نموذج الانحدار الخطي المتعدد

في الواقع الاقتصادي لا يمكن الاستعانة بالنموذج ذي متغيرين لتحليل الظاهرة الاقتصادية حيث أن هذه الأخيرة لا تفسر فقط بحد أو متغير واحد وإنما ينبغي إدماج جميع المحددات أو العوامل المؤثرة في الظاهرة لكي تكون الدراسة أكثر شمولية

1] الصياغة الرياضية للنموذج الخطي العام

يفترض النموذج الانحدار الخطي المتعدد وجود علاقة خطية بين متغير تابع Y_t وعدد من المتغيرات المستقلة X_{kt} على النحو التالي:

$$Y_t = B_0 + B_1 X_{1t} + B_2 X_{2t} + \dots + B_k X_{tk} + \varepsilon_t$$
$$t = 1, \dots, n$$

Y_t: متغير تابع في الفترة t

X_{1t}: المتغير المستقل الأول في الفترة t

X_{kt}: المتغير المستقل k في الفترة t

معلمات النموذج: $B_0, B_1, B_2, \dots B_k$ لدينا هنا $(K+1)$ معلمة

ε_t: حد الخطأ العشوائي النموذج

- من المعادلة السابقة يمكن كتابة ما يلي:

$$\begin{cases} Y_1 = B_0 + B_1 X_{11} + B_2 X_{12} + \dots + B_k X_{1k} + \varepsilon_1 \\ Y_2 = B_0 + B_1 X_{21} + B_2 X_{22} + \dots + B_k X_{2k} + \varepsilon_2 \\ \vdots \\ Y_n = B_0 + B_1 X_{n1} + B_2 X_{n2} + \dots + B_k X_{nk} + \varepsilon_n \end{cases}$$

ومنه يمكن كتابة النظام التالي على الشكل المصفوفي:

$$\underset{(n,1)}{Y} = \underset{(n,k+1)}{X}\ \underset{(k+1,1)}{B} + \underset{(n,1)}{\varepsilon}$$

ومنه يكون الشكل المصفوفي كما يلي:

$$\underset{(n,1)}{Y=\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}} \quad \underset{(n,k+1)}{X=\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & & & & \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}} \; , \; \underset{(k+1,1)}{B=\begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ \vdots \\ B_k \end{bmatrix}}$$

دلالة على وجود الحد الثابت

$$\underset{(n,1)}{\varepsilon=\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}}$$

2] الفرضيات الأساسية للنموذج:

إن بناء نموذج الانحدار الخطي المتعدد يجب أن يكون مستوفيا لعدد من الفرضيات التي يمكن إجمالها كما يلي:

1] الفرضية الأولى: المتغيرات المفسرة (المهملة) في النموذج لها متوسط معدوم

2] الفرضية الثانية: تجانس التباين لمختلف الحدود العشوائية
$$Var(\varepsilon_t) = \sigma^2$$

3] الفرضية الثالثة: الأخطاء العشوائية ليست مرتبطة ببعضها
$$Cov(\varepsilon_i, \varepsilon_j) = 0$$

يمكن كتابة الفرضية 2 و 3 على الشكل التالي:

$$\Omega_\varepsilon = E(\varepsilon\varepsilon') = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & \cdots & 0 \\ 0 & \sigma_\varepsilon^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_\varepsilon^2 \end{bmatrix} = \sigma_\varepsilon^2 I_n$$

تسمى المصفوفة Ω_ε مصفوفة التباين والتباين المشترك للأخطاء العشوائية

4] الفرضية الرابعة: المصفوفة X غير عشوائية وثابتة وتعني بأن قيم المتغيرات المستقلة يمكن مراقبتها وبالإضافة إلى ذلك نفترض X ثابتة لضمان بأن قيم المتغيرات المستقلة

لا تتغير من حين لآخر ويمكن القول بشكل عام
بأن الأخطاء العشوائية مستقلة عن المتغيرات
المستقلة أي $Cov(x_t, \varepsilon_t) = E(X'\varepsilon) = 0$

5) الفرضية الخامسة: التوزيع الطبيعي للأخطاء
ونكتب $\varepsilon_t \sim N(0, \sigma^2_{\varepsilon})$

6) الفرضية السادسة: عدم وجود ارتباط ذاتي بين
المتغيرات المستقلة أي أن محدد المصفوفة $X'X$
موجود ويمكن حسابه ⇐ يوجد معكوس المصفوفة $(X'X)$

7) الفرضية السابعة: عدد المشاهدات n أكبر
من عدد المتغيرات المستقلة k

3) تقدير معلمات النموذج الخطي المتعدد

لدينا في النموذج $Y = X\beta + \varepsilon$ ، المجاهيل
هي β و ε حيث المصفوفة X والشعاع Y
هي معطيات النموذج ، ومنه بنفس طريقة
نموذج الانحدار الخطي البسيط ونفس
المنهجية أي منهجية المربعات الصغرى العادية
نسعى الى تدنية مجموع مربعات البواقي $\sum \varepsilon_t^2$
كالتالي:

$$Y_t = X_t B + \varepsilon_t \rightarrow \varepsilon_t = Y_t - X_t B$$

ومنه تدنية مجموع مربعات البواقي يكون كالتالي

$$\min \sum \varepsilon_t^2 = \min \varepsilon'\varepsilon = \min (Y_t - X_t B)'(Y_t - X_t B)$$
$$= \min S$$

حيث ε' منقول الشعاع ε

لدينا

$$S = (Y - X_t B)'(Y_t - X_t B) = (Y' - X'B')(Y - XB)$$
$$= Y'Y - Y'XB - X'B'Y + X'B'XB$$

ولدينا $Y'XB = X'B'Y$

ومنه يصبح لدينا:

~~$S = Y'Y - 2BX'Y + B'X'X\beta$~~

$$S = Y'Y - 2\beta'X'Y + \beta'X'X\beta$$

=

ومنه من أجل $\min S$ نقوم بحساب المشتقة
الجزئية المساوية لـ 0 أي

$$\frac{\partial S}{\partial \beta} = 0 \Rightarrow 0 - 2X'Y + 2(X'X)\beta = 0$$

أو

$$S = Y'Y - 2\beta'X'Y + \beta'X'X\beta$$
$$= Y'Y - 2B'X'Y + B'B(X'X)$$
$$= Y'Y - 2B'X'Y + B^2(X'X)$$
$$\frac{\partial S}{\partial B} = 0 \Rightarrow 0 - 2X'Y + 2B(X'X) = 0$$

$$\frac{\partial S}{\partial \beta} = 0 \Rightarrow 2(X'X)B - 2X'Y = 0$$

ومنه

$$(X'X)B - X'Y = 0$$
$$(X'X)B = X'Y \quad \text{--- (1)}$$

ولدينا رتبة المصفوفة X هي $K+1$ فإن المصفوفة $(X'X)$
مصفوفة مربعة من الشكل $(K+1, K+1)$ وهذا يعني
وجود المعكوس أي $(X'X)^{-1}$ موجود (حسب الفرضية 6)
بضرب طرفي المعادلة (1) في $(X'X)^{-1}$ نجد

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

ومنه

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

ملاحظة:

1] الشكل المصفوفي للعلاقة $\hat{\beta}=(X'X)^{-1}X'Y$

يكتب على الشكل التالي:

عدد المشاهدات

$$|X'X| = \begin{bmatrix} n & \sum x_{1t} & \sum x_{2t} & \cdots & \sum x_{kt} \\ \sum x_{1t} & \sum x_{1t}^2 & \sum x_{1t}x_{2t} & \cdots & \sum x_{1t}x_{kt} \\ \sum x_{2t} & \sum x_{2t}x_{1t} & \sum x_{2t}^2 & \cdots & \sum x_{2t}x_{kt} \\ \vdots & \vdots & & & \vdots \\ \sum x_{kt} & \sum x_{kt}x_{1t} & \sum x_{kt}x_{2t} & \cdots & \sum x_{kt}^2 \end{bmatrix}$$

$$|X'Y| = \begin{bmatrix} \sum y_t \\ \sum x_{1t}\cdot y_t \\ \sum x_{2t}\, y_t \\ \vdots \\ \sum x_{kt}\, y_t \end{bmatrix}, \quad \hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

2] في حالة المعطيات الممركزة: $x_t = x_t - \bar{x}$

يتم حساب المقدرات كما يلي:

$$= \begin{bmatrix} Var(x_1) & Cov(x_1,x_2) & Cov(x_1,x_3) & \cdots & Cov(x_1,x_k) \\ Cov(x_2,x_1) & Var(x_2) & Cov(x_2,x_3) & \cdots & Cov(x_2,x_k) \\ Cov(x_3,x_1) & Cov(x_3,x_2) & Var(x_3) & \cdots & Cov(x_3,x_k) \\ \vdots & \vdots & \vdots & & \\ Cov(x_k,x_1) & Cov(x_k,x_2) & Cov(x_k,x_3) & \cdots & Var(x_k) \end{bmatrix}^{-1} \times \begin{bmatrix} Cov(x_1,y) \\ Cov(x_2,y) \\ Cov(x_3,y) \\ \vdots \\ Cov(x_k,y) \end{bmatrix} \Rightarrow (X'X)^{-1}(X'Y)$$

ويتم حساب الحد الثابت كما يلي:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}_1 - \hat{\beta}_2\bar{X}_2 - \cdots - \hat{\beta}_k\bar{X}_k$$

3] حالة تقدير نموذج انحدار خطي متعدد ذو متغيرين مستقلين:

لدينا المعادلة المقدرة التالية:

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}_1 x_{1t} + \hat{\beta}_2 x_{2t}$$

ومنه يمكن حساب المقدرات كما يلي:

$$\hat{\beta}_1 = \frac{\sum x_{1t}y_t \cdot \sum x_{2t}^2 - \sum x_{2t}y_t \sum x_{1t}x_{2t}}{\sum x_{1t}^2 \sum x_{2t}^2 - \left(\sum x_{1t}x_{2t}\right)^2}$$

$$\hat{\beta}_2 = \frac{\sum x_{2t}y_t \sum x_{1t}^2 - \sum x_{1t}y_t \sum x_{1t}x_{2t}}{\sum x_{1t}^2 \sum x_{2t}^2 - \left(\sum x_{1t}x_{2t}\right)^2}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}_1\bar{X}_1 - \hat{\beta}_2\bar{X}_2$$

4] تقدير تباين الأخطاء وتباين المقدرات

1] تباين الأخطاء:

$$\hat{\sigma}^2_{\varepsilon_t} = \frac{\varepsilon'\varepsilon}{n-k-1} = \frac{\sum \varepsilon^2}{n-k-1}$$

حيث $\varepsilon = Y - \hat{Y}$ ، k: عدد المتغيرات المستقلة

2] تباين المقدرات:

$$\hat{\Omega}_{\hat{\beta}} = \hat{\sigma}^2_{\varepsilon_t}(X'X)^{-1} = \frac{\sum \varepsilon_t^2}{n-k-1}(X'X)^{-1}$$

حيث $\Omega_{\hat{\beta}}$: مصفوفة التباين والتباين المشترك للمقدرات

مثال 1: لدينا البيانات المبينة في الجدول التالي:

t	Y_t	X_{1t}	X_{2t}
1	4	2	3
2	6	4	2
3	7	5	10
4	9	7	8
5	10	9	8
6	12	10	9
7	14	12	11
8	16	14	13
9	18	15	14
10	20	17	15

- نفرض نموذج الانحدار الخطي المتعدد التالي:

$$y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t$$

المطلوب:

- أحسب المقدرات $\hat{\beta}_0 , \hat{\beta}_1 , \hat{\beta}_2$
- أحسب مصفوفة التباين والتباين المشترك للمقدرات

الحل:

- كتابة البيانات السابقة على الشكل المصفوفي

لدينا $n=10$ مشاهدات و متغيران مستقلان $K=2$ ومنه يمكن كتابة النموذج كما يلي:

$$Y = X\beta + \varepsilon$$

حيث

$$Y=\begin{bmatrix}4\\6\\7\\\vdots\\20\end{bmatrix},\ X=\begin{bmatrix}1&2&3\\1&4&7\\1&5&10\\\vdots&\vdots&\vdots\\1&17&15\end{bmatrix},\ \beta=\begin{pmatrix}\beta_0\\\beta_1\\\beta_2\end{pmatrix},\ \varepsilon=\begin{bmatrix}\varepsilon_1\\\varepsilon_2\\\varepsilon_3\\\vdots\\\varepsilon_{10}\end{bmatrix}$$

عمود الحد الثابت في النموذج

* لدينا قانون حساب المقدرات

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

حساب $(X'X)$

$$(X'X)=\begin{bmatrix}1&1&1&\cdots&1\\2&4&5&\cdots&17\\3&7&10&\cdots&15\end{bmatrix}\begin{bmatrix}1&2&3\\1&4&7\\1&5&10\\\vdots&\vdots&\vdots\\1&17&15\end{bmatrix}=\begin{bmatrix}10&95&98\\95&1129&1081\\98&1081&1078\end{bmatrix}$$

ومنه

$$(X'X)^{-1}=\begin{pmatrix}1.172&0.085&-0.192\\0.085&0.028&-0.036\\-0.192&-0.036&0.054\end{pmatrix}$$

ولدينا أيضا:

$$(X'Y)=\begin{bmatrix}1&1&1&\cdots&1\\2&4&5&\cdots&17\\3&7&10&\cdots&15\end{bmatrix}\begin{bmatrix}4\\6\\7\\\vdots\\20\end{bmatrix}=\begin{bmatrix}116\\1342\\1298\end{bmatrix}$$

ومنه

$$\hat{\beta}=\begin{pmatrix}\hat{\beta}_0\\\hat{\beta}_1\\\hat{\beta}_2\end{pmatrix}=(X'X)^{-1}X'Y=\begin{bmatrix}1.172&0.085&-0.192\\0.085&0.028&-0.036\\-0.192&-0.036&0.054\end{bmatrix}\begin{bmatrix}116\\1342\\1298\end{bmatrix}$$

$$\Rightarrow \begin{pmatrix}\hat{\beta}_0\\\hat{\beta}_1\\\hat{\beta}_2\end{pmatrix}=\begin{pmatrix}1.099\\0.977\\0.123\end{pmatrix}$$

ومنه يكتب النموذج المقدر كما يلي:

$$\boxed{\hat{y}_t = 1.099 + 0.977X_{1t} + 0.123X_{2t}}$$

2) حساب مصفوفة التباين والتباين المشترك للمقدرات:

* نحسب أولا تباين البواقي أو الأخطاء $\hat{\sigma}^2_{\varepsilon_t}$ باستعمال الجدول التالي، حيث

$$\hat{\sigma}^2_{\varepsilon_t}=\frac{\sum_{t=1}^{10}\varepsilon_t^2}{n-K-1}$$

t	Y_t	$\hat{Y}_t$	$\varepsilon_t = Y_t - \hat{Y}_t$	ε_t^2
1	4	3.42	0.57	0.32
2	6	5.87	0.12	0.01
3	7	7.22	-0.22	0.05
4	9	8.93	0.06	0.004
5	10	10.88	-0.88	0.78
6	12	11.98	0.01	0.0001
7	14	14.19	-0.19	0.037
8	16	16.39	-0.39	0.15
9	18	17.49	0.50	0.25
10	20	19.57	0.42	0.17
Σ	/	/	0	1.815

$$\hat{\sigma}^2_{\varepsilon_t}=\frac{\varepsilon'\varepsilon}{n-K-1}=\frac{\sum\varepsilon_t^2}{n-K-1}=\frac{1.815}{n-2-1}=\frac{1.815}{10-3}=0.2593$$

* نحسب الآن $\hat{\Omega}_{\hat{\beta}}$: مصفوفة التباين والتباين المشترك للمقدرات حيث

$$\hat{\Omega}_{\hat{\beta}}=\hat{\sigma}^2_{\varepsilon_t}(X'X)^{-1}$$

ومنه

$$\hat{\Omega}_{\hat{\beta}} = 0.2593 \times \begin{bmatrix} 1.172 & 0.085 & -0.192 \\ 0.085 & 0.028 & -0.036 \\ -0.192 & -0.036 & 0.054 \end{bmatrix}$$

$$= \begin{bmatrix} 0.3040 & 0.022 & -0.049 \\ 0.022 & 0.0073 & -0.0093 \\ -0.049 & -0.0093 & 0.014 \end{bmatrix}$$

ومنه

$$\hat{\sigma}^2_{\hat{\beta}_0} = 0.3040 \Rightarrow \hat{\sigma}_{\hat{\beta}_0} = 0.551$$

$$\hat{\sigma}^2_{\hat{\beta}_1} = 0.0073 \Rightarrow \hat{\sigma}_{\hat{\beta}_1} = 0.085$$

$$\hat{\sigma}^2_{\hat{\beta}_2} = 0.014 \Rightarrow \hat{\sigma}_{\hat{\beta}_2} = 0.119$$

5} اختبار فرضيات النموذج الخطي المتعدد

1} اختبار المعنوية الإحصائية للمعلمات (اختبار T)

$$\begin{cases} H_0: \beta_j = 0 \text{ غ معنوي} & j = 0 \dots k \\ H_1: \beta_j \neq 0 \text{ معنوي} \end{cases}$$

$$t^* = t_c = \frac{|\hat{\beta}_j - \beta_j|}{\hat{\sigma}_{\hat{\beta}_j}} \sim t^{\alpha}_{n-k-1}$$

إذا كانت:

* T_c أكبر من t^{α}_{n-k-1} الجدولية نرفض الفرضية الصفرية (المتغير x_k يؤثر على المتغير y)

* T_c أقل من t^{α}_{n-k-1} الجدولية نقبل الفرضية الصفرية (المتغير x_k لا يؤثر على المتغير y)

2} مجال الثقة

1} مجال الثقة للمعالم:

$$\beta_j \in \left[\hat{\beta}_j \pm \hat{\sigma}_{\hat{\beta}_j} \times t^{\alpha/2}_{n-k-1}\right]$$

2} مجال الثقة لتباين الأخطاء:

$$\hat{\sigma}^2_{\varepsilon_t} \in \left[\frac{(n-k-1)\hat{\sigma}^2_{\varepsilon_t}}{\chi^2_{n-k-1, \frac{\alpha}{2}}} ; \frac{(n-k-1)\hat{\sigma}^2_{\varepsilon_t}}{\chi^2_{n-k-1, 1-\frac{\alpha}{2}}}\right]$$

3} اختبار المعنوية الكلية للنموذج المقدر: اختبار فيشر F

أما في حالة الانحدار الخطي البسيط يمكن أيضا اختبار المعنوية الكلية لنموذج الانحدار الخطي المتعدد وفق الفرضيات التالية:

$$\begin{cases} H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0 \\ H_1: \beta_1 \neq \beta_2 \neq \dots \neq \beta_k \neq 0 \end{cases}$$

الفرضية البديلة هنا نقصد بها أنه توجد على الأقل معلمة واحدة تختلف معنويا عن الصفر

يتم حساب F_c المحسوبة كما يلي:

$$F^* = F_c = \frac{ESS/k}{RSS/(n-k-1)} = \frac{\sum(\hat{y}-\bar{y})^2/k}{\sum e_t^2/(n-k-1)}$$

$$= \frac{R^2/k}{(1-R^2)/(n-k-1)} = \frac{\hat{\beta}'X'Y - n\bar{Y}^2/k}{\varepsilon'\varepsilon/(n-k-1)}$$

إذا كانت:

* $F^* < F^{\alpha}_{k,n-k-1}$ عند مستوى α احتمال: نقبل الفرضية الصفرية ونقول أن النموذج ككل غير معنوي

* $F^* > F^{\alpha}_{k,n-k-1}$ عند مستوى α% احتمال: نرفض الفرضية الصفرية ونقول أن النموذج ككل معنوي

4} حساب معامل التحديد R^2 ومعامل التحديد المصحح $\bar{R}^2$ (أو المرجح أو المعدل)

1} معامل التحديد R^2: يقيس لنا جودة النموذج المقدر وفق القانون التالي:

$$R^2 = \frac{\sum(\hat{y}_t-\bar{y})^2}{\sum(y_t-\bar{y})^2} = 1 - \frac{\sum e_t^2}{\sum(y_t-\bar{y})^2}$$

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

* اذا كان النموذج لا يحتوي على الحد الثابت أو كانت البيانات ممركزة فإن:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = 1 - \frac{\varepsilon'\varepsilon}{Y'Y} = \frac{\hat{Y}'\hat{Y}}{Y'Y}$$

* توجد علاقة بين R^2 وشعاع المقدرات كما يلي

$$R^2 = \frac{\hat{Y}'\hat{Y}}{Y'Y} = \frac{\hat{\beta}'X'X\hat{\beta}}{Y'Y}$$

5) معامل التحديد المصحح $\bar{R}^2$:

في حالة إضافة متغير مستقل للنموذج فإن قيمة معامل التحديد R^2 سترتفع حتى ولو لم تكن هناك أهمية لهذا المتغير المستقل في النموذج، لأن إضافة هذا الأخير سوف تؤدي إلى زيادة مجموع المربعات المشروحة ESS مع ثبات مجموع المربعات الكلية TSS، ولهذا تم حساب معامل التحديد المصحح $\bar{R}^2$ الذي يأخذ في الحسبان النقصان الحاصل في درجات الحرية وقيمته عادة أقل من قيمة معامل التحديد R^2، ويحسب كما يلي:

$$\bar{R}^2 = 1 - \frac{RSS/(n-k-1)}{TSS/(n-1)} = 1 - (1-R^2)\left(\frac{n-k-1}{n-1}\right)$$

K: عدد المتغيرات المستقلة.

- ملاحظة:

إذا كان: $K>1$: $R^2 \geq \bar{R}^2$

نموذج خطي بسيط $K=1$: $R^2 = \bar{R}^2$

يمكن أن يأخذ قيما سالبة $\bar{R}^2$

6) معادلة وجدول تحليل التباين

1) معادلة تحليل التباين: تكتب على الشكل التالي:

$$\sum (y_t - \bar{y})^2 = \sum (\hat{y}_t - \bar{y})^2 + \sum e_t^2$$

$$TSS = ESS + RSS$$

2) جدول تحليل التباين ANOVA:

مصدر التغير	مجموع مربعات الانحرافات	درجات الحرية	متوسط مربعات الانحرافات
المتغيرات المستقلة X	$ESS = \sum (\hat{y}_t - \bar{y})^2$	K	ESS/K
البواقي	$RSS = \sum e_t^2$	$n-K-1$	$RSS/(n-K-1)$
المجموع	$TSS = \sum (y_t - \bar{y})^2$	$n-1$	$F^* = \frac{ESS/K}{RSS/(n-K-1)}$

مثال:

بالاستعانة بمعطيات المثال السابق:

1- اختبر المعنوية الإحصائية للمعلمات عند 5%

2- انشئ مجال ثقة للمعلمات عند مستوى 95%

4- اختبر المعنوية الكلية للنموذج

3- احسب معامل التحديد R^2 المشترك

أحسب معامل التحديد المصحح $\bar{R}^2$

الحل: لدينا $t_7^{0.05} = 2.365$ ، $F_{2,7}^{0.05} = 4.74$

* اختبار معنوية B_0

$$\begin{cases} H_0: B_0 = 0 & \text{غير معنوي} \\ H_1: B_0 \neq 0 & \text{معنوي} \end{cases}$$

$$t^* = \frac{|\hat{B}_0 - B_0|}{\sigma_{\hat{B}_0}} = \frac{|1.099 - 0|}{0.551} = 1.99$$

ومنه $t^* = 1.99 < t_7^{0.05} = 2.365 \Rightarrow$ نقبل الفرضية الصفرية

أي أن B_0 لا يختلف معنويا عن الصفر عند مستوى معنوية 5% أو B_0 ليس لها معنوية إحصائية

* اختبار معنوية B_1

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

نحسب t^*

$$t^* = t_c = \frac{|\hat{B}_1 - B_1|}{\hat{\sigma}_{\hat{B}_1}} = \frac{|0.977 - 0|}{0.085} = 11.39$$

ومنه $t^* = 11.39 > t_7^{0.05} = 2.365$ ⇐ نرفض الفرضية الصفرية أي أن B_1 يختلف معنويا عن الصفر عند المستوى 5% أي أن X_{1t} يؤثر معنويا على Y_t

* اختبار معنوية B_2

$$\begin{cases} H_0: B_2 = 0 \\ H_1: B_2 \neq 0 \end{cases}$$

نحسب t^*

$$t^* = t_c = \frac{|\hat{B}_2 - B_2|}{\hat{\sigma}_{\hat{B}_2}} = \frac{|0.123 - 0|}{0.119} = 1.0394$$

ومنه: $t^* = t_c = 1.0394 < t_7^{0.05} = 2.365$ ⇐ نقبل الفرضية الصفرية H_0: أي أن B_2 لا يختلف معنويا عن 0 عند 5% أي يساوي معنويا الصفر عند $a = 5\%$ وهذا يدل على أن X_{2t} لا يؤثر على Y

* مجالات الثقة:

$$B_0 \in [\hat{B}_0 \pm \hat{\sigma}_{\hat{B}_0} \cdot t_7^{0.05}]$$

$$B_0 \in [1.099 - 0.551 \times 2.365 ; 1.099 + 0.551 \times 2.365]$$

$$B_0 \in [-0.2042 ; 2.4034]$$

$$B_1 \in [\hat{B}_1 \pm \hat{\sigma}_{\hat{B}_1} \cdot t_7^{0.05}]$$

$$B_1 \in [0.977 - 0.085 \times 2.365 ; 0.977 + 0.085 \times 2.365]$$

$$B_1 \in [0.774 ; 1.1805]$$

$$B_2 \in [\hat{B}_2 \pm \hat{\sigma}_{\hat{B}_2} \, t_7^{0.05}]$$

$$B_2 \in [0.123 - 0.119 \times 2.365 ; 0.123 + 0.119 \times 2.365]$$

$$B_2 \in [-0.1577 ; 0.4051]$$

~~* اختبار المعنوية الكلية للنموذج اختبار F~~

~~$H_0: B_0 = B_1 = B_2 = 0$~~

~~$H_1: B_0 \neq B_1 \neq B_2 \neq 0$~~

~~لدينا~~

~~نحسب F^*~~

~~$F^* =$~~

* حساب معامل التحديد R^2 و $\bar{R}^2$

لدينا $\sum e_t^2 = 1.815$ ، نقوم بحساب

• $TSS = \sum (y_t - \bar{y})^2 = 256.4$

باستعمال الجدول المساعد نجد TSS أو ESS

ومنه $R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{1.815}{256.4} = 0.9929$

و $\bar{R}^2 = 1 - (1 - R^2)\left(\frac{n-k-1}{n-1}\right) = 1 - (1 - 0.9929) \times \left(\frac{10-2-1}{10-1}\right)$

$\Rightarrow \bar{R}^2 = 0.9908$

نلاحظ أن R^2 أكبر نسبيا من $\bar{R}^2$ ومن خلال النتائج أيضا نجد أن للنموذج قدرة تفسيرية عالية جدا، أي أن المتغيرات المستقلة تشرح المتغير التابع بنسبة 99.29 %

* اختبار المعنوية الكلية للنموذج: اختبار F^*
لدينا الفرضيات التالية

$$\begin{cases} H_0: B_0 = B_1 = B_2 = 0 \\ H_1: B_0 \neq B_1 \neq B_2 \neq 0 \end{cases}$$

غ معنوي
معنوي

نحسب F^*

$$F^* = \frac{R^2 / K}{(1-R^2)/(n-K-1)} = \frac{0,9929/2}{(1-0,9929)/10-2-1}$$

$$F^* = 490,72$$

ومنه لدينا $F^* = 490,72 > F^{0,05}_{2,7} = 4,74$

ومنه ~~[illegible]~~ نرفض الفرضية الصفرية
ونقبل الفرضية البديلة أي أن النموذج
معنوية إحصائية عند المستوى 5%
أي يوجد على الأقل متغير مستقل تختلف قيمته عن 0

* جدول تحليل التباين

مصدر التغير	مجموع مربعات الانحرافات	درجات الحرية	متوسط مربعات الانحرافات
المتغيرات المستقلة X	$ESS = 254,58$	$K = 2$	$254,58/2 = 127,29$
البواقي ε	$RSS = 1,815$	$n-K-1 = 7$	$1,815/7 = 0,259$
المجموع	$TSS = 256,4$	$n-1 = 9$	$F^* = \frac{127,29}{1,815} = 491,$

7] التنبؤ ضمن نموذج الانحدار الخطي المتعدد

في حالة كانت المتغيرات المستقلة معلومة: فإن
في حالة توفر معلومات عن المشاهدة $t+h$ يمكننا
التنبؤ بالقيم المستقبلية للمتغير التابع عند نفس
الملاحظة y_{t+h} على النحو التالي:

$$y_{t+h} = B_0 + B_1 x_{1t+h} + B_2 x_{2t+h} + \dots + B_k x_{kt+h} + \varepsilon_{t+h}$$

$$\hat{y}_{t+h} = \hat{B}_0 + \hat{B}_1 x_{1t+h} + \hat{B}_2 x_{2t+h} + \dots + \hat{\beta}_k x_{kt+h}$$

$h = 1, 2, \dots, H$

أو نكتب

$$\hat{y}_t(h) = X_{t+h} \hat{\beta}$$

ومنه متوسط خطأ التنبؤ يكتب كما يلي:

$$E(\hat{y}_t(h)) = X_{t+h} \cdot \beta$$

* تباين ~~خطأ~~ خطأ التنبؤ يكتب على الشكل التالي:

$$Var(\hat{y}_{t+h}) = \hat{\sigma}^2_\varepsilon \cdot X_{t+h} (X'X)^{-1} X'_{t+h}$$

* تباين شعاع أخطاء التنبؤ ~~(الانحراف المعياري)~~
~~[illegible]~~

$$\hat{\sigma}^2_{\hat{\varepsilon}_{t+h}} = \hat{\sigma}^2_\varepsilon \left(X'_{t+h} (X'X)^{-1} X_{t+h} + 1 \right)$$

حيث الانحراف المعياري لخطأ التنبؤ هو

$$\hat{\sigma}_{\varepsilon_{t+h}} = \sqrt{\hat{\sigma}^2_{\varepsilon_{t+h}}}$$

* مجال الثقة للقيمة المتنبؤ بها يكتب كما يلي

$$Y_{t+h} \in \left[\hat{y}_{t+h} \pm t^{\alpha}_{n-k-1} \cdot \hat{\sigma}_{\varepsilon_{t+h}} \right]$$

8] اختبارات أخرى في النموذج الخطي المتعدد

A] اختبار إضافة متغيرات أخرى

وهو اختبار يهدف إلى معرفة ما مدى تعزيز
القدرة التفسيرية في النموذج قيد الدراسة
وذلك في حالة إضافة متغير أو أكثر
مستقلين، على سبيل المثال نريد

دراسة إضافة متغيرين إضافيين لنموذج بمتغير مستقل واحد، ونختبر هل إضافة هذين المتغيرين تعزز القدرة التفسيرية للنموذج قيد الدراسة أم لا، ومنه لدينا مراحل إجراء الاختبار على النحو التالي:

1) **المرحلة الأولى:** تتمثل في تقدير النموذج الكلي بثلاثة متغيرات واستخراج جدول تحليل التباين للنموذج المقدر

2) **المرحلة الثانية:** تتمثل في تقدير النموذج ذو المتغير الواحد قبل إضافة المتغيرات الأخرى واستخراج جدول تحليل التباين للنموذج المقدر

3) **المرحلة الثالثة:** نقوم باختبار الفرضيتين التاليتين:

$$\begin{cases} H_0: B_1 = B_2 = \dots = B_k = 0 \\ H_1: B_1 \neq B_2 \neq \dots \neq B_k \neq 0 \end{cases}$$

* نلاحظ أن الفرضيتين تتعلقان بالمعالم دون المعلمة B_0 (معلمة الحد الثابت) التي تعتبر أساسية في النموذج، والاختبار يعمد إلى دراسة هل إضافة هذه المعلمات سيزيد من القدرة التفسيرية للنموذج أم لا

* يتم حساب إحصائية فيشر المحسوبة لهذا الاختبار وفق ما يلي:

$$F^* = \frac{(ESS - ESS_c)/(k-k')}{RSS/(n-k-1)} \quad \} \; ESS - ESS_1$$

$$F^* = \frac{(RSS_c - RSS)/(k-k')}{RSS/(n-k-1)}$$

حيث تمثل RSS و ESS مجموع مربعات البواقي ومجموع مربعات الانحرافات المفسرة للنموذج الكلي الذي يحتوي على كل المعالم (النموذج غير المقيد)

بينما تمثل RSS_c و ESS_c مجموع مربعات البواقي ومجموع مربعات الانحرافات المفسرة للنموذج بدون إضافة المتغيرات الإضافية (النموذج المقيد)

k: عدد المتغيرات المستقلة في النموذج الكلي (نموذج غير مقيد)

k': عدد المتغيرات المستقلة في النموذج بدون إضافة المتغيرات الإضافية

n: عدد المشاهدات

ومنه إذا كانت:

$F^* \leq F^{\alpha}_{k-k', n-k-1}$: نقبل الفرضية العدمية ونقول أن إضافة متغيرات إضافية للنموذج لا تزيد من قدرته التفسيرية عند مستوى احتمال $\alpha\%$

$F^* > F^{\alpha}_{k-k', n-k-1}$: نرفض الفرضية العدمية ونقول أن إضافة متغيرات إضافية للنموذج يحسن من قدرته التفسيرية عند مستوى احتمال $\alpha\%$

2) اختبار استقرار معاملات النموذج (اختبار Chow)

يهتم هذا الاختبار بدراسة ثبات معلمات النموذج في كامل الفترة الزمنية، بعبارة أخرى هل معلمات النموذج على طول الفترة تختلف عن معلمات النماذج المقدرة وفق الفترات الجزئية ضمن الفترة الكلية ويتم هذا الاختبار وفق ثلاث مراحل كالتالي:

1) **المرحلة الأولى:** تقدير النموذج الأساسي على طول الفترة T واستخراج جدول تحليل التباين للنموذج المقدر

2) **المرحلة الثانية:** تقدير النماذج الفرعية

كل طول الفترات البينية الى فترتين جزئيتين من 0 إلى $\frac{t}{2}$ ثم من $\frac{t}{2}$ إلى t مع استخراج جدول التباين لكل نموذج مقدر وتحليل

3) المرحلة الثالثة: اختبار الفرضية التالية:

$$H_0: \begin{cases} \beta_0 = \beta_0^1 = \beta_0^2 \\ \beta_1 = \beta_1^1 = \beta_1^2 \\ \beta_2 = \beta_2^{(1)} = \beta_2^{(2)} \\ \vdots \quad \vdots \quad \vdots \\ \beta_k = \beta_k^{(1)} = \beta_k^{(2)} \end{cases}$$

حيث نقوم بحساب إحصائية فيشر المحسوبة وفق العلاقة التالية

$$F^* = \frac{\left[RSS - \left(RSS^{(1)} + RSS^{(2)}\right)\right] / df_{(1)}}{\left(RSS^{(1)} + RSS^{(2)}\right) / df_{(2)}}$$

حيث RSS: مجموع مربعات البواقي للنموذج الكلي على كامل الفترة

$RSS^{(1)}, RSS^{(2)}$: مجموع مربعات البواقي للنموذجين الضمنيين على الفترات البينية

$df_{(1)}, df_{(2)}$: عدد درجات الحرية وتحسب كما يلي:

$$df_{(1)} = (n-k-1) - \left[(n_1-k-1) + (n_2-k-1)\right] = k+1$$

$$df_{(2)} = (n_1-k-1) + (n_2-k-1) = n - 2(k+1)$$

$$df_{(2)} = ddl_d \ , \ df_{(1)} = ddl_n$$

ومنه اذا كانت:

$F^* \leq F^{a}_{k+1,\, n-2(k+1)}$: نقبل الفرضية الصفرية ونقول أن المعلمات ثابتة عبر الزمن عند مستوى احتمال $a\%$

$F^* > F^{a}_{k+1,\, n-2(k+1)}$: نرفض الفرضية الصفرية ونقول أن المعلمات غير ثابتة على طول الفترة حيث أنها تتغير عبر الفترات الضمنية عند مستوى احتمال $a\%$

مثال 3

+1) بالرجوع للمثال السابق، تنبأ بالمتغير y للفترة 11 ثم قم ببناء فترات ثقة بنسبة معنوية 5%، اذا علمت أن القيمة المستقبلية لكل من x_1 و x_2 على التوالي هي 18 و 16

+2) في هذه الحالة نعرف أن النموذج المراد دراسته هو $y_t = \beta_0 + \beta_1 x_{1t} + \varepsilon_t$، هل إضافة المتغير x_{2t} سيحسن معنويا جودة التقدير؟ بعبارة أخرى أي النموذج نختار؟

حيث نجد مايلي:

$$\hat{y}_t = 1.53 + 1.05 x_t$$
$$(4.21) \quad (31.15)$$

→ نموذج مقدر بمتغيرات نموذج مقيد

$$R^2 = 0.9918 \qquad RSS_c = 2.095$$

الحل:

* التنبؤ للفترة 11 = h ومنه

$$\hat{y}_t = 1.099 + 0.977 x_{1t} + 0.123 x_{2t}$$

$$\hat{y}_{11} = 1.099 + 0.977 x_{1,11} + 0.123 x_{2,11}$$

$$= 1.099 + 0.977(18) + 0.123(16)$$

$$\hat{y}_{11} = 20.6756$$

* حساب الانحراف المعياري لخطأ التنبؤ:

$$\hat{\sigma}^2_{\hat{\varepsilon}_{t+h}} = \hat{\sigma}^2_{\varepsilon}\left(X'_{t+h}(X'X)^{-1}X_{t+h} + 1\right)$$

وجود الثابت في النموذج

$$\hat{\sigma}^2_{\hat{\varepsilon}_{11}} = 0.2593\left[\begin{pmatrix} 1 & 18 & 16 \end{pmatrix}\begin{bmatrix} 1.172 & 0.085 & -0.192 \\ 0.085 & 0.028 & -0.036 \\ -0.192 & -0.036 & 0.054 \end{bmatrix}\begin{bmatrix} 1 \\ 18 \\ 16 \end{bmatrix} + 1\right]$$

ومنه تباين خطأ التنبؤ:

$$\hat{\sigma}^2_{\hat{\varepsilon}_{11}} = 0.2593(0.4461+1) = 0.3749$$

ومنه الانحراف المعياري لخطأ التنبؤ:

$$\hat{\sigma}_{\hat{\varepsilon}_{11}} = \sqrt{0.3749} = 0.6123$$

* مجال الثقة للقيمة المتنبأ بها

$$Y_{t+1} \in [\hat{Y}_{t+1} \pm t_{n-k-1} \cdot \hat{\sigma}_{\varepsilon_{t+1}}]$$

بالتطبيق العددي نجد

$$Y_{11} \in [\hat{Y}_{11} \pm t^{5\%}_{n-3} \cdot \hat{\sigma}_{\varepsilon_{11}}]$$

$$Y_{11} \in [20{,}6756 - 2{,}365(0{,}6123) ; 20{,}6756 + 2{,}365(0{,}6123)]$$

$$Y_{11} \in [19{,}2275 ; 22{,}1236]$$

R/ لدينا النموذج الكلي (غير مقيد) يحتوي على متغيرين مستقلين $\Leftarrow K=2$

النموذج ذو المتغير واحد (نموذج مقيد) يحتوي على متغير مستقل واحد $K'=1$

(نموذج مقيد بدون إضافة متغيرات مستقلة)

* بما أننا ندرس معنوية إضافة X_{2t} فإن طبيعة الاختبار تكون كما يلي:

$$\begin{cases} H_0: B_2 = 0 \\ H_1: B_2 \neq 0 \end{cases}$$

ومنه نحسب F^* حيث

$$F^* = \frac{(RSS_c - RSS)/(K-K')}{RSS/(n-K-1)}$$

$RSS = \sum e_t^2 = 1.815$ (الموجود في تحليل التباين) محسوبة سابقا في النموذج الكلي أو الأصلي

$$TSS_c = \frac{RSS_c}{1-R^2} = \frac{2{,}095}{1-0.9918} = 255{,}48$$

$$ESS_c = TSS_c - RSS_c = 255{,}48 - 2{,}095$$

$$\Rightarrow ESS_c = 253{,}39$$

ومنه

$$F^* = \frac{(2{,}095 - 1{,}815)(2-1)}{1{,}815/(10-2-1)} = 1{,}0802 \quad \leftarrow F^{a}_{k-k', n-k-1}$$

ومنه لدينا $F^* = 1{,}080 < F^{0.05}_{1,7} = 5{,}59$ ∴ نقبل الفرضية الصفرية ونقول أن إضافة المتغير المستقل X_{2t} لا يحسن بصفة معنوية القدرة التفسيرية للنموذج (وذلك يتطابق مع دراسة المعنوية B_2 في السابق)

مثال حول اختبار Chow

نعتبر النموذج المقدر التالي على عينة حجمها $n=14$

$$\hat{Y}_t = 32{,}89 + 0{,}8X_{1t} - 0{,}38X_{2t} - 0{,}03X_{3t}$$
$$(11{,}66) \quad (0{,}24) \quad (0{,}15) \quad (0{,}05)$$

$R^2 = 0{,}702$, $n = 14$, $RSS = 67{,}48$

القيم بين قوسين هي قيم الأخطاء المعيارية

- تم تقدير النموذج السابق على فترتين زمنيتين (عينتين جزئيتين من العينة الأم)، وكانت النتائج كما يلي:

1- الفترة الجزئية الأولى:

$$\hat{Y}_t = 25{,}27 + 0{,}774X_{1t} - 0{,}293X_{2t} - 0{,}012X_{3t}$$
$$(0{,}53) \quad (0{,}31) \quad (0{,}10)$$

$R^2 = 0.692$, $n_1 = 7$, $TSS^1 = 88{,}85$

$ESS^1 = 61{,}54$, $RSS^1 = 27{,}31$

- الفترة الجزئية الثانية:

$$\hat{Y}_t = 62{,}63 + 1{,}228X_{1t} - 0{,}62X_{2t} - 0{,}184X_{3t}$$
$$(0{,}69) \quad (0{,}52) \quad (0{,}15)$$

$R^2 = 0.543$, $n_2 = 7$, $TSS^2 = 45{,}43$

$ESS^2 = 24{,}70$, $RSS^2 = 20{,}73$

* هل معلمات النموذج مستقرة على كامل الفترة الزمنية عند مستوى احتمال 5%

الحل:

نضع الفرضية التالية

$$H_0 : \begin{cases} B_0 = B_0^{(1)} = B_0^{(2)} \\ B_1 = B_1^{(1)} = B_1^{(2)} \\ B_2 = B_2^{(1)} = B_2^{(2)} \\ B_3 = B_3^{(1)} = B_3^{(2)} \\ \cancel{B_4 = B_4^{(1)} = B_4^{(2)}} \end{cases}$$

يمكن وضع

$$H_1 : \begin{cases} B_0 \neq B_0^1 \neq B_0^2 \\ B_1 \neq B_1^1 \neq B_1^2 \\ B_2 \neq B_2^1 \neq B_2^2 \\ B_3 \neq B_3^1 \neq B_3^2 \\ \cancel{B_4 \neq B_4^1 \neq B_4^2} \end{cases}$$

نقوم بحساب F^* حيث:

$$F^* = \frac{[RSS - (RSS^1 + RSS^2)]/K+1}{(RSS^1 + RSS^2)/[n - 2(K+1)]}$$

$$= \frac{\{67.45 - (27.31 + 20.73)\}/3+1}{(27.31 + 20.73)/14 - 2(3+1)}$$

$$= \frac{4.852}{8.00} = 0.60$$

ولدينا $F^{0.05}_{4,6} = 4.53$

ومنه $F^* = 0.60 < F^{0.05}_{4,6} = 4.53$: نقبل الفرضية الصفرية ونقول أن معاملات النموذج مستقرة معنويا على كامل الفترة الزمنية بنسبة خطأ 5%