04/21/2021

2.28. $\underset{b_1}{Y = -1.19X} + \underset{b_0}{156.35}$

a. $\hat{Y}_h = -1.19 \times 60 + 156.35 = 84.95$

$$S^2(\hat{Y}_h) = MSE\left(\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\Sigma(X_i - \bar{X})^2}\right)$$
$$= 1.13$$

$\therefore S(\hat{Y}_h) = 1.055$

$$\hat{Y}_h \pm \underbrace{t(1-\alpha/2, 58)}_{2.002} \cdot S(\hat{Y}_h)$$

$\therefore [82.835, 87.059]$

b. $S^2(\hat{Y}_{h(new)}) = S^2(\hat{Y}_h) + MSE = 67.91$

$\therefore S(\hat{Y}_{h(new)}) = 8.24$

The prediction interval is $[68.45, 101.44]$

ANOVA

$$SSTO = \Sigma(Y_i - \bar{Y})^2 = \Sigma(Y_i - \hat{Y}_i + \hat{Y}_i - \bar{Y})^2$$
$$= \Sigma(Y_i - \hat{Y}_i)^2 + \Sigma(\hat{Y}_i - \bar{Y})^2 + \underbrace{2\Sigma(Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{Y})}_{0 \text{ as } \Sigma\hat{Y}_i = \Sigma Y_i}$$
$$= SSE + SSR$$

df $\quad n-1 \quad n-p \quad \underset{\#\text{predictors}}{p-1}$

$b_1 = \frac{\Sigma(X_i - \bar{X})(Y_i - \bar{Y})}{\Sigma(X_i - \bar{X})^2}$ $\quad \Sigma X_i Y_i - n\bar{X}\bar{Y}$

In simple LR, $p = 1 + \underset{\text{intercept}}{1} = 2$

$b_0 = \bar{Y} - b_1\bar{X}$

$b_1^2\Sigma(X_i - \bar{X})^2$ $\quad \therefore E(SSR) = \Sigma(X_i - \bar{X})^2 \overbrace{E(b_1^2)}^{Var(b_1)+(E(b_1))^2}$

Source of variation	SS	df	MS	E(MS)
Regression	SSR	1	SSR/1	$\sigma^2 + \beta_1^2\Sigma(X_i - \bar{X})^2$
Error	SSE	$n-2$	SSE/(n-2)	σ^2
Total	SSTO	$n-1$		

$\Sigma Y_i^2 - n\bar{Y}^2$

e.g.							mean
X_i	43	39	41	86	72	76	59.5
Y_i	106	106	97	60	70	80	86.5

$\Sigma(X_i - \bar{X})^2$ $\quad$ 2165.5

$\Sigma(Y_i - \bar{Y})^2$ $\quad$ 1887.5 $\quad$ SSTO

$\Sigma(X_i - \bar{X})(Y_i - \bar{Y})$ $\quad$ −1931.5

$$SSR = b_1^2\Sigma(X_i - \bar{X})^2 = \frac{(\Sigma(X_i - \bar{X})(Y_i - \bar{Y}))^2}{\Sigma(X_i - \bar{X})^2} = 1722.786$$

$b_1 = \frac{\Sigma(X_i - \bar{X})(Y_i - \bar{Y})}{\Sigma(X_i - \bar{X})^2}$ $\quad \therefore SSE = SSTO - SSR = 164.7$

Source of variation	SS	df	MS	F	P-value
Regression		1	MSR	$\frac{MSR}{MSE} = \frac{SSR/1}{SSE/(n-2)} \sim F_{1,n-2}$	$P(F \geq F^{(obs)})$
Error		4	MSE		
Total		5			

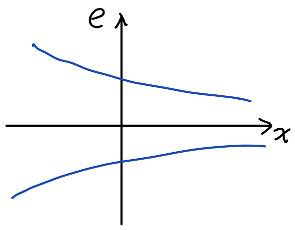
2.29 b

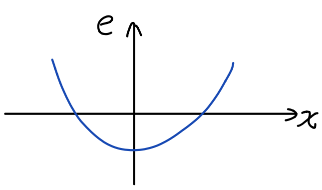
Source of variation	SS	df	MS
Regression	11627.5	1	11627.5
Error	3874.4	58	66.8
Total	15501.9	59	

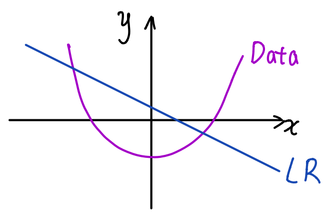
d. $1-R^2 = 1-\frac{SSR}{SSTO} = 1-0.75 = 0.25$

e. $R^2 = \frac{SSR}{SSTO}$, $r = -\sqrt{R^2} = -0.866$

3.2 (a)







3.9 Page 105

Nonlinearity of Regression Function

e.g.

$z = X^2$, $Y = \beta_0 + \beta_1 z + \varepsilon$

Nonindependence of Error Terms

Time series model

or include more predictors.