STA 580 — Spring 2009 — Dr. Charnigo

Written Assignment 6 Solutions

1a. [Using SAS for this and subsequent items is permissible.] We have $\bar{x} = 9.931$, $\bar{y} = 2.637$, $L_{xx} = 5697.9$, $L_{yy} = 490.92$, and $L_{xy} = 1265.2$. The least squares estimate of β is $b = 1265.2/5697.9 = 0.222$. The least squares estimate of α is $a = 2.637 - 0.222 \times 9.931 = 0.432$.

1b. We have Tot SS = 490.92, Reg SS = Reg MS = $1265.2^2/5697.9 = 280.93$, Res SS = $490.92 - 280.93 = 209.99$, and Res MS = $209.99/652 = 0.3221$. Since $t_{652,.975} = 1.964$, the 95% confidence interval for α is

$$0.432 \pm 1.964\sqrt{0.3221(1/654 + 9.931^2/5697.9)} = 0.432 \pm 0.153,$$

or 0.279 to 0.585. The 95% confidence interval for β is

$$0.222 \pm 1.964\sqrt{0.3221/5697.9} = 0.222 \pm 0.015,$$

or 0.207 to 0.237. [Using $z_{.975} = 1.960$ as an approximation to $t_{652,.975}$ is permissible.]

1c. The f statistic is $280.93/0.3221 = 872$, 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 $0.222/\sqrt{0.3221/5697.9} = 29.53$, 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 $t^2 = 29.53^2 = 872 = f$.]

1e. We have $\hat{y} = 0.432 + 0.222 \times 11 = 2.874$. The 95% prediction interval is

$$2.874 \pm 1.964\sqrt{0.3221(1 + 1/654 + (11 - 9.931)^2/5697.9)} = 2.874 \pm 1.116,$$

which is 1.758 to 3.990.

1f. The 95% confidence interval is

$$2.874 \pm 1.964\sqrt{0.3221(1/654 + (11 - 9.931)^2/5697.9)} = 2.874 \pm 0.046,$$

which is 2.828 to 2.920.

1g. The fraction of variability in forced expiratory volume accounted for by age is $R^2 = 280.93/490.92 = 0.572$. This is the square of the Pearson correlation between X and Y, which is $1265.2/\sqrt{5697.9 \times 490.92} = 0.756$.

2a. [Using SAS for this and subsequent items is permissible.] We have $\hat{p}_1 = 70/200 = 0.3500$, $\hat{p}_2 = 85/200 = 0.4250$, $a = 70$ (Table 13.1 notation), $b = 130$, $c = 85$, $d = 115$, $n_1 = 200$, and $n_2 = 200$. The point estimate of the risk difference is $0.3500 - 0.4250 = -0.0750$. The 95% confidence interval is

$$-0.0750 \pm 1.96\sqrt{0.3500(1 - 0.3500)/200 + 0.4250(1 - 0.4250)/200} = -0.0750 \pm 0.0952,$$

which is -0.1702 to 0.0202.

2b. The point estimate of the relative risk is $0.3500/0.4250 = 0.824$. The 95% confidence interval is

$$0.824 \exp[\pm 1.96 \sqrt{130/(70 \times 200) + 115/(85 \times 200)}] = 0.824 \exp[\pm 0.2483],$$

which is 0.643 to 1.056.

2c. The point estimate of the odds ratio is

$$\frac{0.3500/(1 - 0.3500)}{0.4250/(1 - 0.4250)} = \frac{0.3500(1 - 0.4250)}{0.4250(1 - 0.3500)} = 0.729.$$

The 95% confidence interval is

$$0.729 \exp[\pm 1.96 \sqrt{1/70 + 1/130 + 1/85 + 1/115}] = 0.729 \exp[\pm 0.4038],$$

which is 0.487 to 1.092.