Probability

In it's most general usage, probability is the numerical language of uncertainty:

> The chance of rain tommorrow is 35%.
>
> Given their performance to date, the Yankees have a 70% chance of winning the World Series.

The basis for these probabilities is complex, and subjective.

We will look only a probabilities that are based on how frequently things happen over many repetitions:

> A fair coin will land heads half the time, in the long run.
>
> A pair of fair dice produce double-six one time in 36, over a large number of rolls.

When individual outcomes are impossible to predict, we use these frequencies to define probabilities.

Uses of Probability in Statistics

Probabilities get into statistics in (at least) two ways:

- When we introduce randomization into an experiment.
- When we use probability as a model for apparently random phenomena in the real world.

We then use the mathematics of probability to judge how much random variation has affected our results. In particular, we can find the *sampling distribution* of an estimator, and from that, determine its bias and variability.

These mathematical results then *help* us to draw conclusions. But you should't expect to draw conclusions by mathematics alone!

Outcomes, Events, and Probabilities

Probability theory uses the following notions:

- A *sample space* consists of all possible *outcomes* of a random process.

  Example: All possible samples of size 100 drawn from a population of 1000.

- Sets of outcomes, called *events*:

  Example: The event that Fergie Feinbaum is in the sample, which is the set of all outcomes of the sampling that include her.

- An assignment of *probabilities* to events:

  Example: For a simple random sample, all the possible outcomes (samples) have equal probability.

We give events upper-case letters, and use S for the whole sample space. We write the probability of an event A as $P(A)$.

Example: For a roll of a die, $S = \{1, 2, 3, 4, 5, 6\}$. The event of rolling an even number is $E = \{2, 4, 6\}$. If the die is fair, $P(E) = 1/2$.

Basic Rules of Probability

Probabilities must obey the following rules:

Rule 1: For any event A,

$$0 \leq P(A) \leq 1$$

Rule 2: For the entire sample space, S,

$$P(S) = 1$$

Rule 3: The *complement* of an event A, written A^c, is the set of outcomes in S that are not in A. It's probability is

$$P(A^c) = 1 - P(A)$$

Rule 4: If two events A and B are *disjoint* (have no outcomes in common), then

$$P(A \text{ or } B) = P(A) + P(B)$$

Probabilities in a Finite Sample Space

If the sample space has a finite number of outcomes, we can assign probabilities as follows:

- Assign each outcome in S a non-negative probability.

  Ensure that the sum of all these outcome probabilities is one.

- Let the probability of an event, A, be the sum of the probabilities of the outcomes that it consists of.

These event probabilities will obey the rules.

Example: The number of pickles that a canning machine puts in a jar might have probabilities as follows:

0	1	2	3	4	5	6
0.002	0.028	0.070	0.195	0.405	0.296	0.004

The probability of the event $B = \{4, 5, 6\}$ is then

$$P(B) \;=\; 0.405 + 0.296 + 0.004 \;=\; 0.705$$

Assigning Probabilities When Outcomes are Equally Likely

Assigning probabilities is even simpler when all outcomes are equally likely — such as with fair dice, coins, and cards, or when we pick a simple random sample from a population.

Example: If we deal two cards from a well shuffled deck. All pairs of cards are equally likely.

There are 52×51 pairs (if we remember the order of the cards), so each has probability $1/(52 \times 51) = 1/2652$.

What is the probability of the event that the first card is red and the second card is black? There are 26×26 outcomes in this event, so its probability is

$$26 \times 26 \times (1/2652) \;=\; 13/51 \;\approx\; 0.2549$$

Assigning Probabilities Based on Independence of Processes

Sometimes an outcome consists of the results of two independent processes. We can then assign probabilities to combined outcomes by multiplying probabilities for each process.

Recall the probabilities for the number of pickles in a jar:

0	1	2	3	4	5	6
0.002	0.028	0.070	0.195	0.405	0.296	0.004

Suppose we flip a fair coin at the same time as the jar is packed with pickles. What is the probability that the coin lands heads and two pickles are put in the jar?

Since the coin and the jar have no effect on each other, we multiply the probabilities to get $0.070 \times (1/2) \;=\; 0.035$.

Definition of Independence for Events

In general, we say two events are *independent* if finding out whether one occurred tells us nothing about whether the other occurred.

If A and B are independent, we can figure out $P(A \text{ and } B)$, $P(A \text{ and } B^c)$, etc. from $P(A)$ and $P(B)$. First, make a two-way table of probabilities with only marginal probabilities filled in (eg, the table on the left below):

	B	B^c	
A	?	?	0.8
A^c	?	?	0.2
	0.5	0.5	

	B	B^c	
A	0.4	0.4	0.8
A^c	0.1	0.1	0.2
	0.5	0.5	

Then fill in the middle of the table, splitting the marginals in the *same proportions* in each row/column, since knowledge of one event doesn't affect the other.

Mathematically, we *define* independence of A and B to mean that $P(A \text{ and } B) = P(A)P(B)$. It follows that $P(A \text{ and } B^c) = P(A)P(B^c)$, etc.