Assignment #12 contains a few textbook problems from Chapter 12 and some associated R problems.

1. Do Exercise 12.9.

Solution:

(a) 
\[ b_1 = \frac{731.36}{2094.55} = 0.349 \text{cm/mm}, \quad b_0 = 103.99 \cdot (0.349)(149.64) = 51.7 \text{cm} \]

In words,
\[ \text{(max jump in cm)} = 51.7 + 0.349(\text{frog length in mm}) \]

(b) For every additional mm length of the bull frog, we expect its maximum jump to increase by 0.35 cm.

(c) \[ s_y = \sqrt{3218.99/10} = 17.9 \text{cm}, \quad s_{y|x} = \sqrt{2963.61/9} = 18.1 \text{cm} \]. Note, the sum of squares around the regression line must be less than or equal to the sum of squares around \( \bar{y} \). The denominators for \( s_y \) and \( s_{y|x} \) are different (\( n - 1 \) and \( n - 2 \)), making it possible to have the unusual circumstance where \( s_y < s_{y|x} \). This is not typically the case, but it is for this data set.

(d) Given the length of the frog, an estimate of the maximum jump length will differ from the actual value by about 18 cm or so, typically.

2. Use R to find the least squares regression line for the data in Exercise 12.9. Here is the sample code that will do this.

```r
> bullfrog = read.table("ex12-9.txt", header = T)
> attach(bullfrog)
> fit = lm(jump ~ length)
> summary(fit)
```

... 

Coefficients:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| (Intercept) | 51.7416 | 59.5828 | 0.868 | 0.408 |
| length | 0.3492 | 0.3965 | 0.881 | 0.401 |

Residual standard error: 18.15 on 9 degrees of freedom
Multiple R-Squared: 0.07933, Adjusted R-squared: -0.02296
F-statistic: 0.7755 on 1 and 9 DF, p-value: 0.4014

Solution:

\[ (\text{jump}) = 51.7416 + 0.3492(\text{length}) \]
3. Construct a 95% confidence interval for the slope of the regression line in the previous two exercises.

Solution: Part of the summary says the following.

Coefficients:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| Intercept| 51.7416    | 59.5828 | 0.868    | 0.408    |
| length   | 0.3492     | 0.3965  | 0.881    | 0.401    |

A 95% confidence interval for the slope is as follows.

\[ 0.3492 \pm (2.262)(0.3965) \] or \[ 0.35 \pm 0.90 \] or \[ (-0.65, 1.25) \]

We are 95% confident that the estimated change in the maximum jump length per additional mm of length of the bullfrog is between -0.65 and 1.25 cm/mm.


Solution: The estimated maximum jump of a 150 mm bullfrog is \( 51.5 + (0.349)(150) = 103.85 \) cm. The length 150 mm is within the range of data on which the regression line is based, so this is not extrapolation.

5. Exercise 12.31.

Solution: We have \( r^2 = 0.43^2 = 0.185 \). This means that age can explain about 18 percent of the variation in systolic blood pressure in the population of men. Much of the variation cannot be explained by age alone.


Solution:

The first residual plot matches the third scatter plot. This one looks consistent with model assumptions.

The second residual plot matches the second scatter plot. There is non-linearity.

The third residual plot matches the first scatter plot. There is heteroscedasticity (non-constant variance).