PROBLEM SET 5

(Due on Thursday, April 5)

(1) A sample of 20 speedometers of a particular brand is obtained, and each is calibrated to check for accuracy at 55 mph. The resulting sample data set listed as **SPEED** on the Class Homepage (also available as the file **speed.txt** on **f:\sys302** in Towne computer labs). Let $\mu =$ the actual mean reading of these speedometers at 55 mph.

(a) Using JMPIN, determine whether the t-test can be applied to this data.

(b) Construct the appropriate hypothesis and test it using $\alpha = .025$.

(c) Compute the P-value and interpret its meaning for this problem.

(2) [Based on Problem 8.38 in Devore, p.343 (6^{th} ed.)] A university library ordinarily has a complete shelf inventory done once every year. Because of the new shelving rules instituted the previous year, the head librarian believes that it may be possible to save money by postponing the inventory. As a preliminary measure, the librarian decides to select at random 950 books from the library's catalogue and have them searched. If evidence indicates strongly that the actual proportion of misshelved or lost books is less than 5%, then the inventory will be postponed.

(a) Among the 950 books searched, 35 were either misshelved or lost. Using this data, test the relevant hypothesis (with $\alpha = .01$) and advise the librarian what to do. Be sure to *justify* your choice of distribution theory.

(b) If the true proportion of misshelved or lost books is 4%, what is the probability that the test in part (a) will result in an unnecessary inventory?

(c) If the true proportion is actually 6%, what is the probability that the inventory will be postponed?

(3) Recall from the Pollution Example in class that the EPA has determined that the maximum permissible mean level of effluent for the chemical company should be μ =10 gal/day. To enforce this limit they have informed the company that

they will measure the company's effluent level on twenty randomly selected days during the year, and will use this sample to test the hypothesis, $H_1{:}\mu \geq 10$. They will then impose a fine on the company at the end of year unless this hypothesis can be rejected at level $\alpha = .05$. The company has strongly protested the use of this procedure. In particular, while they have no quarrel with either the EPA's stipulated limit of 10 gal/day or their choice of significance level, $\alpha = .05$, they have argued that they should only be considered in violation if the hypothesis, $H_2{:}\mu \leq 10$, can be rejected at level α.

(a) Why is the second procedure preferred by the chemical company?

(b) If this case were brought to court, what argument might be made by the chemical company in favor of the second procedure?

(4) The table below summarizes the data taken from an experiment that studied various characteristics of anchor bolts with differing diameters.

D	N	S	SD
3/8	93	7.85	1.42
1/2	112	10.35	1.50

Here D refers to the *diameters* of the bolts (in inches), N to the *sample size* for each diameter of bolt, S to the *average shear strength* (in kips) for each sample, and SD to the associated *sample standard deviations.*

(a) Carry out a test with $\alpha = .01$ to decide whether the true mean shear strength of 1/2-inch bolts exceeds that of 3/8-inch bolts by more than two kips. Bure sure to *justify* your choice of testing procedure here, and also to include a P-value in your analysis.

(b) Calculate and interpret a 95% CI for the difference in true mean shear strengths.

(5) A new diet product makes the claim that by using their product "you will lose at least 10 lbs. in the first week". A consumer group has decided to test this claim. A sample of 20 volunteer (overweight) women are tested, and the results are listed in the data set labeled **DIET** on the Class Homepage (also available as the file **diet.txt** on **f:\sys302** in Towne computer labs). The first item, X_1, is the weight of each subject prior to the test, and the second item, X_2, is her weight after a week on the diet. Given this data:

(a) Construct an appropriate null hypothesis, and test this hypothesis at the $\alpha = .05$. Be sure to *justify* your choice of testing procedure.

(b) Determine the P-value for this test. What else can you conclude about this particular test in view of the P-value?

(6) Honda is considering a new fuel injection system for the Civic. They have determined that the new system will be profitable only if it can increase mileage by at least 5 mpg under city driving ('worst case') conditions. To reach their decision, Honda has equipped 25Civics with the new system and has tested them against 25 unequipped Civics by driving them in New York city for a period of a week. The average mileage for equipped cars was $\overline{x}_1 = 37.3$ mpg with standard deviation, $s_1 = 1.2$, and for unequipped cars was $\overline{x}_2 = 31.5$ with standard deviation, $s_2 = 1.5$. Assuming the mileage data is normally distributed:

(a) Construct an appropriate null hypothesis for this decision by Honda, and test this hypothesis at the $\alpha = .05$ level.

(b) How confident can Honda be that the increase is at least 5 mpg?

(7) A 1993 study of fuel efficiency in Chicago tested 93 cars of differing *horsepower* (HP), *wheel-base widths* (WB), and *weights* (W). For each car the average *miles per gallon* (MPG) was measured over a week of normal city driving, and the resulting data set is listed as **CARS** on the class web page (and is also available as **cars.txt** in the class directory **f:/sys302** in the Towne computer labs. Here the four columns are MPG, HP, WB, and W, respectively). Using this data, regress MPG on (HP, WB, W) and carry out the following analyses:

(a) Draw conclusions about the significance of these three explanatory variables by examining their *P-values.*

(b) Rerun the regression without HP and determine a 99% CI for β_W.

(a) For the regression in (b) determine a 99% PI for the average miles per gallon (MPG) of a car with $WB = 105$ (inches) and $W = 3000$ (lbs).