Problem Set 11

For this homework, you can use the tool at `http://inst.eecs.berkeley.edu/~cs70/sp07/normcdf/` to compute, for r.v. X with the standard normal distribution, either the value of $\Pr[X < a]$ or, given some b, an a such that $\Pr[X < a] = b$.

1. **Confidence Intervals for Letter Grades** (4 pts)

   The grades on the CS70 final exam are approximately normally distributed with a mean of 75 and a standard deviation of 11, find

   (a) The lowest passing grade if the lowest 5% of the students are given F's[1].

   (b) The highest B if the top 10% of the students are given A's.

2. **Elevator insurance** (4 pts)

   A new hotel is being constructed. The elevator needs to carry up to 25 people at a time, and is designed to withstand a load of up to 5000 lbs. Assume the average weight of guests is 170 lbs., and the standard deviation is 40 lbs.

   (a) Use the Chebyshev bound to bound the probability that a full elevator weighs more than the limit. Assume passenger weights are i.i.d.

   (b) When the elevator is full, what is the confidence level of the weight being within the design limit? (Use the normal approximation.)

3. **Memorylessness** (4 pts)

   (a) Prove that if X is a geometric r.v., it's *memoryless*, that is, $\Pr[X - a = b | X > a] = \Pr[X = b]$.

   (b) Prove that if an r.v. X with positive integer values is memoryless, then its distribution must be geometric.

4. **Putting CLT to the test** (8 pts)

   For any (non-trivial) probability space Ω of your choice, define some random variable X on Ω (do not choose a random variable that is too simple — e.g., a constant r.v. — but you don't have to choose one that's very complicated, either). For this problem, you will experimentally "verify" the idea that "all distributions approach the normal distribution asymptotically."

   Let A_n be the random variable that is the average of n independent samples from the distribution of X. In a reasonable programming language of your choice, write a program that takes input $n \geq 1$ and $k \geq 1$ and does the following: samples n times from Ω, averages the values of X on those sample points, and repeats this k times. This is equivalent to sampling the random variable A_n k times. Plot the distribution for various values of k — in particular, choose k large enough so that the distribution is clear.

   The central limit theorem says that the distribution of A_n should approach the normal distribution as n goes to infinity. How large does n get before you observe a noticeable "convergence"? (You don't need to define a numeric measure of "convergence" for this problem.)

   Submit a printout of your source code as well as all output files, plots, etc that you produce while working on this problem. You may use any programming language for which there's a commpiler or interpreter on EECS instructional machines.

[1] Disclaimer: The numbers in this problem have nothing to do with the grading curve in CS70. No, really.