1. Two different laboratory procedures were used to find the amount of calcium in a collection of orange juice samples. The correlation between the two measures was 0.99. Does this imply that the value given by the first procedure was approximately the same as the value given by the second procedure for each of the samples analyzed? Explain why or why not.

2. The 90% confidence interval for a regression coefficient is (2.1, 8.6). What can you say, if anything, about the $p$-value of the hypothesis test that this coefficient is zero?

3. Assume a one-way ANOVA with the same number of observations for each treatment. Suppose the degrees of freedom for treatments is 4, and the degrees of freedom for error is 20. Find (i) the number of treatments, and (ii) the number of observations for each treatment.

4. When running a quadratic regression, you “center” the explanatory variable by subtracting its mean $\bar{X} = 1$. The estimated coefficients for the model in this form are $b_0 = 2$, $b_1 = 3$, $b_2 = 4$, and $b_3 = -1$. Write the fitted model in terms of $X$, $X^2$, and $X^3$, i.e., give $a_0$, $a_1$, $a_2$, and $a_3$ in $\hat{Y} = a_0 + a_1X + a_2X^2 + a_3X^3$. (Hint: use the fact that $(x - c)^2 = x^2 - 2cx + c^2$ and $(x - c)^3 = x^3 - 3cx^2 + 3c^2x - c^3$.)

5. The parameter estimates given by the SAS generalized inverse method for a one-way ANOVA are intercept = 9, $\alpha_1 = 5$, $\alpha_2 = -3$, $\alpha_3 = 0$. Find $\hat{Y}_{2,1}$.

6. You are using the $C_p$ criterion to select a subset of explanatory variables in a multiple regression. The best subset of size 3 gives $C_p = 8$. Interpret this result.

7. For what value of $n$ is the $AIC$ criterion smaller than the $SBC$ criterion? (Assume all logs are natural.)

8. If you see that the square of the group mean and the group variance are linearly related, which transformation should you use to stabilize the error variance?

9. A one-way model has

$$y_{i,j} = \mu + \tau_i + \epsilon_{i,j},$$

where $\epsilon_{i,j}$ are independent with mean 0. As a result, what is the mean and variance of the $y_{i,j}$? Under the usual null hypothesis, what is the mean and variance of $y_{i,j}$?

10. Data were collected on a sample of 25 women and 25 men. Separate linear regressions were run for each group. For women, the slope was 30 with a standard error of 6; for men the slope was 20 with a standard error of 10. Calculate the value of the $t$-statistic that you would use to compare the two slopes and give its degrees of freedom.

11. In the cell means model parameterization of one-way ANOVA, the regression design matrix $X$ is such that $X'X = nI$. If one applies ridge regression shrinkage to the regression, thereby producing the parameter estimates via the formula

$$\hat{\beta}_{ridge} = (X'X + \lambda I)^{-1}X'y,$$

what will be the resulting cell means estimates $\hat{\mu}_i$ ($i = 1, \ldots, p$) based on the regression parameter estimates $\hat{\beta}_{ridge}$?