Stat 321 - Day 3
Counting Techniques

Activity 3: College Committee Formation
A college professor found herself to be one of two women on a six-person committee. The committee needed to choose two representatives to speak for the committee to other groups, and both women were selected. You will investigate how unusual this would be if the two representatives had in fact been chosen at random.

(a) Let the names of the committee members be: Alice, Barb, Chuck, Dave, Ethan, and Fred. Using initials to represent these people, list the sample space (set of all possible pairs) for this process.

(b) How many outcomes are in this sample space?

(c) If the two representatives had been chosen at random (without regard to gender), all possible pairs of the six members would have been equally likely. What then is the probability for each of the possible pairs?

(d) Assuming equal likeliness, what is the probability that both persons selected would be women?

(e) Does this probability indicate that it would be very surprising to find both representatives women if all possible pairs were equally likely? Explain.

(f) Assuming equal likeliness, what is the probability that both persons selected would be men?

(g) Assuming equal likeliness, what is the probability that one person of each gender would be selected?

(h) What do you notice about the sum of the probabilities in (d), (f), and (g)? Explain why this makes sense.

(i) Now suppose that the committee had been choosing a chair and a secretary rather than just two representatives. In other words, suppose now that order matters, so Alice as chair and Barb as secretary is a different outcome than Barb as chair and Alice as secretary. Without listing them, indicate how many outcomes comprise this sample space. How does this number compare to the previous one where they were simply two representatives? Explain why this makes sense.

(j) In this new scenario, how many of the outcomes consist of two women? What is the probability that the chair and secretary would both be women? Has this probability changed from the one calculated with the original sample space (where order did not matter)?

In many probability calculations, it does not matter whether you count as if order matters or not, as long as you *count consistently* in both the numerator and denominator of the probability calculation. In this case, the probability of choosing two women is 1/15 or 2/30 depending on how you count.

More on Counting:

When selections are made "at random," the equal likeliness principle enables us to calculate probabilities by enumerating the possible outcomes in the sample space and counting how many outcomes comprise the event of interest. Clearly, though, it can be quite cumbersome to list the outcomes in a sample space. It would be very helpful to develop some general techniques for *counting* the number of possibilities without listing them.

We will derive a general rule for counting how many outcomes exist for choosing k people or objects from a group of n people or objects. Our starting point will be the *general product rule*: If a process consists of multiple stages (call the number of stages k) and if stage i can be completed in n_i ways regardless of which outcomes occur in earlier stages, then the process itself can be completed in $n_1 n_2 \cdots n_k$ ways.

(k) Suppose that the committee of six were to choose three officers: chair, secretary, and refreshment provider. Use the general product rule to determine how many different choices are possible. [*Hint*: How many ways are there to select the chair? Once that first person is selected, how many ways are there to choose the secretary? Once those two are selected, how many ways are there to choose the refreshment provider?]

(l) If a committee of size n were to select officers for these three positions, how many choices would they have? [*Hint*: Express your answer in terms of n.]

(m) Write an expression for the number of ways to choose k specific officers from a committee of n members. Then write this as a ratio of two factorial expressions.

You have found that the number of *permutations* (where order matters) of n objects taken k at a time is: $n(n-1)(n-2)\cdots(n-k+1)$. Notice that there are k terms in this product. This can also be expressed as: $P_{k,n} = \dfrac{n!}{(n-k)!}$.

(n) Now suppose that instead of choosing three specific officers, the committee had to simply choose a subcommittee of three representatives. Calculating the number of permutations will overcount the number of such subcommittees, since any subcommittee with the same three people is considered equivalent. To determine the *factor* by which a permutation overcounts, use the general product rule to calculate the number of arrangements that consist of the same three people. In other words, if you were considering order, in how many ways could a subcommittee of three people be selected?

(o) To adjust for this overcount in determining the number of ways of choosing three representatives from the committee of six, divide your answer to (l) by your answer to (n).

(p) If a committee of size *n* were to select three representatives, how many choices would they have? [*Hint*: Follow the previous question to calculate this.]

(q) Write an expression for the number of ways to choose *k* representatives from a committee of *n* members. [*Hint*: Follow the previous two questions and use your answer to (m).]

You have discovered that the number of *combinations* (where order does not matter) of *n* objects taken *k* at a time is: $C_{k,n} = \binom{n}{k} = \frac{n!}{(n-k)!k!}$. This symbol $\binom{n}{k}$ is often pronounced "*n* choose *k*" because it is the number of ways of selecting a group of *k* objects from a collection of *n* objects.

(r) Calculate $\binom{6}{2}$ from this expression, and verify that it equals the number of pairs in your sample space in (a).

<u>Hypergeometric Probabilities:</u>

Suppose that a *population* of *N* objects consists of *r* "successes" and *N-r* "failures." Suppose that a *sample* of *n* objects is to be chosen at random. Consider the probability of drawing *x* successes in the sample.

(s) How many ways are there to choose *n* objects from the population of *N* objects?

(t) How many ways are there to select *x* successes from the population of *r* successes?

(u) How many ways are there to select *n-x* failures from the population of *N-r* failures?

(v) Produce an expression for the probability of obtaining *x* successes in the sample of *n* objects by multiplying your answers to (t) and (u) and then dividing by your answer to (s).

You should have found that the probability of drawing *x* successes in a sample of size *n* from this population is: $\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$. These are called *hypergeometric* probabilities.

Return to considering the population of four men and two women, with a sample of size two to be drawn from that population.

(w) Use the above expression for hypergeometric probabilities to find the probability of selecting two women. Then do the same for the probability of choosing one woman. Then repeat for zero women. Do these agree with your calculations in (d), (f), and (g)?

(x) Use Minitab to verify these calculations by putting the values 0, 1, and 2 into `c1` and then selecting `Calc > Probability distributions > Hypergeometric`. Be sure to click on the "Probability" option and enter the appropriate parameter values.

Activity 1: Friendly Observers (cont.)
Reconsider the psychology experiment about whether an observer with a vested interest inhibits one's performance on a skilled task. To assess whether the experimental results (3/12 successes in group A vs. 8/12 successes in group B) were unlikely to occur just by random chance, you performed a simulation analysis. Now you are ready to calculate this probability exactly using counting techniques, specifically hypergeometric probabilities.

(y) Specify the relevant hypergeometric parameters N, r, and n for calculating the probability of finding three or fewer successes in group A just by chance.

(z) Calculate the probability in (y) either by hand or by using Minitab. Then comment on how closely your empirical estimates of this probability came to the actual value. Finally, summarize whether this probability suggests that the experimental results are very unlikely to have occurred by chance alone if there were no effect of the observer with the vested interest.