Stat 704: Homework 2

1. **Regression through the origin** We will consider a special case of simple linear regression where the intercept is assumed to be zero from the outset (this is often assumed in calibration of certain measuring devices). Let

$$Y_i = x_i\tau + \epsilon_i,$$

where $E(\epsilon_i) = 0$ and $\text{var}(\epsilon_i) = \sigma^2$.

   (a) Define $Q(\tau) = \sum_{i=1}^n (Y_i - x_i\tau)^2$. Show that the minimizer of $Q(\tau)$ is $\hat{\tau} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$.

   (b) Show $E(\hat{\tau}) = \tau$.

   (c) Show $\text{var}(\hat{\tau}) = \frac{\sigma^2}{\sum_{i=1}^n x_i^2}$.

2. A cereal company claims its boxes contain 445 grams of cereal. A random sample of 14 boxes produces the following measurements: 441.82 437.38 445.92 444.17 444.89 445.93 443.97 445.40 445.95 443.35 441.95 444.86 438.96 439.38

   (a) Use a graph to determine whether the assumption of normality is reasonable.

   (b) Using an appropriate test (at the $\alpha = 0.05$ significance level), determine whether the center of the distribution of cereal weights is 445 grams. Use R or SAS and report the p-value of your test.

Do the following problems from the textbook:

- Thought & pencil/paper problems, chapter 1: 1.2, 1.5, 1.8, 1.11, 1.16, 1.18, 1.30, 1.33.
- Thought & pencil/paper problems, chapter 2: 2.10, 2.12.
- **Grade point average** (use SAS): 1.19, 1.23(b), 2.4(a,b,c), 2.13(a,b,c), 2.23.
- **Plastic hardness** (use SAS): 1.22, 1.26(b), 2.7(a,b), 2.16(a,b), 2.26.
- **Property assessments** (use SAS): 2.42(a,b,c), 2.46.

To test, for example, $H_0 : \beta_1 = 2$, add `hyp1: test x=2;` if your predictor is called `x`. This produces the t-statistic discussed in the book and notes.