# HOMEWORK 7

STA 321, Basic Statistical Theory I
Spring Semester, 2014

**Due:** March 27th, 2014

**1.** We would like to see how a random sample of 8th grade children from our school district compares to the national norm. The $L_{\pi Q}$-scores for the 57 children have an average of $\bar{Y} = 255$ and a standard deviation of $s = 16$. The national population mean is 250. Is it plausible that the mean $\mu$ of the $L_{\pi Q}$-scores of the population of 8th grade children in our school district equals 250? Set up null and alternative hypothesis, choose an appropriate statistical method (check its assumptions), and then answer this question using each of the following three equivalent approaches.

(a) Construct a rejection region based on $\alpha = 0.05$ and check whether $\bar{Y}$ is in the rejection region.

(b) Calculate the P-value and check whether it is less than $\alpha = 0.05$.

(c) Calculate the 95%-confidence interval for $\mu$ and check whether it includes 250.

(d) Discuss advantages and disadvantages of the three different methods. Which one would you recommend? Why?

(e) Repeat parts (a)-(c) for $\alpha = 0.01$ (corresponding to 99% confidence).

(f) How large a random sample would we need if a 95% confidence interval should have margin of error $B$ of at most $B = 10$?

**2.** An otherwise moderately difficult exam contains one very difficult multiple-choice question with five possible answers of which exactly one is correct. We would like to see whether more people answer the question correctly than would be expected just by chance (that is, everyone guessing randomly, with 0.2 probability of hitting the correct answer).

(a) Set up the null and alternative hypothesis.

(b) Out of 123 students, 34 answered the question correctly. What statistical method can we use to answer the above question? Are the assumptions for this method satisfied?

(c) Answer the question using the three approaches described in the previous exercise (rejection region, P-value, confidence interval).

(d) How large a random sample would we need if we wanted to achieve 99% power at a true alternative of $p = 0.3$, while maintaining $\alpha = 0.05$?

(e) How large a random sample would we need if a 95% confidence interval should have margin of error $B$ of at most $B = 0.1$?

**3.** IQ Scores are scaled to be approximately normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 15$.

(a) What is the 80th percentile of IQ scores?

(b) What are upper and lower quartile of IQ scores?

(c) What are the middle 90% of the values?

(d) If you take a random sample of 10 people, can you figure out the probability that their average IQ score is above 110?

**4.** A recent survey asked, How many days in the past seven days have you felt sad? The 816 women who responded had a median of 1, mean of 1.81, and standard deviation of 1.98. The 633 men who responded had a median of 1, mean of 1.42, and standard deviation of 1.83.

(a) Find a 95% confidence interval for the population mean for women. Interpret.

(b) Explain why the mean and standard deviation values suggest that this variable does not have a normal distribution. Does this cause a problem with the confidence interval method in (a)? Explain.

**5.** (a) A television network plans to predict the outcome of an election. They will do this with an exit poll on election day and decide to use a sample size for which the margin of error is 0.04 for 95% confidence intervals for population proportions. What sample size should they use?

(b) A public health unit wants to sample death records for the past year in Toronto to estimate the proportion of the deaths that were due to accidents. Health officials want the estimate to be accurate to within 0.02 with probability 0.95. Find the necessary sample size if, based on previous studies, officials believe that this proportion does not exceed 0.10.

(c) Suppose that (in the previous question) in determining the necessary sample size, officials can not rely on previous studies. Compare the result to the answer in part (b), and note the reduction in sample size that occurs by making an educated guess for the population proportion.