Solutions to Supplemental Problems for Stat 206 Review

Note: Solutions to odd-numbered problems from the text are at the back of the text. Some solutions to other problems are listed below. Those not listed below will be available seperately as pdf files.

Sampling distributions

See text for solutions.

Confidence Intervals and Estimation Problems

Problems 1-5: See text for solutions.

Problems 6-9: See seperate solutions.

Hypothesis Testing Problems

Problems 1-5: See text for solutions.

6. $H_0 : \mu = 0$
   $H_a : \mu > 0$

   $$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} = \frac{4.8 - 0}{4.4/\sqrt{19}} = 4.76$$

   The corresponding p-value is less than .005, since the t_{18} distribution has $\Pr(t > 2.878) = .005$. So at $\alpha = .05$ we reject H_0 and conclude that the plant is doing worse on average than the computer.

7. In (B), the significance level α should be smaller.

   α is the chance that if the null hypothesis is correct, you will (mistakenly) decide to reject it. That is, it's the chance that you decide to switch when in fact you shouldn't. The bigger the cost of switching, the more certain you want to be that you're making the right decision to switch. That is, you want the chance of mistakenly switching to be smallest when the cost of mistakenly switching is largest.

8. True. The p-value is the probability that (for example) a standard normal is less than the observed value of the test statistic. The observed value can't be calculated until after the data are observed.

9. $H_0 : \pi = .5$
   $H_a : \pi > .5$
   $n = 218, p = 126/218 = .578, \alpha = .05$
   test statistic $= z_{obs} = (p - \pi_0)/\sqrt{\pi_0(1-\pi_0)/n} = (.578 - .5)/\sqrt{.5(.5)/218} = 2.303$
   p-value $= \Pr(Z > z_{obs}) = \Pr(Z > 2.303) = .0107$ so we don't reject H_0. There's not sufficient evidence to conclude that the returning champ has a better than 50% chance of winning.

10. The results of the first test mean that the test statistic was not a large positive number. In the second test we would reject the null if the test statistic was large enough in either the positive or the negative direction. This might or might not have happened, and we have no way to tell without knowing the observed value of the test statistic. So it's not possible to tell whether we'd reject the null in second hypothesis test.

11. The p-value is calculated with the data.

    For example, if your test statistic has a standard normal distribution and the observed value of the test statistic is $z = 5.30$, the p-value would be one of $\Pr(Z > 5.30), \Pr(Z < 5.30)$ or $2\Pr(Z > 5.30)$, depending on whether the alternative hypothesis was $" > ", " < ", " \neq "$.

12. See seperate solutions

13. See seperate solutions

Regression problems

1. SS_1 will be smaller. This is because a and b are chosen to minimize the sum of squared distances between the Y_i values and the fitted line. Because $a \neq \beta_0$ and $b \neq \beta_1$, the constants β_0 and β_1 will not minimize the sum of squared distances for this data set.

2. Output 1 matches plot <u>1</u> (plot 1 has smallest s)
   Output 2 matches plot <u>4</u> (plot 4 many more data points, and same s as plots 2 and 3. Thus s_a and s_b are smaller)
   Output 3 matches plot <u>3</u> (after identifying output 4 this is the only one left)
   Output 4 matches plot <u>2</u> (plot 2 has an intercept of 4)

3. (Note - I used 2 instead of 1.96 in solutions below for a z-critical value corresponding to 95%)

   (a) $a \pm 2s_a = 4.1191 \pm 2(0.5152) = 4.1191 \pm 1.0304 = (3.0887, 5.1495)$

(b) $b \pm 2s_b = -0.28563 \pm 2(0.06771) = -0.28563 \pm 0.13542 = (-0.42105, -0.15021)$.

(c) $z = (-0.28563 - (-0.2))/0.06771 = -1.26$
We do not reject H_0 at $\alpha = 0.05$, since $-2 < z < 2$. Thus we conclude that $\beta_1 = -0.2$ is possible.

(d) $a = 4.12$ means that a notebook with no weight would last 4.12 hours.

$b = -0.286$ means that as weight increases, battery life decreases. Specifically, for every extra pound, the battery life decreases by 0.286 hours.

(e) In the formulas below, I use the following facts

- $s_z = 16s_x$, i.e. the standard deviation of the new predictor is 16 times as large.
- $\overline{z} = 16\overline{x}$, i.e. the mean of the new predictor is 16 times as large.
- $S_{zy} = 16S_{xy}$
- $S_{zz} = 16^2 S_{xx}$

i. $d = S_{zy}/S_{zz} = 16S_{xy}/(16^2 S_{xx}) = S_{xy}/(16S_{xx}) = b/16 = -0.28563/16 = -.01785$

ii. $c = \overline{y} - d\overline{x} = \overline{y} - \frac{b}{16}(16\overline{x}) = a = 4.1191$

iii. $s = \sum_{i=1}^{n}(Y_i - \hat{y}_i)^2/(n-2) = 0.5476$. That is, s doesn't change from the original model. The new fitted model has the same predicted values because $\hat{y} = c + dZ = a + \frac{b}{16}(16X_i) = a + bX_i$. Since the predictions are the same, the sum of squared errors is the same.

iv. $s_d = \frac{s}{\sqrt{S_{zz}}} = \frac{s}{\sqrt{16^2 S_{xx}}} = \frac{s}{16\sqrt{S_{xx}}} = s_b/16 = 0.06771/16 = 0.00423$.

4.

$$30 \pm 1.96(1)\sqrt{1 + \frac{5.2^2}{1.96^2}}$$