Problem Set 5

Assigned: July 18
Due: August 1

Use MATLAB for all these except problem 3.

The point of problems 1 and 2 is only in part that you should know how to do these calculations. You should also think about what these mean, intuitively, about the shape of the distribution. I'll discuss these when I give the solutions.

Problem 1

You have a box with 4 coins.
2 of the coins are weighted so that they come up heads with probability 0.1.
2 are weighted so that they come up heads with probability 0.9.

A. Suppose that you pick a coin at random and flip it. Let D be a random variable which is 1 if heads and 0 if tails. What is $\mathrm{Var}(D)$? What is $\mathrm{Std}(D)$?

B. Suppose that you pick a coin, flip it, put it back, pick again at random, and flip. Let E be the number of heads. What is $\mathrm{Var}(E)$? What is $\mathrm{Std}(E)$?

C. Suppose that you pick a coin, and flip it twice. Let F be the number of heads. What is $\mathrm{Var}(F)$? What is $\mathrm{Std}(F)$?

D. Suppose that you pick two coins, and flip each once. Let G be the number of heads. What is $\mathrm{Var}(G)$? What is $\mathrm{Std}(G)$?

Problem 2

You have a weighted coin that comes up heads with probability 3/4. You flip it 40 times. Let H be the number of heads; thus H follows the binomial distribution $B_{40,0.75}$.

A. What are $\mathrm{Exp}(H), \mathrm{Var}(H)$, and $\mathrm{Std}(H)$? (Use the formulas at the end of section 9.6.2.)

B. What are $P(H = 25), P(H = 30), P(H = 31), P(H = 35)$?

C. Compute $P(25 \leq H \leq 30)$ exactly. Compute the estimate given by the Gaussian approximation (Use 24.5 and 30.5 as the corresponding bounds on the Gaussian.) How close are they?

D. What is $P(H \leq 15 | H \leq 30)$? What is $P(H \leq 5 | H \leq 10)$?

Problem 3

Using paper and pencil, but no computer, estimate confidence intervals for polls over sample sizes with $N = 2500$, and $N = 160,000$; $\bar{f} = 0.4$ and $\bar{f} = 0.8$; and confidence levels of 0.95 and 0.999. (Consider all combinations; thus your answer should have 8 parts.)

Problem 4

Using Monte Carlo sampling, estimate the volume of the three dimensional region

$$|x|^{1/2} + 2|y|^{1/3} + 4|z|^{1/4} \leq 4$$

with a 95% confidence interval of 1% of the computed value. That is, you are 95% sure that the difference between the computed value and the true value is less than 1% of the true value. Note that the region fits inside the box $[-16, 16] \times [-8, 8] \times [-1, 1]$. How many sample points do you need?