المحور الثالث: الانحدار الخطي البسيط

أولا:تقديمالنموذج

يسمىهذاالنموذجبالبسيطلأنهيتضمنمتغيرينأحدهمامستقل(مؤثر)والآخرتابع(مفسر)،

كمايسمىأيضاب"الخطي "لأنهيعطىبصيغةخطيةمنالشكل $y=\alpha+\beta x$

وهيمنالدوال،التيتمثلبيانيابخطمستقيم،لأنميلهاثابت b

يعتبر الانحدار الخطي البسيط أبسط أنواع نماذج الانحدار، بحيث يوجد الكثير من العلاقاتالاقتصادية التي يمكن قياسها باستخدام هذا الأسلوبلتحليل 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.

ويمكنالتعبيرعنالعلاقةبينالمتغيرين بالعلاقة التالية:

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

X_i: المتغيرالمستقل

Y_i: المتغيرالتابع

α وβ هامامعلمتاأومعاملاالنموذج.

ε : فيعبرعنالخطأأوباقيالمتغيراتالمفسرةل xالتيلميتضمنهاالنموذج.

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 المستقلXوالمتغير 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

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

الفرضية الثالثة : تجانس (ثبات) تباين الأخطاء Homoscedasticity

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

الفرضية الرابعة: عدم وجود ارتباط ذاتي بين الأخطاء

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

-الفرضية الخامسة: استقلال الأخطاء عن المتغير المستقل

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

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: تقديرمعالمنموذجالإنحدارالخطيالبسيط

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**

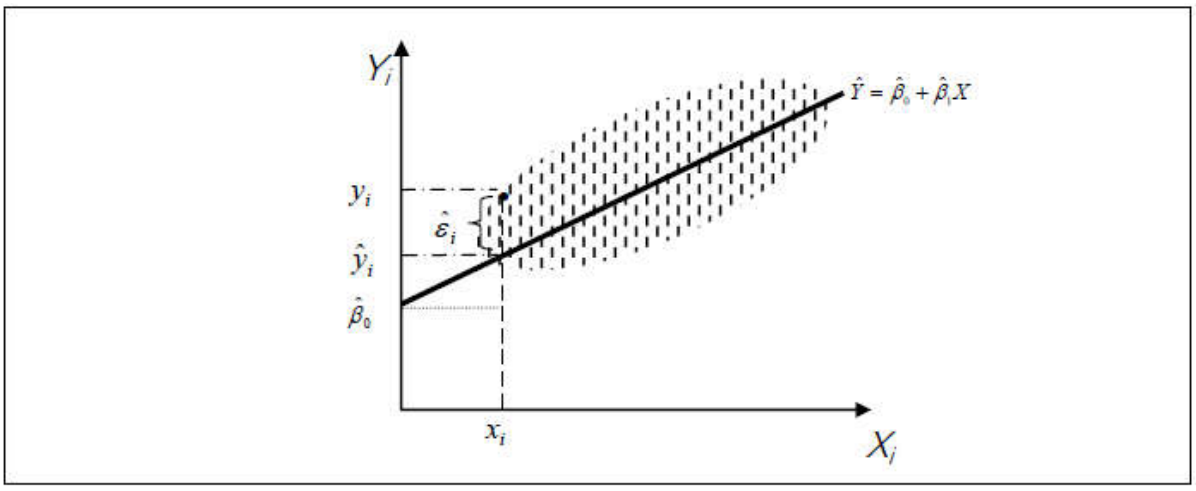
تعتمد **طريقة المربعات الصغرى العادية** على فكرة تقليل أو تدنية مجموع مربعات الانحرافات بين القيم الفعلية للمتغير التابع والقيم المقدرة له من خلال نموذج الانحدار. أي أن الهدف من هذه الطريقة هو إيجاد معاملات الانحدار التي تجعل مجموع مربعات الأخطاء أصغر ما يمكن.

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 Yiوالقيمة المقدرة له Yiمن 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:

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

بعد حل جملة المعادلتين نتحصل على الأتي:

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

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

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:

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

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

	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 Cوالدخل Rلمجموعة من العائلات، وذلك حسب نموذج الانحدار الخطي البسيط التالي:

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

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:

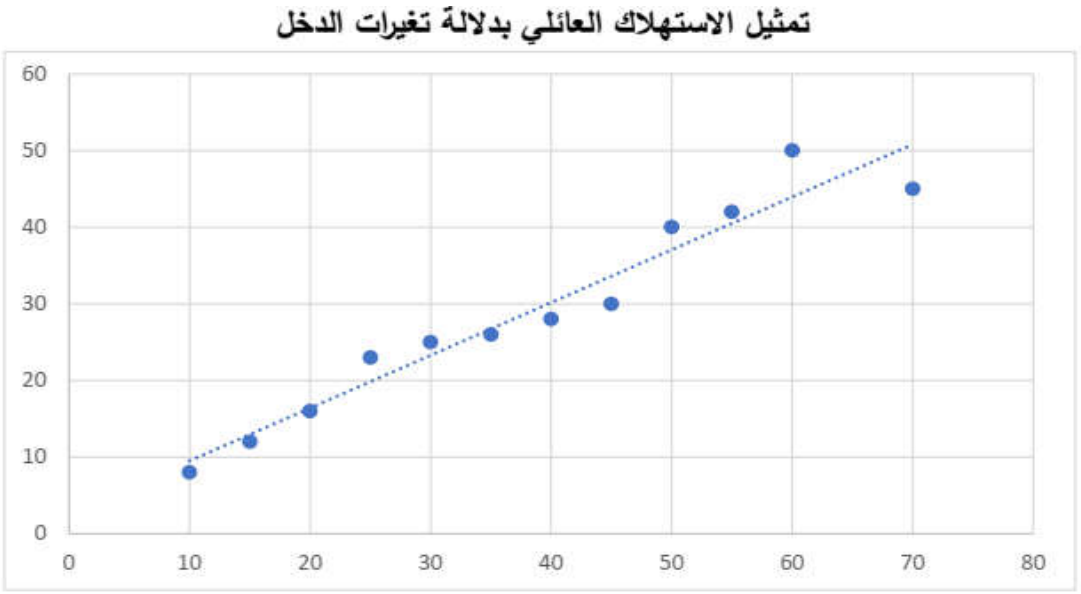
- مثل سحابة النقاط بيانيا.
- قدر معامل معادلة الانحدار
- ماذا تمثل معاملات الانحدار اقتصادياً؟ وهل نتائج النموذج المقدر موافقة للنظرية الاقتصادية؟
- باستعمال معادلة الانحدار، احسب القيم الجديدة المقدرة للاستهلاك.
- احسب البواقي المقدرة

الحل

1. تمثيل سحابة النقاط بيانياً

لتمثيل النقاط، نضع الدخل R_iعلى المحور الأفقي والاستهلاك C_iعلى المحور الرأسي.

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



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 وموافقته للنظرية الاقتصادية

- $\hat{\beta}$=**0.8173** يعني أن زيادة الدخل بمقدار ألف وحدة نقدية واحدة تؤدي إلى زيادة الاستهلاك بمقدار **0.8173 ألف 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.
- **موافقة 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:

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

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	$\hat{Ci}$	$\hat{e_i}$
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 $\hat{\beta}$هي 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 $\hat{\alpha}$و $\hat{\beta}$، 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
- **التباين والانحراف المعياري لـــ β**

$v(\hat{B}) = \frac{\frac{(y_i-\bar{y})^2}{n-2}}{\Sigma(x_i-\bar{x})^2} = \frac{S^2}{\Sigma(x_i-\bar{x})^2}$	تباين $\hat{B}$
$\delta(\hat{B}) = \sqrt{\frac{\frac{(y_i-\bar{y})^2}{n-2}}{\Sigma(x_i-\bar{x})^2}} = \sqrt{\frac{S^2}{\Sigma(x_i-\bar{x})^2}}$	الانحراف المعياري $\hat{B}$

- وتساوي ايضا :

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

- تقدير التباين والانحراف المعياري لـ α

$$v(\hat{\alpha}) = = \frac{\frac{(y_i - \hat{y})^2}{n-2}}{n} + \bar{X}^2 \cdot v(\hat{B}) \qquad (\hat{\alpha}) \text{ تباين}$$

$$\delta(\hat{\alpha}) = \sqrt{\frac{\frac{(y_i - \hat{y})^2}{n-2}}{n} + \bar{X}^2 \cdot v(\hat{B})} \qquad (\hat{\alpha}) \text{ الانحراف المعياري}$$

- وتساوي ايضا:

- 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

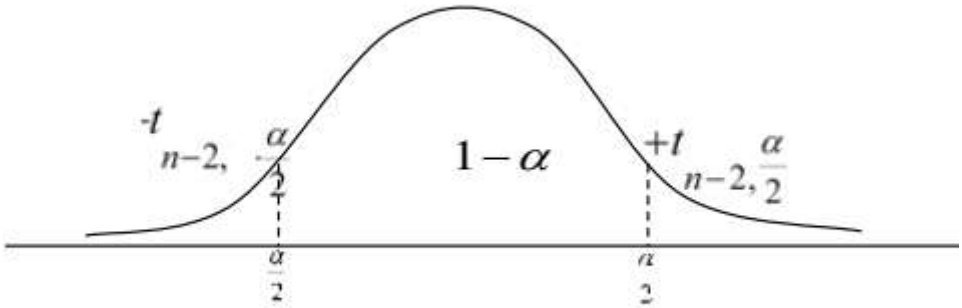
يعطى مجال ثقة بالنسبة للمعالم α و β كما يلي:

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

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

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 αوβإلى 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:

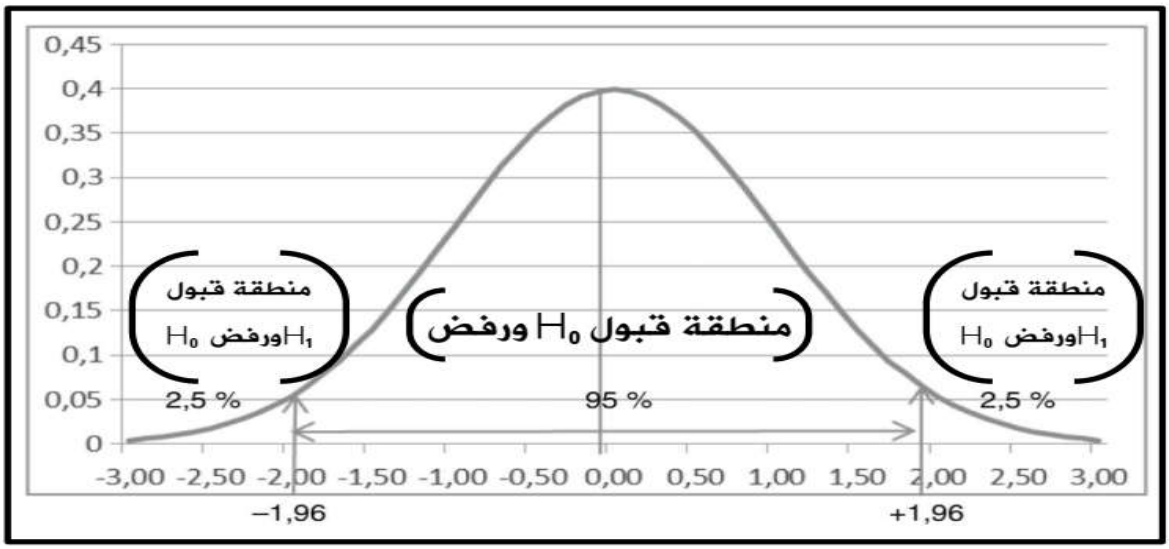
$$T_c = \left| \frac{\hat{\beta}}{\delta_{\hat{\beta}}} \right|$$

حيث $\hat{\beta}$ القيمة المقدرة ل β

$\hat{\beta}$ δ :الانحراف المعياري لـ

- 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_0تتواجد في طرفي المنحنى، أي أن قيم Studentالمحسوبة مثلاالتي تكون أكبر من 1.96أو أقل من 1.96-، أي $\left[|t_{Cal}| > t_{n-2}^{0.05}\right]$ تعبر عن قبول الفرضيةH_1 ورفض H_0أي عندها نقول أن المعلمة معنوية أو دالة إحصائيا عند مستوى المعنوية5%.

دراسة صلاحية النموذج

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

معامل التحديد R^2يقيس ويشرح نسبة التغيرات التي تحدث في المتغير التابع Y_i والمشروحة بواسطة تغيرات المتغير المستقل Xiفهي 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

$$R^2 = \frac{SSR}{SST} = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = \frac{\hat{B}^2\sum(X_i - \bar{X})^2}{\sum(y_i - \bar{y})^2}$$

$$R^2 = 1 - \frac{SSE}{SST} = \frac{\sum\varepsilon_i^2}{\sum(y_i - \bar{y})^2} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

علما ان: $\sum(y_i - \bar{y})^2 = \sum(\hat{y}_i - \bar{y})^2 + \sum(y_i - \hat{y}_i)^2$

وتكون قيمة معامل التحديد محصورة بين 0 و1

$$0 \le R^2 \le 1$$

2. 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:

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

سابعا: اختبار المعنوية الكلية للنموذج

يمكن اختبار المعنوية الكلية لنموذج الانحدار البسيط من خلال تشكيل جدول التباين، وهو تحليل مجموع مربعات الصغرى إلي مجموع مربعات البواقي ومجموع مربعات الانحدار. الغرض من هذا التحليل لاختبار معنوية مجموع مربعات الانحدار وهذا أيضا يدخل في اختبار معنوية المعامل β. ونمثل هذا التحليل في جدول تحليل التباين:

مصدر التغير	التغير	درجات الحرية	المتوسط التربيعي
الانحدار	$SSR = \sum(\hat{y}_i - \bar{y})^2$	$K-1$	$\frac{SSR}{K-1}$
البواقي	$SSE = \sum(y_i - \hat{y}_i)^2$	$n-K$	$\frac{SSE}{n-K}$
المجموع	$SST = \sum(y_i - \bar{y})^2$	$n-1$	$F = \frac{\frac{SSR}{K-1}}{\frac{SSE}{n-k}}$

حيث : K: يمثل عدد المتغيرات

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

ويختبر هذا الإختبار الفرضيتين التاليتين:

فرضية العدم

$$H_0 : \beta = 0$$

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 ،2 $-n$)ويقيس معنوية النموذج ككل، حيث تُحسب إحصائية الاختبار كما يلي:

$$F = \frac{\frac{SSR}{SSE}}{\frac{}{n-k}}$$

وتساوي ايضا:

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

وكما رأينا سابقا فإن قاعدة اتخاذ القرار هي كما يلي:

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

$\alpha\%$

●في حالة $F_{Cal}<F_{(1,n-2)}$ 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_0 فيمكن إيجاد قيمة Y_{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 التنبؤية** $\hat{Y}_{PR}$ **هي:**

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

2. التنبؤ بمجال

ندرس التنبؤ بمجال الثقة لأن التنبؤ النقطي هو تنبؤ مبني على معادلة مقدرة، وبالتالي فمن المحتمل أنتكون 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:

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

مثال

نفس معطيات المثال السابق رقم 1

1) احسب جدول تحليل التباين ANOVA
2) أوجد R^2 وr
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)=	1−(−0.357141)=	$(1.357141)^2$	$(1-8)^2$=49

		−0.357141	1.357141	=1.84183≈1.842	
				∑=10.715	∑=228

جدول تحليل التباين

مصدر التغير	مجموع مربعات الانحرافات	درجة الحرية	متوسط مربعات الانحرافات
الانحدار (المتغيرX)	$\overline{SSR = \sum(\hat{y}_i - \bar{y})^2}$ SSR=SST-SSE= 228-10 ,715=217 .285	1	217 .285
البواقي	SSE=$\sum \hat{\varepsilon}_i^2$ =10.715	n-2=5	2،143
المجموع	$\overline{SST} = \sum(y_i - \bar{y})^2 = 228$	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

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

ط2لدينا :

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

$$n = 7$$

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

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

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

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

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

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

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

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

$$\sqrt{196} = 14$$

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

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

$$14 \times 39.94997 \approx 559.2996$$

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

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** $(\hat{\sigma}^2)$

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

- **تقدير تباين المقدرات**

تباين $\hat{\beta}$

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

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

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

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

تباين $\hat{\alpha}$

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

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

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

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

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

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

المعلمة (β)معنوية إحصائيًا عند مستوى دلالة 5 %

➢ اختبار معنوية α

الفرضيات:

$H_0: \alpha = 0$

$H_1: \alpha \neq 0$

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 $t_{5,0.025}=2.571$

$|t_{\alpha^\wedge}|=15.47>2.571 \Longrightarrow$ 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(اختبارF)

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

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

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

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

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