STA 570 — Spring 2012 — Dr. Charnigo

Written Assignment 6 Solutions

1a. [Using SAS for this and subsequent items is permissible.] We have $\bar{x} = 38.99$, $\bar{y} = 3419$, $L_{xx} = 3800$, $L_{yy} = 2.171 \times 10^8$, and $L_{xy} = 4.757 \times 10^5$. The least squares estimate of β is $b = 4.757 \times 10^5/3800 = 125.2$. The least squares estimate of α is $a = 3419 - 125.2 \times 38.99 = -1462$.

1b. We have Tot SS $= 2.171 \times 10^8$, Reg SS $=$ Reg MS $= (4.757 \times 10^5)^2/3800 = 5.955 \times 10^7$, Res SS $= 2.171 \times 10^8 - 5.955 \times 10^7 = 1.575 \times 10^8$, and Res MS $= 1.575 \times 10^8/698 = 2.257 \times 10^5$. Since $t_{698,.975} = 1.963$, the 95% confidence interval for α is

$$-1462 \pm 1.963\sqrt{2.257 \times 10^5(1/700 + 38.99^2/3800)} = -1462 \pm 591,$$

or -2053 to -871. The 95% confidence interval for β is

$$125.2 \pm 1.963\sqrt{2.257 \times 10^5/3800} = 125.2 \pm 15.1,$$

or 110.1 to 140.3. [Using $z_{.975} = 1.960$ as an approximation to $t_{698,.975}$ is permissible.]

1c. The f statistic is $5.955 \times 10^7/2.257 \times 10^5 = 263.9$, which easily exceeds $f_{1,698,.95} = 3.855$, so we reject $H_0 : \beta = 0$. [The p-value is less than 0.0001. Using $f_{1,120,.95} = 3.92$ or $f_{1,\infty,.95} = \chi^2_{1,.95} = 3.84$ as an approximation to $f_{1,698,.95}$ is permissible.]

1d. The t statistic is $125.2/\sqrt{2.257 \times 10^5/3800} = 16.24$, which easily exceeds $t_{698,.975} = 1.963$, so we reject $H_0 : \beta = 0$. [The p-value is less than 0.0001. Also, to check our work, we can note that $t^2 = 16.24^2 = 263.9 = f$.]

1e. We have $R^2 = 5.955 \times 10^7/2.171 \times 10^8 = 0.274$. This is the fraction of variability in birthweight accounted for by gestational age.

1f. We have $\hat{y} = -1462 + 125.2 \times 37 = 3170$. The 95% confidence interval for the mean birthweight among all infants with gestational age 37 weeks is

$$3170 \pm 1.963\sqrt{2.257 \times 10^5(1/700 + (37 - 38.99)^2/3800)} = 3170 \pm 46,$$

which is 3124 to 3216.

2a. [Using SAS for this and subsequent items is permissible.] We have $\hat{p}_1 = 48/127 = 0.3780$, $\hat{p}_2 = 26/73 = 0.3562$, $a = 48$ (Table 13.1 notation), $b = 79$, $c = 26$, $d = 47$, $n_1 = 127$, and $n_2 = 73$. The point estimate of the risk difference is $0.3780 - 0.3562 = 0.0218$. The 95% confidence interval is

$$0.0218 \pm 1.96\sqrt{0.3780(1 - 0.3780)/127 + 0.3562(1 - 0.3562)/73} = 0.0218 \pm 0.1385,$$

which is -0.1167 to 0.1603.

2b. The point estimate of the relative risk is $0.3780/0.3562 = 1.061$. The 95% confidence interval is

$$1.061 \exp[\pm 1.96\sqrt{79/(48 \times 127) + 47/(26 \times 73)}] = 1.061 \exp[\pm 0.3807],$$

which is 0.725 to 1.553.

2c. The point estimate of the odds ratio is

$$\frac{0.3780/(1-0.3780)}{0.3562/(1-0.3562)} = \frac{0.3780(1-0.3562)}{0.3562(1-0.3780)} = 1.098.$$

The 95% confidence interval is

$$1.098 \exp[\pm 1.96\sqrt{1/48 + 1/79 + 1/26 + 1/47}] = 1.098 \exp[\pm 0.5985],$$

which is 0.604 to 1.998.

2d. The 95% confidence interval for the risk difference includes 0 and those for the relative risk and odds ratio include 1, meaning that $p_1 = p_2$ is not implausible. This accords with acceptance of the null hypothesis at level 0.05 in exercise 1 of Written Assignment 4.