Probability measures the likelihood that an event will occur, quantified as a number between 0 and 1 (where 0 indicates impossibility and 1 indicates certainty). The simplest model involves phenomena where every outcome is equally likely, e.g. throwing a dice or flipping coins.

As with every mathematical model, it is not an exact copy of the physical world, but is designed to replicate some aspect of it. There are two major schools of thought on what probability should model.

1. Objectivists assign numbers to describe some objective or physical state of affairs. In particular, the proponents of frequentist probability claim that the probability of a random event denotes the relative frequency of occurrence of an experiment’s outcome, when repeating the experiment, i.e. “in the long run” of outcomes.

2. Subjectivists assign numbers per subjective probability, i.e., as a belief system about the chance of an event occurring. In particular the proponents of Bayesian probability claim that all experts will eventually have sufficiently similar assessments of probabilities of a single event, given enough evidence.

Later the effect of these differences will become clear.

As probability matured, mathematicians started looking at experiments with infinitely many outcomes. For example, one could look at all the points a dart could hit on a solid circle. The probability of hitting any given point is zero since the point is arbitrarily small, but the probability of hitting some point on a circle is strictly positive. A rigorous framework for handling such experiments was established in 1930s by Kolmogorov (building on the work of Lebesgue), and this framework is still used today. While we will not study this framework (measure theory and Lebesque integration) in this class, we will see a
few examples demonstrating when such framework becomes necessary.

The probability where every outcome is equally likely amounts to counting. Therefore, we will go over the counting principles in Chapter 1 of Ross. The theory of counting is called *combinatorial analysis*.

**Theorem 1.** *(The generalized basic principle of counting, Ross, Section 1.2, a/k/a Multiplication Principle)* Each distinct element of a set is determined by first making one of \( n_1 \) choices. For each of these choices, you then must make one of \( n_2 \) choices, and so forth until you have made your \( r \)-th choice. Then the set contains \( n_1 \cdot \ldots \cdot n_r \) elements.

*Proof.* In the base case, we make one of \( n_1 \) choices and then one of \( n_2 \) choices. Then all the elements can be enumerated as matrix of \( n_1 \) rows and \( n_2 \) columns, where the entry in the \( i \)-th row and \( j \)-th column represents the element determined by making the \( i \)-th of \( n_1 \) choices followed by the \( j \)-th of \( n_2 \) choices. There are \( n_1 \cdot n_2 \) elements.

For the inductive case, the set is determined by making \( r - 1 \) choices followed by the \( r \)-th choice, and by the inductive hypothesis we have \( n_1 \cdot \ldots \cdot n_{r-1} \) outcomes of the \( r - 1 \) choices. Similarly to the above, we can enumerate these elements as a \( (n_1 \cdot \ldots \cdot n_{r-1}) \times n_r \) matrix, which has \( n_1 \cdot \ldots \cdot n_r \) elements. \( \square \)

This procedure can be extended to count more complex objects. Fix a set \( S \) of size \( n \).

(1) **Order matters**

(a) The number of ordered lists of size \( k \) chosen from \( S \) with repeats allowed is \( n^k \).

(b) The number of ordered lists of size \( k \) chosen from \( S \) without repeats allowed is \( n(n - 1)\ldots(n - k + 1) \).

(c) Each different ordered arrangement of a set \( S \) is called a *permutation* (Ross, 1.3).

(i) If \( S \) has \( n \) distinguishable objects we can see that \( S \) has \( n(n - 1)\ldots1 = n! \) permutations.
Proof: The first object is determined by making \( n \) choices, the second object is determined by making \((n-1)\) choices, etc. Therefore, by the generalized principle of counting, there are \( n(n-1)...1 = n! \) permutations.

(ii) If \( S \) has \( n \) objects, of which \( n_1, n_2, ..., n_r \) are alike, we see that \( S \) has \( \frac{n!}{n_1!n_2!...n_r!} \) permutations.

Proof: Label each class of \( n_i \) objects that are alike by labels from 1 to \( n_i \) so that to get a new set \( S' \) of \( n \) distinguishable objects. By the previous result, \( S' \) has \( n! \) permutations. However, by permuting the objects within each newly labeled class, we don’t get any new permutations of \( S \). By the permutation formula we have \( n_i! \) permutations within each such class, and by the general counting principle, there are \( n_1! \cdot ... \cdot n_r! \) total permutations of \( S' \) that don’t result in permutations of \( S \). Therefore, we have \( \frac{n}{n_1!...n_r!} \) permutations of \( S \).

(2) Order does not matter /Combinations (Ross 1.4) Now suppose we want to choose a subcollection of \( r \) objects from \( S \). The number of ways we can do this is given by \( \frac{n!}{r!(n-r)!} = \binom{n}{r} \)

Proof: (Counting) We can prove the immediately preceding result by counting the same quantity twice. By the multiplication principle, there are \( n(n-1)...(n-r+1) \) ordered sequences of length \( r \) taken from a set of size \( n \). Counted differently, first select one of the \( \binom{n}{r} \) (unordered) subsets of size \( r \) to use. Then count all of the \( r! \) permutations for these \( r \) elements. So we have \( n(n-1)...(n-r+1) = \binom{n}{r}r! \)

For applications, see Ross, Examples 1.4 (a), 1.4(b) & 4(c).

Theorem 2. (Ross, Chapter 1, (4.1)) If \( 1 \leq r \leq n-1 \) then

\[
\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}
\]

Proof. (Counting) Choose an arbitrary element of \( n \) and label it as ”1”. For each possible subcollection of \( r \) elements, ”1” is either included or not. Therefore,
\[ \binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r} \]

where the first term represents the number of subcollections of \( r \) elements that contain "1" and the second term represents the number of subcollections that don’t.

□

References