Assignment 1: Selected solutions

September 30, 2021

- Problem 1: Omitted (reasonable explanation will be credited)
- Problem 2: sample mean $\bar{x} = \frac{1}{20}\sum_{i=1}^{20} x_i = 311.15$, sample variance $S^2 = \frac{1}{19}\sum_{i=1}^{20}(x_i - \bar{x})^2 = 4146.45$, then the t-statistics:
$$|t| = \left|\frac{\bar{x} - 200}{S/\sqrt{20}}\right| = 7.719448 > t_{19,0.01} = 2.539$$
Hence we reject the null hypothesis H_0.
- Problem 3:
  (a) Disagree. $|Cor(Y, X)| \leq 1$
  (b) Disagree. This only implies no linear relationship between X and Y.
  (c) Agree. This follows from the property of least square solutions.
  (d) Agree. Consider the least square line fitted to $\{(\hat{Y}_i, Y_i)\}_{i=1}^n$ with intercept $\hat{\alpha}_0$ and slop $\hat{\alpha}_1$:
$$\hat{\alpha}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2} = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{y})}{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2} = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{\beta}_0 + \hat{\beta}_1 x_i - \hat{\beta}_0 - \hat{\beta}_1 \bar{x})}{\sum_{i=1}^n (\hat{\beta}_0 + \hat{\beta}_1 x_i - \hat{\beta}_0 - \hat{\beta}_1 \bar{x})^2}$$
$$= \frac{\hat{\beta}_1}{\hat{\beta}_1^{\,2}} \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = 1$$
$$\hat{\alpha}_0 = \bar{y} - \hat{\alpha}_1 \bar{\hat{y}} = \bar{y} - \bar{y} = 0$$
(However, the reasoning above holds only if $\hat{\beta}_1 \neq 0$. So it's okay if you disagree the statement with the counterexample when $\hat{\beta}_1 = 0$)
- Problem 4:
  (a) Since
$$|t_1| = \left|\frac{\hat{\beta}_1 - 15}{\text{s.e.}(\hat{\beta}_1)}\right| = 1.007921 < t_{12,0.025} = 2.178813$$
Hence we can not reject H_0.
  (b) Since
$$t_1 = \frac{\hat{\beta}_1 - 15}{\text{s.e.}(\hat{\beta}_1)} = 1.007921 < t_{12,0.05} = 1.782288$$
Hence we can not reject H_0.

(c) Since

$$|t_1| = \left| \frac{\hat\beta_0}{\text{s.e.}(\hat\beta_0)} \right| = 1.240537 < t_{12,0.025} = 2.178813$$

Hence we can not reject H_0.

(d) Since

$$|t_1| = \left| \frac{\hat\beta_0 - 5}{\text{s.e.}(\hat\beta_1)} \right| = 0.2497765 < t_{12,0.025} = 2.178813$$

Hence we can not reject H_0.

- Problem 5:

$$\hat\beta_0 \pm t_{12,0.01} \times \text{s.e.}(\hat\beta_0) = [-4.833, 13.157]$$

- Problem 6:

(a) Define

$$S(\beta_0) = \sum_{i=1}^{n} (y_i - \beta_0)^2 = n\beta_0^2 - 2\beta_0 \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} y_i^2$$

Hence

$$\hat\beta_0 = \min_{\beta_0} S(\beta_0) = \bar y$$

(b)

$$e_i = y_i - \hat y_i = y_i - \bar y$$

(c)

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} (y_i - \bar y_i) = n\bar y - n\bar y = 0$$

- Problem 7: Omitted.

- Problem 8:

(a) SST $=$ SSR $+$ SSE $= 0.0902$, hence $\text{Var}(Y) = \frac{1}{n-1}\text{SST} = 0.00501$. $|\text{Cor}(Y, X)| = \sqrt{R^2} = 0.630$. Since $\text{Cor}(Y, X)$ and $\hat\beta_1$ have the same sign, we know $\text{Cor}(Y, X) = 0.630$.

(b) At $x_0 = 0.45$,

$$\hat y_0 = \hat\beta_0 + \hat\beta_1 x_0 = 0.498529$$

(c) Confidence interval corresponds to the mean response.

$$\text{s.e.}(\hat y_0) = \hat\sigma \sqrt{\frac{1}{n} + \frac{(x_0 - \bar x)^2}{\sum_{i=1}^{n} (x_i - \bar x)^2}} = 0.0160498$$

Hence

$$\hat y_0 \pm t_{17,0.025} \times \text{s.e.}(\hat y_0) = [0.4646669, 0.5323911]$$

(d)

$$\hat\beta_1 \pm t_{17,0.025} \times \text{s.e.}(\hat\beta_1) = [0.2423052, 1.069775]$$

(e)

$$t_1 = \frac{\hat{\beta}_1 - 1}{\text{s.e.}(\hat{\beta}_1)} = -1.754003 < t_{17,0.01}$$

We can not reject H_0.

(f) $R^2 = 0.397$.

- Problem 9:

(a) Omitted.

(b) Define

$$S(\beta_1) = \sum_{i=1}^{n}(y_i - \beta_1 x_i)^2 = \beta_1^2 \sum_{i=1}^{n} x_i^2 - 2\beta_1 \sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} y_i^2$$

Hence

$$\hat{\beta}_1 = \min_{\beta_1} S(\beta_1) = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

(c) Consider a special case, for example, $\bar{x} = 0$ but $\bar{y} \neq 0$, then

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n}(y_i - \hat{\beta}_1 x_i) = \sum_{i=1}^{n}(y_i - \frac{\sum_{j=1}^{n} x_j y_j}{\sum_{j=1}^{n} x_j^2} x_i) = \sum_{i=1}^{n} y_i - \frac{\sum_{j=1}^{n} x_j y_j}{\sum_{j=1}^{n} x_j^2} \sum_{i=1}^{n} x_i \neq 0$$

- Problem 10: See R file.