Final	5 Questions	Purdom
5-2-08	Analysis of Algorithms	B503

1. Consider a randomly selected integer in base b with n digits, where each such integer is equally likely. What is the probability that the leading (most significant) digit of the number is zero.

2. Consider two integers in base b with n digits where the leading digit is not zero that are selected independently at random (each possible integer equally likely). The product of the two integers will need either $2n$ or $2n-1$ digits. For the cases below, how likely is the product to need only $2n-1$ digits (in base b).
   2a. $b = 2$, $n = 1$;

   2b. $b = 2$, $n = 2$;

   2c. $b = 4$, $n = 1$;

   2d. with general b (with b integer and greater than 1), in the limit as n becomes large.

3. Suppose you have a data set with P people, S of the people with family name Smith, and J of the people with first name John.
   3a. Give the best (tightest) lower bound for the number of people that have both family name Smith and first name John.

   3b. Give the best upper bound for the number of people that have both family name Smith and first name John.

   3c. If first names are selected independently of family names (not quite true in practice), what is the expected number of people that have both family name Smith and first name John.

4. Let $x = y + y^2 + y^3 + y^4 + y^5$. Find an approximate formula that gives y and a function of x where your approximation is good when x and y are both near zero.

5. Suppose you run Quicksort and it happens on every split (partition) that 2/3rds of the numbers are in one part and the other $1/3rd$ are in the other part (not counting the splitting number which is in a part by itself).

   5a. Write a recurrence that gives the running time for Quicksort in this particular case.

   5b. Solve the recurrence.

   5c. Compare the solution (question 5b) with the time when the split happens to divide the file exactly evenly (once again, except for the part for the splitting number).