Homework 10 - STAT 231

Suggested Solution [*,†]

Due in class: 21 Nov 2019

Problem 1-5 are from the Devore textbook.

1. **(Chapter 12: Problem 7)**

   (a)
   $$\mu_{y|x=2500} = 1800 + 1.3 * 2500 = 5050$$

   (b) We expect 28-strendth (y) to increase by 1.3 psi when accelerated strength (x) increases by 1 psi
   $$dY/dX = 1.3$$

   (c) We expect 28-strendth (y) to increase by $1.3 * 100 = 130$ psi when accelerated strength (x) increases by 100 psi

   (d) We expect 28-strendth (y) to decrease by $1.3 * 100 = 130$ psi when accelerated strength (x) decreases by 100 psi

2. **(Chapter 12: Problem 8 (a)(b))**

   (a)
   $$\hat{y}|_{x=2000} = 1800 + 1.3 * 2000 = 4400$$
   $$P(y > 5000|x = 2000) = P((y - \mu_{y|x=2000})/350 > (5000 - \mu_{y|x=2000})/350)$$
   $$= P(Z > (5000-4400)/350)) = 1-N(600/350) = 1-N(1.714286) = 0.043238$$

   (b)
   $$\hat{y}|_{x=2500} = 1800 + 1.3 * 2500 = 5050$$

[*]Zhiling GU (zlgu@iastate.edu)

[†]Please inform me if there is any error or possible improvement

$$P(Y > 5000|x = 2500) = P((y - \mu_{y|x=2500})/350 > (5000 - \mu_{y|x=2500})/350)$$

$$= P(Z > (5000-5050)/350)) = 1-N(-50/350) = N(0.1428571) = 0.5567985$$

3. **(Chapter 12: Problem 14)**

   (a) Let y denote Ratio, let x denote temperature, from the JMP output:

   $$y = -15.24497 + .0942361x$$

   (b)

   $$\hat{y}|_{x=182} = -15.24497 + .0942361 * 182 \approx 1.906$$

   (c) Residuals $= y_i|x = 182 - E(y|x = 182) = (.90, .181, 1.94, 2.68) - 1.906 = (-1.006, -1.725, 0.034, 0.774)$ respectively. They are not on the same side of zero because there is random effect that can take both posotove and negative values.

   (d) $R^2 = SSR/SST = 4.4757442/9.9152958 \approx 0.451398$. Therefore around 45.14% of the variation in efficiency ratio can be attributed to the simple linear regression relationship between the two variables.

4. **(Chapter 12: Problem 19)** JMP output:

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	398030.26	398030	294.7380
Error	12	16205.45	1350	**Prob > F**
C. Total	13	414235.71		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-45.55191	25.46779	-1.79	0.0989
x	1.7114323	0.099688	17.17	<.0001*

   (a) From the JMP output:

   $$y = -45.5519 + 1.7114x$$

   (b) $\hat{y}|_{x=225} = -45.5519 + 1.7114 * 225 = 339.5131$

   (c) $dy/dx = 1.7114$. We expect the NO_x to decrease by $50 * 1.7114 = 85.57$ when x is decreased by 50.

(d) No. 500 is out of the range of x observations. Extrapolation is dangerous.

5. **(Chapter 12: Problem 52 (b) - (f))**

(a) From the JMP output, we see the F-statistic for

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

is 112.7559, and the corresponding P-value $< .0001$. Therefore there is sufficient evidence to reject H_0, i.e. the simple linear regression model specifies a useful relationship between chlotine flow and etch rate.

(b) From the JMP output, we have

$$y = 6.4487 + 10.6026x$$

therefore a unit change in flow rate (x), we expect β_1 unite change in etch rate. The 95% confidence interval for β_1 is $\hat{\beta}_1 \pm t_{n-2,\alpha/2}\hat{SE}(\beta_1) = 10.6025 \pm 2.4 * .998484 = (8.206, 12.999)$

(c) 95% CI for $\mu_{Y|x^*=3.0}$ is

$$\hat{y}|_{x=x^*} \pm t_{n-2,\alpha/2}\hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum(x_i - \bar{x})^2}}$$

where $\hat{y} = 6.4487 + 10.6026 * 3.0 = 38.2565$, $\hat{\sigma} = \sqrt{MSE} = \sqrt{6.48} = 2.545584$, $\sum(x_i - \bar{x})^2 = s_x^2 * (n-1) = .9013878^2 * (9-1) = 6.5$.

$$95\%CI = 38.2565 \pm 2.4 * 2.545584 * \sqrt{\frac{1}{9} + \frac{(3.0 - 2.6667)^2}{.9013878^2 * (9-1)}} = (36.06, 40.44)$$

(d) 95% Prediction Interval for $y|_{x^*=3.0}$ is $\hat{y} \pm t_{n-2,\alpha/2}\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(x^*-\bar{x})^2}{\sum(x_i-\bar{x})^2}}$, similary we get

$$95\%PI = 38.2565 \pm 2.4 * 2.545584 * \sqrt{1 + \frac{1}{9} + \frac{(3.0 - 2.6667)^2}{.9013878^2 * (9-1)}} = (31.76729, 44.74571)$$

(e) The 95% CI and PI when $x^* = 2.5$ will be narrower, since it is closer to

$\bar{x} = 2.6667$

(f) No. 6 is not within the range of the observationo of x. Extrapolation can be dangerous, and the prediction out of range can be meaningless.