Problem Set 11

1. **Machine failures**

   Two faulty machines, M_1 and M_2, are repeatedly run synchronously in parallel (i.e., both machines execute one run, then both execute a second run, and so on). On each run, M_1 fails with probability p_1 and M_2 with probability p_2, all failure events being independent. Let the random variables X_1, X_2 denote the number of runs until the first failure of M_1, M_2 respectively; thus X_1, X_2 have geometric distributions with parameters p_1, p_2 respectively.

   Let $X = \min\{X_1, X_2\}$ denote the number of runs until the first failure of *either* machine. Show that X also has a geometric distribution, with parameter $p_1 + p_2 - p_1 p_2$.

2. **Geometric Distribution**
   James Bond is imprisoned in a cell from which there are three possible ways to escape: an air-conditioning duct, a sewer pipe and the door (which is unlocked). The air-conditioning duct leads him on a two-hour trip whereupon he falls through a trap door onto his head, much to the amusement of his captors. The sewer pipe is similar but takes five hours to traverse. Each fall produces temporary amnesia and he is returned to the cell immediately after each fall. Assume that he always immediately chooses one of the three exits from the cell with probability $\frac{1}{3}$. On the average, how long does it take before he realizes that the door is unlocked and escapes?

3. **Poisson Distribution** A textbook has on average one misprint per page. What is the chance that you see exactly 4 misprints on page 1? What is the chance that you see exactly 4 misprints on some page in the textbook *of 250 pages*?

4. **Normal Distribution**
   The average life of a certain type of engine is 10 years, with a standard deviation of 3.5 years. The manufacturer replaces free all engines that fail while under guarantee. If he is willing to replace only 2% of the engines, how long a guarantee should he offer? Assume a normal distribution.

5. **Multiple Choice**
   A multiple-choice quiz has 100 questions each with four possible answers of which only one is correct. What is the probability that sheer gusswork yields between 10 and 30 correct answers for the 40 of the 100 problems about which the student has no knowledge?

6. **Load balancing in action**

   This problem asks you to try out in practice the load balancing scheme discussed in class and to compare the results with the theory. It will also lead you on to try out a more sophisticated strategy that (we hope!) works even better.

   (a) You will find on the web page a Scheme program (and a java program if you prefer) that simulates the load balancing experiment (i.e., balls-and-bins with n balls and n bins). The function called bb on input n will run the experiment with n balls and n bins, and output a single number which is the maximum load.

   Perform 20 simulations with $n = 1000$ and 20 simulations with $n = 10^6$, and draw up a histogram of the maximum loads you observe. How well do these values correspond to those that we derived in class? Report the sample mean and variance of your trials.

   (b) Now consider the following alternative scheme, which uses a minimal amount of communication: jobs arrive in sequence as before, but instead of simply choosing a single processor at random, each job now chooses *two* processors at random, inspects their current loads, and goes to the less heavily loaded of the two. (If both loads are the same, the job chooses one of the two arbitrarily.)

Modify the program to implement this strategy. Again, perform 20 simulations with $n = 1000$ and $n = 10^6$ and tabulate the maximum loads you observe. Are you surprised by your results? What are the sample mean and variance this time?

(c) [**Extra credit**] Would you expect a similarly dramatic effect if the jobs were allowed three choices rather than two? Modify your program again and see what happens.

7. **Non-transitive Dice** (**Extra Credit**)
   Consider three six-sided dice A, B, and C, whose sides are labeled with any numbers of your choice. Say that A beats B if $P[A$ shows a number greater than $B] > 1/2$.

   Choose numbers to label each of the sides of A, B and C such that A beats B, B beats C, and C beats A. Try to create your labels to maximize the probabilties in each case.

   Notice that armed with a set of three non-transitive dice, you could challenge someone to a game of dice, where to ensure fairness, you would give your opponent first choice of die to play with!