

32.2: Probability and Statistics

A random variable X can have more than one value x as an outcome. Which value the variable has in a particular case is a matter of *chance* and cannot be predicted other than that we associate a probability to the outcome. **Probability** p is a number between 0 and 1 that indicates the likelihood that the variable X has a particular outcome x . The set of outcomes and their probabilities form a **probability distribution**. There are two kinds of distributions:

1. discrete ones
2. continuous ones

Total probability should always add up to unity.

Discrete Distributions

A good example of a discrete distribution is a *true coin*. The random variable X can have two values:

1. heads (0)
2. tails (1)

Both have equal probability and as the sum must equal unity, the probability must be $\frac{1}{2}$ for each. 'The probability that X =heads' is written formally as:

$$Pr(X = heads) = Pr(X = 0) = 0.5$$

The random function is written as a combination of three statements.

- $Pr(X=0) = \frac{1}{2}$
- $Pr(X=1) = \frac{1}{2}$
- elsewhere $Pr = 0$

Continuous Distributions

Now consider a spherical die. One could say it has an infinite number of facets that it can land on. Thus the number of outcomes $n = \infty$, this make each probability

$$p = 1/\infty = 0.$$

This creates a bit of a mathematical problem, because how can we get a total probability of unity by adding up zeros? Also, if we divide the sphere in a northern and a southern hemisphere clearly the probability that it lands on a point in say, the north should be $\frac{1}{2}$. Still, $p = 0$ for all points. We introduce a new concept: **probability density** over which we *integrate* rather than *sum*. We assign an equal density to each point of the sphere and make sure that if we integrate over a hemisphere we get $\frac{1}{2}$. (This involves two angles θ and φ and an integration over them and I won't go into that).

A bit simpler example of a continuous distribution than the spherical die is the 1D **uniform** distribution. It is the one that the Excel function =RAND() produces to good approximation. Its probability density is defined as

- $f(x) = 1$ for $0 < x < 1$
- $f(x) = 0$ elsewhere

The figure shows a (bivariate) uniform distribution.

The probability that the outcome is smaller than 0.5 is written as $Pr(X < 0.5)$ and is found by integrating from 0 to 0.5 over $f(x)$.

$$Pr(X < 0.5) = \int f(\mathbf{x}) \cdot dx \text{ from } 0 \text{ to } 0.5 = \int 1 \cdot dx \text{ from } 0 \text{ to } 0.5 = [x]^{0.5} - [x]_0 = 0.5$$

Notice that in each individual outcome b the probability is indeed zero because an integral from b to b is always zero, even if the probability density $f(b)$ is not zero. Clearly the probability and the probability density are not the same. Unfortunately the distinction between probability and probability density is often not properly made in the physical sciences. Moments can also be computed for continuous distributions by integrating over the probability density .

Another well-known continuous distribution is the **normal** (or **Gaussian**) distribution, defined as:

$$f(x) = 1/[\sqrt{(2\pi)\sigma}] * \exp(-\frac{1}{2}[(x-\mu)/\sigma]^2)$$

(Notice the normalization factor $1/[\sqrt{(2\pi)\sigma}]$)

We can also compute moments of continuous distribution. Instead of using a summation we now have to evaluate an integral:

$$\langle X \rangle = \int [f(x) * x] dx$$

$$\langle X^2 \rangle = \int [f(x) * x^2] dx$$

For the normal distribution $\langle X \rangle = \mu$

? Exercise

Compute $\langle X^2 \rangle$ and $\langle X^3 \rangle$ for the uniform distribution. answer >

Indistinguishable Outcomes

When flipping two coins we could get four outcomes: two heads (0), heads plus tails (1), tails plus heads (1), two tails (2)

Each outcome is equally likely, this implies a probability of $1/4$ for each:

$$X_{\text{tot}} = X_1 + X_2 = 0 + 0 = 0 \quad p=1/4$$

$$X_{\text{tot}} = X_1 + X_2 = 0 + 1 = 1 \quad p=1/4$$

$$X_{\text{tot}} = X_1 + X_2 = 1 + 0 = 1 \quad p=1/4$$

$$X_{\text{tot}} = X_1 + X_2 = 1 + 1 = 2 \quad p=1/4$$

The probability of a particular outcome is often abbreviated simply to p . The two middle outcomes collapse into one with $p=1/4+1/4=1/2$ if the coins are **indistinguishable**. We will see that this concept has very important consequences in statistical thermodynamics.

If we cannot distinguish the two outcomes leading to $X_{\text{tot}}=1$ we get the following random function:

- $\text{Pr}(X_{\text{tot}}=0) = 1/4$
- $\text{Pr}(X_{\text{tot}}=1) = 1/2$
- $\text{Pr}(X_{\text{tot}}=2) = 1/4$
- elsewhere $\text{Pr} = 0$

Notice that it is quite possible to have a distribution where the probabilities differ from outcome to outcome. Often the p values are given as $f(x)$, a function of x . An example:

X_3 defined as:

- $f(x) = (x+1)/6$ for $x=0,1,2$;
- $f(x) = 0$ elsewhere;

The factor $1/6$ makes sure the probabilities add up to unity. Such a factor is known as a **normalization factor**. Again this concept is of prime importance in statistical thermodynamics.

Another example of a discrete distribution is a die. If it has 6 sides (the most common die) there are six outcomes, each with $p=1/6$. There are also dice with $n=4, 12$ or 20 sides. Each outcome will then have $p=1/n$.

Moments of Distributions

In important aspect of probability distributions are the **moments** of the distribution. They are values computed by summing over the whole distribution.

The **zero** order moment is simply the sum of all p and that is unity:

$$\langle X^0 \rangle = \sum X^0 * p = \sum 1 * p = 1$$

The **first** moment multiplies each outcome with its probability and sums over all outcomes:

$$= \sum X * p$$

This moment is known as the **average** or **mean**. (It is what we have done to your grades for years...)

For one coin is $\frac{1}{2}$, for two coins is 1. (Exercise: verify this)

The **second** moment is computed by summing the product of the square of X and p:

$$\langle X^2 \rangle = \sum X^2 * p$$

For one coin we have $\langle X^2 \rangle = \frac{1}{2}$,

For two coins $\langle X^2 \rangle = [0 * \frac{1}{4} + 1 * \frac{1}{2} + 4 * \frac{1}{4}] = 1.5$

What is $\langle X^3 \rangle$ for X3? answer

The notation is used a lot in quantum mechanics, often in the form $\langle \psi * \psi \rangle$ or $\langle \psi * |h| \psi \rangle$. The $\langle ..$ part is known as the **bra**, the $.. \rangle$ part as the **ket**. (Together **bra(c)ket**)

Intermezzo: The strange employer

You have a summer job but your employer likes games of chance. At the end of every day he rolls a die and pays you the square of the outcome in dollars per hour. So on a lucky day you'd make \$36.- per hour, but on a bad day \$1.-. Is this a bad deal? What would you make on the average over a longer period?

To answer this we must compute the second moment $\langle X^2 \rangle$ of the distribution:

$$\langle X^2 \rangle = 1/6 * [1+4+9+16+25+36] = 91/6 = \$15.17 \text{ per hour.}$$

(I have taken $p=1/6$ out of brackets because the value is the same for all six outcomes)

As you see in the intermezzo, the value of the second moment is in this case what you **expect** to be making on the long term. Moments are examples of what is known as **expectation values**. Another term you may run into that of a **functional**. A functional is a *number* computed by some operation (such as summation or integration) over a whole function. Moments are clearly an example of that too.

32.2: Probability and Statistics is shared under a [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/) license and was authored, remixed, and/or curated by LibreTexts.