Homework 6

Math 445, Spring 2017

Due Wednesday, May 24 by 4:30 p.m.

Instructions

- All assignments should be stapled.

Assignment

Problem 1

Suppose that $X_1, \ldots, X_n$ form a random sample from the normal distribution with unknown mean μ and unknown variance σ^2.

(a) What distribution does $\dfrac{\sum_{i=1}^{n}(X_i - \overline{X})}{\sigma^2}$ have?

(b) Why is $\dfrac{\sum_{i=1}^{n}(X_i - \overline{X})}{\sigma^2}$ a pivotal quantity?

(c) Describe a method for constructing a confidence interval for σ^2 with a specified confidence level $1 - \alpha$. Hint: Determine constants c_1 and c_2 such that

$$P\left(c_1 < \frac{\sum_{i=1}^{n}(X_i - \overline{X})}{\sigma^2} < c_2\right) = 1 - \alpha$$

Problem 2

Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample of size n from a population with uniform distribution on the interval $(0, \theta)$. Let $Y_{max} = \max(Y_1, Y_2, \ldots, Y_n)$ and $U = Y_{max}/\theta$.

(a) Show that U has CDF

$$F(u) = \begin{cases} 0 & u < 0 \\ u^n & 0 \leq u \leq 1 \\ 1 & u > 1 \end{cases}$$

(b) Why is U a pivotal quantity?

(c) Find a 95% confidence interval for θ.

Problem 3

An investigator is interested in determining if the average amount of sleep for Lawrentians is less than the recommended amount of at least 8 hours per night. For a random sample of $n = 91$ students in a survey given the first week of class, the average amount of sleep was calculated to be 7.319 hours per night on and the standard deviation in this sample was 1.029.

(a) Find the 95% confidence interval for the average amount of nightly sleep for all Lawrence students. (Remember that `qt()` can find quantiles from a t-distribution in R.)

(b) Based on your confidence interval, is there sufficient evidence to show that the average amount of sleep is less than the recommended amount?

Problem 4

Financial statisticians are interested in the volatility of a market. When volatility is low, prices change very little over time. When volatility is high, prices change a lot. The direction of the change (up or down) is not important. All that matters is the magnitude of the change. The following data are a small part of a larger study on volatility of stock prices.

There are three major stock exchanges in the US. This small study concerns the volatility on two exchanges (NYSE and NASDAQ) during the week of 12-16 Feb 2007. On each exchange, 40 stocks were randomly selected from all stocks listed on the exchange that week. The volatility was calculated for each stock as the absolute value of the proportional change in stock price. A value of 0% indicates no change in stock price. An exchange with an average change of 5% is less volatile than one with an average change of 10%.

Summary statistics for each exchange are:

Exchange	n	average	s.d.
NYSE	40	2.91%	2.23%
NASDAQ	40	4.39%	3.37%

(a) Estimate the pooled standard deviation, s_p.
(b) How many degrees of freedom do we have?
(c) Calculate and interpret a 95% confidence interval for the difference in mean volatility (NASDAQ - NYSE).
(d) Assume that the sample size was 20 in each group, but the observed mean and s.d. are unchanged. Calculate a 95% confidence interval for the difference in mean volatility (NASDAQ - NYSE). (Hint to save some computation: the pooled s.d. does not change.)
(e) Assume that the sample size was 100 in each group, but the observed mean and s.d. are unchanged. Calculate a 95% confidence interval for the difference in mean volatility (NASDAQ - NYSE).
(f) How does your interval change as the sample size increases? Explain (briefly) why this is reasonable.

Problem 5

The data set `TXBirths` in the `resampledata` package contains data on babies born in Texas in 2004 (see Case Study 1.2 in your textbook for more information).

(a) Create plots to investigate the distribution weights of babies in each group. Comment on the shape of the distributions.

(b) Calculate and interpret a 90% confidence interval for the difference in mean weight between the groups (No - Yes).

(c) Calculate a 90% bootstrap percentile interval for the difference in mean weight between the groups (No - Yes).

(d) How do the confidence intervals from parts (b) and (c) compare?