Pharmacy 2012 - Winter 2020 - Biostatistics - Assignment 4 solutions

(out of 40 points)

1. Data from a sample of 10 pharmacies are used to examine the relationship between prescription sales volume and the percentage of prescription ingredients purchased directly from the supplier. The sample data are shown below:

Pharmacy	Sales Volume, y (in \$10,000)	% Ingredients Purchased directly, x
1	25	10
2	55	18
3	50	25
4	75	40
5	110	50
6	138	63
7	90	42
8	60	30
9	10	5
10	100	55

The data summaries are $\bar{y} = 71.3$, $\bar{x} = 33.8$, $SXY = 6714.6$, $SXX = 3407.6$, $SYY = 13882.1$.

(3) (a) Calculate Pearson's correlation coefficient. What proportion of the variation in Sales is explained by % Ingredients purchased directly?

- The correlation coefficient is

$$r = \frac{SXY}{\sqrt{SXX \times SYY}} = \frac{6714.6}{\sqrt{3407.6(13882.1)}} = .9762$$

- The proportion of variation explained is $r^2 = .9762^2 = .9530$

(7) (b) Find the equation of the least squares line.

- The slope estimate is

$$b = \frac{SXY}{SXX} = \frac{6714.6}{3407.6} = 1.970478$$

- The intercept estimate is

$$a = \bar{y} - b\bar{x} = 71.3 - 1.97(33.8) = 71.3 - 66.586 = 4.714$$

- The equation of the regression (least squares) line is

$$\hat{y} = 4.714 + 1.9705\hat{x}$$

or:

$$\hat{Sales} = 4.714 + 1.9705\%purchased$$

(3 points for each of slope and intercept; 1 point for writing the equation).

2. A study involving 42 subjects found that bone mineral density (BMD) in g/cm^2, measured at the left femural neck, was related to Weight (in kg) according to the least squares equation

$$BMD = 0.47 + .0049Weight$$

(1) (a) What is the predicted BMD for a subject with weight 80kg?

- The predicted BMD is

$$.47 + .0049(80) = .862$$

(2) (b) What increase in BMD is expected with an increase of weight of 5 kg?

- The expected increase is $5 * b = 5(.0049) = .0245$

(2) (c) What weight is predicted with a BMD of .8 g/cm^2?

- We turn the equation around, so

$$Weight = \frac{BMD - .47}{.0049} = \frac{.8 - .47}{.0049} \approx 67.35$$

(d) Assess the hypothesis that there is no association between Weight and BMD. Use the fact that the standard error of the slope estimate is .0020

(2) i. State the hypotheses.

- $H_0 : \beta = 0$
- $H_a : \beta \neq 0$

(2) ii. Calculate the test statistic.

- The test statistic is

$$t = \frac{b - 0}{\hat{se}(b)} = \frac{.0049}{.0020} = 2.45$$

(1) iii. What are the degrees of freedom? 40 = 42-2

(1) iv. Determine the P value as accurately as possible.
The p-value is $2P(t_{40} > 2.45)$. From the tables, $P(t_{40} > 2.423) = .01$, and $P(t_{40} > 2.704) = .005$, so that the p-value is $> 2(.005)$ and $< 2(.01)$, or $.01 < p - value < .02$.

(4) (e) Calculate a 95% confidence interval for the slope coefficient.
The confidence interval is

$$b \pm t^{n-2}_{\alpha/2}\hat{se}(b)$$

where $t^{40}_{.025} = 2.021$.

so the interval is

$$(.0049 - 2.021(.0020) \quad , \quad .0049 + 2.021(.0020))$$

or, approximately, $(.00086, .00894)$.

3. Investigators wish to assess whether there is a difference in the efficacy of salbutamol and ipratropium bromide in the treatment of asthma. They will measure forced expiratory volume in 1 second (FEV_1) after two weeks of treatment. They wish to detect a difference of 0.2 liters with a two-sided alternative using $\alpha = .05$. Assume that the SD is 1.0 liter in both groups. (Note: for a two-sided alternative test with $\alpha = .05$, use $z_{\alpha/2} = 1.96$; for having a 80% power (i.e., $\beta = .2$), use $z_\beta = 0.84$; and for having a 90% power (i.e., $\beta = .1$), use $z_\beta = 1.28$ in calculation).

(5) (a) What would be the power of the test if they included 100 subjects in each group?

- $\delta = 0.2$ and $\sigma = 1.0$ are given.
- Use $z_{\alpha/2} = 1.96$ for a two-sided alternative in equation (3) on page 5 of the lecture note, so

$$\begin{aligned} Power &= 1 - \Phi(1.96 - \frac{0.2}{1.0\sqrt{2/100}}) \\ &= 1 - \Phi(0.546) \\ &= 1 - 0.7088 = 0.29. \end{aligned}$$

(5) (b) How large a sample in each group should they take to obtain 80% power?

- Use equation (4) in the lecture note.
- $z_{\alpha/2} = 1.96$ and $z_\beta = 0.84$

$$\begin{aligned} n &= \frac{2(1.0^2)(1.96 + 0.84)^2}{(0.2)^2} \\ &= 392 \end{aligned}$$

so they should take 392 in each group.

(5) (c) How large a sample in each group should they take to obtain 90% power?

- Again use equation (4) in the lecture note.
- $z_{\alpha/2} = 1.96$ and $z_\beta = .84$

$$\begin{aligned} n &= \frac{2(1.0^2)(1.96 + 1.28)^2}{(0.2)^2} \\ &= 524.88 \end{aligned}$$

so they should take 525 in each group (rounding up).

- (For marking: if the formula used is correct, but the calculated value is quite wrong, e.g., not like due to an rounding issue, then deduct 2 points for each answer that has this problem).