المحور01: الانحدار الخطي البسيط

المحاضرة 02 :

ثالثا: خصائص مقدرات OLS

المقدرات $\hat{\alpha}$ و $\hat{\beta}$ هي متغيرات عشوائية، تحسب بدلالة كل من المتغير التابع Y_t والمتغير المستقل X_t ، وتأخذ قيما حقيقية في العينة بعد تعويض Y_t و X_t بقيمها الحقيقية، حيث تعتبر المعلمات دالة خطية للمتغير التابع Y_t والمتغير المستقل X_t ، والتي تعتبر أول خاصية من خصائص مقدرات طريقة المربعات الصغرى العادية، بالإضافة إلى خصائص أخرى سنوجزها فيما يلي:

1- خاصية عدم التحيز (Unbiased):

يقصد بالتحيز الفرق الموجود بين مقدر معين وتوقعه الرياضي(أمله الرياضي)، فإذا كان هذا الفرق مختلفا عن الصفر نقول عن ذلك المقدر بأنه مقدر متحيز، وإذا كان هذا الفرق معدوما فإن هذا المقدر مقدر غير متحيز. فيكون $\hat{\theta}$ مقدرا غير متحيز لـ: θ إذا حقق الشرط التالي: $E(\hat{\theta}) = \theta$.

بنفس التعريف نقول أن مقدرات طريقة المربعات الصغرى العادية $\hat{\alpha}$ و $\hat{\beta}$ هي مقدرات غير متحيزة إذا حققت الشرطين التاليين:

$$E(\hat{\alpha}) = \alpha \qquad \text{and} \qquad E(\hat{\beta}) = \beta$$

1-1- بالنسبة لـــ: $\hat{\beta}$:

$$E(\hat{\beta}) = E\left[\frac{\Sigma(Y_t - \overline{Y})(X_t - \overline{X})}{\Sigma(X_t - \overline{X})^2}\right]$$

نضع: $\Sigma(X_t - \overline{X}) = \Sigma x_t$

ومنه:

$$E(\hat{\beta}) = E\left[\frac{\sum(Y_t - \overline{Y})\cdot x_t}{\sum x_t^2}\right] = E\left[\frac{\sum x_t Y_t - \overline{Y}\sum x_t}{\sum x_t^2}\right]$$
$$= E\left[\frac{\sum x_t \cdot (\alpha + \beta \cdot X_t + \varepsilon_t)}{\sum x_t^2}\right] = E\left[\frac{\alpha\sum x_t + \beta\sum x_t \cdot X_t + \sum x_t\varepsilon_t}{\sum x_t^2}\right]$$
$$= E\left[\frac{\beta\sum x_t^2 + \sum x_t\varepsilon_t}{\sum x_t^2}\right]$$
$$= \beta + E\left[\frac{\sum x_t\varepsilon_t}{\sum x_t^2}\right] = \beta + \left[\frac{\sum x_t E(\varepsilon_t)}{\sum x_t^2}\right]$$

β مقدر غير متحيز لـ: $\hat{\beta}$ ، إذن $E(\hat{\beta}) = \beta$ ومنه:

2-1- بالنسبة لـــ: $\hat{\alpha}$

$$E(\hat{\alpha}) = E(\overline{Y} - \hat{\beta}\cdot\overline{X})$$

لدينا: $\overline{Y} = \alpha + \beta\cdot\overline{X} + \overline{\varepsilon}$

ومنه:

$$E(\hat{\alpha}) = E(\alpha + \beta\cdot\overline{X} + \overline{\varepsilon} - \hat{\beta}\cdot\overline{X})$$
$$= E(\alpha - (\hat{\beta} - \beta)\cdot\overline{X} + \overline{\varepsilon})$$
$$= \alpha - E(\hat{\beta} - \beta)\cdot\overline{X} + E(\overline{\varepsilon})$$
$$= \alpha$$

إذن: $E(\hat{\alpha}) = \alpha$ ، ومنه: $\hat{\alpha}$ مقدر غير متحيز لـ: α

2- أفضل مقدرات خطية غير متحيزة 'BLUE' Best Linear Unbiased Estimators :

تكون مقدرتا طريقة المربعات الصغرى العادية $\hat{\alpha}$ و $\hat{\beta}$ أفضل مقدرتين خطيتين وغير متحيزتين، حيث أن لها أصغر تباين ممكن مقارنة مع بقية المقدرات الخطية وغير المتحيزة الأخرى. ويمكن إثبات هذه الخاصية بعد ايجاد تباين مقدرات طريقة المربعات الصغرى العادية كمايلي:

2-1- تباين $\hat{\beta}$:

$$V(\hat{\beta}) = E(\hat{\beta} - E(\hat{\beta}))^2 = E(\hat{\beta} - \beta)^2$$
$$= E\left[\beta + \frac{\sum x_t \varepsilon_t}{\left[\sum x_t^2\right]^2} - \beta\right]^2 = E\left[\frac{\sum x_t \varepsilon_t}{\left[\sum x_t^2\right]^2}\right]^2$$
$$= E\left[\frac{\left[\sum x_t \varepsilon_t\right]^2}{\left[\sum x_t^2\right]^2}\right] = E\left[\frac{\sum x_t^2 \varepsilon_t^2 + 2\sum\sum x_i \varepsilon_i x_j \varepsilon_j}{\left[\sum x_t^2\right]^2}\right]$$
$$= \left[\frac{\sum x_t^2 \cdot E(\varepsilon_t^2)}{\left[\sum x_t^2\right]^2}\right] = \frac{\delta_\varepsilon^2 \sum x_t^2}{\left[\sum x_t^2\right]^2}$$

ومنه:

$$V(\hat{\beta}) = \delta_\varepsilon^2 \left[1 \Big/ \sum x_t^2 \right]$$

2-2- تباين $\hat{\alpha}$:

$$V(\hat{\alpha}) = E(\hat{\alpha} - E(\hat{\alpha}))^2 = E(\hat{\alpha} - \alpha)^2$$
$$= E\left[\alpha - \left[\hat{\beta} - \beta\right]\overline{X} + \bar{\varepsilon} - \alpha\right]^2 = E\left[-\left[\hat{\beta} - \beta\right]\overline{X} + \bar{\varepsilon}\right]^2$$
$$= E\left[\left[\hat{\beta} - \beta\right]^2 \overline{X}^2 + \bar{\varepsilon}^2 - 2\overline{X}\bar{\varepsilon}\left[\hat{\beta} - \beta\right]\right]$$
$$= \overline{X}^2 \cdot E\left[\hat{\beta} - \beta\right]^2 + E\left[\bar{\varepsilon}^2\right] - 2\overline{X} \cdot E\left[\bar{\varepsilon}\right] \cdot E\left[\hat{\beta} - \beta\right]$$
$$= \overline{X}^2 \left[\frac{\delta_\varepsilon^2}{\sum x_t^2}\right] + \frac{\delta_\varepsilon^2}{n} = \delta_\varepsilon^2 \left[\frac{\overline{X}^2}{\sum x_t^2} + \frac{1}{n}\right]$$

ومنه:

$$V(\hat{\alpha}) = \delta_\varepsilon^2 \left[\frac{\overline{X}^2}{\sum x_t^2} + \frac{1}{n}\right]$$

بعد حساب تباين المقدرات سنعمل على إثبات خاصية أقل تباين كمايلي:

- بالنسبة لـ: $V(\hat{\beta})$

$$V(\hat{\beta}) = \frac{\delta_\varepsilon^2}{\sum x_t^2} = \frac{\delta_\varepsilon^2 \Big/ n}{\sum x_t^2 \Big/ n}$$

بما أن: δ_ε^2 ثابت حسب فرضيات النموذج فإن:

$$\lim_{n\to+\infty}\left(\delta_\varepsilon^2 / n\right)=0$$

من جهة أخرى لدينا:

$$\lim_{n\to+\infty}\left(\Sigma x_t^2 / n\right)=\lim_{n\to+\infty}\left(\Sigma(X_t-\overline{X})^2 / n\right)=\lim_{n\to+\infty} V(X)\neq 0$$

وبالتالي نجد أن:

$$\lim_{n\to+\infty} V(\hat{\beta})=\lim_{n\to+\infty}\left(\frac{\delta_\varepsilon^2}{\Sigma x_t^2}\right)=0 \qquad \cdots\cdots\cdots(01)$$

- **بالنسبة لـ:** $V(\hat{\alpha})$

$$V(\hat{\alpha})=\delta_\varepsilon^2\left[\frac{\overline{X}^2}{\Sigma x_t^2}+\frac{1}{n}\right]=\delta_\varepsilon^2\left[\frac{\overline{X}^2 / n}{\Sigma x_t^2 / n}+\frac{1}{n}\right]=\delta_\varepsilon^2\left[\frac{\overline{X}^2 / n}{V(X_t)}+\frac{1}{n}\right]$$

لدينا:

$$\lim_{n\to+\infty}\left(\overline{X}^2 / n\right)=0 \quad \Rightarrow \quad \lim_{n\to+\infty}\left(\frac{\overline{X}^2 / n}{V(X_t)}\right)=0$$

$$\lim_{n\to+\infty}\left(\frac{1}{n}\right)=0$$

إذن:

$$\lim_{n\to+\infty} V(\hat{\alpha})=\delta_\varepsilon^2\left[\lim_{n\to+\infty}\left(\frac{\overline{X}^2 / n}{V(X_t)}\right)+\lim_{n\to+\infty}\left(\frac{1}{n}\right)\right]=\delta_\varepsilon^2[0+0]=0 \qquad \cdots\cdots\cdots(02)$$

من (01) و(02) نستنتج أن مقدرات طريقة المربعات الصغرى العادية تمتاز بخاصية أقل تباين ممكن.

3- خاصية الاتساق (Consistent):

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

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

مما سبق نجد أن:

المقدر $\hat{\alpha}$ **يحقق الشرطين:**

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

المقدر $\hat{\beta}$ يحقق الشرطين:

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

إذن نستنتج أن المقدرات $\hat{\alpha}$ و $\hat{\beta}$ هي مقدرات متسقة للمعلمتين α و β على الترتيب.

4- التباين المشترك بين $\hat{\alpha}$ و $\hat{\beta}$:

$$\begin{aligned}
Cov(\hat{\alpha}.\hat{\beta}) &= E\left[\left[\hat{\alpha} - E(\hat{\alpha})\right]\cdot\left[\hat{\beta} - E(\hat{\beta})\right]\right] \\
&= E\left[\left[-\overline{X}(\hat{\beta}-\beta) + \overline{\varepsilon}\right]\cdot\left[\hat{\beta}-\beta\right]\right] \\
&= E\left[\left[-\overline{X}(\hat{\beta}-\beta)^2\right] + \left[\overline{\varepsilon}(\hat{\beta}-\beta)\right]\right] \\
&= -\overline{X}\cdot E\left[\hat{\beta}-\beta\right]^2 = -\overline{X}\cdot V(\hat{\beta}) \\
&= -\overline{X}\left[\delta_\varepsilon^2 \Big/ \sum x_t^2\right]
\end{aligned}$$

ومنه:

$$Cov(\hat{\alpha}.\hat{\beta}) = \delta_\varepsilon^2\left[-\overline{X} \Big/ \sum x_t^2\right]$$

1- تقدير تباين المتغير العشوائي ε_t:

من خلال النتائج السابقة يتضح لنا أنه يمكننا تقدير كل من α و β ، إلا أنه لا يمكننا إيجاد تباين كل مقدر، لأنه بدلالة تباين المتغير العشوائي المجهول.

لدينا:

$$\begin{cases} Y_t = \alpha + \beta X_t + \varepsilon_t \\ \\ \overline{Y} = \alpha + \beta \overline{X} + \overline{\varepsilon} \end{cases} \Rightarrow Y_t - \overline{Y} = \beta(X_t - \overline{X}) + (\varepsilon_t - \overline{\varepsilon})$$

ولدينا:

$$\begin{cases} Y_t = \hat{\alpha} + \hat{\beta} X_t + e_t \\ \\ \overline{Y} = \hat{\alpha} + \hat{\beta} \overline{X} \end{cases} \Rightarrow Y_t - \overline{Y} = \hat{\beta}(X_t - \overline{X}) + e_t \Rightarrow e_t = Y_t - \overline{Y} - \hat{\beta}(X_t - \overline{X})$$

بالتعويض نجد:

$$e_t = \beta(X_t - \overline{X}) + (\varepsilon_t - \overline{\varepsilon}) - \hat{\beta}(X_t - \overline{X}) = -(\hat{\beta} - \beta)(X_t - \overline{X}) + (\varepsilon_t - \overline{\varepsilon})$$
$$e_t^2 = (\beta - \hat{\beta})^2 (X_t - \overline{X})^2 + (\varepsilon_t - \overline{\varepsilon})^2 - 2\left[(\hat{\beta} - \beta)(X_t - \overline{X})(\varepsilon_t - \overline{\varepsilon})\right]$$
$$\sum e_t^2 = (\beta - \hat{\beta})^2 \sum (X_t - \overline{X})^2 + \sum (\varepsilon_t - \overline{\varepsilon})^2 - 2\sum (\hat{\beta} - \beta)(X_t - \overline{X})(\varepsilon_t - \overline{\varepsilon})$$
$$E(\sum e_t^2) = E(\beta - \hat{\beta})^2 \sum (X_t - \overline{X})^2 + E\left(\sum (\varepsilon_t - \overline{\varepsilon})^2\right) - 2E\left(\sum (\hat{\beta} - \beta)(X_t - \overline{X})(\varepsilon_t - \overline{\varepsilon})\right)$$

نقوم بحساب التوقع لكل طرف على حدى:

$$E(\beta - \hat{\beta})^2 \sum (X_t - \overline{X})^2 = v(\hat{\beta}) \sum (X_t - \overline{X})^2 = \frac{\delta_\varepsilon^2}{\sum (X_t - \overline{X})^2} \sum (X_t - \overline{X})^2 = \delta_\varepsilon^2$$

$$E\left(\sum (\varepsilon_t - \overline{\varepsilon})^2\right) = E\left(\sum (\varepsilon_t^2 + \overline{\varepsilon}^2 - 2\overline{\varepsilon}\varepsilon_t\right) = E\left(\sum \varepsilon_t^2 + n\overline{\varepsilon}^2 - 2\overline{\varepsilon}\sum \varepsilon_t\right) = E\left(\sum \varepsilon_t^2 + n\overline{\varepsilon}^2 - 2n\overline{\varepsilon}^2\right)$$
$$= \sum E(\varepsilon_t^2) - nE(\overline{\varepsilon}^2) = n\delta_\varepsilon^2 - nE\left(\frac{\sum \varepsilon_t}{n}\right)^2 = n\delta_\varepsilon^2 + n\left(\frac{E(\sum \varepsilon_t)^2}{n^2}\right)$$
$$= n\delta_\varepsilon^2 - n\left(\frac{n\delta_\varepsilon^2}{n^2}\right) = n\delta_\varepsilon^2 - \delta_\varepsilon^2 = (n-1)\delta_\varepsilon^2$$

$$E\left(\sum(\hat{\beta}-\beta)(X_t-\overline{X})(\varepsilon_t-\overline{\varepsilon})\right)=E\sum\left[\frac{\sum(X_t-\overline{X})\varepsilon_t}{\sum(X_t-\overline{X})^2}(X_t-\overline{X})(\varepsilon_t-\overline{\varepsilon})\right]$$

$$=E\left[\left(\frac{\sum(X_t-\overline{X})^2\varepsilon_t^2}{\sum(X_t-\overline{X})^2}\right)\right]=\frac{\sum(X_t-\overline{X})^2\delta_\varepsilon^2}{\sum(X_t-\overline{X})^2}=\delta_\varepsilon^2$$

إذن:

$$E(\sum e_t^2)=\delta_\varepsilon^2+(n-1)\delta_\varepsilon^2-2\delta_\varepsilon^2=(n-2)\delta_\varepsilon^2$$

وبالتالي يكون $\sum e_t^2$ **مقدرا متحيزا لـ:** δ_ε^2 **. ويمكن استنتاج المقدر غير المتحيز لـ:** δ_ε^2 **كمايلي:**

$$\hat{\delta}_\varepsilon^2=\frac{\sum e_t^2}{n-2}$$

ومنه يكون التباين المقدر لـ: $\hat{\alpha}$ **و** $\hat{\beta}$ **كمايلي:**

$$\hat{V}(\hat{\alpha})=\hat{\delta}_{\hat{\alpha}}^2=\hat{\delta}_\varepsilon^2\left[\frac{\overline{X}^2}{\sum x_t^2}+\frac{1}{n}\right]$$

$$\hat{V}(\hat{\beta})=\hat{\delta}_{\hat{\beta}}^2=\hat{\delta}_\varepsilon^2\left[1\Big/\sum x_t^2\right]$$

<u>مثال</u>: من المثال السابق (في المحاضرة 1) قدر تباين المتغير العشوائي، ثم استنتج التباين المقدر للمقدرات.

e_t	e_t^2
3.19	10.23
-3.53	12.48
1.73	3.00
-2.26	5.13
-4.99	24.99
0.26	0.07
3.00	9.00

	-2.46	6.08
	4.26	18.20
	0.80	0.64
$\sum$	0	89.86

وبالتالي نجد:

$$\hat{\delta}_{\varepsilon}^2 = \frac{\sum e_t^2}{n-2} = \frac{89.86}{10-2} = 11.23$$

أما التباينات المقدرة للمقدرات فتكون كمايلي:

$$\hat{V}(\hat{\alpha}) = \hat{\delta}_{\hat{\alpha}}^2 = \hat{\delta}_{\varepsilon}^2\left[\frac{\overline{X}^2}{\sum x_t^2} + \frac{1}{n}\right] = 11.23\left[\frac{7^2}{30} + \frac{1}{10}\right] = 19.47 \Rightarrow \hat{\delta}_{\hat{\alpha}} = \sqrt{19.47} = 4.41$$

$$\hat{V}(\hat{\beta}) = \hat{\delta}_{\hat{\beta}}^2 = \hat{\delta}_{\varepsilon}^2\left[1 \Big/ \sum x_t^2\right] = 11.23\left[\frac{1}{30}\right] = 0.3744 \Rightarrow \hat{\delta}_{\hat{\beta}} = \sqrt{0.3744} = 0.6119$$

2- بناء مجالات ثقة لمعلمات النموذج:

بعد تحديد توزيع كل من المقدرات $\hat{\alpha}$ و $\hat{\beta}$، سنعمل على تكوين مجالات ثقة تنتمي إليها معلمات النموذج، وهذا عند مستوى معنوية معين، حيث احتمال أن تنتمي معلمة النموذج إلى مجال الثقة يساوي الواحد مطروحا منه مستوى المعنوية.

- **بالنسبة للمعلمة α:**

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

- **بالنسبة للمعلمة β:**

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

مثال: من المثال السابق يمكن بناء مجالات ثقة لمعلمات النموذج عند مستوى معنوية $\alpha = 5\%$ كمايلي:

- **بالنسبة للمعلمة α:**

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

$$\Rightarrow P\left(\alpha \in \left[21.86 - St_{n-2}^{\frac{\alpha}{2}} \cdot \sqrt{19.47} \quad . \quad 21.86 + St_{n-2}^{\frac{\alpha}{2}} \cdot \sqrt{19.47}\right]\right) = 1 - 0.05 = 0.95$$

من جدول Student نجد أن القيمة المجدولة عند درجة حرية (10-2=8) وعند مستوى معنوية $\alpha = 5\%$ هي:

$$St_{n-2}^{\frac{\alpha}{2}} = St_{10-2}^{\frac{5}{2}} = St_{8}^{2.5\%} = 2.306$$

وبالتالي يكون مجال الثقة للمعلمة α كما يلي:

$$P\left(\alpha \in \left[21.86 - 2.306 \cdot \sqrt{19.47} \quad . \quad 21.86 + 2.306 \cdot \sqrt{19.47}\right]\right) = 0.95$$
$$P\left(\alpha \in \left[11.68 \quad . \quad 32.03\right]\right) = 0.95$$

- **بالنسبة للمعلمة β:**

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

$$\Rightarrow P\left(\beta \in \left[4.73 - 2.306 \cdot \sqrt{0.3744} \quad . \quad 4.73 + 2.306 \cdot \sqrt{0.3744}\right]\right) = 0.95$$

$$\Rightarrow P\left(\beta \in \left[3.31 \quad . \quad 6.14\right]\right) = 0.95$$

رابعا: تقييم النموذج الخطي البسيط

يتم تقييم نموذج الانحدار الخطي البسيط باستعمال المعايير التالية:

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 نما 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:

تعتبر بواقي التقدير e_t مقياسا لمدى تمثيل النموذج المقدر للنموذج الحقيقي، وهو ما يعرف بجودة التوفيق، فكلما كانت البواقي كبيرة قلت وضعفت جودة التمثيل والعكس صحيح.

لدينا:

$$Y_t = \hat{Y}_t + e_t$$
$$\Rightarrow Y_t - \overline{Y} = \hat{Y}_t - \overline{Y} + e_t$$
$$\Rightarrow (Y_t - \overline{Y})^2 = (\hat{Y}_t - \overline{Y})^2 + 2(\hat{Y}_t - \overline{Y}) \cdot e_t + e_t^2$$
$$\Rightarrow \Sigma(Y_t - \overline{Y})^2 = \Sigma(\hat{Y}_t - \overline{Y})^2 + 2\Sigma(\hat{Y}_t - \overline{Y}) \cdot e_t + \Sigma e_t^2$$

بالمقابل لدينا: $\Sigma(\hat{Y}_t - \overline{Y}) \cdot e_t = 0$ **، وبالتالي نجد:**

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

هذه المعادلة الأخيرة تعرف بمعادلة تحليل التباين، وهي تتكون من ثلاث أجزاء كما يلي:

$\Sigma(Y_t - \overline{Y})^2$: مجموع مربعات الانحرافات الكلية (Total Sum of Squares (TSS)).

$\Sigma(\hat{Y}_t - \overline{Y})^2$: مجموع مربعات الانحرافات المفسرة (Explained Sum of Squares (ESS)).

Σe_t^2 : مجموع مربعات البواقي (Residual Sum of Squares (RSS))

ومنه يمكن إعادة صياغة معادلة تحليل التباين على النحو التالي:

$$TSS = ESS + RSS$$

انطلاقا من معادلة تحليل التباين يمكن استخلاص مؤشر تقاس به القدرة التفسيرية للنموذج، والمتمثل في ما يعرف بمعامل التحديد، والذي نرمز له بالرمز R^2 ، وهو مؤشر يقيس النسبة المفسرة من التغير الكلي بدلالة خط الانحدار، أي نسبة مجموع مربعات الانحرافات المفسرة إلى مجموع مربعات الانحرافات الكلية، تتراوح قيمته بين الصفر والواحد، أي: $0 \le R^2 \le 1$ ، فكلما اقترب R^2 من الواحد، تكون للنموذج قدرة تفسيرية عالية، وكلما اقترب من الصفر، دل ذلك على ضعف القدرة التفسيرية للنموذج. بصيغة رياضية يكتب معامل التحديد كما يلي:

$$R^2 = \frac{ESS}{TSS} = \frac{\sum(\hat{Y}_t - \overline{Y})^2}{\sum(Y_t - \overline{Y})^2}$$

$$\frac{TSS}{TSS} = \frac{ESS}{TSS} + \frac{RSS}{TSS} \Rightarrow 1 = R^2 + \frac{RSS}{TSS} \Rightarrow R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum e_t^2}{\sum(Y_t - \overline{Y})^2}$$

كذلك رياضيا يمكن إثبات أن معامل التحديد في النموذج الخطي البسيط (متغيرين فقط، أحدهما تابع والآخر مستقل) ما هو إلا مربع معامل الارتباط الخطي البسيط بين متغيري النموذج.

أما جدول تحليل التباين Analysis Of Variance (ANOVA) فيأخذ الشكل التالي:

مصدر التغير	مجموع المربعات	درجة الحرية	متوسط المربعات
المتغير المستقل X_t	$ESS = \sum(\hat{Y}_t - \overline{Y})^2$	1	$\sum(\hat{Y}_t - \overline{Y})^2 / 1$
البواقي e_t	$RSS = \sum e_t^2$	$n-2$	$\sum e_t^2 / n-2$
المجموع	$TSS = \sum(Y_t - \overline{Y})^2$	$n-1$	

مثال: من المثال السابق أوجد معامل التحديد R^2 ، ثم كون جدول ANOVA

بعد تقدير النموذج وبعد تقدير تباين المتغير العشوائي، فإن أسرع علاقة لإيجاد معامل التحديد هي:

$$R^2 = 1 - \frac{\sum e_t^2}{\sum(Y_t - \overline{Y})^2} = 1 - \frac{89.86}{762} = 1 - 0.1179 = 0.8820$$

$$R^2 = \frac{\hat{\beta}^2 \sum(X_t - \overline{X})^2}{\sum(Y_t - \overline{Y})^2} = \frac{(4.73)^2(30)}{762} = \frac{672.13}{762} = 0.8820$$

$R^2 = 0.8820$ ، أي أن للنموذج جودة توفيق عالية، وأن قدرته التفسيرية مرتفعة، كما نستنتج من معامل التحديد أيضا أن 88.20% من تغيرات عوائد المبيعات مفسر بتغيرات الانفاق على الاعلانات، أما النسبة المتبقية والمقدرة بـ: 11.80% فهي مفسرة بعوامل أخرى غير مدرجة بالنموذج.

أما جدول تحليل التباين ANOVA فيكون كمايلي:

مصدر التغير	مجموع المربعات	درجة الحرية	متوسط المربعات
المتغير المستقل X_t	$ESS = 672.13$	1	$672.13/1 = 672.13$
البواقي e_t	$RSS = 89.86$	$10-2=8$	$89.86/8 = 11.23$
المجموع	$TSS = 762$	$10-1=9$	

2- 2- اختبارات المعنوية:

تتمثل هذه الاختبارات فيما يلي:

2-2-1- اختبار Student:

يستعمل هذا الاختبار لدراسة المعنوية الجزئية لمعلمات النموذج عند مستوى معنوية معين، حيث نختبر المعنوية الاحصائية لمعامل الانحدار (β)، والتي تسمح بالحكم على معنوية العلاقة بين المتغير التابع والمتغير المستقل، كما نختبر المعنوية الاحصائية للحد الثابت (α)، والتي تسمح بالحكم على جدوى وجود الحد الثابت في النموذج من عدمها.

- **بالنسبة لـــ: β**

يأخذ اختبار Student الشكل التالي:

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

تشير فرضية العدم H_0 إلى عدم وجود علاقة ذات دلالة احصائية بين المتغير التابع والمتغير المستقل، بينما تشير الفرضية البديلة H_1 إلى وجود علاقة ذات دلالة احصائية.

مما سبق لدينا:

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

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

أما قرار الاختبار فيكون كما يلي:

- نرفض الفرضية H_0 إذا كانت $St_{cal} \geq St_{tab}$، و منه $\beta \neq 0$، وبالتالي وجود علاقة ذات دلالة احصائية بين المتغير التابع والمتغير المستقل.
- نقبل الفرضية H_0 إذا كانت $St_{cal} < St_{tab}$، و منه $\beta = 0$، وبالتالي عدم وجود علاقة ذات دلالة احصائية بين المتغير التابع والمتغير المستقل.
- **بالنسبة لـ:** α

يأخذ اختبار Student الشكل التالي:

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

تشير فرضية العدم H_0 إلى عدم وجود دلالة احصائية لإدراج الحد الثابت في النموذج ، بينما تشير الفرضية البديلة H_1 إلى أن وجود الحد الثابت في النموذج له دلالة احصائية.

مما سبق لدينا:

$$\frac{\hat{\alpha} - \alpha}{\sqrt{\hat{\delta}^2_{\hat{\alpha}}}} \rightarrow St(n-2)$$

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

أما قرار الاختبار فيكون كما يلي:

- نرفض الفرضية H_0 إذا كانت $St_{cal} \geq St_{tab}$ ، و منه $\alpha \neq 0$ ، وبالتالي وجود الحد الثابت في النموذج له دلالة احصائية.
- نقبل الفرضية H_0 إذا كانت $St_{cal} < St_{tab}$ ، و منه $\alpha = 0$ ، وبالتالي عدم وجود دلالة احصائية لإدارج الحد الثابت في النموذج.

<u>مثال:</u> من المثال السابق يمكننا اجراء اختبار Student عند مستوى معنوية $\alpha = 5\%$ كمايلي:

- **بالنسبة لـــ:** β

يأخذ اختبار Student الشكل التالي:

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

إحصائيةStudent المحسوبة:

$$St_{cal} = \left| \frac{\hat{\beta}}{\sqrt{\hat{\delta}^2_{\hat{\beta}}}} \right| = \left| \frac{4.73}{0.6119} \right| = 7.73$$

إحصائيةStudent المجدولة:

$$St^{\frac{\alpha}{2}}_{n-2} = St^{\frac{5}{2}}_{10-2} = St^{2.5\%}_{8} = 2.306$$

القرار: نلاحظ أن $St_{cal} \geq St_{tab}$ ، وبالتالي نقبل الفرضية $\beta \neq 0$ ، أي وجود علاقة ذات دلالة احصائية بين الانفاق على الاعلانات وعوائد المبيعات في هذه الشركة.

- **بالنسبة لـ:** α

يأخذ اختبار Student الشكل التالي:

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

إحصائيةStudent المحسوبة:

$$St_{cal} = \left| \frac{\hat{\alpha}}{\sqrt{\hat{\delta}^2_{\hat{\alpha}}}} \right| = \left| \frac{21.86}{4.4126} \right| = 4.95$$

إحصائيةStudent المجدولة:

$$St^{\frac{\alpha}{2}}_{n-2} = St^{\frac{5}{2}}_{10-2} = St^{2.5\%}_{8} = 2.306$$

القرار: نلاحظ أن $St_{cal} \geq St_{tab}$ ، وبالتالي نقبل الفرضية $\alpha \neq 0$ ، أي وجود الحد الثابت في النموذج له دلالة احصائية.

2-2-2- اختبار Fisher:

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

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

نقوم بحساب إحصائية Fisher التي تعطى بالعلاقة التالية:

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

الإحصائية F_{cal} تتبع توزيع Fisher بدرجة حرية $\nu_1 = 1$ و $\nu_2 = n-2$ ، أي $F_{(1.n-2)}^{\alpha=5\%}$.

و يكون قرار الاختبار كمايلي:

- نرفض الفرضية H_0 إذا كانت $F_{cal} \geq F_{(1.n-2)}^{\alpha=5\%}$ ، و منه $\alpha \neq 0$ أو $\beta \neq 0$.
- نرفض الفرضية H_1 إذا كانت $F_{cal} < F_{(1.n-2)}^{\alpha=5\%}$ ، و منه $\alpha = 0$ و $\beta = 0$.

مثال: من المثال السابق يكون اختبار Fisher عند مستوى معنوية $\alpha = 5\%$ كما يلي:

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

إحصائيةFisher المحسوبة:

$$F_{cal} = \frac{R^2 / 2-1}{(1-R^2) / n-2} = \frac{0.8820 / 1}{(1-0.8820) / 10-2} = 59.83$$

إحصائيةFisher المجدولة:

$$F_{(1.n-2)}^{\alpha=5\%} = F_{(1.8)}^{\alpha=5\%} = 5.318$$

خامسا: التنبؤ في النموذج الخطي البسيط

بعد تقدير معلمات النموذج الخطي البسيط، يكون في الامكان حساب التنبؤ لأفق معين، فالنموذج المقدر للفترة $t = 1 \cdots\cdots n$ يعطى بالعلاقة التالية:

$$Y_t = \hat{\alpha} + \hat{\beta}X_t + e_t$$

يمكن بناء مجال ثقة لـ: Y_{n+1} عند مستوى معنوية معين كما يلي:

$$Y_{n+1} \in \left[\hat{Y}_{n+1} \pm St^{\alpha/2}_{(n-2)} \sqrt{\left(\frac{(X_{n+1} - \overline{X})^2}{\sum (X_t - \overline{X})^2} + \frac{1}{n} + 1 \right) \hat{\delta}^2_\varepsilon} \right]$$

$$Y_{n+1} \in \left[\hat{Y}_{n+1} - St^{\alpha/2}_{(n-2)} \sqrt{\left(\frac{(X_{n+1} - \overline{X})^2}{\sum (X_t - \overline{X})^2} + \frac{1}{n} + 1 \right) \hat{\delta}^2_\varepsilon} \quad . \quad \hat{Y}_{n+1} + St^{\alpha/2}_{(n-2)} \sqrt{\left(\frac{(X_{n+1} - \overline{X})^2}{\sum (X_t - \overline{X})^2} + \frac{1}{n} + 1 \right) \hat{\delta}^2_\varepsilon} \right]$$

مثال: من المثال السابق أوجد عائد المبيعات عند تخصيص انفاق على الاعلانات بمقدار 12 مليون دج سنة 2019

$$\hat{Y}_{2019} = \hat{Y}_{10+1} = \hat{\alpha} + \hat{\beta}X_{10+1} = 21.86 + 4.73(12) = 78.62$$

أما مجال الثقة فيكون كمايلي:

لدينا:

$$n = 10$$
$$\hat{\delta}^2_\varepsilon = 11.23$$
$$\sum (X_t - \overline{X})^2 = 30$$
$$\overline{X} = 7$$
$$St^{2.5\%}_8 = 2.306$$
$$X_{11} = 12$$
$$\hat{Y}_{11} = 78.62$$

إذن:

$$Y_{2019} \in \left[78.62 \pm 2.306 \sqrt{\left(\frac{(12-7)^2}{30} + \frac{1}{10} + 1 \right) \cdot 11.23} \right]$$
$$Y_{n+1} \in \left[78.62 - 2.306\sqrt{21.71} \quad . \quad 78.62 + 2.306\sqrt{21.71} \right]$$
$$Y_{n+1} \in \left[78.62 - 10.71 \quad . \quad 78.62 + 10.71 \right]$$
$$Y_{n+1} \in \left[67.87 \quad . \quad 89.36 \right]$$