاقة التالية:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

فإذا كانت قيمة(t_c المحسوبة)أكبر من قيمة(tT المجدولة أو النظرية) tT<t_c فإننا في هذه الحالة

نرفض الفرضية H_0 لصالح الفرضية البديلةH_1 والعكس صحيح

ثامنا: اختبار المعنوية الكلية للنموذج

يهدف اختبار المعنوية الكلية في نموذج الانحدار الخطي البسيط إلى التحقق من مدى قدرة النموذج على تفسير العلاقة بين متغير مستقل واحد X ومتغير تابعY وبما أن هذا النموذج يتضمن معاملًا واحدا فقط للميلβ ، فإن اختبار المعنوية الكلية يعادل في مضمونه اختبار معنوية هذه المعلمة.

ويختبر هذا الإختبار الفرضيتين التاليتين:

<u>فرضية العدم</u>

$$H_0 : \beta = 0$$

أي لا توجد علاقة خطية بين X وY، مقابل **الفرضية البديلة**:

$$H_1 : \beta \neq 0$$

ويُعتمد في هذا السياق على اختبار (F) بدرجتي حرية ($1 ،n-2$) ويقيس معنوية النموذج ككل، حيث تُحسب إحصائية الاختبار كما يلي:

$$F_{Cal} = \frac{\frac{R^2}{1}}{\frac{(1-R^2)}{n-2}}$$

وكما رأينا سابقا فإن قاعدة اتخاذ القرار هي كما يلي:

- في حالة $F_{Cal} > F_{(1,n-2)}^{\alpha\ \%}$ نقبل الفرضية H_1 ونرفض H_0 أي أن النموذج معنوي وبمعنى أخر المتغير المستقل X يفسر المتغير التابع Y.
- في حالة $F_{Cal} < F_{(1,n-2)}^{\alpha\ \%}$ نرفض الفرضية H_1 ونقبل H_0 أي أن النموذج غير معنوي وبمعنى أخر أن المتغير المستقل X لا يفسر المتغير التابع Y.

ملاحظة:

ويجدر التنويه إلى أنه في حالة الانحدار الخطي البسيط، فإن اختبار F للمعنوية الكلية يعادل اختبار t الخاص بالمعلمة β الميل، حيث:

$$F_{Cal} = \left(t_{tab}\left(\hat{\beta}\right)\right)^2$$

مما يؤكد أن معنوية النموذج ككل ترتبط مباشرة بمعنوية معامل β.

تاسعا: التنبؤ باستخدام نموذج الانحدار الخطي البسيط

1. **التنبؤ النقطي**

ويقصد به إيجاد قيمة المتغير التابع في المستقبل أو عند معلومية قيمة المتغير المستقل، حيث بعد القيام بتقدير نموذج الانحدار الخطي البسيط، و علمنا قيمة المتغير المستقل X_0فيمكن إيجاد قيمةY_{PR} التنبؤية بالتعويض في معادلة الإنحدار مباشرة

مثل إذا كانت لديك $X_{PR}=100$ **ومعادلة الإنحدار هي:** $\hat{Y}_i=100+0.1X$ **فإن قيمة** $\hat{Y}_{PR}$ **التنبؤية هي:**

$$\hat{Y}_{PR}=100+0.1(X_{PR}=100)=110$$

2. التنبؤ بمجال

ندرس التنبؤ بمجال الثقة لأن التنبؤ النقطي هو تنبؤ مبني على معادلة مقدرة، وبالتالي فمن المحتمل أن تكون القيمة التنبؤية المقدرة $\hat{Y}_{PR}$=110 مختلفة عن القيمة التنبؤية الحقيقية Y_{PR} ويسمى الفرق بين القيمة المقدرة والقيمة الحقيقية للمتغير التابع بخطأ التقدير.

وفي هذه الحالة يتطلب الأمر معرفة توزيع المعاينة لـ: $\hat{Y}_{PR}$:حيث أن توزيع المعاينة ل $\hat{Y}_{PR}$ هو توزيع طبيعي بمتوسط $[\alpha+\beta X_{PR}]$وتباين:

$$Var\left(\hat{Y}_{PR}\right)=\hat{\sigma}_{\varepsilon}^{2}\left[1+\frac{1}{n}+\frac{\left(X_{PR}-\overline{X}\right)^2}{\sum\left(X_i-\overline{X}\right)^2}\right]$$

حيث يكتب مجال الثقة لـ Y_{PR} كما يلي:

$$Y_{PR}\in\hat{Y}_{PR}\pm t_{n-2}^{\alpha}\times\sqrt{\hat{\sigma}_{\varepsilon}^{2}\left[1+\frac{1}{n}+\frac{\left(X_{PR}-\overline{X}\right)^2}{\sum\left(X_i-\overline{X}\right)^2}\right]}$$

<u>مثال</u>

نفس معطيات المثال السابق رقم 1

1) احسب جدول تحليل التباين ANOVA
2) أوجد R^2وr
3) ادرس معنوية المعلمات والمعنوية الكلية للنموذج
4) اوجد حدود الثقة لمعالم معادلة الانحدار لدالة الطلب

الحل:

1. حساب جدول تحليل التباين ANOVA

X_i	Y_i	$\hat{Y}_i$	$Y_i - \hat{Y}_i$	$\hat{\varepsilon}_i^2$	$(Y_i - \bar{Y})^2$
1	18	19.142857−2.785714(1)= 16.357143	18−16.357143= 1.642857	$(1.642857)^2$ =2.69898≈2.699	$(18-8)^2$ =100
2	14	19.142857−2.785714(2)= 13.571429	14−13.571429= 0.428571	$(0.428571)^2$ =0.183673	$(14-8)^2$=36
3	9	19.142857−2.785714(3)= 10.785715	9−10.785715= −1.785715	$(-1.785715)^2$ =3.18878≈3.189	$(9-8)^2$=1
4	7	19.142857−2.785714(4)= 8.000001	7−8.000001= −1.000001	$(-1.000001)^2$ =1.000002≈1.000	$(7-8)^2$=1
5	4	19.142857−2.785714(5)= 5.214287	4−5.214287= −1.214287	$(-1.214287)^2$ =1.47449≈1.474	$(4-8)^2$=16
6	3	19.142857−2.785714(6)= 2.428573	3−2.428573= 0.571427	$(0.571427)^2$ =0.32653≈0.327	$(3-8)^2$=25
7	1	19.142857−2.785714(7)= −0.357141	1−(−0.357141)= 1.357141	$(1.357141)^2$ =1.84183≈1.842	$(1-8)^2$=49
				∑=10.715	∑=228

جدول تحليل التباين

مصدر التغير	مجموع مربعات الانحرافات	درجة الحرية	متوسط مربعات الانحرافات
الانحدار (المتغير X)	$ESS = \sum(\hat{Y}_i - \bar{Y})^2$ ESS=TSS-RSS= 228-10 ,715=217 .285	1	217 .285
البواقي	RSS=$\sum \hat{\varepsilon}_i^2$ =10.715	n-2=5	2،143

المجموع	$TSS = \sum (y_i - \bar{y})^2 = 228$	n-1=6	/

2. حساب معامل التحديد ومعامل الارتباط

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} = 1 - \frac{\sum \hat{\varepsilon}_i^2}{\sum (Y_i - \bar{Y})^2}$$

$$R^2 = 1 - \frac{10.715}{228} = 1 - 0.0469956 \approx 0.9530044$$

$$\boxed{R^2 \approx 0.9530}$$

حساب معامل الارتباط

ط1

$$r = \sqrt{0.95} = -0.97$$

ط2 لدينا :

$$\sum X_i = 28, \quad \sum Y_i = 56, \quad \sum X_i Y_i = 146, \quad \sum X_i^2 = 140$$

$$n = 7$$

$$\bar{X} = \frac{28}{7} = 4, \quad \bar{Y} = \frac{56}{7} = 8$$

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2] \cdot [n\sum Y^2 - (\sum Y)^2]}}$$

$$n\sum XY = 7 \times 146 = 1022$$

$$(\sum X)(\sum Y) = 28 \times 56 = 1568$$

$$n\sum X^2 - (\sum X)^2 = 7 \times 140 - 28^2 = 980 - 784 = 196$$

$$n\sum Y^2 - (\sum Y)^2 = 7 \times 676 - 56^2 = 4732 - 3136 = 1596$$

$$r = \frac{1022 - 1568}{\sqrt{196 \times 1596}}$$

$$= \frac{-546}{\sqrt{196 \times 1596}}$$

$$\sqrt{196} = 14$$

$$r = \frac{-546}{14 \times \sqrt{1596}}$$

$$\sqrt{1596} \approx 39.94997$$

$$14 \times 39.94997 \approx 559.2996$$

$$r \approx \frac{-546}{559.2996} \approx -0.9762$$

95% من التغير الذي يحدث في الكميات المطلوبة يرجع إلى التغيرات التي تحدث في سعر السلعة، بينما ترجع النسبة المتبقية 5% إلى محددات الطلب الأخرى مثل الأذواق، أسعار السلع البديلة.....وغيرها.

r=−0.97 يعني أن **العلاقة عكسية وقوية جداً بين الكمية المطلوبة وسعر السلعة**

3. **حساب معنوية المعلمات**

أ. **حساب التباين**

- **تقدير تباين الخطأ** $(\hat{\sigma}^2)$

$$\hat{\sigma}^2_{\varepsilon} = \frac{\sum \hat{\epsilon}_i^2}{n-2} = \frac{10.714308}{5} \approx 2.1428616$$

➢ **تقدير تباين المقدرات**

تباين $\hat{\beta}$

$$(\hat{\sigma}^2)_{\hat{\beta}} = \mathrm{Var}(\hat{\beta}) = \frac{\hat{\sigma}^2_{\varepsilon}}{\sum X_i^2 - n\bar{X}^2}$$

$$\mathrm{Var}(\hat{\beta}) = \frac{\hat{\sigma}^2_{\varepsilon}}{\sum X_i^2 - n\bar{X}^2}$$

$$\sum X_i^2 - n\bar{X}^2 = 140 - 7 \times 16 = 140 - 112 = 28$$

$$\mathrm{Var}(\hat{\beta}) = \frac{2.1428616}{28} \approx 0.07653077$$

تباين $\hat{\alpha}$

$$(\hat{\sigma}^2)_{\hat{\alpha}} = \mathrm{Var}(\hat{\alpha}) = \hat{\sigma}^2_{\varepsilon}\left(\frac{1}{n} + \frac{\bar{X}^2}{\sum X_i^2 - n\bar{X}^2}\right)$$

$$= 2.1428616\left(\frac{1}{7} + \frac{16}{28}\right)$$

$$\frac{1}{7} \approx 0.142857, \quad \frac{16}{28} \approx 0.5714286, \quad \text{المجموع} = 0.7142856$$

$$\mathrm{Var}(\hat{\alpha}) \approx 2.1428616 \times 0.7142856 \approx 1.530615$$

ب. **اختبار المعنوية الإحصائية لمقدرات النموذج**

لدراسة **المعنوية الإحصائية** لمقدرات نموذج الانحدار الخطي البسيط $\hat{\alpha}$ و $\hat{\beta}$ نستخدم **اختبار الفرضيات** (اختبارt)

لدينا البيانات السابقة:

$$n = 7, \quad \hat{\alpha} = \frac{268}{14} \approx 19.142857, \quad \hat{\beta} = -\frac{39}{14} \approx -2.785714$$

$$\mathrm{Var}(\hat{\alpha}) = \frac{75}{49} \approx 1.530612, \quad \mathrm{Var}(\hat{\beta}) = \frac{15}{196} \approx 0.076531$$

$$\hat{\sigma}^2_{\varepsilon} = \frac{15}{7} \approx 2.142857$$

- **اختبار معنوية** $\hat{\beta}$

الفرضيات:

H_0:β=0(لا توجد علاقة خطية)

H_1:β≠0(توجد علاقة خطية)

إحصائية الاختبار:

$$t_{\hat{\beta}} = \frac{\hat{\beta}}{\sqrt{\mathrm{Var}(\hat{\beta})}} = \frac{-39/14}{\sqrt{15/196}}$$

$$\sqrt{\mathrm{Var}(\hat{\beta})} = \sqrt{\frac{15}{196}} = \frac{\sqrt{15}}{14} \approx \frac{3.872983}{14} \approx 0.276642$$

$$t_{\hat{\beta}} = \frac{-2.785714}{0.276642} \approx -10.07$$

القيمة الحرجة: $t_{(n-2,\alpha/2)}=t_{(5,0.025)}=2.571$

$|t_{\beta\wedge}|=10.07>2.571 \implies$ نرفض H_0

المعلمة (β) معنوية إحصائيًا عند مستوى دلالة 5 %

- **<u>اختبار معنوية α</u>**

الفرضيات:

$H_0:\alpha=0$

$H_1:\alpha\neq0$

إحصائية الاختبار:

$$t_{\hat{\alpha}} = \frac{\hat{\alpha} - 0}{\sqrt{\mathrm{Var}(\hat{\alpha})}} = \frac{268/14}{\sqrt{75/49}}$$

$$\sqrt{\mathrm{Var}(\hat{\alpha})} = \sqrt{\frac{75}{49}} = \frac{\sqrt{75}}{7} = \frac{5\sqrt{3}}{7} \approx \frac{8.660254}{7} \approx 1.237179$$

$$\hat{\alpha} = \frac{268}{14} \approx 19.142857$$

$$t_{\hat{\alpha}} = \frac{19.142857}{1.237179} \approx 15.47$$

القيمة الحرجة: نفسها $t_{5,0.025}=2.571$

$|t_{\alpha^\wedge}|=15.47>2.571 \implies$ نرفض H0|

االمعلمة (α) معنوية إحصائيًا عند مستوى دلالة 5%.

جدول ملخص نتائج الاختبارات

المعامل	التقدير	الخطأ المعياري (SE)	قيمة t	القيمة الحرجة (5%)	القرار
$\alpha^\wedge$	19.143	1.237	15.47	2.571	✓ معنوي
$\beta^\wedge$	-2.786	0.277	-10.07	2.571	✓ معنوي

ج. اختبار معنوية النموذج ككل (اختبار F)

في الانحدار الخطي البسيط، اختبار F يساوي مربع قيمة t للمعلمة β :

$$F = t_{\hat{\beta}}^2 \approx (10.07)^2 \approx 101.4$$

$$F_{(1,5,0.05)} = 6.61$$

$101.4 > 6.61 \implies$ النموذج ككل معنوي

ط2/

$$F_{Cal} = \frac{\frac{R^2}{1}}{\frac{(1-R^2)}{n-2}}$$

F_{cal}=0 ,9530 /((1-0.9530) /(7-2))=0 ,9530/0.0094=**101 ,4**

F_{cal}=101 ,4

$$F_{(1,5,0.05)} = 6.61$$

$$101.4 > 6.61 \implies \text{النموذج ككل معنوي}$$

4. حساب مجال الثقة للمعلمات

يمكن بناء مجالات ثقة للمعالم:

$$\beta_0 \in \left[\hat{\beta}_0 - t_{0.025}\hat{\sigma}_{\hat{\beta}_0}, \hat{\beta}_0 + t_{0.025}\hat{\sigma}_{\hat{\beta}_0}\right]$$
$$\beta_1 \in \left[\hat{\beta}_1 - t_{0.025}\hat{\sigma}_{\hat{\beta}_1}, \hat{\beta}_1 + t_{0.025}\hat{\sigma}_{\hat{\beta}_1}\right]$$

حيث $t_{0.025}$ هي القيمة الحرجة لتوزيع ستيودنت بنسبة معنوية 5% ودرجة حرية $n-2=5$ والتي تعادل القيمة 2.570 في جدول توزيع ستيودنت. بالتطبيق العددي لدينا :

$$\beta_0 \in [19.14 - 2.570\times 1.23,\ 19.14 + 2.570\times 1.23]$$

أي :

$$\beta_0 \in [15.963\ ,\ 22.32]$$

و

$$\beta_1 \in [-2.78 - 2.570\times 0.27\ ,\ -2.78 + 2.57\times 0.27]$$

هذا يعني :

$$\beta_1 \in [-3.49\ ,\ -2.07]$$

تمارين

يريد آحد الباحثين إيجاد آثر التضخم على قيمة الناتج الوطني في السنوات 2010 – 2019

السنوات	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
التضخم	1	3	4	4	6	8	10	10	11	13
قيمة الناتج الوطني	80	97	92	102	103	111	119	123	117	136

المطلوب:

1. **حدد المتغير التابع والمتغير المستقل.**
2. **قدر معاملات نموذج الإنحدار الخطي البسيط باستخدام طريقة المربعات الصغرى العادية .**
3. **أوجد مقدر تباين الخطأ $\hat{\sigma}_{\varepsilon}^{2}$.**
4. **أحسب مقدرات الإنحرافات المعيارية للمعلمات $\hat{\sigma}_{\hat{\beta}_0}$ و $\hat{\sigma}_{\hat{\beta}_1}$.**
5. **أوجد t المحسوبة لـ: $\hat{\beta}_0$ و $\hat{\beta}_1$، ماذا تستنتج حول معنوية المعلمات وذلك عند مستوى معنوية 5٪.؟**
6. **أرسم جدول تحليل التباين.**
7. **أوجد معامل التحديد R^2 و F المحسوبة، ماذا تستنتج؟**
8. **قدر مجال الثقة لقيمة الناتج الوطني عند مستوى 15 من التضخم .**
9. **قدر مجال الثقة لقيمة الناتج الوطني في سنة 2021.**
10. **استعرض النتائج السابقة باستخدام برنامج 10 EViews.**

الحل

ملاحظة: في حل التمرين تم استخدام β1 و β2 α=β1 β =β2

1. المتغير المستقل هو المتغير الذي نفترض تأثيره على المتغير التابع، وبالتالي فإن:
   - المتغير المستقل: X : هو مستوى التضخم.
   - المتغير التابع: Y : هو قيمة الناتج الوطني.
2. تقدير معاملات نموذج الإنحدار الخطي البسيط: تعطى الصيغة العامة لنموذج الإنحدار الخطي البسيط بالصيغة: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ ويمكن إيجاد المقدرات $\hat{\beta}_0$ و $\hat{\beta}_1$ باستخدام العلاقات التالية:

$$\begin{cases} \hat{\beta}_1 = \dfrac{n\sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{n\sum_i X_i^2 - \left(\sum_i X_i\right)^2} = \dfrac{\sum X_i Y_i - n\overline{XY}}{\sum X_i^2 - n\overline{X}^2} = \dfrac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \\ \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X} \end{cases}$$

ونلخص الحسابات في الجدول التالي:

السنوات	X	Y	XY	X^2	Y^2	$(X-\overline{X})^2$	$\hat{Y}$	$(Y_i-\hat{Y}_i)^2$	t	tX	t^2
2010	1	80	80	1	6400	36	84	16	1	1	1
2011	3	97	291	9	9409	16	92	25	2	6	4
2012	4	92	368	16	8464	9	96	16	3	12	9
2013	4	102	408	16	10404	9	96	36	4	16	16
2014	6	103	618	36	10609	1	104	1	5	30	25
2015	8	111	888	64	12321	1	112	1	6	48	36
2016	10	119	1190	100	14161	9	120	1	7	70	49
2017	10	123	1230	100	15129	9	120	9	8	80	64
2018	11	117	1287	121	13689	16	124	49	9	99	81
2019	13	136	1768	169	18496	36	132	16	10	130	100
المجموع	70	1080	8128	632	119082	142	1080	170	55	492	385

$$\begin{cases} \hat{\beta}_1 = \dfrac{n\sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{n\sum_i X_i^2 - \left(\sum_i X_i\right)^2} = \dfrac{10(8128) - (70)(1080)}{10(632) - (70)^2} = \dfrac{5680}{1420} = 4 \\ \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X} = 108 - (4)(7) = 80 \end{cases}$$

ومنه تكتب معادلة التقدير على النحو التالي: $\boxed{\hat{Y}_i = 80 + 4X_i}$

3. إيجاد مقدر تباين الخطأ $\hat{\sigma}_\varepsilon^2$: يعطى مقدر تباين الخطأ بالعلاقة التالية:

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_i e_i^2}{n-2} = \frac{\sum_i (Y_i - \hat{Y})^2}{n-2} = \frac{\sum (Y_i - \overline{Y})^2 - \hat{\beta}_1^2 \sum (X_i - \overline{X})^2}{n-2}$$

نقوم في الجدول بحساب المقدار $\sum_i \left(Y_i - \hat{Y}_i\right)^2$ بحيث: Y_i هي القيم المعطاة، و $\hat{Y}_i$ هي القيم المقدرة، والقيمة المحصل عليها هي: $\boxed{\sum_i \left(Y_i - \hat{Y}_i\right)^2 = 170}$، (أنظر الجدول).

كما بإمكاننا استخدام العبارة $\sum_i \left(Y_i - \hat{Y}_i\right)^2 = \sum \left(Y_i - \overline{Y}\right)^2 - \hat{\beta}_1^2 \sum \left(X_i - \overline{X}\right)^2 = \sum Y_i^2 - n\overline{Y}^2 - \hat{\beta}_1^2 \sum \left(X_i - \overline{X}\right)^2$

$$\boxed{\sum_i \left(Y_i - \hat{Y}_i\right)^2 = 119082 - (10)(108)^2 - (4)^2(142) = 170}$$

مع ملاحظة أن المقدار: $\sum \left(X_i - \overline{X}\right)^2$ يمكن أن يكتب كما يلي: $\sum \left(X_i - \overline{X}\right)^2 = \sum X_i^2 - n\overline{X}^2$ ونفس الملاحظة بالنسبة للمقدار: $\sum \left(Y_i - \overline{Y}\right)^2$ إذ يمكن أن يكتب كما يلي: $\sum \left(Y_i - \overline{Y}\right)^2 = \sum Y_i^2 - n\overline{Y}^2$

ومنه فإن مقدر تباين الخطأ هو:

$$\boxed{\hat{\sigma}_\varepsilon^2 = \frac{\sum_i \left(Y_i - \hat{Y}\right)^2}{n-2} = \frac{170}{10-2} = 21.25}$$

وبالتالي فإن مقدر الإنحراف المعياري للخطأ هو الجذر التربيعي لـ: $\hat{\sigma}_\varepsilon^2$ ، ونكتب $\hat{\sigma}_\varepsilon = \sqrt{21.25} = 4.61$

4. حساب مقدرات الإنحرافات المعيارية للمعلمات $\hat{\sigma}_{\hat{\beta}_0}$ و $\hat{\sigma}_{\hat{\beta}_1}$:

تجدر الإشارة إلى أن مقدرات الإنحرافات المعيارية للمعلمات تساوي الجذر التربيعي لمقدرات تباينات المعلمات، لهذا فإنه لإيجاد المقدرات المطلوبة نستخدم العبارتين التاليتين:

$$\hat{\sigma}_{\hat{\beta}_1} = \sqrt{\frac{\hat{\sigma}_\varepsilon^2}{\sum_i \left(X_i - \overline{X}\right)^2}} = \sqrt{\frac{21.25}{142}} = 0.39$$

$$\hat{\sigma}_{\hat{\beta}_0} = \sqrt{\hat{\sigma}_\varepsilon^2 \left[\frac{1}{n} + \frac{\overline{X}^2}{\sum_i \left(X_i - \overline{X}\right)^2}\right]} = \sqrt{21.25\left[\frac{1}{10} + \frac{7^2}{142}\right]} = 3.07$$

5. إيجاد t المحسوبة لـ: $\hat{\beta}_0$ و $\hat{\beta}_1$، ثم اختبار معنوية المعلمات.

◄ **إيجاد t المحسوبة لـ: $\hat{\beta}_0$ و $\hat{\beta}_1$:**

- t المحسوبة لـ: $\hat{\beta}_0$: $t_{Cal} = \dfrac{\left|\hat{\beta}_0\right|}{\hat{\sigma}_{\hat{\beta}_0}} = \dfrac{|80|}{3.07} \approx 26.01$
- t المحسوبة لـ: $\hat{\beta}_1$: $t_{Cal} = \dfrac{\left|\hat{\beta}_1\right|}{\hat{\sigma}_{\hat{\beta}_1}} = \dfrac{|4|}{0.39} \approx 10.34$

◄ **اختبار معنوية المعلمات**

الفرضيات التي نود اختبارها هي: $\begin{cases} H_0 : \beta_1 = 0 \\ H_1 : \beta_1 \neq 0 \end{cases}$ و $\begin{cases} H_0 : \beta_0 = 0 \\ H_1 : \beta_0 \neq 0 \end{cases}$

- لا بد أولا من تحديد t الجدولية، لدينا: $t_{Tab} = t_{n-2}^{0.05} = t_8^{0.05} = 2.306$

- نقارن الآن t المحسوبة و t الجدولية، بحيث إذا كانت t المحسوبة أكبر من t الجدولية نقبل الفرضية H_1. نلاحظ أن t المحسوبة أكبر من t الجدولية في كلتا المعلمتين، ومنه نقبل الفرضية H_1، ونقول أن المعلمتان "الحقيقيتان" مختلفتان إحصائيا عن الصفر.

6. جدول تحليل التباين:

منشأ التباين	مجموع المربعات	درجة الحرية	متوسط المربعات
X	$ESS = \sum\left(\hat{Y}_i - \overline{\hat{Y}}\right)^2 = \hat{\beta}_1^2 \sum\left(X_i - \overline{X}\right)^2 = 2272$	1	$2272/1$
البواقي	$RSS = \sum\left(Y_i - \hat{Y}_i\right)^2 = 170$	$10-2=8$	$170/8$
الكلي	$TSS = \sum\left(Y_i - \overline{Y}\right)^2 = \sum Y_i^2 - n\overline{Y}^2 = 2442$	$10-1=9$	

7. إيجاد معامل التحديد R^2 و F المحسوبة:

- معامل التحديد R^2:

$$R^2 = \frac{ESS}{TSS} = \frac{2272}{2442} = 0.93$$

التعليق: 93٪ من تغيرات الناتج الوطني راجعة إلى تغير التضخم و 7٪ من تغيرات الناتج الوطني تعود إلى عوامل أخرى غير مُدرجة في النموذج.

- F المحسوبة:

$$F = \frac{ESS/1}{RSS/n-2} = \frac{2272/1}{170/8} = 106.92$$

لدينا F الجدولية هي: $F_{(1,n-2)}^{0.05} = F_{(1,8)}^{0.05} = 3.32$

يقوم اختبار Fisher على الفرضية التالية: $\begin{cases} H_0 : ESS = 0 \\ H_1 : ESS \neq 0 \end{cases}$، ونقبل الفرضية H1 إذا كانت F المحسوبة أكبر من F الجدولية ونقول أن ESS يختلف معنويا عن الصفر.

مع ملاحظة أنه في نموذج الإنحدار الخطي البسيط فإن الفرضيات التالية كلها تؤدي نفس الغرض.

الإختبار	H_0	H_1
Student	$\beta_1 = 0$	$\beta_1 \neq 0$
Student	$r_{X,Y} = 0$	$r_{X,Y} \neq 0$
Fisher	$ESS = 0$	$ESS \neq 0$

8. تقدير مجال الثقة لقيمة الناتج الوطني عند مستوى 15 من التضخم .

لدينا: $X_{PR} = 15$. و لدينا عبارة مجال الثقة كما يلي:

$$Y_{PR} \in \hat{Y}_{PR} \pm t_{n-2}^{\alpha} \times \sqrt{\hat{\sigma}_{\varepsilon}^2 \left[1 + \frac{1}{n} + \frac{\left(X_{PR} - \overline{X}\right)^2}{\sum \left(X_i - \overline{X}\right)^2} \right]}$$

- $\hat{Y}_{PR} = 80 + 4(15) = 140$
- $t_{n-2}^{\alpha} = t_{10-2}^{0.05} = t_8^{0.05} = 2.306$
- $\hat{\sigma}_{\varepsilon}^2 = 21.25$
- $\left(X_{PR} - \overline{X}\right)^2 = (15-7)^2 = 64$
- $\sum \left(X_i - \overline{X}\right)^2 = 142$

$$Y_{PR} \in (140) \pm (2.306) \sqrt{21.25 \left[1 + \frac{1}{10} + \frac{64}{142} \right]}$$

$$Y_{PR} \in (140) \pm (13.24)$$

$$Y_{PR} \in [126.76 , 153.24]$$

9. تقدير مجال الثقة لقيمة الناتج الوطني في سنة 2021:

في البداية لا بد من إيجاد القيمة التنبؤية للمتغير المستقل وهي القيمة: X_{PR} ، ويتم إيجادها بإجراء انحدار لقيم X على قيم معطاة للزمن وتسمى المعادلة المحصل عليها بمعادلة الإتجاه العام، وصيغتها: $\hat{X}_i = \hat{\alpha}_0 + \hat{\alpha}_1 t, \quad t = 1.....n$ ، بنفس منهجية تقدير معاملات الإنحدار الخطي البسيط يتم تقدير معاملات معادلة الإتجاه العام(انظر الحسابات في الجدول)، فنجد المعادلة:

$$\hat{X}_i = -0.13 + 1.30t \, , t = 1....10$$

للحصول على X_{PR} نعوض بالقيمة 12 (إذا كانت 10 تقابل السنة 2019 فإن 12 تقابل السنة 2021)

$$X_{PR} = -0.13 + 1.30(12) = 15.47$$

بعد حصولنا على $X_{PR} = 15.47$ نعوض في معادلة الإنحدار الخطي البسيط للحصول على Y_{PR} ثم ننشئ مجال التنبؤ، كما في الإجابة السابقة:

$$Y_{PR} \in \hat{Y}_{PR} \pm t_{n-2}^{\alpha} \times \sqrt{\hat{\sigma}_{\varepsilon}^2 \left[1 + \frac{1}{n} + \frac{\left(X_{PR} - \overline{X}\right)^2}{\sum \left(X_i - \overline{X}\right)^2} \right]}$$

- $\hat{Y}_{PR} = 80 + 4(15.47) = 141.88$
- $t_{n-2}^{\alpha} = t_{10-2}^{0.05} = t_8^{0.05} = 2.306$
- $\hat{\sigma}_{\varepsilon}^2 = 21.25$
- $\left(X_{PR} - \overline{X}\right)^2 = (15.47-7)^2 = 71.74$
- $\sum \left(X_i - \overline{X}\right)^2 = 142$