STA 580 — Fall 2008 — Dr. Charnigo

Written Assignment 6 Solutions

1a. [Using SAS for this and subsequent items is permissible.] We have $\bar{x} = 61.144$, $\bar{y} = 2.637$, $L_{xx} = 21242$, $L_{yy} = 490.92$, and $L_{xy} = 2803.4$. The least squares estimate of β is $b = 2803.4/21242 = 0.132$. The least squares estimate of α is $a = 2.637 - 0.132 \times 61.144 = -5.434$.

1b. We have Tot SS = 490.92, Reg SS = Reg MS = $2803.4^2/21242 = 369.98$, Res SS = $490.92 - 369.98 = 120.94$, and Res MS = $120.94/652 = 0.1855$. Since $t_{652,.975} = 1.964$, the 95% confidence interval for α is

$$-5.434 \pm 1.964\sqrt{0.1855(1/654 + 61.144^2/21242)} = -5.434 \pm 0.356,$$

or -5.790 to -5.078. The 95% confidence interval for β is

$$0.132 \pm 1.964\sqrt{0.1855/21242} = 0.132 \pm 0.006,$$

or 0.126 to 0.138. [Using $z_{.975} = 1.960$ as an approximation to $t_{652,.975}$ is permissible.]

1c. The f statistic is $369.98/0.1855 = 1995$, which easily exceeds $f_{1,652,.95} = 3.858$, 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 $f_{1,652,.95}$ is permissible.]

1d. The t statistic is $0.132/\sqrt{0.1855/21242} = 44.67$, which easily exceeds $t_{652,.975} = 1.964$, so we reject $H_0 : \beta = 0$. [The p-value is less than 0.0001. Also, to check our work, we can note that $t^2 = 44.67^2 = 1995 = f$.]

1e. We have $\hat{y} = -5.434 + 0.132 \times 57 = 2.090$. The 95% prediction interval is

$$2.090 \pm 1.964\sqrt{0.1855(1 + 1/654 + (57 - 61.144)^2/21242)} = 2.090 \pm 0.847,$$

which is 1.243 to 2.937.

1f. The 95% confidence interval is

$$2.090 \pm 1.964\sqrt{0.1855(1/654 + (57 - 61.144)^2/21242)} = 2.090 \pm 0.041,$$

which is 2.049 to 2.131.

1g. The fraction of variability in forced expiratory volume accounted for by height is $R^2 = 369.98/490.92 = 0.754$. This is the square of the Pearson correlation between X and Y, which is $2803.4/\sqrt{21242 \times 490.92} = 0.868$.

2a. [Using SAS for this and subsequent items is permissible.] We have $\hat{p}_1 = 74/576 = 0.1285$, $\hat{p}_2 = 27/533 = 0.0507$, $a = 74$ (Table 13.1 notation), $b = 502$, $c = 27$, $d = 506$, $n_1 = 576$, and $n_2 = 533$. The point estimate of the risk difference is $0.1285 - 0.0507 = 0.0778$. The 95% confidence interval is

$$0.0778 \pm 1.96\sqrt{0.1285(1 - 0.1285)/576 + 0.0507(1 - 0.0507)/533} = 0.0778 \pm 0.0331,$$

which is 0.0447 to 0.1109.

2b. The point estimate of the relative risk is $0.1285/0.0507 = 2.535$. The 95% confidence interval is

$$2.535 \exp[\pm 1.96 \sqrt{502/(74 \times 576) + 506/(27 \times 533)}] = 2.535 \exp[\pm 0.4246],$$

which is 1.66 to 3.88.

2c. The point estimate of the odds ratio is

$$\frac{0.1285/(1-0.1285)}{0.0507/(1-0.0507)} = \frac{0.1285(1-0.0507)}{0.0507(1-0.1285)} = 2.761.$$

The 95% confidence interval is

$$2.761 \exp[\pm 1.96 \sqrt{1/74 + 1/502 + 1/27 + 1/506}] = 2.761 \exp[\pm 0.4576],$$

which is 1.75 to 4.36.