

Question 4.2

- a. Yes there is evidence for a significant statistical change in the proportion of male births over the study period because the p-value for the slope (and the equivalent overall F p-value) is very small ($p < 0.00005$; Table 1).

Table 1. 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

```

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

Table 2. 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

```

- c. The very small slope coefficient is statistically different from zero because the SE for the slope coefficient is very small (0.000009; Table 1) and the overall scale of the measurements is very small.

Question 4.3

- a. Yes, there is a significant relationship between t-cell response and mass ($p = 0.0061$; Table 3). Specifically, as mass increases by 1 g the t-cell response increases between 0.011 and 0.055, on average (Table 4).

Table 3. 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  0.27996
mass         0.03282     0.01064   3.084  0.00611
---
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 4. 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

```

- b. The mean t-cell response for all birds that carried a mean stone mass of 5 g is between 0.190 and 0.313.
 c. The t-cell response for a bird that carried a mean stone mass of 5 g is between 0.071 and 0.432.

- d. 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. Variability for predicting an individual includes both sampling and natural variability, whereas variability for the mean includes only sampling variability.

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,
               0.5122,0.5126,0.5122,0.5120,0.5121,0.5120,0.5120)
> d <- data.frame(year,propmale)
> lm1 <- lm(propmale~year,data=d)
> summary(lm1)
> confint(lm1)
```

```
> 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,
             0.342,0.471,0.431,0.312,0.304,0.370,0.381,0.430,0.430,
             0.213,0.508,0.411)
> d <- data.frame(mass,t.cell)
> lm2 <- lm(t.cell~mass,data=d)
> summary(lm2)
> confint(lm2)
> predict(lm2,data.frame(mass=5),interval="c")
> predict(lm2,data.frame(mass=5),interval="p")
```

Notes from the Professor

- The data are probably best entered into Excel and the loaded into R via a tab-delimited text file.
- When discussing whether there is a relationship between the response and the explanatory variable you must explicitly note that you are referring to the slope p-value. You cannot just refer the reader to the “p-value” in the table from `summary()` because there are three p-values in that table. Be precise with your language!!
- Remember to use CI when describing rates of change (i.e., slopes) or predictions; don’t just use the best estimate.