

ISyE403 Regression and Forecasting
Practice Problems 1 Solutions
Spring 2016

1.a. $a = 2.7738/0.1846 = 15.02$.

b. $b = 11392/25.5833 = 445.29$.

c. $s = \sqrt{25.5833} = 5.05$.

d. $R^2 = 22784/23091 = 0.987$.

2. a. $\hat{\beta}_1 = SS_{xy}/SS_{xx} \Rightarrow SS_{xx} = 16.22/3.4 = 4.77$.

b. $SSE = SS_{yy} - \hat{\beta}_1 S_{xy} = 4.062 \Rightarrow SS_{yy} = 4.062 + 3.4(16.22) = 59.21$.

3. a. We test $H_0: \beta_1 = \beta_2 = \beta_3 = 0$ vs. H_a : at least one β is not 0. Since

$F(\text{model}) = 35.51 > F_{0.05, 3, 15} = 3.29$, we reject H_0 . Rejecting H_0 implies the linear regression model as a whole is useful. Corresponding p value $= 0 < 0.05$ (or any α) \Rightarrow reject H_0 . It confirms the conclusion.

b. $\hat{y} = -40.7 + 0.00362(1531) + 1.23(21.3) + 4.76(7.6) = 27.217$

$y - \hat{y} = 29 - 27.25 = 1.783$ (or 1.756).

c. $R\text{-Sq} = 87.7\%$ of the total variability in homicide rate is explained by the regression.

d. We test $H_0: \beta_2 = 0$ vs. $H_a: \beta_2 \neq 0$. Test statistic, $t = 2.6 > t_{0.025, 15} = 2.131$. So, we reject H_0 which implies X_2 is a significant predictor.

e. When $\alpha = 0.01$, X_1 and X_3 are statistically significant, since their p -values < 0.01 . The predictor X_2 is not significant, since its p -value > 0.01 .

4. a. True

b. False

c. False

d. False

e. True

5. a. Proved.

$$\begin{aligned} \sum_{i=1}^n e_i &= \sum_{i=1}^n (y_i - \hat{y}_i) = \sum_{i=1}^n (y_i - \bar{y} - \hat{\beta}_1(x_i - \bar{x})) = \sum_{i=1}^n y_i - \sum_{i=1}^n \bar{y} - \hat{\beta}_1 \sum_{i=1}^n (x_i - \bar{x}) \\ &= n\bar{y} - n\bar{y} - \hat{\beta}_1(n\bar{x} - n\bar{x}) = 0. \end{aligned}$$

b. Disproved.

$$\begin{aligned} \sum_{i=1}^n e_i &= \sum_{i=1}^n (y_i - \hat{y}_i) = \sum_{i=1}^n (y_i - \bar{y} - \hat{\beta}_1 x_i) = \sum_{i=1}^n y_i - \sum_{i=1}^n \bar{y} - \hat{\beta}_1 \sum_{i=1}^n x_i \\ &= n\bar{y} - n\bar{y} - \hat{\beta}_1 n\bar{x} \neq 0 \text{ (true only if } \bar{x} = 0). \end{aligned}$$

6. Short-answer questions.

a. iv. R^2 goes up.

b. ii. The length of the estimated prediction interval would be decreased as the value of x_p gets closer to \bar{x} .

c. True

d. True

e. True

f. False

g. False