STA 580 — Spring 2011 — Dr. Charnigo

Written Assignment 6 Solutions

1a. [Using SAS for this and subsequent items is permissible.] We have $\bar{u} = 2.249$, $\bar{v} = 0.915$, $L_{uu} = 63.90$, $L_{vv} = 72.53$, and $L_{uv} = 54.07$. The least squares estimate of β is $b = 54.07/63.90 = 0.846$. The least squares estimate of α is $a = 0.915 - 0.846 \times 2.249 = -0.988$.

1b. We have Tot SS $= 72.53$, Reg SS $=$ Reg MS $= 54.07^2/63.90 = 45.75$, Res SS $= 72.53 - 45.75 = 26.77$, and Res MS $= 26.77/652 = 0.04106$. Since $t_{652,.975} = 1.964$, the 95% confidence interval for α is

$$-0.988 \pm 1.964\sqrt{0.04106(1/654 + 2.249^2/63.90)} = -0.988 \pm 0.113,$$

or -1.101 to -0.875. The 95% confidence interval for β is

$$0.846 \pm 1.964\sqrt{0.04106/63.90} = 0.846 \pm 0.050,$$

or 0.796 to 0.896. [Using $z_{.975} = 1.960$ as an approximation to $t_{652,.975}$ is permissible.]

1c. [The elements of the ANOVA table are documented in the answer to exercise 1b.] The f statistic is $45.75/0.04106 = 1114$, 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 $0.846/\sqrt{0.04106/63.90} = 33.38$, 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 $t^2 = 33.38^2 = 1114 = f$.]

1e. The fraction of variability in the logarithm of forced expiratory volume explained by the logarithm of age is $R^2 = 45.75/72.53 = 0.631$.

1f. We have $\hat{y}_i = \exp[\hat{v}_i] = \exp[a + bu_i] = \exp[a + b\log x_i] = \exp[-0.988 + 0.846\log x_i]$. Using standard properties of exponentials and logarithms, we simplify this to $\exp[-0.988] \times \exp[0.846\log x_i] = \exp[-0.988] \times \exp[\log(x_i^{0.846})] = \exp[-0.988]x_i^{0.846} = 0.372x_i^{0.846}$.

1g. We have $\hat{v} = -0.988 + 0.846 \times 2.398 = 1.041$. The 95% prediction interval for the logarithm of forced expiratory volume is

$$1.041 \pm 1.964\sqrt{0.04106(1 + 1/654 + (2.398 - 2.249)^2/63.90)} = 1.041 \pm 0.398,$$

which is 0.643 to 1.440. As such, the 95% prediction interval for forced expiratory volume is $\exp[0.643] = 1.902$ to $\exp[1.440] = 4.219$. This interval is of comparable width to that obtained by the students in Spring 2009 (1.758 to 3.990), but both the left and right endpoints are shifted rightward.

2a. [Using SAS for this and subsequent items is permissible.] We have $\hat{p}_1 = 13/72 = 0.1806$, $\hat{p}_2 = 7/48 = 0.1458$, $a = 13$ (Table 13.1 notation), $b = 59$, $c = 7$, $d = 41$, $n_1 = 72$, and $n_2 = 48$. The point estimate of the risk difference is $0.1806 - 0.1458 = 0.0348$. The 95% confidence interval is

$$0.0348 \pm 1.96\sqrt{0.1806(1 - 0.1806)/72 + 0.1458(1 - 0.1458)/48} = 0.0348 \pm 0.1337,$$

which is -0.0989 to 0.1685.

2b. The point estimate of the relative risk is $0.1806/0.1458 = 1.239$. The 95% confidence interval is

$$1.239 \exp[\pm 1.96\sqrt{59/(72 \times 13) + 41/(48 \times 7)}] = 1.239 \exp[\pm 0.8432],$$

which is 0.533 to 2.879.

2c. The point estimate of the odds ratio is

$$\frac{0.1806/(1-0.1806)}{0.1458/(1-0.1458)} = \frac{0.1806(1-0.1458)}{0.1458(1-0.1806)} = 1.291.$$

The 95% confidence interval is

$$1.291 \exp[\pm 1.96\sqrt{1/13 + 1/59 + 1/7 + 1/41}] = 1.291 \exp[\pm 1.0016],$$

which is 0.474 to 3.515.