STAT 830

Problems: Assignment 2

1. Consider the empirical distribution funct $\hat{F}_n(x)$ for a sample $X_1, \ldots, X_n$ from a cdf F. In this problem I want you to compare several confidence limits for $F(x)$:

   - The pointwise interval based on the approximately normal pivot

   $$\frac{\sqrt{n}(\hat{F}_n(x) - F(x))}{\sqrt{\hat{F}_n(x)[1 - \hat{F}_n(x)]}}$$

   - The pointwise interval based on the approximately normal pivot

   $$\frac{\sqrt{n}(\hat{F}_n(x) - F(x))}{\sqrt{F(x)[1 - F(x)]}}$$

   - The simultaneous interval based on the Dvoretsky-Kiefer-Wolfowitz inequality as described in the text.
   - The simultaneous interval based on the assertion that

   $$\sup_x\{\sqrt{n}|\hat{F}_n(x) - F(x)|\} \xrightarrow{d} \sup_x\{|B_0(x)|\}$$

   where B_0 is a Brownian Bridge.

   I want you to do the following to make the comparison for 95% intervals:

   (a) Generate a sample of size 20 from the Uniform[0,1] distribution. Plot, on one graph, the 4 intervals above along with the true cdf F for x running from 0 to 1.

   (b) Generate 1000 samples of size 20 from the same distribution and for each x in $\{0.1, 0.2, \ldots, 0.9\}$ estimate the pointwise coverage probability for each procedure.

   (c) For the same samples estimate the simultaneous coverage probability of all 4 intervals for the set of 9 x values in the previous problem. Please do the same for the 99 values $i/100$ for $i = 1, 2, \ldots, 99$.

Then I want you to summarize in a paragraph the conclusions of the comparisons. In making the comparisons you need to know that

$$P\left(\sup_x\{|B_0(x)|\} \geq 1.358\right) = 0.05.$$

Your summary will take the form of a paragraph or two written in LaTeXin which you discuss the comparisons as if you were advising people on which procedure to use in which circumstances.

You might like to look at the `R` function `ecdf` in case you find it useful.

2. From the text Chapter 6 # 2, p 95.

3. From the text Chapter 6 # 3, p 95.

4. From the text Chapter 7 # 5, p 104.

5. From the text Chapter 7 # 6, p 104.

6. From the text Chapter 8 # 4, p 117. In addition please look back at my results in the notes with $n = 5$ and compute the exact distribution of the "pivots", $\bar{X}^* - \bar{X}$ and $\sqrt{n}(\bar{X}^* - \bar{X})/s^*$ for the particular data set in the on-line notes. Each of these two quantities has 126 possible values and you should plot the probabilities of each of the 126 values of the statistic against the corresponding value. (Some values involve division by 0 and may be omitted from the graphs.) The distribution I have in mind is discrete; it is a conditional given the original data set.

7. From the text Chapter 8 # 5, p 117.

8. From the text Chapter 8 # 7, p 117.

Due date: 3 October 2013.