8- حساب تباينات الأخطاء المقدرات ودراسة المعنوية :

بعد حساب المقدرات بطريقة MCO، لابد من الانتقال إلى مرحلة مهمة في عملية دراسة العلاقة بين المتغيرين قيد الدراسة، وهي مرحلة اختبار المعنوية الإحصائية للمقدرات، حيث لابد من اختبار تباين هذه المعلمات ومدى ابتعادها عن القيم الحقيقية، وذلك من خلال اختبار الفرضيات. وتتم العملية من خلال اختبار فرضية التوزيع الطبيعي للأخطاء بناء على تباين الأخطاء، وتباين المقدرات كل على حدى، حيث تحسب التباينات الثلاثة وفق العلاقة التالية:

* تباين الأخطاء المقدرة : $\hat{\delta}_u^2 = \frac{\sum u_i^2}{n-2}$

* تباين الميل الحدي $\hat{\beta}$: $\hat{\delta}_\beta^2 = \frac{\delta_u^2}{\sum (x_i - \bar{x})^2}$

* تباين الحد الثابت $\hat{\alpha}$: $\hat{\delta}_{\hat{\alpha}}^2 = \delta_u^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2} \right)$

من المعلوم أن العلاقتين $\frac{\hat{\alpha} - \alpha}{\delta_{\hat{\alpha}}}$ و $\frac{\hat{\beta} - \beta}{\delta_{\hat{\beta}}}$ تتبعان توزيع طبيعي $N(0,1)$ انحرافه 1، متوسطه 0

كما لدينا: $(n-2)\frac{\hat{\delta}_u^2}{\delta_u^2} = \frac{\sum u_i^2}{\delta_u^2}$ ، وتجدر الإشارة إلى أن العلاقة الأخيرة تتبع لتوزيع كاي تربيع χ^2 بدرجة حرية $(n-2)$، ونلاحظ من هذه العلاقة أن:

$$(n-2)\frac{\hat{\delta}_u^2}{\delta_u^2} = \frac{\sum u_i^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

وهذه العلاقة أيضا تتبع توزيع كاي تربيع ب $(n-2)$ درجات حرية، ومنه نستنتج أن كلا من العلاقتين $\frac{\hat{\alpha} - \alpha}{\hat{\delta}_{\hat{\alpha}}}$

و $\frac{\hat{\beta}-\beta}{\hat{\delta}_{\hat{\beta}}}$ يتبعان توزيع Student ب $(n-2)$ درجة حرية :

$$\frac{\hat{\beta}-\beta}{\hat{\delta}_{\hat{\beta}}} = \frac{\frac{\hat{\beta}-\beta}{\delta_{\hat{\beta}}}}{\left[(n-2)\frac{\hat{\delta}_{\beta}^2}{\delta_{\beta}} \cdot \frac{1}{(n-2)}\right]}$$

وهذا التوزيع يساعدنا على بناء اختبار الفرضيات.

من هذه العلاقات يمكننا القيام باختبار فرضيات للتأكد منا :

1- مقارنة المقدرات المحسوبة مع قيمة ثابتة.

2- مقارنة مقدرتين محسوبتين من عينتين مختلفتين.

3- تحديد مجال ثقة أو قبول للمقدرتين.

وأخيرا تستعمل هذه التوزيعات والعلاقتين $\frac{\hat{\alpha}-\alpha}{\hat{\delta}_{\hat{\alpha}}}$ و $\frac{\hat{\beta}-\beta}{\hat{\delta}_{\hat{\beta}}}$

لاختبار معنوية المقدرتين $\hat{\alpha}$, $\hat{\beta}$ ، حيث يمكننا اختبار فرضيتين لكل معلمة:

بالنسبة لـ $\hat{\alpha}$

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

بالنسبة لـ $\hat{\beta}$

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

$\{H_1 : \beta \neq 0$ $\{H_1 : \alpha \neq 0$

حيث تمثل H_0: الفرضية العدمية، و H_1: الفرضية البديلة، ولغرض اختبار كل فرضية صفرية نقوم بتعويض كل من $\hat{\alpha}$ و $\hat{\beta}$ بالصفر وحساب القيمة المتحصل عليها والتي تمثل إحصائية ستيودنت المحسوبة $(t^*_{\hat{\beta}} / t^*_{\hat{\alpha}})$ ومقارنتها مع ستيودنت الحرجة (الجدولية) بـ $n-2$ درجات حرية عند مستوى احتمال $\alpha\%$، وتكون المقارنة كالتالي:

✓ $t^* < t^{\alpha}_{n-2}$: نقبل الفرضية العدمية، ونقول أن المقدرة غير معنوية (لا تختلف معنويا عن 0) عند مستوى احتمال $\alpha\%$.

✓ $t^* > t^{\alpha}_{n-2}$: نرفض الفرضية العدمية ونقول أن المقدرة معنوية (تختلف معنويا عن 0) عند مستوى ثقة $\alpha\%$.

9- بناء مجالات الثقة للمقدرات المحسوبة:

بعد تحديد توزيع كلا المقدرتين (Student)، نستطيع الآن تحديد مجال الثقة لكل مقدرة عند $\alpha\%$ مستوى معنوية وذلك على النحو التالي:

$$P\left(\hat{\alpha} - t^{\alpha/2}_{n-2}\hat{S}_{\hat{\alpha}} < \alpha < \hat{\alpha} + t^{\alpha/2}_{n-2}\hat{S}_{\hat{\alpha}}\right) = 1-\alpha$$

$$P\left(\hat{\beta} - t^{\alpha/2}_{n-2}\hat{S}_{\hat{\beta}} \leq \beta \leq \hat{\beta} + t^{\alpha/2}_{n-2}\hat{S}_{\hat{\beta}}\right) = 1-\alpha$$

وبالتالي تصحيح المعلمة α, β ضمن المجال التالي:

$$\alpha = \hat{\alpha} \pm \hat{\delta}_{\hat{\alpha}} \; t^{\alpha/2}_{n-2}$$

$$\beta = \hat{\beta} \pm \hat{\delta}_{\hat{\beta}} \; t^{\alpha/2}_{n-1}$$

وبنفس الطريقة يمكننا تحديد مجال الثقة لتباين الأخطاء δ_u^2 عند مستوى معنوية $\alpha\%$.

$$P\left(\frac{(n-2)\,\hat{\delta}_u^2}{\chi^2_{1-\frac{\alpha}{2}}} \leq \delta_u^2 < \frac{(n-2)\,\hat{\delta}_u^2}{\chi^2_{\alpha/2}}\right) = 1 - \alpha$$

III- معادلة وجدول تحليل التباين:

2- معادلة تحليل التباين:

معادلة تحليل التباين هي اختبار لمدى جودة النموذج المقدر من خلال مدى مطابقة القيم المقدرة للقيم الحقيقية، ويمكننا استخراج المعادلة على النحو التالي:

1- تبيين صحة العلاقتين: $\sum_{i=1}^{n} u_i = 0$

و $\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \hat{y}_i$

$$y_i = \underbrace{\hat{\alpha} + \hat{\beta} x_i}_{\hat{y}_i} + u_i \quad *$$

$$y_i = \hat{y}_i + u_i$$ (1

$$u_i = y_i - \hat{y}_i$$

$$\sum u_i = \sum y_i - (n\hat{\alpha} + \hat{\beta}\sum x_i)$$

$$\sum u_i = \sum y_i - n\bar{y} + n\hat{\beta}\bar{x} - \hat{\beta}\sum x_i \quad (\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x})$$

$$\sum u_i = \cancel{n\bar{y}} - \cancel{n\bar{y}} + \cancel{n\hat{\beta}\bar{x}} - \cancel{\hat{\beta}n\bar{x}}$$

$$\sum_{i=1}^{n} u_i = 0$$

(2

$$y_i - \hat{y}_i = u_i$$

$$\sum y_i - \sum \hat{y}_i = \sum u_i = 0$$

$$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \hat{y}_i$$

لدينا المعادلة $y_i = \hat{y}_i + u_i$ ، في حالة ما طرحنا $\bar{y}$

من كلا طرفي المعادلة نجد:

$$y_i - \bar{y} = \hat{y}_i - \bar{y} + u_i$$

وإذا ما قمنا بتربيع الطرفين مع جمع كل الأطراف

بالنسبة لكل سنة i نتحصل على:

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

$$\sum (\hat{y}_i - \bar{y})^2 u_i$$

ولدينا $\bar{y}_i = \bar{\hat{y}}_i$ ، وكما نعلم أن $\sum (\hat{y}_i - \bar{\hat{y}}) = 0$

Scanné avec CamScanner

فتصبح العلاقة كالتالي :

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

$$SST = SSR + SSE$$

وتسمى هذه المعادلة بمعادلة تحليل التباين، حيث نلاحظ كلما انخفض مربع الأخطاء اتجه مربع القيم المقدرة نحو مساواة مربع القيم الحقيقية مما يجعل المعادلة التقديرية قريبة ومعبرة عن الواقع.

SST : مجموع مربعات الانحرافات الكلية

SSR : " " " المفسرة

SSE : " " " البواقي

وبالتالي كلما اتجه SSR نحو SST فهذا يعني أن جودة النموذج عالية نظرا لاتجاه SSE نحو الصفر.

ب- جدول تحليل التباين :

يمكننا تلخيص وتمثيل معادلة تحليل التباين في جدول يلخص كل المعطيات وكيفية استخراج معامل التحديد وكذا اختبار معنوية النموذج المقدر عند مستوى احتمال ما من أجل الحكم على قبول النموذج المقدر إحصائيا، وذلك من خلال اختبار فيشر fischer حيث يكون الجدول كالتالي :

مصدر التغير	مج مربعات الانحرافات	درجة الحرية	متوسط مربعات الانحرافات
المتغير المستقل x	$SSR = \sum (\hat{y}_i - \bar{y})^2$	1	$SSE/1$
البواقي	$SSE = \sum u_i^2$	$n-2$	$SSR/n-2$
المجموع	$SST = \sum (y_i - \bar{y})^2$	$n-1$	—

ملاحظة: في إطار نموذج الانحدار البسيط فإن اختبار معنوية β هو نفس اختبار معنوية مربعات الانحرافات المستقلة، حيث يتم اختبار فرضيتي β من خلال توزيع Student، في حين تختبر فرضيتي SSR من خلال توزيع Fischer ب $(n-k-1)$ درجة حرية:

n: عدد المشاهدات، k: عدد المتغيرات المفسرة والتي تساوي 1، فيكون عدد درجات الحرية هو $n-2$ عند مستوى احتمال 5%، وإحصائية اختبار فيشر يرمز لها ب F^* في هذه الحالة تحسب وفق العلاقات التالية:

$$F^* = \frac{SSR}{ddl_{SSR}} \Big/ \frac{SSE}{ddl_{SSE}} = \frac{\sum (\hat{y}_i - \bar{y})^2 / 1}{\sum u_i^2 / n-2}$$

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

بعد حساب إحصائية فيشر نقارنها مع قيمة فيشر الجدولية

بـ $n-k-1$ درجة حرية عند مستوى احتمال $\alpha\%$، حيث يتم القرار على النحو التالي:

* إذا كانت $F^* < F^{\alpha}_{1,n-2}$: نقبل الفرضية العدمية $(SSR = 0)$، ونقول أن مجموع مربعات الانحرافات المفسّرة غير معنوي أو لا يختلف معنويا عن 0 عند مستوى احتمال $\alpha\%$، وهذا ما يدل على عدم جودة النموذج المقدّر.

* إذا كانت $F^* > F^{\alpha}_{1,n-2}$: نرفض الفرضية العدمية، ونقول أن مجموع مربعات الانحرافات المفسّرة يختلف معنويا عن 0 وهذا يدل على جودة النموذج المقدّر.

11- التنبؤ في إطار نموذج الانحدار الخطي البسيط:

عندما نقوم بتقدير نموذج انحدار بسيط وحساب المعلمتين $\hat{\alpha}$ و $\hat{\beta}$ يمكن القيام بحسابات تنبؤية خلال الفترة المستقبلية h. إذا كانت قيمة المتغير المفسّر x_i معروفة خلال الفترة $i+1$ (x_{i+1}) يمكن كتابة التنبؤ على الشكل

$$y_{i+1} = \hat{\alpha} + \hat{\beta} x_{i+1}$$

ولدينا:
$$\frac{\hat{\alpha} + \hat{\beta} x_{i+1} - y_{i+1}}{\hat{\delta}_u \sqrt{1 + \frac{1}{n} + \frac{(x_{i+1} - \bar{x})^2}{\sum (x_i - \bar{x})^2}}}$$
تتبع توزيع Student

عند $n-2$ درجة حرية، مجال الثقة للقيمة المراد

التنبؤ على الشكل $\hat{y}_{i+1} = \hat{\alpha} + \hat{\beta} x_{i+1}$

وليكن : $\dfrac{\hat{\alpha} + \hat{\beta} x_{i+1} - y_{i+1}}{\hat{S}_u \sqrt{1 + \frac{1}{n} + \frac{(x_{i+1} - \bar{x})^2}{\sum (x_i - \bar{x})^2}}}$ تتبع توزيع Student

عند $n-2$ درجة حرية، مجال الثقة للقيمة المراد

التنبؤ بها يكون على النحو التالي:

$$y_{i+1} = \hat{y}_{i+1} \pm t_{n-2}^{\alpha/2} \hat{S}_u \sqrt{1 + \frac{1}{n} + \frac{(x_{i+1} - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$$

$$P\left(\hat{y}_{i+1} - t_{n-2}^{\alpha/2} \hat{S}_u \sqrt{1 + \frac{1}{n} + \frac{(x_{i+1} - \bar{x})^2}{\sum (x_i - \bar{x})^2}} < y_{i+1} < \hat{y}_{i+1} + t_{n-2}^{\alpha/2} \hat{S}_u \sqrt{1 + \frac{1}{n} + \frac{(x_{i+1} - \bar{x})^2}{\sum (x_i - \bar{x})^2}} \right) = 1 - \alpha$$