MATH 252-01: Probability and Statistics II

Problem Set 3

Assigned 2016 September 6
Due 2016 September 13

Show your work on all problems! If you use a computer to assist with numerical computations, turn in your source code as well.

1 Devore Chapter 7, Problem 4

2 Devore Chapter 7, Problem 22

3 Devore Chapter 7, Problem 34

4 Computational Exercise (Extra Credit)

Suppose that that a proportion p of the electorate prefers candidate A to candidate B, and a poll is conducted which selects a truly random sample of $n = 1000$ voters. The polling firm wishes to quote a 95% CL (i.e., $\alpha = .05$) interval on their estimate of p. If they find $n\hat{p}$ voters expressing a preference for candidate A, the lower and upper ends p_- and p_+ of the CI are given by

$$p_\pm = \frac{\hat{p} + z^2_{\alpha/2}/(2n)}{1 + z^2_{\alpha/2}/n} \pm \frac{z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n + z^2_{\alpha/2}/(4n^2)}}{1 + z^2_{\alpha/2}/n}$$

a. Plot p_- and p_+ versus $\hat{p}$ for each possible value of $\hat{p}$ between 0 and 1, keeping in mind that $n\hat{p}$ must be an integer.

b. The distance between the midpoint of the confidence interval and one end is sometimes called the "margin of error" associated with the finite sample size. Plot the margin of error $\frac{p_+ - p_-}{2}$ as a function of $\hat{p}$ for each possible value of $\hat{p}$ between 0 and 1.

c. Check the coverage of the confidence intervals when the true value of p is .5 by generating $N = 1{,}000{,}000$ binomial random variables each with $n = 100$ and $p = .5$. (You can do this in numpy with the command `binomial(n,p,N)`.) Determine what fraction lie below, within, and above the confidence interval (p_-, p_+).

d. Repeat part (c) for a true value of $p = .7$.