Due at the beginning of class on Thursday, December 5, 2013. Reminder: Show your reasoning!

Chapter 8, Sections 8.1-8.4, and first part of Section 8.5 (on one-sided Chebyshev); handout “Notes on the Chebyshev Inequality”.

Examples to read carefully:

Chapter 8: 2a, 2b, 3a, 3b, 3c, 3d, 5a.

(1). (20 points) The average number of home foreclosures in Stony Brook is 2 per month.

(a). Estimate the probability, \( p \), that at least 5 foreclosures occur in Stony Brook next month. (What inequality are you using?)

(b). Assume now (for parts (b), (c), and (d)) that you are told that the variance of the number of foreclosures in Stony Brook in any one month is 5. Now give an improved estimate of \( p \) (using an inequality).

(c). Give a Central Limit Theorem estimate for the probability \( q \) that during the next 4 years there are more than 105 foreclosures in Stony Brook.

(d). Use an inequality to get the best bounds you can on the probability \( q \) estimated in part (c).

(2). (15 points) The cash register is broken at the bagel shop, so transactions are done by hand and recorded. Your lazy employee, Joe, charges customers exactly (e.g., if they owe $3.61, he makes sure they pay exactly $3.61); however, lazy Joe only records in the notebook the dollar amount of each transaction (e.g., $3 for the $3.61 transaction), omitting the cents. Assume that Joe serves 20 customers today. Find an upper bound on the probability that the total shown in Joe’s notebook is at least $15 less than the actual amount of money those 20 customers paid to Joe. (Hint: Let \( X_i \) be the amount (in dollars) Joe fails to record when he writes the \( i \)th customer’s purchase in the notebook. Then, a reasonable model is that \( X_i \) is Uniform(0,1). Possibly more realistic is to model \( X_i \) as a discrete uniform on the set \{0, 0.01, \ldots, 0.99\}. You are welcome to use either model (they give virtually the same result).)

(3). (15 points) You flip a fair coin \( n \) times and keep track of the sample mean, \( \bar{X}^{(n)} \) (the fraction of heads among the \( n \) flips). Of course, when \( n \) is very large, you expect that the random variable \( \bar{X}^{(n)} \) will be very close to 0.5 (since the coin is fair).

(a). (8 points) Use the Central Limit Theorem to estimate how large \( n \) must be in order for you to be 95\% confident that \( \bar{X}^{(n)} \) is between 0.45 and 0.55.

(b). (7 points) Use an inequality to obtain a number \( K \) such that you can guarantee that if \( n \) is at least \( K \), then the probability that \( \bar{X}^{(n)} \) is between 0.45 and 0.55 is at least 0.95.

(4). (15 points) A group of 100 music fans is going to a concert. Each fan gets 1 ticket plus (perhaps) some extra tickets. The number of extra tickets each fan gets is a random integer between 0 and 4 (inclusive). Thus, including his own ticket, a fan gets \( j \) tickets with
probability $1/5$, for $1 \leq j \leq 5$. We are interested in estimating the probability, $p$, that the fans in total receive at least 310 tickets. (a). Give an estimate based on CLT. (b). Give the best guaranteed bounds you can on $p$ (i.e., find $a$ and $b$ so that $a \leq p \leq b$, and $b - a$ is as small as possible).

(5). (15 points) A lake contains 5 distinct types of fish. Suppose that each fish caught is equally likely to be any of these types. Let $Y$ denote the number of fish that need be caught to obtain at least one of each type. (a). Give an interval $(a, b)$ such that $P(a \leq Y \leq b) \geq .95$. (b). Using the one-sided Chebyshev inequality, how many fish need we plan on catching so as to be at least 95 percent certain of obtaining at least one of each type?

(Hint: Refer to the “coupon collecting” example we did in class (and you read about in the text); this lets you write $Y$ as a sum of 5 geometric random variables, so you can easily compute the mean and variance of $Y$. In part (a), use a two-sided Chebyshev, in part (b) use a one-sided Chebyshev.)

(6). (20 points) A student is learning to spell some difficult words. She estimates that it takes her 7 attempts to learn to spell each word, with a standard deviation of 1.9 (in units of “attempts”). We want to estimate the probability that she will require 400 or more attempts to learn to spell 50 words. (a). Estimate the probability using the CLT. (b). Compute the tightest bounds (upper and lower) that you can guarantee on the probability.