

المحور 03: الانحدار الخطي المتعدد

المحاضرة 03:

قبل الدخول في تفاصيل تطبيق النموذج الخطي المتعدد على البرنامج الإحصائي EViews، نذكر ببعض المفاهيم والنتائج المتعلقة بهذا النموذج والتي تعتبر مهمة في استيعاب الطالب لكيفية بناء هذا النموذج، تقدير معلماته وتفسير مخرجاته على البرنامج.

أولاً: تقديم نموذج الانحدار الخطي المتعدد

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

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \dots + \beta_k X_{kt} + \varepsilon_t$$

ويمكن كتابته على شكل المصفوفات كالتالي:

$$\begin{aligned} Y_1 &= \beta_1 + \beta_2 X_{21} + \beta_3 X_{31} + \dots + \beta_k X_{k1} + \varepsilon_1 \\ Y_2 &= \beta_1 + \beta_2 X_{22} + \beta_3 X_{32} + \dots + \beta_k X_{k2} + \varepsilon_2 \\ &\vdots \quad \vdots \quad \vdots \quad \quad \quad \vdots \quad \vdots \\ &\vdots \quad \vdots \quad \vdots \quad \quad \quad \vdots \quad \vdots \\ Y_n &= \beta_1 + \beta_2 X_{2n} + \beta_3 X_{3n} + \dots + \beta_k X_{kn} + \varepsilon_n \end{aligned}$$

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & X_{21} & X_{31} & \dots & X_{k1} \\ 1 & X_{22} & X_{32} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & X_{3n} & \dots & X_{kn} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

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

حيث: Y: شعاع مشاهدات المتغير التابع (n×1).

X: مصفوفة مشاهدات المتغيرات المستقلة (n×k).

β: شعاع المعلمات (k×1).

ε: شعاع المتغير العشوائي (n×1).

2- الفرضيات الاحتمالية التي يقوم عليها النموذج الخطي المتعدد:

$$E(\varepsilon) = 0 = \begin{pmatrix} E(\varepsilon_1) \\ E(\varepsilon_2) \\ \vdots \\ \vdots \\ E(\varepsilon_n) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} \quad -1$$

$$\Omega_\varepsilon = E(\varepsilon\varepsilon') = \delta_\varepsilon^2 I_n \quad -2$$

$$\text{Cov}(X, \varepsilon) = 0 \quad -3$$

$$\text{تؤول إلى مصفوفة منتهية وغير أحادية.} \left(\frac{X'X}{n} \right) \quad -4$$

5- أشعة المصفوفة X مستقلة، هذا ما يسمح بالتخلص من مشكل التعدد الخطي (MULTICOLINEARITY) وحساب $(X'X)^{-1}$.

ثانيا: مقدرات معالم النموذج الخطي المتعدد:

مقدرات النموذج الخطي المتعدد (بدون برهان) بطريقة OLS هي:

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

ثالثا: نتائج وخواص مقدرات طريقة OLS

1- خاصية عدم التحيز (UNBIASED):

β مقدر غير متحيز لـ $\hat{\beta}$ حيث:

$$E(\hat{\beta}) = \beta$$

2- مقدرات OLS أفضل مقدرات خطية غير متحيزة 'BLUE' BEST LINEAR UNBIASED ESTIMATORS:

مقدرات طريقة المربعات الصغرى العادية هي أفضل مقدرات خطية غير متحيزة، حيث أن لها أصغر تباين ممكن

مقارنة بباقي المقدرات الخطية وغير المتحيزة الأخرى:

$$\Omega_{\hat{\beta}} = V(\hat{\beta}) = \delta_\varepsilon^2 (X'X)^{-1}$$

وبذلك خاصية أقل تباين ممكن تكون كمايلي:

$$\Omega_{\hat{\beta}} = \frac{\delta_{\varepsilon}^2}{n} \left(\frac{X'X}{n} \right)^{-1} \Rightarrow \lim_{n \rightarrow +\infty} \Omega_{\hat{\beta}} = 0$$

3- خاصية الاتساق (CONSISTENT):

نقول عن مقدر $\hat{\theta}$ بأنه مقدر متسق (CONSISTENT ESTIMATOR)، إذا حقق الشرطين التاليين:

$$\text{i/ } E(\hat{\theta}) = \theta \quad \text{ii/ } \lim_{n \rightarrow +\infty} V(\hat{\theta}) = 0$$

مما سبق نجد أن:

المقدر $\hat{\beta}$ يحقق الشرطين:

$$\text{i/ } E(\hat{\beta}) = \beta \quad \text{ii/ } \lim_{n \rightarrow +\infty} V(\hat{\beta}) = 0$$

إذن نستنتج أن المقدر $\hat{\beta}$ هو مقدر متسق لشعاع المعلمات β

4- مقدر تباين المتغير العشوائي ε_i :

يعطى مقدر أو تقدير تباين المتغير العشوائي (بدون برهان) وهو مقدر غير متحيز، كما يلي:

$$\hat{\delta}_{\varepsilon}^2 = e'e / (n - k)$$

5- بناء مجالات ثقة لمعلمات النموذج:

تعطى مجالات الثقة لمعلمات النموذج الخطي المتعدد بالصيغة التالية:

$$P\left(\beta_i \in \left[\hat{\beta}_i - St_{n-k}^{\frac{\alpha}{2}} \cdot \sqrt{\hat{\delta}_{\beta_i}^2} \quad \hat{\beta}_i + St_{n-k}^{\frac{\alpha}{2}} \cdot \sqrt{\hat{\delta}_{\beta_i}^2} \right] \right) = 1 - \alpha$$

رابعا: تقييم النموذج المقدر للنموذج الخطي المتعدد

يتم تقييم النموذج المقدر للنموذج الخطي المتعدد بنفس المراحل السابقة في تقييم النموذج الخطي البسيط:

1- المعايير الاقتصادية:

تتعلق بحجم وإشارة المعلمات المقدرة، لأن النظرية الاقتصادية تضع قيودا مسبقة على حجم وإشارة المعلمات، فإذا

ما جاءت هذه المعلمات على عكس ما تقرره النظرية مسبقا فإن هذا يمكن أن يكون مبررا كافيا لرفض هذه المعلمات.

2- المعايير الإحصائية:

تتمثل هذه المعايير فيما يلي:

1-2- تحليل التباين ومعامل التحديد:

يمكن صياغة معادلة تحليل التباين على الشكل التالي:

$$(Y - \bar{Y})'(Y - \bar{Y}) = (\hat{Y} - \bar{Y})'(\hat{Y} - \bar{Y}) + e'e$$

$$\text{TSS} = \text{ESS} + \text{RSS}$$

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

حيث:

$\sum(Y_t - \bar{Y})^2 = (Y - \bar{Y})'(Y - \bar{Y})$: مجموع مربعات الانحرافات الكلية ((TOTAL SUM OF SQUARES (TSS)).

$\sum(\hat{Y}_t - \bar{Y})^2 = (\hat{Y} - \bar{Y})'(\hat{Y} - \bar{Y})$: مجموع مربعات الانحرافات المفسرة ((EXPLAINED SUM OF SQUARES (ESS)).

$\sum e_t^2 = e'e$: مجموع مربعات البواقي ((RESIDUAL SUM OF SQUARES (RSS)).

أما جدول تحليل التباين (ANALYSIS OF VARIANCE (ANOVA) فيأخذ الشكل التالي:

متوسط المربعات	درجة الحرية	مجموع المربعات	مصدر التغير
$\sum(\hat{Y}_t - \bar{Y})^2 / k$	k	$\text{ESS} = \sum(\hat{Y}_t - \bar{Y})^2$	المتغيرات المستقلة
$\sum e_t^2 / n - k$	n - k	$\text{RSS} = \sum e_t^2$	البواقي e_t
	n - 1	$\text{TSS} = \sum(Y_t - \bar{Y})^2$	المجموع

في حالة النموذج الخطي المتعدد يمكن قياس القدرة التفسيرية للنموذج وجودة توفيقه من خلال معامل التحديد المتعدد R^2 ، حيث يشير هذا المعامل إلى النسبة التي يمكن تفسيرها من التغير الكلي في المتغير التابع بدلالة المتغيرات التفسيرية المدرجة في نموذج الانحدار المتعدد، ويمكن حسابه انطلاقا من معادلة تحليل التباين التي تعطى بالشكل التالي:

$$\begin{aligned}
R^2 &= \frac{(\hat{Y} - \bar{Y})'(\hat{Y} - \bar{Y})}{(Y - \bar{Y})'(Y - \bar{Y})} = \frac{\sum(\hat{Y}_t - \bar{Y})^2}{\sum(Y_t - \bar{Y})^2} \\
&= \frac{\hat{\beta}'X'X\hat{\beta} - \bar{Y}^2}{\sum(Y_t - \bar{Y})^2} \\
&= 1 - \frac{(e - \bar{e})'(e - \bar{e})}{(Y - \bar{Y})'(Y - \bar{Y})} = 1 - \frac{\sum e_{t2}}{\sum(Y_t - \bar{Y})^2}
\end{aligned}$$

وبما أنّ قيمة معامل التحديد تتأثر بعدد المتغيرات المستقلة المدرجة في النموذج، فإنه يمكن أن نصحح قيمة معامل التحديد عن طريق أخذ درجات الحرية في الحسبان عند حسابه، حيث أن درجة الحرية (n-k) تقل مع زيادة عدد المتغيرات التفسيرية وثبات حجم العينة.

وتصبح قيمة معامل التحديد المعدل \bar{R}^2 كما يلي:

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

تتراوح قيمة معامل التحديد بين الصفر والواحد، فإذا كان يساوي الواحد فهذا يعني أن القدرة التفسيرية جيدة، وأن جودة التوفيق عند حدها الأقصى، أما إذا كان يساوي الصفر فهذا يعني أن القدرة التفسيرية للنموذج منعدمة، وأن جودة التوفيق عند حدها الأدنى.

2-2- اختبارات المعنوية:

1-2-2- اختبار STUDENT: يستعمل هذا الاختبار لدراسة المعنوية الجزئية لمعاملات النموذج عند مستوى معنوية معين. لاختبار العلاقة الموجودة بين المتغير التابع Y_t والمتغير المستقل X_{it} (معنوية كل معلمة على حدى)، نقوم بإجراء الاختبار التالي:

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

مما سبق لدينا:

$$\frac{\hat{\beta}_i - \beta_i}{\sqrt{\hat{\delta}_{\beta_i}^2}} \rightarrow St(n-k)$$

تحت ظل الفرضية $H_0: \beta_i = 0$ نجد أن $\frac{\hat{\beta}_i - 0}{\sqrt{\hat{\delta}_{\beta_i}^2}}$ تتبع أيضا توزيع STUDENT بدرجة حرية تساوي $(n - k)$ ، حيث يقوم هذا

الاختبار على مقارنة إحصائية STUDENT المحسوبة $St_{cal} = \left| \frac{\hat{\beta}_i}{\sqrt{\hat{\delta}_{\beta_i}^2}} \right|$ مع الإحصائية المجدولة من جدول STUDENT عند درجة

حرية $(n - k)$ ومستوى معنوية $\frac{\alpha}{2}$ ، أي $St_{tab} = St_{n-k}^{\frac{\alpha}{2}}$. (في حالة $(n - k) > 30$ فإن $St_{tab} = 1.96$).

أما قرار الاختبار فيكون كما يلي:

- نرفض الفرضية H_0 إذا كانت $St_{cal} \geq St_{tab}$ ، ومنه $\beta_i \neq 0$ ، وبالتالي وجود علاقة ذات دلالة إحصائية بين المتغير التابع Y_i والمتغير المستقل X_{it} .

- نقبل الفرضية H_0 إذا كانت $St_{cal} < St_{tab}$ ، ومنه $\beta_i = 0$ ، وبالتالي عدم وجود علاقة ذات دلالة إحصائية بين المتغير التابع Y_i والمتغير المستقل X_{it} .

2-2-2- اختبار FISHER: يوضح لنا هذا الاختبار المعنوية الكلية للنموذج بصورة عامة، و يأخذ الشكل التالي:

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

نقوم بحساب إحصائية FISHER التي تعطى بالعلاقة التالية:

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

الإحصائية F_{cal} تتبع توزيع FISHER بدرجة حرية $v_1 = k - 1$ و $v_2 = n - k$ ، أي $F_{tab} = F_{(k-1, n-k)}^{\alpha=5\%}$.

ويكون قرار الاختبار كما يلي:

- نرفض الفرضية H_0 إذا كانت $F_{cal} \geq F_{(k-1, n-k)}^{\alpha=5\%}$ ، ومنه: $\exists \beta_i \neq 0$ ، وبالتالي فالنموذج ككل له معنوية إحصائية.

- نرفض الفرضية H_1 إذا كانت $F_{cal} < F_{(k-1, n-k)}^{\alpha=5\%}$ ، ومنه: $\beta_i = 0 \quad \forall i = 1, \dots, k$ ، وبالتالي فالنموذج ككل ليس له معنوية

إحصائية.

3-2- اختبار WALD (فيشر للقيود المتعددة):

يستعمل اختبار STUDENT لاختبار فرضية من قيد واحد، أما في حالة القيود المتعددة فالواجب تطبيق اختبار فيشر . لتكن عندما تكون فرضية العدم والفرضية البديلة في شكل مصوفات والتي تضع قيودا على مجموعة من المعلمات، فإننا نستخدم اختبار WALD الذي يعتمد على إحصائية فيشر. ويقوم هذا الاختبار على اختبار الفرضية التالية:

$$\begin{cases} H_0 : R\beta = r \\ H_1 : R\beta \neq r \end{cases}$$

حيث:

R: مصفوفة بعدها $(q \times k)$ ، β : شعاع المعلمات $(k \times 1)$.

r: شعاع بعده $(q \times 1)$ ، ويمثل عدد القيود، وهو أيضا عدد أسطر المصفوفة R .

أما إحصائية الاختبار فهي:

$$F_{cal} = \frac{(R\hat{\beta} - r)'(R'(X'X)^{-1}R)^{-1}(R\hat{\beta} - r) / q}{e'e / (n - k)} \rightarrow F_{(q, n-k)}^{\alpha}$$

القرار ويكون:

- نرفض الفرضية H_0 إذا كانت $F_{cal} \geq F_{(q, n-k)}^{\alpha=5\%}$ ، ومنه: $R\beta \neq r$.

- نرفض الفرضية H_1 إذا كانت $F_{cal} < F_{(q, n-k)}^{\alpha=5\%}$ ، ومنه: $R\beta = r$.

2-4- إختبارات التغير الهيكلي (أو إختبارات استقرارية معلمات النموذج عبر الزمن)

عند استخدام نموذج انحدار على بيانات سلاسل زمنية، يمكن أن يحدث تغير هيكلي في العلاقة بين المتغير التابع والمتغيرات المفسرة، ويقصد بالتغير الهيكلي في هذه الحالة أن قيمة معلمات النموذج لا تبقى كما هي خلال كل الفترة الزمنية، فقد يحدث التغير الهيكلي نتيجة لقوة خارجية، أو نتيجة لتغير السياسات الاقتصادية، كالتحول من نظام اقتصادي إلى آخر، أو أي أسباب أخرى.

من أهم هذه الإختبارات نذكر:

2-4-1- اختبار "CHOW FORECAST TEST":

يسمح هذا الاختبار بمعرفة إذا ما كانت معلمات النموذج تتغير مع الزمن أم لا، ولتطبيق هذا الاختبار يجب معرفة وتحديد زمن التغير في حالة بيانات السلاسل الزمنية، أو معرفة المفردة التي حصل عندها التغير في حالة البيانات المقطعية. وبالتالي فهذا الاختبار يسمح بمعرفة إذا كانت المعلمات المقدرة قبل التغير هي نفسها بعد التغير. ويمر هذا الاختبار بالمراحل التالية:

- المرحلة الأولى:

لـ تقسيم المشاهدات إلى عينتين، العينة الأولى طولها n_1 مشاهدة، والعينة الثانية طولها n_2 مشاهدة، حيث: $n_1 + n_2 = n$.
لـ تقدير نموذجين لكل عينة بطريقة المربعات الصغرى العادية:

$$Y_t = \beta_1^{(1)} + \beta_2^{(1)} X_{2t} + \beta_3^{(1)} X_{3t} + \dots + \beta_k^{(1)} X_{kt} + \varepsilon_t \quad / t = 1 \dots n_1$$

$$Y_t = \beta_1^{(2)} + \beta_2^{(2)} X_{2t} + \beta_3^{(2)} X_{3t} + \dots + \beta_k^{(2)} X_{kt} + \varepsilon_t \quad / t = n_1 + 1 \dots n_2$$

✓ حساب مجموع مربعات بواقي التقدير للنموذجين السابقين، أي حساب RSS_1 و RSS_2 .

✓ تقدير النموذج على طول الفترة الزمنية الكلية والمقدرة بـ n مشاهدة:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \dots + \beta_k X_{kt} + \varepsilon_t \quad / t = 1 \dots n$$

✓ حساب مجموع مربعات بواقي التقدير للنموذج السابق، أي حساب RSS .

- المرحلة الثانية: نقوم باختبار الفرضيات التالية:

$$\left\{ \begin{array}{l} H_0 : \begin{cases} \beta_1 = \beta_1^{(1)} = \beta_1^{(2)} \\ \beta_2 = \beta_2^{(1)} = \beta_2^{(2)} \\ \quad \quad \quad M \quad M \quad M \\ \beta_k = \beta_k^{(1)} = \beta_k^{(2)} \end{cases} \\ H_1 : \exists i / \beta_i \neq \beta_i^{(1)} \neq \beta_i^{(2)} \end{array} \right.$$

مقارنة إحصائية FISHER المحسوبة (برنامج EVIEWS يقوم بحسابها أوتوماتيكيا)، والتي تعطى بالعلاقة التالية:

$$F_{cal} = \frac{[RSS - (RSS_1 + RSS_2)] / df_1}{(RSS_1 + RSS_2) / df_2}$$

حيث:

$$df_1 = (n - k) - [(n_1 - k) + (n_2 - k)] = k$$

$$df_2 = (n_1 - k) + (n_2 - k) = n - 2k$$

نقوم بمقارنة الإحصائية المحسوبة مع إحصائية FISHER المجدولة بدرجة حرية $v_1 = k$ و $v_2 = n - 2k$ أي $F_{tab} = F_{(k, n-2k)}^{\alpha=5\%}$.

القرار:

- نرفض الفرضية H_0 إذا كانت $F_{cal} \geq F_{(k, n-2k)}^{\alpha=5\%}$ ، ومنه: $\exists i / \beta_i \neq \beta_i^{(1)} \neq \beta_i^{(2)}$ ، وبالتالي فالنموذج غير مستقر، أي هناك تغير هيكل.

2-4-2- اختبارات الاستقرار المعتمدة على البواقي المتكررة:

يفترض اختبار التغير الهيكلي لـ CHOW أن نقطة التغير معلومة، بالمقابل فإن الاختبارات المعتمدة على البواقي المتكررة فهي تسمح بتحديد هل هناك تغير هيكل أم لا من جهة، كما تسمح بتحديد نقطة التغير الهيكلي من جهة أخرى. ومن أهمها نجد: اختبار البواقي المتكررة RECURSIVE RESIDUALS:

تعتمد فكرة البواقي المتكررة على التنبؤ بالقيم المستقبلية للمتغير التابع Y_t لما نستعمل $r-1$ مشاهدة فقط، ثم حساب بواقي التقدير، أي:

$$e_r = Y_r - [\hat{\beta}_1^{r-1} + \hat{\beta}_2^{r-1} X_{2r} + \hat{\beta}_3^{r-1} X_{3r} + L + \hat{\beta}_k^{r-1} X_{kr}]$$

حيث: $\hat{\beta}_i^{r-1}$ هي مقدرات طريقة المربعات الصغرى العادية للنموذج من خلال عينة مشاهدات حجمها $r-1$.

بداية التقدير تكون من $r = k+1$.

$$V(e_r) = \hat{\delta}_{\varepsilon_{r-1}}^2 [1 + X_r' (X_{r-1}' X_{r-1})^{-1} X_r]$$

البواقي المتكررة w_r تعطى بالعلاقة التالية:

$$w_r = \frac{e_r}{\sqrt{1 + X_r' (X_{r-1}' X_{r-1})^{-1} X_r}}$$

تحت فرضية الاستقرار البواقي المتكررة تتبع التوزيع الطبيعي، أي: $w_r \rightarrow N(0, \delta_{w_r}^2)$

ويكون قرار الاختبار كمايلي:

$$\text{si } w_r \in \left[-2\sqrt{\hat{\delta}_{\varepsilon_{r-1}}^2 [1 + X_r' (X_{r-1}' X_{r-1})^{-1} X_r]} \cdot + 2\sqrt{\hat{\delta}_{\varepsilon_{r-1}}^2 [1 + X_r' (X_{r-1}' X_{r-1})^{-1} X_r]} \right] \quad \forall r = k+1 \text{L } n$$

فيكون النموذج مستقرا، وبالتالي عدم وجود تغير هيكل.

فيكون $\text{si } \exists r / w_r \notin \left[-2\sqrt{\hat{\delta}_{\varepsilon_{r-1}}^2} \left[1 + X_r' (X_{r-1}' X_{r-1})^{-1} X_r \right] \cdot + 2\sqrt{\hat{\delta}_{\varepsilon_{r-1}}^2} \left[1 + X_r' (X_{r-1}' X_{r-1})^{-1} X_r \right] \right] \quad \forall r = k+1, \dots, n$
 النموذج غير مستقر، وبالتالي وجود تغير هيكلية عند النقطة r .

مثال تطبيقي على برمجية EViews:

لتكن لديك البيانات الافتراضية التالية الخاصة بدالة انتاج في اقتصاد ما :

t	Y_t	X_{2t}	X_{3t}
1	4	4	0.9
2	4.5	5	0.8
3	5	6	0.9
4	5.5	7	0.8
5	6	9	0.7
6	7	8	0.6
7	6.5	10	0.6
8	6.5	11	0.8
9	7.5	12	0.5
10	7.5	13	0.5

- 1- إدخال هذه البيانات يدويا في برمجية EViews، موضحا مختلف التعليمات التي تم اتباعها.
- 2- إعادة إدخال هذه البيانات في برمجية EViews من خلال إستيراد ملف بصيغة EXCEL موضحا التعليمة التي تم اتباعها.
- 3- إعادة تسمية المتغيرات المستقلة في برمجية EViews، حيث المتغير المستقل الأول يمثل عدد ساعات العمل، والمتغير المستقل الثاني يمثل رأس المال المستخدم في الإنتاج.
- 4- تقدير النموذج، كتابته في شكله المقدر وتفسير النتائج، موضحا التعليمات والأوامر المستخدمة في برمجية EViews.
- 5- إيجاد القيم المقدرة للمتغير التابع وأيضا البواقي، مع توضيح مختلف التعليمات والأوامر المستخدمة على برمجية EViews.
- 6- إختبر فرضية وجود تغير هيكلية في النموذج موضحا مختلف التعليمات والأوامر التي تم استخدامها على برمجية EViews.

