553.633/433

Homework #8

Due Wed. 10/25/17

Three problems:

4.8 (textbook) (Note: The reference to "simulation time" in part (c) refers to a value of T potentially different than T in parts (a) and (b) (not to, say, computer CPU time)).

A. The Kantorovich inequality mentioned in class lecture (and handout slides) relates $1/E(X)$ to $E(1/X)$ for a positive random variable X. In particular, if $0 < m \le X \le M < \infty$ for some positive constants $m \le M$, the Kantorovich inequality states that

$$1 \le E(X)E\left(\frac{1}{X}\right) \le \frac{(m+M)^2}{4mM}.$$

Suppose that for all i it is known that the samples τ_i in a regenerative process satisfy $0 < m \le \tau_i \le M$ and that the samples $R_i \ge 0$. Further, suppose that the τ_i and R_i are independent of each other for all i (so the sample path for collecting the τ_i is distinct and independent of the sample path for collecting the R_i). Do the following:

(a) Use the Kantorovich inequality to derive an upper bound to the magnitude of the bias in the regenerative estimate $\hat{\ell}$ in terms of m and M.

(b) What are the numerical values of the upper bound in part (a) as a function of the ratio M/m for M/m = 1.1, 1.5, and 2.0? Given the answers as a fraction of the true value ℓ. Also give a brief (~2 sentence) interpretation of the results relative to potential practical applications.

B. Reproduce and expand on the results in slide 17 of the online class handout, `Chap4_bootstrap_handout.pdf` (use current version at the website). In particular do parts (a) and (b):

(a) Generate an original sample of 5000 replicates of sample means (each the average of 25 points) and give a histogram (like the left plot on slide 17). Then resample from the original sample (i.e., from the 5000 sample means) according to the EDF and produce a histogram of the bootstrap sample (like the right plot on slide 17). Note that the bootstrap samples do *not* need to be built from the individual points going into the sample means; rather, the bootstrap samples should come directly from the sample means themselves.

(b) Compare the two-sided 95% uncertainty bounds (i.e., 2.5% in each tail) centered around 0 of: (i) original sample, (ii) bootstrap sample, (iii) central limit theorem. Provide a brief discussion on how the three bounds differ and why they are like they are. Use the true variance, not the estimated variance, for part (iii). Note: The problem asks you to center around 0, which is convenient when considering the central limit theorem. However, it is trivial to modify the bounds to shift to the true mean 0.5. Please submit your answers centered around 0.