Assignment #9 — Due Friday, November 12, 2010, by 4:00 P.M.
Turn in homework to your TA's mailbox using this sheet as the cover page.

Fill in your name and also circle the *lecture section in which you are registered* and circle the *discussion section you expect to attend* to pick up this assignment.

Name:

Lecture 1 (Larget).   **311:** Tu 1:00 - 2:15pm   **312:** Th 8:00 - 9:15am   **313:** We 1:00 - 2:15pm

Lecture 2 (Hanlon).   **321:** Tu 1:00 - 2:15pm   **322:** We 2:30 - 3:45pm   **323:** We 1:00 - 2:15pm

Please answer the following questions.

1. The Luso variety of wheat is resistant to the Hessian fly. In order to understand the genetic mechanism controlling the resistance, an agronomist plans to examine the progeny of a certain cross invovling Luso and a nonresistant variety. Each progeny plant will be classified as resistant or susceptible and the agronomist will estimate the proportion of progeny that are resistant and provide a 95% confidence interval for this proportion. How many progeny does she need to classify in order to guarantee that the margin of error for this confidence interval is less than 3%? How many progeny would she need to classify in order to guarantee that the margin of error for a 99% confidence interval would be less than 3%?

2. The diameter of the stem of a wheat plant is an important trait because of its relationship to breakage of the stem, which interferes with harvesting the crop. An agronomist wants to compute a 90% confidence interval for the mean stem diameter. A pilot study indicates that the standard deviation among stem diameters is 0.25 mm. How many plants should the agronomist sample so that the margin of error for this confidence interval is less than 0.01 mm?

3. Use the data set from p.353, Problem 19, but answer these questions. In this version, you are asked to separately analyze the climbing data and the non-climbing data.

   (a) Using the climbing data, answer the following. (i) Plot a density plot and normal quantile plot of the data; (ii) Comment on the normality of the data and decide if a transformation is needed; and (iii) Compute a 95% confidence interval for the mean number of species in the climbing genera (making use of a transformation if appropriate).

   (b) Answer the same set of questions for the non-climbing data.

4. Long-run average percent butterfat in milk at a farm is 3.35 and the standard deviation is 0.15, with measurements taken by the load. How large should a sample size be to have an 80% chance of detecting a change to $\mu = 3.30$ at a significance level of $\alpha = 0.05$ with a two-sided test?

   (a) Let a be the percentage of butterfat that is the 0.025 quantile of the sampling distribution of $\bar{x}$ assuming that $\mu = 3.35$. Express a as a function of n. This is the boundary of the lower rejection region when $\alpha = 0.05$. (You need to find the critical z value for the 0.025 quantile).

   (b) If the power is 0.8 when $\mu = 3.30$, then a will also be (approximately) the 0.8 quantile for the $N(3.30, (0.15)^2/n)$ distribution. Express a as a function of n (also using the critical z value for the 0.8 quantile).

   (c) Set these expressions for a equal to one another and solve for n algebraically. How large does the sample size need to be to meet the criteria?

   (d) Use R to find both a and the power (when $\mu = 3.30$) accurate to three decimal places.

5. The R function `graphPower()` is in the file `power.R` on the course web page. This function allows you to graph the power function for a two-sided alternative hypothesis test versus the null hypothesis that the mean is equal to μ_0. Examples on how to use the function are included in the file.

   (a) Graph the power versus μ for the hypothesis $\mu = 100$ versus the alternative $\mu \neq 100$ with $\sigma = 15$ and $\alpha = 0.1$ for each $n = 10$, $n = 40$, $n = 160$, $n = 640$. Describe qualitatively how the power curve changes.

   (b) Repeat the previous problem, but keep $n = 40$ fixed and vary $\alpha = 0.1$, $\alpha = 0.05$, and $\alpha = 0.01$. Describe qualitatively how the power curve changes.

   (You do not need to include the graphs with your turned in assignment.)