DAY 2 ACTIVITIES

Activity 1 (Slide 7):

1. Go to the website: http://www.rossmanchance.com/applets/NewConfsim/Confsim.html
2. In the first pull-down menu, change from "proportions" to "means."
3. In the next pull-down menu, select "t".
4. Put in these values: $\mu = 100$, $\sigma = 10$ (default), n = 25, intervals = 20.
5. Hit "Sample." Repeat this a few times to illustrate the concept (see question in #6).
6. Notice how many intervals are green (they cover the "true mean" of 100) and how many are red. With confidence level = .95, about how many of the 20 intervals would you expect to be red?

Activity 2: Instructions on Slide 19 (Import DavisRain, then *Statistics → Means → Single sample t-test*)
Activity 3: Instructions on Slide 25 (Import Benzene, histogram, log transform, find CI, transform back)
Activity 4: Instructions on Slide 27 (Saddle R. data, look at shape, CI for mean, transform, CI for median)

Activity 5 (Slide 31):
Find a 90% prediction interval for a new value based on the San Francisco effluent nickel data. Here are the lowest and highest 3 values: 1.9, 2.0, 2.1, ..., 3.7, 3.9, 4.4

Activity 6 (Slide 61):

1. Go to statpages.org, click or scroll to the section labeled "Power, sample size and experimental design" then scroll to the section on "For two-group tests."
2. Click on "Sample size" in the first bullet.
3. Use the Walker Creek Delta means (1.6 and .5) and standard deviations (1.1 and .3) to find sample sizes required for power of .95 (still using $\alpha = 5\%$). (Answer should be 12 for each.)
4. Go back to the page in #1 above, and this time click on "Statistical power" in the first bullet.
5. Put in the same mean and standard deviation values. Find what the power would be if both sample sizes were 5. (Note that the button to click erroneously says "Calculate sample size.") Answer: 69.6%.
6. Repeat #5 using samples of 20 for each sample. (Answer: 99.6%)
7. Go back to the "Sample size" version (Steps 1 and 2 above). Leave the default standard deviation values of 1.0 for both samples. Play with various differences in means to see what sample sizes you need to get 90% power (beta error of 10%). Notice that it is the *difference* in means that matter, not their absolute size. For example, try putting in means of 20 and 21, then means of 100 and 101.

Activity 7 (Afternoon, Slide 15):

1. Generate 5 random samples of n = 25 from a normal distribution with $\mu = 100$ and $\sigma = 10$:

Distributions → Continuous distributions → Normal distribution →Sample from normal distribution
Then enter mu = 100, sigma = 10, (rows) = 25, (columns) = 5, and uncheck the "sample mean" box. (The box is set up to put the samples into rows, but we are putting each sample of 25 in its own column.)

2. Go to "View data set" to make sure you have 5 columns of 25 numbers each.
3. Create a histogram for each sample to see shape. (*Graphs → Histogram*). No need to save them.
4. For each sample, do a t-test of H_0: $\mu = 100$. Do you get any *p*-values < .05? Do all C.I.s cover 100?
5. Summarize class results. We expect about 1 out of 20 samples to reject H_0, even though it is true.

Activity 8 (Afternoon, Slide 25):

1. Import the Pilots dataset
2. Do a paired t-test (*Statistics → Means → Paired t-test*). Make sure you use the correct alternative.
3. Create a new variable with the differences (NoAlcohol – Alcohol). (*Data → Manage.. → Compute...*)
4. Look at the shape of the differences.
5. Do a one-sample t-test of H_0: $\mu = 0$ for the differences. Make sure you use the correct alternative.
6. Compare the results in Steps 2 and 5.

Activity 9: Slide 8 of the Case Study