مصادر التباين التعاقبية Sequential source of variation

وهي عبارة عن تقسيم درجات الحرية ومجموع المربعات العائد إلى الانحدار إلى مصادرها ،ويتم ذلك بشرط تحقيق ثلاثة قيود:

1- أن يكون عدد المصادر التعاقبية بعدد المتغيرات المستقلة

2- كل مصدر تعاقبي له درجة حرية واحدة أي ان مجموع درجات الحرية لهذه المصادر تساوي درجات الحرية العائد للانحدار

3- مجموع مربعات هذه المصادر تساوي مجموع المربعات العائد للانحدار.

ويجب ملاحظة مايحتاج كل مصدر تعاقبي من خطا خاص به للتوصل إلى اختبار F وكمثال على ذلك

s.o.v	df	SS
$SSR(x_1x_2x_3x_4)$	**4**	$SSR(x_1)+SSR(x_2/x_1)+SSR(x_3/x_1x_2)+SSR(x_4/x_1x_2x_3)$
$SSR(x_1)$	1	
$SSR(x_2/x_1)$	1	$SSR(x_2x_3x_4/x_1)-SSR(x_3x_4/x_1x_2)$
$SSR(x_3/x_1x_2)$	1	$SSR(x_3x_4/x_1x_2)-SSR(x_4/x_1x_2x_3)$
$SSR(x_4/x_1x_2x_3)$	1	$\frac{\widehat{\beta}_4^2}{C_{44}}$
$E(x_1)$		
$E(x_1x_2)$		
$E(x_1x_2x_3)$		
$E(x_1x_2x_3x_4)$		

إن الهدف من مصادر التباين التعاقبية هو لمعرفة إمكانية حذف المتغيرات غير المهمة من النموذج الكلي

طريقة دولتل لإيجاد $\hat{\beta}$ Doolittle Method to Find

باستخدام هذه الطريقة يمكن إيجاد $(X'X)^{-1}$ ، $\hat{\beta}$ ، SS،SSR التعاقبية وتحتاج هذه الطريقة إلى $(X'X)$ والمتجه $(X'Y)$ ، وكمثال على ذلك نفرض أن لدينا:

$$(XX)=\begin{bmatrix}10 & 10 & 5\\ & 20 & 15\\ & & 22.5\end{bmatrix} \qquad X'Y=\begin{bmatrix}20\\40\\45\end{bmatrix} \qquad S_{yy}=122.5$$

N	$(X'X)$			$X'Y$	I			Total
1	10	10	5	20	1	0	0	46
2		20	15	40	0	1	0	86
3			22.5	45	0	0	1	88.5
4	10	10	5	20	1	0	0	46
5	1	1	0.5	2	0.1	0	0	4.6
6		10	10	20	−1	1	0	20
7		1	1	2	−0.1	0.1	0	2
8			10	15	0.5	−1	1	25.5
9			1	1.5	0.05	−0.1	0.1	2.55

$$(X'X)^{-1}=\begin{bmatrix}0.225 & -0.15 & 0.05\\ & 0.2 & -0.1\\ & & 0.1\end{bmatrix}$$

$$\hat{\beta}_0 = 1*2+-1*2+0.5*1.5=0.75$$

$$\hat{\beta}_1 = 0*2+1*2+(-1)*1.5=0.5$$

$$\hat{\beta}_2 = 1.5$$

$$SSR(X_0)=n\bar{y}^2=20*2=40$$

$$SSR(X_1)=20*2=40$$

$$\mathrm{SSR}(x_2/x_1)=15*1.5=22.5$$

$$\mathrm{SSR}(x_0x_1x_2)=20*2+20*2+15*1.5=102.5$$

$$\mathrm{SSR}(x_1x_2)=SSR(X_1)+\mathrm{SSR}(x_2/x_1)=40+22.5=62.5$$

S.O.V.	DF	SS	MS	F
$SSR(x_1, x_2)$	2	62.5	31.25	
$SSR(x_1/x_0)$	1	40		3.65
$SSR(x_2 / x_1x_0)$	1	22.5		
Error	7	60	8.57	

S.O.V	D.F
$R(X_1X_2X_3)$	3
$\boldsymbol{R(X_3/X_1X_2)}$	1
$\boldsymbol{R(X_3X_2/X_1)}$	2
$E(X_1X_2X_3)$	n−4
Total	**n−1**

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$$
$$H_1 : \beta_1 \neq \beta_2 \neq \beta_3 \neq 0$$
$$H_0 : \beta_3 / \beta_1\beta_2 = 0$$
$$H_1 : \beta_3 / \beta_1\beta_2 \neq 0$$
$$H_0 : (\beta_3 = \beta_2) / \beta_1 = 0$$
$$H_1 : (\beta_3 = \beta_2) / \beta_1 \neq 0$$

العلاقة بين حدود الثقة واختبار الفرضيات

$$\hat{\beta}_j - t(\frac{\alpha}{2}, n-p)S(\hat{\beta}_j) \leq \beta_j \leq \hat{\beta}_j + t(\frac{\alpha}{2}, n-p)S(\hat{\beta}_j)$$

$$\hat{y}_0 - t(\frac{\alpha}{2}, n-p)S(\hat{y}_0) \leq y_0 \leq \hat{y}_0 + t(\frac{\alpha}{2}, n-p)S(\hat{y}_0)$$

$$\frac{(n-m-1)\hat{\sigma}^2}{\chi^2(\frac{\alpha}{2}, n-m-1)} \leq \sigma^2 \leq \frac{(n-m-1)\hat{\sigma}^2}{\chi^2(1-\frac{\alpha}{2}, n-m-1)}$$

أن العلاقة بين الثقة C.I واختبار الفرضيات H.t تقبل فرضية العدم

مثلا كانت الفرضية التالية

$$H_0 : \beta_3 / \beta_1\beta_2 = 6$$
$$H_1 : \beta_3 / \beta_1\beta_2 \neq 6$$

وكانت حدود الثقة

$$1 \le \beta_3 / \beta_1\beta_2 \le 8$$

مثال: لنفس بيانات المثال السابق

1- جد حدود الثقة ل β_1 و β_3 واستخدمها لاختبار الفرضيتين

$$H_0 : \beta_1 = 0 \quad H_1 : \beta_1 \neq 0$$

وكذلك

$$H_0 : \beta_3 = 0 \quad H_1 : \beta_3 \neq 0$$

2-ثم بين نسبة تأثير المتغيرات

أ- X_1 على التغيرات الحاصلة في y

ب- X_1X_2على التغيرات الحاصلة في y باستخدام معيار التباين

ت- $X_1X_2X_3$ على التغيرات الحاصلة في y باستخدام مصادر التباين التعاقبية التعاقبية

Sol:

$$\hat{\beta}_j - t(\frac{\alpha}{2}, n-p)S(\hat{\beta}_j) \le \beta_j \le \hat{\beta}_j + t(\frac{\alpha}{2}, n-p)S(\hat{\beta}_j)$$

$$H_0 : \beta_1 = 0$$
$$H_1 : \beta_1 \neq 0$$

$$\hat{\beta}_1 = 1.4$$

$$V(\hat{\beta}_1) = \hat{\sigma}^2 C_{11} = 2.013(0.053541) = 0.108$$
$$S(\hat{\beta}_1) = \sqrt{0.108} = 0.33$$

$$t(\frac{O.O5}{2}, 6) = 2.447$$

$$\hat{\beta}_1 - t(\frac{O.O5}{2}, 6)S(\hat{\beta}_1) \le \beta_1 \le \hat{\beta}_1 + t(\frac{O.O5}{2}, 6)S(\hat{\beta}_1)$$

$$-2.20751 \le \beta_1 \le -0.59$$

وبهذا ترفض فرضية العدم وتقبل البديلة أي ان لx اهمية في تفسير التغيرات الحاصلة في y

$H_0 : \beta_3 = 0$

$H_1 : \beta_3 \neq 0$

$$\hat{\beta}_3 - t(\frac{O.O5}{2},6)S(\hat{\beta}_3) \leq \beta_3 \leq \hat{\beta}_3 + t(\frac{O.O5}{2},6)S(\hat{\beta}_3)$$

$V(\hat{\beta}_3) = \hat{\sigma}^2 C_{33} = 0.0742$

$S(\hat{\beta}_3) = 0.272$

$2.5 \leq \beta_3 \leq 3.866$

إذا تقبل فرضية العدم لوقوعها داخل حدود الثقة أي أن $\hat{\beta}_3$ تعود فعلا إلى مجتمع $\hat{\beta}_3$ فيه تسوي (3)

2-

$$\therefore \text{SSR}(x_2 x_3 / x_1) = \begin{bmatrix} \hat{\beta}_2 & \hat{\beta}_3 \end{bmatrix} * \begin{bmatrix} C_{22} & C_{23} \\ c_{32} & C_{33} \end{bmatrix}^{-1} * \begin{bmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} =$$

$$= 320.9582$$

$$\therefore \text{SSR}(x_3 / x_1 x_2) = \frac{\hat{\beta}_3^2}{C_{33}} = 277.73257$$

$$\therefore \text{SSR}(x_2 x_3 / x_1) = \text{SSR}(x_2 x_3 / x_1) - \text{SSR}(x_3 / x_1 x_2)$$

$$= 43.2256$$

$$\text{SSR}(x_1 x_2 x_3) = \text{SSR}(x_1) + \text{SSR}(x_2 / x_1) + \text{SSR}(x_3 / x_1 x_2)$$

=

$$\text{SSR}(x_1) = \text{SSR}(x_1 x_2 x_3) - \text{SSR}(x_2 x_3 / x_1)$$

$$= 406.216 - 32O.74074784 = 85.26816$$

$SST = \underline{y}' \underline{y} - n\bar{y}^2 = 418.716$

$R^2 = \frac{SSR(X_1)}{SST} = \frac{85.67818}{418.716} = 0.201$ نسبة تأثير X_1

$R^2 = \frac{SSR(X_1 X_2)}{SST} = \frac{128.48336}{418.716} = 0.31$ نسبة تأثير $X_1 X_2$

$R^2 = \frac{SSR(X_1 X_2 X_3)}{SST} = \frac{406.216}{418.716} = 0.97$ نسبة تأثير $X_1 X_2 X_3$