الإنحدار الخطي المتعدد:

مقدمة: يعتبر الإنحدار الخطي المتعدد من بين الأدوات المهمة المستعملة في التنبؤ، حيث يستعمل في إتخاذ القرار والتوقع والرقابة، فهو يهتم بتحديد العلاقة بين متغير تابع (Y) وعدة متغيرات مستقلة ($X_1, X_2 \dots X_k$)، وتفترض هذه الطريقة وجود علاقة خطية بين المتغير التابع والمتغيرات المستقلة ($Y = B_0 + B_1 X_1 + B_2 X_2 \dots B_k X_k$).

لكن في الواقع لا توجد علاقة خطية تامة بين المتغير التابع والمتغيرات المستقلة وذلك لعدة أسباب منها أخطاء القياس، ووجود متغيرات أخرى تؤثر على المتغير التابع لم نأخذها بعين الإعتبار... إلخ، لذلك نضيف لهذه العلاقة مقدار نرمز له بـ e ويسمى حد الخطأ أو معامل الإزعاج، وبالتالي تصبح علاقة (خط) الإنحدار بالشكل التالي:

$$y_i = \underbrace{b_0 + b_1 x_{i1} + b_2 x_{i2} \dots b_p x_{iK}}_{\text{جزء الإنحدار}} + \underbrace{e_i}_{\text{جزء الخطأ}}$$

حيث: ($i = 1 \dots n$) وعدد المشاهدات لكل متغير

K: وعدد المتغيرات.

$B_1, B_2 \dots B_k$ هي ثوابت أو معاملات النموذج

e_i: حد الخطأ.

ملاحظة: يجب أن نعلم أن طريقة الإنحدار الخطي المتعدد لا يمكن تطبيقها إلا على المتغيرات التي توجد بينها علاقة خطية أو تقبل التحويل إلى علاقة خطية.

شكل الإنحدار الخطي المتعدد

في الواقع المعادلة التي كتبناها سابقا هي مجرد معادلة واحدة من بين n معادلة وذلك كما يلي

$$Y_1 = B_0 + B_1x_{11} + B_2x_{12} + \ldots + B_Kx_{1K} + e_1$$
$$Y_2 = B_0 + B_1x_{21} + B_2x_{22} + \ldots B_Kx_{2K} + e_2$$
$$\ldots \quad \ldots \quad \ldots\ldots\ldots$$
$$\ldots \quad \ldots \quad \ldots\ldots\ldots$$
$$. \quad \ldots \quad \ldots\ldots\ldots$$
$$p \quad Y_n = B_0 + B_1x_{n1} + B_2x_{n2} + \ldots + B_Kx_{nK} + e_n$$

وهذه المعادلة تتضمن (K+1) من المعلومات المراد تقديرها علما أن الحد الأول منها (B_0) يمثل الحد الثابت، الأمر الذي يتطلب اللجوء إلى المصفوفات والأشعة لتقدير تلك المعلومات. وعليه يمكن صياغة هذه المعادلات في صورة مصفوفات كالتالي

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1k} \\ 1 & x_{21} & x_{22} & \ldots & x_{2k} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ i & x_{i1} & x_{i2} & \ldots & x_{ik} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ i & x_{n1} & x_{n2} & \ldots & x_{nk} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_i \\ \vdots \\ b_p \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_i \\ \vdots \\ e_n \end{bmatrix}$$

ويمكن كتابة المعادلات السابقة إختصارا

$$Y = XB + \varepsilon$$

حيث: Y = شعاع عمودي أبعاده ($n \times 1$) يحتوي مشاهدات المتغير التابع

X: مصفوفة أبعادها ($n*(K+1)$) يحتوي العمود الأول على الواحد وباقي الأعمدة قيم المتغيرات المستقلة.

$Y(n\times1)$: المتغير التابع أو المفسَّر،

$X(n\times(k+1))$: مصفوفة المتغيرات المُفَسِّرة أو المستقلة،

$\beta((k+1)\times1)$: شعاع المعالم،

$\varepsilon(n\times1)$: شعاع الأخطاء

وبما المعادلة السابقة هي معادلة حقيقية لكنها مجهولة المعالم والمراد تقديرها باستخدام الاحصاءات المتوفرة عن المتغير التابع والمتغيرات المستقلة فإنه يجب توفر مجموعة من الفرضيات حول هذا النموذج

فرضيات النموذج

$$E(\varepsilon) = E\begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \vdots\\ \varepsilon_n\end{bmatrix} = \begin{bmatrix}E(\varepsilon_1)\\ E(\varepsilon_2)\\ E(\varepsilon_3)\\ \vdots\\ E(\varepsilon_n)\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix}$$

2- تباين العناصر العشوائية ثابت، والتباين المشترك بينها يساوي صفر أي:

$$cov(\varepsilon\varepsilon) = E(\varepsilon * \varepsilon') = \begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \vdots\\ \varepsilon_n\end{bmatrix} * [\varepsilon_1\ \varepsilon_2\ \varepsilon_3\ \dots\ \varepsilon_n] = \sigma^2 * I$$

تسمى المصفوفة أعلاه بمصفوفة التباين والتباين المشترك لحد الخطأ ε، حيث تشكل العناصر القطرية في المصفوفة قيمة التباين لـ ε، بينما تشكل بقية العناصر قيم التباين المشترك لها وهي مساوية للصفر.

3- ليست هناك علاقة خطية تامة بين المتغيرات المستقلة وأن عدد المشاهدات يجب أن تزيد على عدد المعالم المراد تقديرها أي أن $Ran(x) = K+1 < n$ حيث: ran (هي رتبة المصفوفة) $K+1$ يمثل عدد المتغيرات المستقلة زائد 1، n هو حجم العينة المأخوذة (عدد المشاهدات).

4- مصفوفة المتغيرات المستقلة X محددة ومعطاة، فهي مقاسة بدون أخطاء.

5- يوجد استقلال إحصائي بين المشاهدات المستقلة ($x_1, x_2 \dots x_k$) والخطأ العشوائي ε_i، أي $cov(X, \varepsilon) = 0$

6. الأخطاء العشوائية ε_i لها توزيع طبيعي متوسطه معدوم وتباين ثابت σ^2 أي $\varepsilon_i \sim N(0, \sigma^2)$

بعدما تم تعرفنا على فرضيات النموذج سوف نتعرف على:

1 - تقدير معالم النموذج بطريقة المربعات الصغرى (MCO).

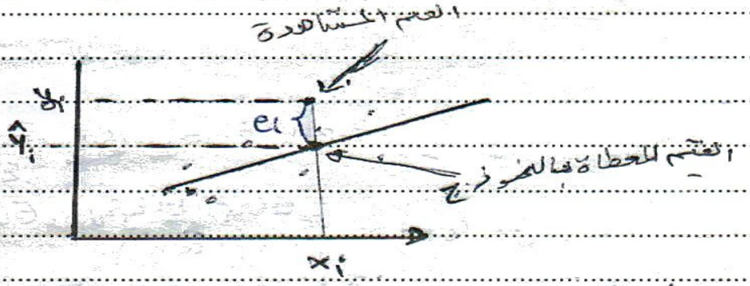
2 - إختبار معنوية المعالم B_i

3 - تقييم (إختبار) القوة التفسيرية للنموذج

4 - إختبار معنوية النموذج الخطي

5 - تقدير مجال لمعالم النموذج ② إيجاد مجال الثقة للقيمة الحقيقية Y

- تقدير معالم النموذج بطريقة المربعات الصغرى



في ضوء الفرضيات السابقة يمكن إستخدام طريقة المربعات الصغرى في تقدير معالم النموذج، حيث الهدف الأساسي لهذه الطريقة هو تصغير البواقي قدر الإمكان. أي

$$= \min \sum e_i^2 = \min \sum (Y_i - \hat{Y}_i)^2$$

ولكي تكون القيمة S في أقل قيمة لها عندما يكون

$$\frac{\partial S}{\partial B} = 0 \quad , \quad \frac{\partial^2 S}{\partial^2 B} > 0$$

لسوف نأخذ متغيرين مستقلين فقط ونعوض

عند تعويض قيمة $(\hat{Y} = \hat{B}_0 + \hat{B}_1 x_{i1} + \hat{B}_2 x_{i2})$ في القيمة S نجد

$$S = \sum (Y_i - \hat{B}_0 - \hat{B}_1 x_{i1} - \hat{B}_2 x_{i2})^2$$

$$\frac{\partial S}{\partial B_0} = 2 \sum (Y_i - \hat{B}_0 - \hat{B}_1 x_{i1} - \hat{B}_2 x_{i2})(-1) = 0$$

$$= \sum Y_i - n\hat{B}_0 - \hat{B}_1 \sum x_{i1} - \hat{B}_2 \sum x_{i2} = 0$$

$$\Rightarrow \boxed{\sum Y_i = n\hat{B}_0 + \hat{B}_1 \sum x_{i1} + \hat{B}_2 \sum x_{i2}} \quad (1)$$

$$\frac{\partial S}{\partial \hat{B}_1} = 2\sum\left(y_i - \hat{B}_0 - \hat{B}_1 x_{i1} - \hat{B}_2 x_{i2}\right)(-x_{i1}) = 0$$

$$= \sum\left(y_i x_{i1} - \hat{B}_0 x_{i1} - \hat{B}_1 x_{i1}^2 - \hat{B}_2 x_{i1} x_{i2}\right) = 0$$

$$\Rightarrow \boxed{\sum x_{i1} y_i = \sum \hat{B}_0 x_{i1} + \sum \hat{B}_1 x_{i1}^2 + \sum \hat{B}_2 x_{i1} x_{i2}} \quad - (2)$$

$$\frac{\partial S}{\partial \hat{B}_2} = 2\sum\left(y_i - \hat{B}_0 - \hat{B}_1 x_{i1} - \hat{B}_2 x_{i2}\right)(-x_{i2}) = 0$$

$$\sum\left(y_i x_{i2} - \hat{B}_0 x_{i2} - \hat{B}_1 x_{i1} x_{i2} - \hat{B}_2 x_{i2}^2\right) = 0$$

$$\Rightarrow \boxed{\sum y_i x_{i2} = \hat{B}_0 \sum x_{i2} + \hat{B}_1 \sum x_{i1} x_{i2} + \hat{B}_2 \sum x_{i2}^2} \quad - (3)$$

نلاحظ أننا تحصلنا على ثلاث معادلات بثلاثة مجاهيل ($\hat{B}_0, \hat{B}_1, \hat{B}_2$)
سوف نستخدم طريقة المحددات لحل هذه الجملة

$$\begin{cases} \hat{B}_0 n + \hat{B}_1 \sum x_{i1} + \hat{B}_2 \sum x_{i2} = \sum y_i \\ \hat{B}_0 \sum x_{i1} + \hat{B}_1 \sum x_{i1}^2 + \hat{B}_2 \sum x_{i1} x_{i2} = \sum y_i x_{i1} \\ \hat{B}_0 \sum x_{i2} + \hat{B}_1 \sum x_{i1} x_{i2} + \hat{B}_2 \sum x_{i2}^2 = \sum y_i x_{i2} \end{cases}$$

حيث (نعرف)

$$D = \begin{vmatrix} n & \sum x_{i1} & \sum x_{i2} \\ \sum x_{i1} & \sum x_{i1}^2 & \sum x_{i1} x_{i2} \\ \sum x_{i2} & \sum x_{i1} x_{i2} & \sum x_{i2}^2 \end{vmatrix}$$

معاملات $\hat{B}_0$ معاملات $\hat{B}_1$ معاملات $\hat{B}_2$

• نستعمل D_0 لحساب B_0 لذلك نعوض معاملات $\hat{B}_0$ في D بمعاملات C

$$D_0 = \begin{vmatrix} \sum Y_i & \sum x_{i1} & \sum x_{i2} \\ \sum x_{i1} Y_i & \sum x_{i1}^2 & \sum x_{i1} x_{i2} \\ \sum x_{i2} Y_i & \sum x_{i1} x_{i2} & \sum x_{i2} \end{vmatrix}$$

• نستعمل D_1 لحساب $\hat{B}_1$ لذلك نعوض معاملات $\hat{B}_1$ في D بمعاملات C

$$D_1 = \begin{vmatrix} n & \sum Y_i & \sum x_{i2} \\ \sum x_{i1} & \sum x_{i1} Y_i & \sum x_{i1} x_{i2} \\ \sum x_{i2} & \sum x_{i2} Y_i & \sum x_{i2} \end{vmatrix}$$

• نستعمل D_2 لحساب $\hat{B}_2$ لذلك نعوض معاملات $\hat{B}_2$ في D بمعاملات C،

$$D_2 = \begin{vmatrix} n & \sum x_{i1} & \sum Y_i \\ \sum x_{i1} & \sum x_{i1}^2 & \sum Y_i x_{i1} \\ \sum x_{i2} & \sum x_{i1} x_{i2} & \sum Y_i x_{i2} \end{vmatrix}$$

• ومنه يكون:

كما يمكن حساب $\hat{B}_0$ من العلاقة $\left(\hat{B}_0 = \overline{Y} - \hat{B}_1 \overline{X}_1 - \hat{B}_2 \overline{X}_2\right)$

$$\hat{B}_0 = \frac{|D_0|}{|D|} \qquad \hat{B}_1 = \frac{|D_1|}{|D|}$$

$$\hat{B}_2 = \frac{|D_2|}{|D|}$$

وبصفة عامة نجد تقدير المعالم ($B_0, B_1, \ldots, B_k$) نعطي بالعلاقة التالية

$$\boxed{\hat{B} = (X'X)^{-1} X'Y}$$

حيث

$$(X'X)^{-1} = \frac{adj(X'X)}{det(X'X)} \quad \text{مع} \quad adj(X'X) = Com(X'X)'$$

مثال: تعطى الكمية المباعة من السلعة (Y) بدلالة سعرها (X_1)، دخل المستهلك (X_2)، سعر السلعة البديلة (X_3).

وفق الجدول:

Σ 385	70	65	60	55	50	45	40	y
53	6	6	7	8	9	8	9	x_1
3200	500	500	500	500	400	400	400	x_2
91	16	15	11	13	12	14	10	x_3

حسب النظرية الاقتصادية فإنه توجد علاقة بين الكمية المباعة من السلعة (Y) والمتغيرات التفسيرية (المستقلة) وهي (سعر السلعة، الدخل، سعر السلعة البديلة) ويمكن معرفة هذا الأثر من خلال تقدير العلاقة التالية

$$\widehat{\beta} = (X'X)^{-1}X'Y$$

حيث

$$X = \begin{vmatrix} 1 & 9 & 400 & 10 \\ 1 & 8 & 400 & 14 \\ 1 & 9 & 400 & 12 \\ 1 & 8 & 500 & 13 \\ 1 & 7 & 500 & 11 \\ 1 & 6 & 500 & 15 \\ 1 & 6 & 500 & 16 \end{vmatrix} , \quad X' = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 9 & 8 & 9 & 8 & 7 & 6 & 6 \\ 400 & 400 & 400 & 500 & 500 & 500 & 500 \\ 10 & 14 & 12 & 13 & 11 & 15 & 16 \end{vmatrix}$$

$$Y = \begin{vmatrix} 40 \\ 45 \\ 50 \\ 55 \\ 60 \\ 65 \\ 70 \end{vmatrix} \qquad \hat{\beta} = \begin{vmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{vmatrix}$$

حساب $X'Y$

$$X'Y = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 9 & 8 & 9 & 8 & 7 & 6 & 6 \\ 400 & 400 & 400 & 500 & 500 & 500 & 500 \\ 10 & 14 & 12 & 13 & 11 & 15 & 16 \end{vmatrix} \begin{vmatrix} 40 \\ 45 \\ 50 \\ 55 \\ 60 \\ 65 \\ 70 \end{vmatrix} = \begin{vmatrix} 385 \\ 2840 \\ 179000 \\ 5100 \end{vmatrix}$$

$$X'X = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 9 & 8 & 9 & 8 & 7 & 6 & 6 \\ 400 & 400 & 400 & 500 & 500 & 500 & 500 \\ 10 & 14 & 12 & 13 & 11 & 15 & 16 \end{vmatrix} \begin{vmatrix} 1 & 9 & 400 & 10 \\ 1 & 8 & 400 & 14 \\ 1 & 9 & 400 & 12 \\ 1 & 8 & 500 & 13 \\ 1 & 7 & 500 & 11 \\ 1 & 6 & 500 & 15 \\ 1 & 6 & 500 & 16 \end{vmatrix} = \begin{vmatrix} 7 & 53 & 3200 \\ 53 & 411 & 23900 \\ 3200 & 23900 & 148\ldots \\ 91 & 677 & 419\ldots \end{vmatrix}$$

حساب مقلوب المصفوفة $(X'X)$

يعطى مقلوب المصفوفة بالعلاقة

$$(X'X)^{-1} = \frac{adj(X'X)}{det(X'X)}$$

$$adj(X'X) = (com(X'X))'$$

$$com(X'X) = \begin{vmatrix} 918280000 & -48340000 & -800000 & -14300000 \\ -48340000 & 6730000 & 39200 & 750000 \\ -800000 & 39200 & 896 & 7200 \\ -14300000 & 750000 & 7200 & 410000 \end{vmatrix}$$

نلاحظ $com(X'X)$ أنه متناظر وعليه

$$com(X'X) = adj(X'X)$$

$$det(X'X) = 4640000$$

ومنه

$$(X'X)^{-1} = \begin{bmatrix} 197,91 & -10,42 & -0,17 & -3,08 \\ -10,42 & 0,59 & 0,01 & -0,16 \\ -0,17 & 0,01 & 0,0002 & 0,0016 \\ -3,08 & 0,16 & 0,0016 & 0,09 \end{bmatrix}$$

وعليه

$$\begin{bmatrix} \hat{B}_0 \\ \hat{B}_1 \\ \hat{B}_2 \\ \hat{B}_3 \end{bmatrix} = (X'X)^{-1}X'Y = \begin{bmatrix} 197,91 & -10,42 & -0,17 & -3,08 \\ -10,42 & 0,59 & 0,01 & -0,16 \\ -0,17 & 0,01 & 0,0002 & 0,0016 \\ -3,08 & 0,16 & 0,0016 & 0,09 \end{bmatrix} \begin{bmatrix} 385 \\ 2846 \\ 179000 \\ 5100 \end{bmatrix} = \begin{bmatrix} 26,34 \\ -3,43 \\ 0,09 \\ 0,93 \end{bmatrix}$$

ومنه نجد

$\hat{B}_0 = 26,34$, $\hat{B}_1 = -3,43$, $\hat{B}_2 = 0,09$, $\hat{B}_3 = 0,93$

وعليه معادلة خط الانحدار

$$\hat{Y}_i = 26,34 - 3,43x_{i1} + 0,09x_{i2} + 0,93x_{i3}$$

من خلال معادلة خط الانحدار نلاحظ أن المتغير X_1 له علاقة عكسية مع المبيعات (Y) أي كلما زاد سعر السلعة قل الطلب عليها.

• معامل التحديد R^2 (التقييم الإجمالي للانحدار):

يعتبر R^2 مقياساً يحدد مساهمة المتغيرات المستقلة في تفسير التغير الحاصل في المتغير التابع، ويمكن اشتقاقه باستخدام المصفوفات بالانحرافات التالية:

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

$$SCT = SCE + SCR$$

$$R^2 = \frac{SCE}{SCT}$$

بحيث

$$\sum(Y_i - \bar{Y})^2 = \sum Y_i^2 - n\bar{Y}^2 = Y'Y - n\bar{Y}^2$$

$$SCE = \sum(\hat{Y} - \bar{Y})^2 = \hat{B}'X'Y - n\bar{Y}^2$$

$$SCR = \sum e_i^2 = e'e = Y'Y - \hat{B}'X'Y = \sum Y_i^2 - \hat{B}'X'Y$$

ومنه

$$R^2 = \frac{\hat{B}'X'Y - n\bar{Y}^2}{Y'Y - n\bar{Y}^2} = \frac{\hat{B}'X'Y - n\bar{Y}}{\sum Y_i^2 - n\bar{Y}^2}$$

$$(= 21174341 - \ldots = 614,6341$$

$$R^2 = \frac{614,6341}{\ldots} = 0,8780$$

مثال: حسب المثال السابق نجد معامل التحديد:

$$SCT = Y'Y - n\bar{Y}^2 = \sum Y_i^2 - n\bar{Y}^2$$

$$\sum Y^2 = 40^2 + 45^2 + 50^2 + 55^2 + 60^2 + 65^2 + 70^2 = 21875$$

$$n\bar{Y}^2 = n\left(\frac{\sum Y_i}{n}\right)^2 = \frac{[\sum(Y_i)]^2}{n} = (40+45+50+55+60+65+70)^2/7$$

$$= 21175 \Rightarrow SCT = 21875 - 21175 = 700$$

$$SCE = B'X'Y - n\bar{Y}^2 = B'X'Y - 21175$$

$$B'X'Y = [26,34 \quad -3,43 \quad 0,09 \quad 0,93] \begin{bmatrix} 385 \\ 2840 \\ 173000 \\ 5100 \end{bmatrix} = 21799,35$$

ومنه

$$R^2 = \frac{21799,35 - 21175}{21875 - 21175} = 0,892$$

ومنه يمكننا القول أن المتغيرات المستقلة تفسر بنسبة 89,2% من المتغيرات.

نلاحظ من خلال صيغة R^2 نلاحظ أنه كلما زاد عدد المتغيرات المستقلة سوف يزيد معامل التحديد حتى ولو لم يكن لهذا المتغير أي تأثير على Y لأن درجة الحرية سوف تقل $(n-K-1)$

لذلك يتطلب استخراج معامل التصحيح المعدل أو المصحح $\bar{R}^2$ وذلك على النحو التالي

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-K-1}$$

لذلك من أجل مقارنة عدة نماذج تختلف في عدد المتغيرات يجب مقارنة $\bar{R}^2$ وليس R^2.

- معامل التحديد المعدل

$$\bar{R}^2 = 1 - (1 - 0,892)\frac{7-1}{7-3-1} = 0,784$$

- اختبار معنوية المعالم:

يستخدم اختبار "T" لتقييم معنوية تأثير المتغيرات المستقلة $X_1, X_2 \ldots X_k$ في المتغير التابع Y. في نموذج الانحدار المتعدد يعتمد على نوعين من الفروض:

$$\begin{cases} H_0: B_j = 0 & \text{لا يوجد أثر لـ } X_j \\ H_1: B_j \neq 0 & \text{يوجد أثر لـ } X_j \end{cases}$$

الدالة المحسوبة

حيث

$$T = \frac{\hat{B}_j}{\hat{\sigma}_{B_j}}$$

$$\hat{\sigma}^2_{\hat{B}} = \hat{\sigma}^2 * (X'X)^{-1}$$

$$\hat{\sigma}^2 = \frac{Y'Y - \hat{B}'X'Y}{n-K-1} = \frac{\sum Y_i^2 - \hat{B}'X'Y}{n-K-1}$$

ونقارن هذه القيمة مع القيمة الجدولية

وحسب المثال السابق فيما إذا أردنا اختبار وجود أثر على المتغيرات المستقلة على المبيعات Y والدخل لدى مكونات؟

1) هل يوجد أثر لسعر السلعة على مبيعاتها عند مستوى معنوية 5%

$$\begin{cases} H_0: B_1 = 0 \\ H_1: B_1 \neq 0 \end{cases}$$

$$T_c = \left| \frac{\hat{B}_1}{\hat{\sigma}_{\hat{B}_1}} \right|$$

$$\hat{\sigma}^2_{\hat{B}} = \hat{\sigma}^2 * (X'X)^{-1}$$

حيث

$$\hat{\sigma}^2 = \frac{\sum Y_i^2 - B'X'Y}{7-3-1} = \frac{21875 - 21799,35}{3} = 25,21$$

$$\hat{\sigma}^2_{\hat{B}} = 25,21 \begin{vmatrix} 197,91 & -10,42 & -0,17 & -3,08 \\ -10,42 & 0,59 & 0,01 & -0,16 \\ -0,17 & 0,01 & 0,0002 & 0,0016 \\ -3,08 & 0,16 & 0,0016 & 0,09 \end{vmatrix}$$

$$\hat{\sigma}^2_{\hat{B}} = \begin{vmatrix} \boxed{4989,19} & -262,64 & -4,347 & -77,7 \\ -262,64 & \boxed{14,83} & 0,213 & 4,077 \\ -4,35 & 0,21 & \boxed{0,005} & 0,039 \\ -77,69 & 4,07 & 0,039 & \boxed{2,228} \end{vmatrix}$$

$$\hat{\sigma}^2_{B_0} = 4989,19 \Rightarrow \hat{\sigma}_{B_0} = 70,63$$ ومنه

$$\hat{\sigma}^2_{B_1} = 14,83 \Rightarrow \hat{\sigma}_{\hat{B}_1} = 3,85$$

$$\hat{\sigma}^2_{B_2} = 0,005 \Rightarrow \hat{\sigma}_{B_2} = 0,07$$

$$\hat{\sigma}^2_{B_3} = 2,228 \Rightarrow \hat{\sigma}_{B_3} = 1,5$$

$T_c = \left| \frac{-3.427}{3,85} \right| = 0,89$ ومنه

والمجدولة $T_{1-\frac{\alpha}{2}, n-k-1} = T_{0,975, 3} = 3,182$

نلاحظ أن $T_c < T_t$ ومنه نرفض H_1 ونقبل H_0 أي لا يوجد أثر لسعر السلعة على مبيعات هذه السلعة عند مستوى معنوية 5%.

أما بالنسبة للمتغير الثاني (الدخل):

$T_c = \left| \frac{0,09}{0,007} \right| = 1,285$

أما القيمة المجدولة

$T_t = 3,182$

نلاحظ أن $T_c < T_t$ أي نقبل H_0 ونرفض H_1

أي لا يوجد أثر عند مستوى معنوية ($\alpha = 0,05$).

هل يوجد أثر لسعر السلعة البديلة (X_3) أثر على المبيعات عند مستوى معنوية 5%.

$$\begin{cases} H_0: \hat{B}_3 = 0 \\ H_1: \hat{B}_3 \neq 0 \end{cases}$$

$T_c = \left| \frac{0,93}{1,5} \right| = 0,62$

لما نقارن T_c و T_t نجد أن $T_c < T_t$ ومنه نرفض H_1 ونقبل H_0 أي لا يوجد أثر لسعر السلعة البديلة على المبيعات عند مستوى معنوية 5%.

اختبار مدى ملائمة النموذج:

4- اختبار مدى ملائمة العلاقة الخطية لتمثيل العلاقة بين المتغيرات المستقلة والمتغير التابع - جدول تحليل التباين

يهدف اختبار F لمعرفة مدى معنوية العلاقة الخطية بين المتغيرات المستقلة ($x_1, x_2 \dots x_k$) على المتغير التابع Y، وهو يعتمد على الفرضيتين

$$\begin{cases} H_0: B_1 = B_2 = B_3 \dots = B_k \\ H_1: \exists\, B_j \neq 0 \quad j = 1 \dots k \end{cases}$$

جدول تحليل التباين

مصدر الانحراف	مجموع مربعات الأخطاء	درجة الحرية	متوسط مربعات الأخطاء	الاختبار F
الانحدار (مفسرة)	$\hat{B}'x'y - n\bar{y}^2$ (SCE)	K	$\dfrac{\hat{B}'x'y - n\bar{y}^2}{K}$	$F = \dfrac{\hat{B}'x'y - n\bar{y}^2 / K}{y'y - \hat{B}'x'y / n-k-1}$
البواقي	$y'y - \hat{B}'x'y$ [SCR]	$n-k-1$	$\dfrac{y'y - \hat{B}'x'y}{n-k-1} = \hat{\sigma}^2$	
الانحراف الكلي	$y'y - n\bar{y}^2$ (SCT)	$n-1$		

بعد حساب F نقوم بمقارنتها بالقيمة المجدولة ($F_{1-\alpha, k, n-k-1}$)
فإذا كان $F_c > F_{1-\alpha, k, n-k-1}$ فإن النموذج الخطي عند ذلك ملائم لتمثيل هذه العلاقة أما إذا كان العكس ($F_c < F_t$) فإن النموذج الخطي لا يلائم لتمثيل العلاقة وذلك عند مستوى دلالة α

$$F = \frac{R^2/K}{(1-R^2)/n-k-1}$$

ملاحظة: يمكن حساب F بالعلاقة التالية

حسب المثال: $F_c = \dfrac{0{,}892/3}{(1-0{,}892)/3} = 8{,}26$ $F_{0.95,3,3} = \mathbf{9.276}$

* مجال الثقة لمتوسط القيمة الحقيقية Y_h عند مستوى معنوية α

* يعطى مجال الثقة لأي قيمة حقيقية Y_h عند القيمة X_h عند مستوى معنوية α بالعلاقة التالية

$$E(Y_h) = \hat{Y}_h \mp t_{1-\frac{\alpha}{2}, n-k-2} \cdot \hat{\sigma}(\hat{Y}_h)$$

حيث:

$$\hat{\sigma}^2(Y_0) = \hat{\sigma}^2 X_h (X'X)^{-1} X_h'$$

ففي المثال السابق إذا أردنا إيجاد متوسط المبيعات عندما يكون سعر السلعة (10) ، الدخل (400) ، سعر السلعة البديلة (13 = X_h) عند 5%

نلاحظ أن $X_h = (1 \;\; 10 \;\; 400 \;\; 13)$

ومنه

$\hat{\sigma}^2(Y_h) = 25,21 * 1,9 = 47,9 \Rightarrow \hat{\sigma}(Y_0) = 6,92$

ومنه

$T_{1-\frac{\alpha}{2}, n-k-1} = 3,182$

ومنه $E(Y_h) \in \{40,13 \pm \underbrace{3,182 * 6,92}_{= 22,02}\}$

$E(Y_h) \in [18,11 \quad 62,15]$

نحن واثقون بنسبة 95٪ أن متوسط المبيعات سيكون بين 18,11 و 62,15 ...

مجال الثقة للمعاملات:

نعطي مجال الثقة لأي معامل $\hat{\beta}_j$ بمستوى دلالة α بالعلاقة التالية

$$\hat{\beta}_j \in \{\hat{\beta}_j \pm T_{1-\frac{\alpha}{2}, n-k-1} * \hat{\sigma}_{\hat{\beta}_j}\}$$

مجال الثقة لـ $\hat{\beta}_1$

$\hat{\beta}_1 \in \{\hat{\beta}_1 \pm T_{0,975, 13} * \hat{\sigma}_{\hat{\beta}_1}\}$

$\hat{\sigma}_{\hat{\beta}_1} = 3,85 \ , \ T_{0,975 \, 13} = 3,182 \ , \ \hat{\beta}_1 = -3,43$

ومنه $\beta_1 \in \{-3,43 \pm \underbrace{3,182 * 3,85}_{12,25}\}$

$\hat{\beta}_1 \in [-15,68 \quad 8,82]$