المحور02: الانحدار الخطي المتعدد

المحاضرة 04 :

ثالثا: نتائج وخواص مقدرات طريقة OLS

1- خاصية عدم التحيز (Unbiased):

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

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

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 <u>أصغر تباين ممكن</u> 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:

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

3- خاصية الاتساق (Consistent):

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

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

مما سبق نجد أن:

المقدر $\hat{\beta}$ يحقق الشرطين:

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

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

4- مقدّر تباين المتغير العشوائي ε_t:

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

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

5- بناء مجالات ثقة لمعلمات النموذج:

تعطى مجالات الثقة لمعلمات النموذج الخطي المتعدد بالصيغة التالية:

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

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:

2-1- تحليل التباين ومعامل التحديد:

يمكن صياغة معادلة تحليل التباين على الشكل التالي:

$$\begin{aligned}(Y-\overline{Y})'(Y-\overline{Y}) &= (\hat{Y}-\overline{Y})'(\hat{Y}-\overline{Y}) + e'e \\ TSS &= ESS + RSS \\ \sum(Y_t-\overline{Y})^2 &= \sum(\hat{Y}_t-\overline{Y})^2 + \sum e_t^2\end{aligned}$$

حيث:

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

$\sum(\hat{Y}_t-\overline{Y})^2 = (\hat{Y}-\overline{Y})'(\hat{Y}-\overline{Y})$: 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:

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

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:

$$R^2 = \frac{(\hat{Y}-\bar{Y})'(\hat{Y}-\bar{Y})}{(Y-\bar{Y})'(Y-\bar{Y})} = \frac{\sum(\hat{Y}_t-\bar{Y})^2}{\sum(Y_t-\bar{Y})^2}$$

$$= \frac{\hat{\beta}'X'X\hat{\beta}-\bar{Y}^2}{\sum(Y_t-\bar{Y})^2}$$

$$= 1-\frac{(e-\bar{e})'(e-\bar{e})}{(Y-\bar{Y})'(Y-\bar{Y})} = 1-\frac{\sum e_{t2}}{\sum(Y_t-\bar{Y})^2}$$

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:

2-2-1- اختبار STUDENT: يستعمل هذا الاختبار لدراسة المعنوية الجزئية لمعلمات 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:

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

تحت ظل الفرضية $H_0 : \beta_i = 0$ نجد أن $\frac{\hat{\beta}_i - 0}{\sqrt{\hat{\delta}^2_{\hat{\beta}_i}}}$ تتبع أيضا توزيع STUDENT بدرجة حرية تساوي $(n-k)$، حيث يقوم هذا الاختبار على مقارنة إحصائيةSTUDENT المحسوبة $St_{cal} = \left| \frac{\hat{\beta}_i}{\sqrt{\hat{\delta}^2_{\hat{\beta}_i}}} \right|$ مع الإحصائية المجدولة من جدول STUDENT عند درجة حرية $(n-k)$ ومستوى معنوية $\alpha/2$، أي $St_{tab} = St^{\alpha/2}_{n-k}$. (في حالة $(n-k) > 30$ فإن $St_{tab} = 1.96$).

أما قرار الاختبار فيكون كما يلي:

- نرفض الفرضية H_0 إذا كانت $St_{cal} \geq St_{tab}$، و منه $\beta_i \neq 0$، وبالتالي وجود علاقة ذات دلالة احصائية بين المتغير التابع Y_t والمتغير المستقل X_{it}.
- نقبل الفرضية H_0 إذا كانت $St_{cal} < St_{tab}$، و منه $\beta_i = 0$، وبالتالي عدم وجود علاقة ذات دلالة احصائية بين المتغير التابع Y_t والمتغير المستقل X_{it}.

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

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

نقوم بحساب إحصائيةFISHER التي تعطى بالعلاقة التالية:

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

الإحصائية F_{cal} تتبع توزيع FISHER بدرجة حرية $\nu_1 = k-1$ و $\nu_2 = n-k$ ، أي $F_{tab} = F^{\alpha=5\%}_{(k-1.n-k)}$.

ويكون قرار الاختبار كمايلي:

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

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