Chapter 4 - Simple Linear Regression

- 4.1 **[9 pts]**
  1. Results in table below.

```
> summary(cricket)
       chirps            temp
 Min.   :14.0   Min.   :69.00
 1st Qu.:15.5   1st Qu.:75.50
 Median :16.0   Median :81.00
 Mean   :16.6   Mean   :_80.13_
 Max.   :20.0   Max.   :93.00
 StDev  : 1.7   StDev  : 6.72

> lm1 <- lm(temp~chirps)
> summary(lm1)
Call: lm(formula = temp ~ chirps)
Coefficients:
            Estimate Std. Error t value  Pr(>|t|)
(Intercept)  26.7420   _10.1807_  2.627  0.020917
chirps        _3.2163_    0.6102  _5.271_ 0.000151

Residual standard error: 3.936 on 13 degrees of freedom
Multiple R-Squared: 0.6812,     Adjusted R-squared: 0.6567
F-statistic: 27.78 on 1 and 13 DF,  p-value: _0.0001513_

> predict(lm1,data.frame(chirps=15),interval="c")
           fit        lwr       upr
[1,] _74.98717_ _71.94262_ 78.03173

> predict(lm1,data.frame(chirps=15),interval="p")
           fit        lwr       upr
[1,] _74.98717_  65.95557 _84.01879_
```

  2. Temp = 26.742+3.2163Chirps
  3. Yes, the p-value for testing that the slope is equal to zero is very small leading to a conclusion that the slope is different than zero implying a significant relationship.

- 4.2 **[6 pts]**
  1. YES, slope and F p-values are both very small (p<0.00005; **Table B.15**).

Table B.15: Summary of simple linear regression results of proportion of males on year.

```
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.201e-01  1.860e-02  33.340  < 2e-16
year        -5.429e-05  9.393e-06  -5.779 1.44e-05
---
Residual standard error: 0.0002607 on 19 degrees of freedom
Multiple R-squared: 0.6374,        Adjusted R-squared: 0.6183
F-statistic:  33.4 on 1 and 19 DF,  p-value: 1.439e-05
```

  2. The proportion of males **declined** between 0.000035 and 0.000074 per year **(Table B.16)**.

Table B.16: Coefficient confidence intervals from simple linear regression results of proportion of males on year.

```
                    2.5 %        97.5 %
(Intercept)  5.811580e-01  6.590134e-01
year        -7.394606e-05 -3.462537e-05
```

  3. The very small coefficients are statistically different from zero because the SE for the coefficients are very small (0.000009; **Table B.15**)

R Commands

```
> year <- 1970:1990
> propmale <- c(0.5134, 0.5126, 0.5125, 0.5128, 0.5133, 0.5132, 0.5128,
+     0.5128, 0.5129, 0.5127, 0.5129, 0.5126, 0.5123, 0.5127, 0.5122,
+     0.5126, 0.5122, 0.512, 0.5121, 0.512, 0.512)
> d <- data.frame(year, propmale)
> attach(d)
> lm1 <- lm(propmale ~ year)
> summary(lm1)
> confint(lm1)
```

- 4.3 **[8 pts]**
  1. Yes, there is a signficant relationship between t-cell response and mass (p=0.0061; **Table B.17**). Specifically, as mass increases by 1g the t-cell response increases between 0.011 and 0.055, on average **(Table B.18)**.

Table B.17: Summary of simple linear regression results of t-cell response on mass.

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.08750    0.07868   1.112   0.2800
mass         0.03282    0.01064   3.084   0.0061
---
Residual standard error: 0.08102 on 19 degrees of freedom
Multiple R-squared: 0.3336,        Adjusted R-squared: 0.2986
F-statistic: 9.513 on 1 and 19 DF,  p-value: 0.006105
```

Table B.18: Coefficient confidence intervals from simple linear regression results of t-cell response on mass.

```
                  2.5 %     97.5 %
(Intercept) -0.07717487 0.25216884
mass         0.01054860 0.05509438
```

  2. The mean t-cell response for all birds that carried a mean stone mass of 5 g is between 0.190 and 0.313.
  3. The t-cell response for a bird that carried a mean stone mass of 5 g is between 0.071 and 0.432.
  4. The prediction interval for the individual is wider than the confidence interval for the mean because there is more variability in predicting an individual as compared to a mean.

R Commands

```
> mass <- c(3.33, 4.62, 5.43, 5.73, 6.12, 6.29, 6.45, 6.51, 6.65, 6.75,
+     6.81, 7.56, 7.83, 8.02, 8.06, 8.18, 9.08, 9.15, 9.35, 9.42, 9.95)
> t.cell <- c(0.252, 0.263, 0.251, 0.251, 0.183, 0.213, 0.332, 0.203,
+     0.252, 0.342, 0.471, 0.431, 0.312, 0.304, 0.37, 0.381, 0.43, 0.43,
+     0.213, 0.508, 0.411)
> d <- data.frame(mass, t.cell)
> attach(d)
> lm2 <- lm(t.cell ~ mass)
> summary(lm2)
> confint(lm2)
> predict(lm2, data.frame(mass = 5), interval = "c")
> predict(lm2, data.frame(mass = 5), interval = "p")
```