HOMEWORK 6 (4653, FALL 2008)

Read: *Tijms*, pages 169 – 177.

Solve: Problems 5.16, 5.18, 5.20, 5.22, 5.24, 5.28, and the following problems

I. (20 points)
a) In a pill clinical study by a CORP, 150 people were tested, 70 were cured. Construct the 90% and 95% confidence intervals.
b) The FDA did its own testing on 400 patients and discovered that 175 were cured. Construct the 90% and 95% confidence intervals in this case.
c) Assume that in reality the pill helps 50% of the patients. Was CORP study flawed? How about the FDA study?
d) Assume now that the FRA study determined the true pill success rate. Was CORP study flawed?

II. (20 points) The SAT is administered to 1 million US graduating senior students in one year, with the average 500 and s.d. 100. In one town there are two high schools HS1 and HS2, and town School Board randomly assigns students to each. Each high school has exactly 1000 graduating senior students. After average SAT scores were calculated, HS1 had 520 on average, while HS2 had 530.
a) Can one conclude that students in HS1 are representative of the whole US student population, and the difference of 20 points is negligible and due to randomness?
b) Can one conclude that students in HS1 and HS2 have similar skill, and the difference of 10 points is negligible and due to randomness?
c) Estimate how many points a random US student needs to be better than 83% students who took the test. Same with 94%.
d) Estimate how many points a random HS1 student needs to be in the top 100 graduates of her/his class. Same with the top 50 graduates.

III. (20 points) Use the *Income of Households by State Using 3-Year-Average Medians* . data given in (the link is clickable)

`http://www.census.gov/hhes/www/income/statemedfaminc.html`

a) Assume everywhere in this problem that median = mean, error = s.d. Use the data to construct 95%-confidence intervals for each state.
b) Interpret the last three columns. Modify them according to 95%-confidence intervals. Can you still conclude that New Jersey households have more money on average?

c) Assume the s.d. of an individual household is about 60% of the income, and that the s.d. in the table is determined by a survey of random taxpayers. Estimate the number of survey participants for New Jersey, Minnesota and Mississippi.

IV. (20 points) Use the *National Estimates by Age, Sex, Race: 1900* (top left file) data given in

`http://www.census.gov/popest/archives/pre-1980/PE-11.html`

a) Someone made a hypothesis that a random person (see table all races, year 1900) at age 0 is equally likely to me male or female. Use the CLT to accept or reject this hypothesis with 95% confidence. Same with age 20, 30, 40, 60 and 75+.
b) Do the same for the 1940 data and compare the results.

⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆

This Homework is due Thursday November 6 at 4:40 pm. (right before the class). You MUST box all the answers. Remember that the answers are not enough, you also need to provide all intermediate calculations exhibiting your logic. However, no written explanations are necessary unless the problem asks for them.