The result of many thousands of chance outcomes can be known with near certainty. A phenomenon is random if individual outcomes are uncertain, but pattern from many individual outcomes is known. It is assumed that the phenomenon can be repeated indefinitely under essentially the same conditions. Examples of random phenomenon include tossing coins, dice, and cards, and we can use a computer to simulate many repetitions of these types of events.

Example: Rolling Fair Dice

If we roll a single die, we can expect to get one of six possible outcomes: 1,2,3,4,5,6.

This set of all possible outcomes is the sample space S.

An event is the outcome of a random phenomenon, in this example, rolling the die.

A probability model is a description of a random phenomenon consisting of

- the sample space S,
- a way of assigning probabilities to an event.

If we assume the die is fair, that is, all outcomes are equally likely, then the probability model for rolling a single die consists of the sample space $S = \{1, 2, 3, 4, 5, 6\}$ where each element of the sample space has probability $1/6 \sim 0.16666667$.

Event	1	2	3	4	5	6
Probability	1/6	1/6	1/6	1/6	1/6	1/6

Let $P(A)$ represent the probability that event A occurs. For this case, we can write $P(6) = 1/6$, which reads "the probability of rolling a six is one sixth". Since the die is fair, we can write $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$.

The probability model can be expressed as a probability histogram, where the height of each bar shows the probability of the outcome at the base. The heights of the bars all add to 1 by Rule 1 of the probability rules.

Example: Rolling Two Fair Dice

For two dice, there are 36 possible ways to roll the dice (see Fig 8.2 in text), all with equal probability since each of the die are fair:

$$
\begin{array}{cccccc}
\boxed{1}\,\boxed{1} & \boxed{1}\,\boxed{2} & \boxed{1}\,\boxed{3} & \boxed{1}\,\boxed{4} & \boxed{1}\,\boxed{5} & \boxed{1}\,\boxed{6} \\
\boxed{2}\,\boxed{1} & \boxed{2}\,\boxed{2} & \boxed{2}\,\boxed{3} & \boxed{2}\,\boxed{4} & \boxed{2}\,\boxed{5} & \boxed{2}\,\boxed{6} \\
\boxed{3}\,\boxed{1} & \boxed{3}\,\boxed{2} & \boxed{3}\,\boxed{3} & \boxed{3}\,\boxed{4} & \boxed{3}\,\boxed{5} & \boxed{3}\,\boxed{6} \\
\boxed{4}\,\boxed{1} & \boxed{4}\,\boxed{2} & \boxed{4}\,\boxed{3} & \boxed{4}\,\boxed{4} & \boxed{4}\,\boxed{5} & \boxed{4}\,\boxed{6} \\
\boxed{5}\,\boxed{1} & \boxed{5}\,\boxed{2} & \boxed{5}\,\boxed{3} & \boxed{5}\,\boxed{4} & \boxed{5}\,\boxed{5} & \boxed{5}\,\boxed{6} \\
\boxed{6}\,\boxed{1} & \boxed{6}\,\boxed{2} & \boxed{6}\,\boxed{3} & \boxed{6}\,\boxed{4} & \boxed{6}\,\boxed{5} & \boxed{6}\,\boxed{6}
\end{array}
$$

The sample space is $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ and the probabilities are different for each element in the sample space because there are different ways to roll the elements in the sample space.

The event of rolling a 3, call it A, is possible in the following two ways:

$$A = \left\{\boxed{1}\boxed{2}, \boxed{2}\boxed{1}\right\}$$

So $P(A) = 2/36 = 1/18 \sim 0.055556$ assuming the dice are fair.

The event of rolling a 6, call it B, is possible in the five following ways:

$$B = \left\{\boxed{1}\boxed{5}, \boxed{2}\boxed{4}, \boxed{3}\boxed{3}, \boxed{4}\boxed{2}, \boxed{5}\boxed{1}\right\}$$

So $P(B) = 5/36 \sim 0.138889$ assuming the dice are fair.

The works because the dice were assumed to be fair, and the probability of rolling any of the 36 possibilities was the same.

If the die are not fair, then the above will not work. We need to be able to combine the probabilities for single die in a systematic way.

Probability Rules

1. **Any probability is a number between 0 and 1 inclusive.** If A is an event in the sample space S: $0 \leq P(A) \leq 1$.

   An event with 0 probability never occurs; an event with probability 1 occurs on every trial.

2. **All possible outcomes together must have probability 1.** If S is the sample space, then $P(S) = 1$.

3. **The probability that an event does not occur is 1 minus the probability that an event does occur.** $P(A \text{ does not occur}) = P(A^C) = 1 - P(A)$.

4. **If two events A and B are independent, then the probability that one event and the event both occur is the product of the probabilities for each event.** $P(A \text{ and } B) = P(A)P(B)$.

5. **The probability that one event or the other occurs is the sum of their individual probabilities minus the probability of their intersection.** $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.

6. **Two events A and B are disjoint if they have no outcomes in common and so can never occur simultaneously.** If A and B are disjoint, then $P(A \text{ or } B) = P(A) + P(B)$ since $P(A \text{ and } B) = 0$.

Example: Rolling Two 1/6 flat Unfair Dice (see text for parallel discussion on fair dice)

A pair of dice have been altered so they are not fair. Each has been altered so that

Event	1	2	3	4	5	6
Probability	1/4	1/8	1/8	1/8	1/8	1/4

For two dice, the sample space is $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

The probability of rolling a 2 is:

$$\begin{aligned}
P(2) &= P\left(\boxed{1}\boxed{1}\right) \\
&= P\left(\boxed{1}\right) P\left(\boxed{1}\right) \text{ the roll of the two dice are independent of each other} \\
&= \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16} = 0.0625
\end{aligned}$$

Similarly,

$$\begin{aligned}
P(3) &= P\left(\boxed{1}\boxed{2} \text{ or } \boxed{2}\boxed{1}\right) \\
&= P\left(\boxed{1}\boxed{2}\right) + P\left(\boxed{2}\boxed{1}\right) \text{ rule 6.} \\
&= P(\boxed{1})P(\boxed{2}) + P(\boxed{2})P(\boxed{1}) \text{ rule 4.} \\
&= \frac{1}{4} \cdot \frac{1}{8} + \frac{1}{4} \cdot \frac{1}{8} = \frac{1}{16} = 0.0625
\end{aligned}$$

In this way we can calculate the probability model for rolling two unfair dice:

Event	2	3	4	5	6	7	8	9	10	11	12
Probability	1/16	1/16	5/64	3/32	7/64	3/16	7/64	3/32	5/64	1/16	1/16

We have constructed a probability model for a finite sample space, which is assigning a probability for each individual outcome in the sample space.

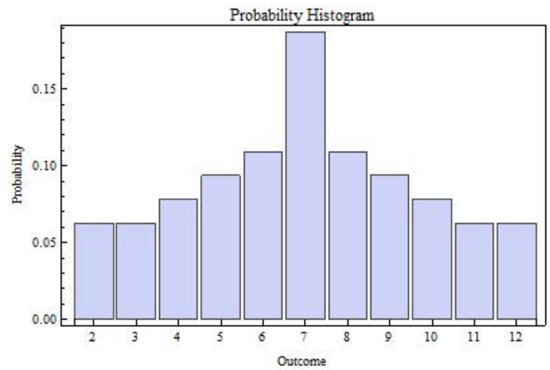
The probability that rolling two unfair dice with probability model given above and getting an even number is

$$\begin{aligned}
P(\text{outcome is even}) &= P(2) + P(4) + P(6) + P(8) + P(10) + P(12) \\
&= \frac{1}{16} + \frac{5}{64} + \frac{7}{64} + \frac{7}{64} + \frac{5}{64} + \frac{1}{16} \\
&= \frac{1}{2}
\end{aligned}$$

The probability of rolling sevens three times in a row is:

$$\begin{aligned}
P(\text{three straight rolls of 7}) &= P(7 \text{ and then } 7 \text{ and then } 7) = P(7)P(7)P(7) \\
&= \left(\frac{3}{16}\right)^3 \\
&= \frac{27}{4096} \sim 0.0065918 = 0.659\%
\end{aligned}$$

Here is the probability histogram for the rolling two unfair dice probability model:



Counting and Equally Likely Outcomes

For fair dice, we could arrive at the probability model by counting rather than explicit use of the probability rules. This is the case if we have equally likely outcomes, and relates to the concepts of combinations and permutations. You can use basic counting principles to answer simple questions.

Example (order matters) A computer assigns three-character login IDs that may contain digits 0-9, the letters a-z.

1. If repeats are allowed, what is the total number of login IDs possible?
2. If repeats are not allowed, what is the total number of login IDs?

In both these cases, the order does matter since the ID 12z is different from z12.

(1) There are 10 digits and 26 letters, so $n = 36$ possible characters at each of $k = 3$ positions.

There are $36 \times 36 \times 36 = 36^3 = 46,656 = n^k$ different login IDs. This is Rule A in the text (order matters, repeats allowed).

(2) The number of logins IDs where repeats are not allowed would be $36 \times 35 \times 34 = 42,840 = \frac{n!}{(n-k)!}$.

This case is what is called a permutation (Rule B: order matters, repeats are not allowed).

Example (order does not matter) You are choosing three members from a committee of thirty-six members to join a subcommittee. How many different ways can the three members be chosen for the subcommitee?

We can choose $36 \times 35 \times 34 = 42,840 = \frac{n!}{(n-k)!}$ groups of 3 people from the committee.

However, this set of groups would differentiate Alice-Bob-Chuck from Chuck-Bob-Alice (the order would matter). Since order should NOT matter, we have to divide out all the possible ways we can combine A-B-C, which is $3 \times 2 \times 1 = 6 = k!$.

Therefore, there are $\frac{42,840}{6} = 7140 = \frac{n!}{k!(n-k)!}$ ways to choose the 3 members of the subcommittee.

This case is called a combination (Rule D: order does not matter, repeats are not allowed).

There is a fourth case the text mentions that is not as useful (Rule C: order does not matter, repeats allowed) but don't worry about that one.

Notice in answering these questions I did not look for a rule and apply it (Rule A, B, or D), I just just used what I knew about counting to figure out the solutions. That is a much better process that memorizing a formula.

Mean of a Discrete Probability Model

We can find the mean of a probability model in the following manner. The mean can be thought of as the weighted average of the outcomes.

Suppose the sample space S has outcomes x_i with probabilities p_i. The mean of the probability distribution is given the greek symbol "mu"

$$\text{mean } = \mu = x_1p_1 + x_2p_2 + \cdots + x_kp_k.$$

Example At a store, the number of people in checkout lines varies. The probability model for the number of people in a randomly chosen line is

Number of People Waiting in Line	0	1	2	3	4	5
Probability	0.08	0.15	0.20	0.22	0.15	0.20

What is the mean number of people in a line?

$$\begin{aligned}\text{mean number people in line} &= 0(0.08) + 1(0.15) + 2(0.20) + 3(0.22) + 4(0.15) + 5(0.20)\\ &= 2.81\end{aligned}$$

The Law of Large Numbers

We have already seen this in action, when we looked at how the long run trials of rolling dice lead to proportions which were close to 1/6 for all faces. This is the law of large numbers, which states that a random phenomenon repeated a large number of times will

- have proportion of trials on which each outcome occurs gets closer and closer to the probability of that outcome, and
- the mean $\bar{x}$ of the outcomes gets closer and closer to the mean of the probability model μ.

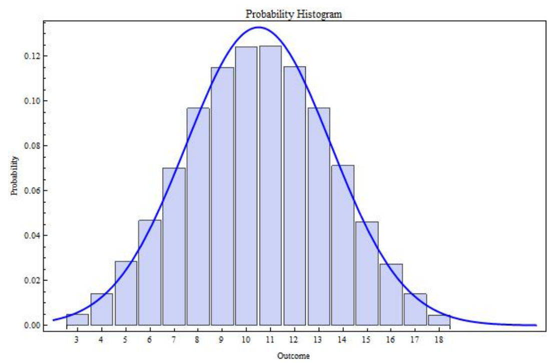
Discrete and Continuous Variables

The probability models we have looked at so far have involved discrete events. You can think of discrete as meaning the variable can only have certain numerical values, with no intermediate values in between.

A variable is continuous if between any two values for the variable there exists another possible value for the variable.

Example of Rolling Three Fair Dice

Here is the histogram that was created by rolling three fair dice 100,000 times and recording the occurrences (computer experiment).



- The height of each bar is the probability of rolling the number labeled at the base of the bar.
- If the width of the base is 1, then the area of each rectangle is the probability of rolling the number at the base.
- Since the sample space is discrete, there is no way to get any other outcomes, like 8.5 for example.
- The sum of all the areas in all the rectangles is 1.

If we were to roll more and more dice, the histogram looks more and more like a continuous distribution.

The Central Limit Theorem

The Central Limit Theorem states in part that a distribution on n random trials has a distribution that is approximately normal when n is large (this is what we see if we roll n dice and take n very large). This fact combined with the fact that a normal distribution is completely described by its mean and standard deviation is what leads to the use of the mean and standard deviation to describe so many distributions we see. People assume the distribution would be a normal distribution if they took enough measurements. But we must be careful–in the real world, not all distributions underlying a probability event are normal distributions! So the mean and standard deviation might not be the best way to describe the distribution you find in a particular situation. You really need to look at each on a case-by-case basis.