STAT W4201 001, Homework 5

Brian Weinstein (bmw2148)

March 2, 2016

Code is attached here and also posted at https://github.com/BrianWeinstein/advanced-data-analysis. Where relevant, code snippets and output are are included in-line.

Problem 1: Ramsey 7.18

(a) *Find the standard error of prediction for the prediction of pH at 5 hours after slaughter.*

The calculations in Display 7.12 give us $\hat{\beta}_0 = 6.9836$, $\hat{\beta}_1 = -0.7257$, $\hat{\sigma} = 0.08226$, $n = 10$, $\overline{X} = 1.190$, $s_X^2 = 0.6344$.

Therefore, the standard error of prediction for the pH at 5 hours is

$$\begin{aligned}
\text{SE}\left[\text{Pred}\{Y|X_0 = \log(5) = 1.609438\}\right] &= \hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(X_0 - \overline{X})^2}{(n-1)s_X^2}} \\
&= (0.08226)\sqrt{1 + \frac{1}{10} + \frac{(1.609438 - 1.190)^2}{9 \cdot 0.6344}} \\
&= 0.0875
\end{aligned}$$

(b) *Construct a 95% prediction interval at 5 hours after slaughter.*

The prediction of pH level at 5 hours is,

$$\begin{aligned}
Pred\{Y|X_0 = \log(5) = 1.609438\} &= \beta_0 + \beta_1 \cdot \log(5) \\
&= 6.9836 - 0.7257 \cdot 1.609438 \\
&= 5.8156.
\end{aligned}$$

A 95% prediction confidence interval at 5 hours is given by

$$\begin{gathered}
5.8156 \pm t_8(0.975) \cdot \text{SE}\left[\text{Pred}\{Y|X_0 = \log(5)\}\right] \\
5.8156 \pm 2.3060 \cdot 0.0875 \\
5.8156 \pm 0.2017 \\
\Rightarrow [5.6139, 6.0173].
\end{gathered}$$

Problem 2: Ramsey 7.24

(a) *With a statistical computer package and the data in the file ex0724, obtain the least squares fits to the four simple regressions, individually, to confirm the estimates and standard errors presented in Display 7.17.*

Confirming the estimates and standard errors from Display 7.17:

i. Denmark

```
> lmDenmark <- lm(formula=Denmark~Year, data=birthData)
> summary(lmDenmark)$coefficients
                 Estimate   Std. Error   t value     Pr(>|t|)
(Intercept)  0.59872329381 0.0408047207 14.672893 2.395722e-18
Year        -0.00004288538 0.0000206916 -2.072598 4.423828e-02
```

ii. The Netherlands

```
> lmNetherlands <- lm(formula=Netherlands~Year, data=birthData)
> summary(lmNetherlands)$coefficients
                 Estimate   Std. Error   t value     Pr(>|t|)
(Intercept)  0.67239837505 0.0279195810 24.083398 1.365923e-26
Year        -0.00008084321 0.0000141577 -5.710196 9.636921e-07
```

iii. Canada

```
> lmCanada <- lm(formula=Canada~Year, data=birthData)
> summary(lmCanada)$coefficients
                Estimate    Std. Error   t value     Pr(>|t|)
(Intercept)  0.7337857143 0.05480068278 13.390083 3.983523e-11
Year        -0.0001111688 0.00002767698 -4.016653 7.375947e-04
```

iv. United States

```
> lmUsa <- lm(formula=USA~Year, data=birthData)
> summary(lmUsa)$coefficients
                 Estimate     Std. Error   t value     Pr(>|t|)
(Intercept)  0.62008571429 0.018598766807 33.340152 2.523643e-18
Year        -0.00005428571 0.000009393273 -5.779212 1.439109e-05
```

(b) *Obtain the t-statistic for the test that the slopes of the regressions are zero, for each of the four countries. Is there evidence that the proportion of male births is truly declining?*

The t-statistics and associated two-sided p-values are computed and shown in the output of part (a). For the `Year` variable:

i. Denmark
The t-statistic is -2.0726, with a one-sided p-value of 0.0221. The data provides moderate, but not convincing evidence, that the proportion of male births is truly declining in Denmark.

ii. The Netherlands
The t-statistic is -5.7102, with a one-sided p-value of 0.000000482. The data provides overwhelming evidence that the proportion of male births is truly declining in the Netherlands.

iii. Canada
The t-statistic is -4.0167, with a one-sided p-value of 0.000369. The data provides convincing evidence that the proportion of male births is truly declining in Canada.

iv. United States
The t-statistic is -5.7792, with a one-sided p-value of 0.00000720. The data provides overwhelming evidence that the proportion of male births is truly declining in the United States.

Problem 3: Ramsey 7.28

Problem 4: Ramsey 8.17

Problem 5: Ramsey 8.20

Problem 6: Ramsey 9.12