Stat 321 - Day 5
"Birthday Problem"
(An Application of Counting Methods and Basic Probability Rules)

Activity 5: Birthday Problem
Situation: What is the probability that at least two people in your class share the same birthday? How does this probability vary as the number of people in the group varies? How many people must be present for this probability to exceed one-half? How about to exceed a 90% chance?

(a) Record the number of people in class today, and then take a guess for the probability of finding at least two people with the same birthday in our class.

(b) Now take a guess for the smallest number of people who must be present in order for this probability of matching birthdays to exceed one-half. Then do the same for exceeding .9.

Analysis (Simple Case):
Let n denote the number of people present. In order to tackle this problem, we need to make some assumptions about people's birthdays. First, let us ignore the complication of February 29 birthdays. Second, we will assume that people are equally likely to be born on any of the (remaining) 365 days of the year.

First consider the fairly simple case of n=4 people. Because of our "equally likely" assumption, the probability of matching birthdays will be the number of ways in which there could be at least one birthday match divided by the total number of ways that four people could have birthdays. But this numerator is very hard to calculate, so we will first count the number of ways in which four people can have *different* birthdays from each other. We will use the "general product rule" of counting to determine these, starting with the denominator.

(c) How many choices are there for the first person's birthday? How many for the second person (with no restrictions)? For the third? For the fourth? Thus, how many total ways are there for four people to have birthdays?

(d) Now consider the number of birthdays for which there is *no* match. Again, how many choices are there for the first person's birthday? Then, in order for their birthdays not to match, how many choices remain for the second person? For the third? For the fourth? Thus, how many ways are there for four birthdays not to match?

(e) Determine the probability of o matching birthdays in the n=4 case by dividing your answer to (d) by your answer to (c).

(f) Use your answer to (e) and the relevant probability rule to find the probability that there *is* a birthday match among a group of n=4 people.

(g) Follow this same procedure to find the probability of at least one birthday match in a group of n=8 people.

Analysis (General Case):

(h) For what values of n will the probability of {at least one match} be equal to one (so that the group is *certain* to have matching birthdays)? Explain. [*Hint*: Remember to consider our simplifying assumptions.]

(i) Now consider the general case of a group of size n. Start with the denominator: how many possibilities are there for these n birthdays?

(j) Now consider the numerator for the complement: how many possibilities are there for which there are *no matches* among the n birthdays?

(k) Use the complement rule with your answers to (i) and (j) to write an expression for the probability of matching birthdays in a group of n people.

Evaluating Probabilities:

We will turn to the spreadsheet package Excel to evaluate this function. First, let A_n denote the event that a matching birthday exists among a group of n people, so A_n' denotes its complement. We can write $P(A_n') = \frac{365}{365} \cdot \frac{364}{365} \cdots \frac{(365-n+1)}{365} = \frac{P_{n,365}}{365^n}$, so we can write the following iterative expression $P(A_n') = \frac{(365-n+1)}{365} P(A_{n-1}')$ for calculate the probability of no birthday match.

(l) Open the Excel file "birthday.xls." Confirm that column B contains this iterative expression. Then use the "fill down" feature to complete column B, and then enter the appropriate formula and complete column C. Record the probabilities of matching birthdays for the values of n presented in the table:

n	3	5	10	20	30	50	75	100
Pr(match)								

(m) From the spreadsheet identify the smallest value of n for which the probability of matching birthdays exceeds one-half. What is this probability? How was your guess in (b)? Then repeat this for the probability to exceed .9.

(n) Use Excel's "chart wizard" to create a graph of the probability of matching birthdays as a function of n. Comment on its behavior.