Statistics for Computing
MA4413

Lecture 6

Random Variables, Expected Value and Variance

Kevin Burke

kevin.burke@ul.ie

Random Variables

In probability theory, a **random variable** is a *numerical* quantity whose value is determined by an *experiment*.

For example, consider the experiment of flipping two coins.

Now define a *random variable* $X =$ "the number of heads" whose value will clearly be 0, 1 or 2 heads:

Outcome	HH	HT	TH	TT
Value assigned to X	2	1	1	0

Distribution of a Random Variable

The **probability distribution** of X is:

x	0	1	2
$\Pr(X = x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

This describes how likely each of the values are, i.e., how the probability gets *distributed* to each possible value of X.

Note that upper case X denotes the random variable whereas lower case x represents a specific value.

$\Pr(X = x)$ means *"the probability that the random variable X attains the specific value x"* where $x \in \{0, 1, 2\}$, e.g., $\Pr(X = 0) = \frac{1}{4}$.

Probability Function

$\Pr(X = x)$ is called the **probability function** - it maps each value of X to a probability value.

This is often shortened to $\boxed{p(x)}$ - pronounced “p - of - x”.

The probability values of this function *must* sum to one:

$$\boxed{\sum p(x_i) = 1}.$$

In the previous example, $p(0) = \frac{1}{4}$, $p(1) = \frac{1}{2}$ and $p(2) = \frac{1}{4}$.
$\Rightarrow p(0) + p(1) + p(2) = 1$.

Random Variable Vs Event

Previously we encountered *events* - *not* the same as *random variables*.

For the sake of clarity consider:

1. The event $A =$ "two heads showing".
   - An *event* which either occurs or does not occur following the experiment.
   - It refers to *one* specific event; we can calculate $\Pr(A)$.
2. The random variable $X =$ "the number of heads".
   - A numeric *variable* whose value is assigned following the experiment.
   - Related to X are *three* events: $X = 0$, $X = 1$ and $X = 2$; we can calculate $\Pr(X = 0)$, $\Pr(X = 1)$ and $\Pr(X = 2)$.

Note: $X = 2$ *is* the event A.

Example: Flipping Two Coins

Continuing the example of flipping two coins, we could define another random variable Y = "the number of unique faces showing".

The possible values for this random variable are 1 (if the faces are the same) or 2 (if the faces are different):

Outcome	HH	HT	TH	TT
Value assigned to Y	1	2	2	1

From the above we get the *probability distribution* of Y:

y	1	2
$\Pr(Y = y)$	$\frac{1}{2}$	$\frac{1}{2}$

Expected Value

Just as we calculated the *mean* as a measure of centrality for a distribution of data, we can calculate the **expected value** for a *probability distribution*.

The expected value is:

$$\boxed{E(X) = \sum x_i \, p(x_i)}.$$

In words: multiply each possible value of X by its probability value and then sum the results.

This is the value we would *expect* to get on average if we carried out the experiment a large number of times.

The Second Moment

We will also need to calculate $E(X^2)$ which is called the *second moment* ($E(X)$ is the first):

$$\boxed{E(X^2) = \sum x_i^2 \, p(x_i)}.$$

Note that $E(X^2)$ is not directly of interest but is used to calculate the *variance* of X.

Example: Flipping Two Coins

The random variable $X =$ “the number of heads” has a probability distribution given by:

x	0	1	2
$\Pr(X = x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$$\Rightarrow E(X) = \left(0 \times \frac{1}{4}\right) + \left(1 \times \frac{1}{2}\right) + \left(2 \times \frac{1}{4}\right) = \frac{1}{2} + \frac{2}{4}$$
$$= \frac{4}{4} = 1.$$

On average there will be *one* head showing.

Example: Flipping Two Coins

We can also calculate:

$$
\begin{aligned}
\Rightarrow E(X^2) &= \left(0^2 \times \frac{1}{4}\right) + \left(1^2 \times \frac{1}{2}\right) + \left(2^2 \times \frac{1}{4}\right) \\
&= \left(0 \times \frac{1}{4}\right) + \left(1 \times \frac{1}{2}\right) + \left(4 \times \frac{1}{4}\right) \\
&= \frac{1}{2} + \frac{4}{4} \\
&= 1.5.
\end{aligned}
$$

This value will be used later to calculate the *variance*.

Question 1

We continue with the experiment of flipping two coins.

We had the random variable $Y =$ "the number of unique faces".

a) Calculate $E(Y)$.

b) Calculate $E(Y^2)$.

Expectation of Functions of X

It is sometimes of interest to calculate the expectation of various functions of X, for example, we have seen how to calculate $E(X^2)$.

Continuing with the previous example, let's say we wanted to know:

$$
\begin{aligned}
E(X^3) &= \left(0^3 \times \frac{1}{4}\right) + \left(1^3 \times \frac{1}{2}\right) + \left(2^3 \times \frac{1}{4}\right) = \frac{1}{2} + \frac{8}{4} \\
&= 2.5,
\end{aligned}
$$

or

$$
\begin{aligned}
E(e^X) &= \left(e^0 \times \frac{1}{4}\right) + \left(e^1 \times \frac{1}{2}\right) + \left(e^2 \times \frac{1}{4}\right) = \frac{1}{4} + \frac{e}{2} + \frac{e^2}{4} \\
&\approx 3.46.
\end{aligned}
$$

Entropy

Later in the course we will discuss *entropy*:

$$E[-\log_2 p(X)] = \sum [-\log_2 p(x_i)]\, p(x_i),$$

i.e., the expectation of the negative log (base 2) of the probability values.

Don't worry about this for now - just be aware of its existence.

Variance and Standard Deviation

Just as we calculated the *variance* of a set of data, we can calculate the **variance** for a *probability distribution*.

Recall that variance is the average squared distance from the mean:

$$\Rightarrow Var(X) = E[(X - E(X))^2] = \sum (x_i - E(X))^2 \, p(x_i).$$

The above formula can be simplified to

$$\boxed{Var(X) = E(X^2) - [E(X)]^2}.$$

The **standard deviation** is then

$$\boxed{Sd(X) = \sqrt{Var(X)}}.$$

(reminder: variance is measured in units-squared and standard deviation is in units)

Example: Flipping Two Coins

We continue the example of flipping two coins where $X =$ "the number of heads".

We have calculated $E(X) = 1$ and $E(X^2) = 1.5$.

$$\Rightarrow Var(X) = E(X^2) - [E(X)]^2 = 1.5 - (1)^2 = 1.5 - 1 = 0.5 \text{ heads}^2,$$

and the *standard deviation* is

$$Sd(X) = \sqrt{Var(X)} = \sqrt{0.5} \approx 0.707 \text{ heads}.$$

Question 2

We had the random variable $Y =$ "the number of unique faces" based on flipping a coin twice.

a) Calculate $Var(Y)$.

b) Calculate $Sd(Y)$.

Question 3

Consider the experiment of rolling two dice. Define the random variable $X =$ “the sum of the two numbers showing”.

a) Construct the probability distribution of X.

b) Calculate $E(X)$.

c) Calculate $E(X^2)$.

d) Calculate $Sd(X)$.

Joint Distributions

We can also construct a **joint distribution** for two random variables.

From the two coin example we had $X =$ “the number of heads” and $Y =$ “the number of unique faces”:

Outcome	HH	HT	TH	TT
X	2	1	1	0
Y	1	2	2	1

Clearly we have the following *joint probabilities*:
$\Pr(X = 2 \cap Y = 1) = \frac{1}{4}$, $\Pr(X = 1 \cap Y = 2) = \frac{2}{4} = \frac{1}{2}$ and $\Pr(X = 0 \cap Y = 1) = \frac{1}{4}$.

The remaining joint probabilities have the value zero.

Example: Flipping Two Coins

Using the information from the previous slide we can construct the *joint distribution*:

		X		
		0	1	2
Y	1	$\frac{1}{4}$	0	$\frac{1}{4}$
	2	0	$\frac{1}{2}$	0

Note that the above probabilities sum to one as we would expect.

Example: Flipping Two Coins

Using the *law of total probability* we can calculate:

$$\begin{aligned}\Pr(X = 0) &= \Pr(X = 0 \cap Y = 1) + \Pr(X = 0 \cap Y = 2)\\ &= \frac{1}{4} + 0 = \frac{1}{4},\end{aligned}$$

$$\begin{aligned}\Pr(Y = 1) &= \Pr(Y = 1 \cap X = 0) + \Pr(Y = 1 \cap X = 1)\\ &\qquad + \Pr(Y = 1 \cap X = 2)\\ &= \frac{1}{4} + 0 + \frac{1}{4} = \frac{1}{2},\end{aligned}$$

$\ldots$ etc.

Example: Flipping Two Coins

We can get these total probabilities by summing across each row and each column:

	X: 0	1	2	$p(y)$
Y = 1	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
2	0	$\frac{1}{2}$	0	$\frac{1}{2}$
$p(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

Thus, if we have a joint distribution, we can calculate the distributions of X and Y by summing across rows / columns.

Checking Independence

Once we have these total probabilities we can check if the variables are *independent* since

$$\boxed{\Pr(X = x \cap Y = y) = p(x) \times p(y) \ \Rightarrow \ \text{independent}}$$

From our example, if X and Y were independent then we would have

$$\begin{aligned}
\Pr(X = 0 \cap Y = 1) &= p(0) \cdot p(1) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}, \\
\Pr(X = 0 \cap Y = 2) &= p(0) \cdot p(2) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}, \\
\Pr(X = 1 \cap Y = 1) &= p(1) \cdot p(1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4},
\end{aligned}$$

$\ldots$ etc.

Example: Flipping Two Coins

If X and Y were independent the joint distribution *would* be

Y \ X	0	1	2	$p(y)$
1	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$
2	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$
$p(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

Since the above does *not* match the real joint distribution (as calculated previously) we conclude that X and Y are *dependent*.

Question 4

Let's assume that X and Y are two random variables with joint distribution:

		X = 0	X = 1
Y	0	0.4	0.2
	1	0.1	?

a) What is the value of $\Pr(X = 1 \cap Y = 1)$?

b) Construct the distribution of X and the distribution of Y.

c) Are X and Y independent?

d) Calculate $\Pr(X = 1 \mid Y = 0)$.

e) Calculate $E(X)$ and $Sd(X)$.