

$$\hat{\alpha} - \hat{\beta} = \bar{y} - \frac{\sum xy}{n} = \frac{1}{n} \sum y_i (1 - x_i) \quad \blacksquare$$

مثال ۴) مدل رگرسیون خطی ساده $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ با فرض $\varepsilon_i : N(0, \sigma^2)$ و $\sum x_i = 0$ توزیع $\hat{\beta}_0$ به روش کمترین مربعات خطا کدام است ؟

$$N\left(\beta_0, \frac{\sigma^2}{\sum x_i^2}\right) \quad (۲)$$

$$N\left(\beta_0, \frac{\sigma^2}{n}\right) \quad (۱)$$

$$N\left(0, \frac{\sigma^2}{\sum x_i^2}\right) \quad (۴)$$

$$N\left(\beta_0, \sigma^2 \left(\frac{1}{n} + \frac{1}{\sum x_i^2}\right)\right) \quad (۳)$$

پاسخ : گزینه ۱

زیرا $\hat{\beta} = \bar{y} - \hat{\beta}_1 \bar{x} = \bar{y}$ و $N\left(\beta_0 + \beta_1 \bar{x}, \frac{\sigma^2}{n}\right)$ است و چون $\sum x_i = 0$ می باشد گزینه ۱ صحیح است. $\blacksquare$

۱-۳ برآورد واریانس در مدل رگرسیون خطی ساده

در مدل رگرسیون خطی ساده مجموع مربعات^۱ خطا یا مجموع مربعات باقیمانده ها به صورت زیر تعریف می شود :

$$SS_{Res} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum y_i^2 - n\bar{y}^2 - \hat{\beta}_1 S_{xy}$$

و از طرفی $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$ است ، بنابر این می توان مجموع مربعات خطا را به صورت زیر تعریف کرد :

^۱ Sum of squares (SS)

$$SS_{Res} = \sum y_i^2 - n\bar{y}^2 - \hat{\beta}_1 S_{xx}$$

که درجه آزادی آن برابر $n-2$ است و ثابت می شود که

$$E\left(\frac{SS_{Res}}{n-2}\right) = \sigma^2$$

و از طرفی از تقسیم مجموع مربعات خطا به درجه آزادی، میانگین مربعات خطا^۱ حاصل می شود که آن را با MS_E یا MS_{Res} نشان می دهند. بنابر این MS_{Res} یک برآورد کننده نااریب برای σ^2 می باشد.

مثال ۵) براساس اطلاعات زیر برآورد σ^2 را بدست آورید؟

$$n=20 \quad \sum Y = 42627.15 \quad \sum Y^2 = 92547433.45 \quad S_{XY} = -41112.65 \quad S_{XX} = 1106.56$$

پاسخ:

$$SS_{Res} = \sum Y^2 - n\bar{Y}^2 - \hat{\beta}_1 S_{XY} = 166402.65$$

$$\hat{\sigma}^2 = \frac{SS_{Res}}{n-2} = 9244.59$$

۱-۴-۱ آزمون فرضیه ها و فواصل اطمینان در رگرسیون

۱-۴-۱ آزمون رگرسیونی

در مدل رگرسیون خطی ساده $i = 1, 2, \dots, n$ $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ آزمون رگرسیونی به صورت زیر تعریف می شود:

$$\begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$$

^۱ Mean squares error

برای انجام آزمون فرض بالا جدول تحلیل واریانس را به صورت زیر تشکیل می شود:

جدول ۱ آنالیز واریانس در رگرسیون خطی ساده

منبع تغییرات	مجموع مربعات	درجه آزادی	میانگین مربعات	آماره آزمون
رگرسیون	SS_R	۱	$MS_R = \frac{SS_R}{1}$	$F = \frac{MS_R}{MS_{Res}}$
خطا	SS_{Res}	$n - 2$	$MS_{Res} = \frac{SS_{Res}}{n - 2}$	
کل	SS_T	$n - 1$		

$$SS_T = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n}$$

$$SS_R = \sum (\hat{y}_i - \bar{y})^2 = \hat{\beta}_1 S_{xy}$$

$$SS_{Res} = \sum (y_i - \hat{y}_i)^2 = SS_T - SS_R$$

چنانچه آماره آزمون F_{obs} بزرگتر از مقدار بحرانی $F_\alpha(1, n-2)$ باشد، فرض H_0 رد می شود و آزمون رگرسیونی معنادار می باشد.

۱-۴-۲ آزمون فرض ها عرض از مبدا و شیب

اگر مقدار a معلوم باشد آزمون فرضهای عرض از مبدا به صورتهای زیر

می باشند:

(الف)

$$\begin{cases} H_0 : \beta_0 = a \\ H_1 : \beta_0 < a \end{cases}$$

آماره آزمون عبارت است از :

$$t_{obs} = \frac{\hat{\beta}_0 - a}{\sqrt{MS_{Res} \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)}}$$

اگر $t_{obs} < -t_{\alpha}(n-2)$ باشد فرض H_0 رد می شود.

(ب)

$$\begin{cases} H_0: \beta_0 = a \\ H_1: \beta_0 > a \end{cases}$$

آماره آزمون عبارت است از:

$$t_{obs} = \frac{\hat{\beta}_0 - a}{\sqrt{MS_{Res} \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)}}$$

اگر $t_{obs} > +t_{\alpha}(n-2)$ باشد، فرض H_0 رد می شود.

(ج)

$$\begin{cases} H_0: \beta_0 = a \\ H_1: \beta_0 \neq a \end{cases}$$

اگر $|t_{obs}| > t_{\alpha/2}(n-2)$ باشد، فرض H_0 رد می شود.

اگر مقدار b معلوم باشد آزمون فرضهای شیب به صورتهای زیر می باشند:

(الف)

$$\begin{cases} H_0: \beta_1 = b \\ H_1: \beta_1 < b \end{cases}$$

آماره آزمون عبارت است از :

$$t_{obs} = \frac{\hat{\beta}_1 - b}{\sqrt{\frac{MS_{Res}}{S_{xx}}}}$$

اگر $t_{obs} < -t_{\alpha}(n-2)$ باشد H_0 رد می شود .

(ب)

$$\begin{cases} H_0 : \beta_1 = b \\ H_1 : \beta_1 > b \end{cases}$$

اگر مقدار آماره آزمون t_{obs} بزرگتر $+t_{\alpha}(n-2)$ از باشد ، فرض H_0 رد می شود .

(د)

$$\begin{cases} H_0 : \beta_1 = b \\ H_1 : \beta_1 \neq b \end{cases}$$

اگر $|t_{obs}| > t_{\alpha/2}(n-2)$ باشد ، فرض H_0 رد می شود .

نکته : در آزمون فرض $\begin{cases} H_0 : \beta_1 = 0 \\ H_1 : \beta_1 \neq 0 \end{cases}$ مقدار مجذور آماره آزمون t_{obs} با آماره آزمون

فیشر F_{obs} در جدول تحلیل واریانس برابر است .

$$t_{obs}^2 = \left(\frac{\hat{\beta}_1 - 0}{\sqrt{\frac{MS_{Res}}{S_{xx}}}} \right)^2 = \frac{MS_R}{MS_{Res}} = F_{obs}$$