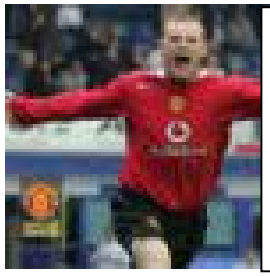
DISCRETE PROBABILITY DISTRIBUTIONS



I have monitored my team over a long period and observed the number of goals they scored in each match.
The "number of goals scored in a match" is a discrete random variable taking only values 0, 1, 2, 3, and 4.
Based on the ratios of goals, I can describe the model for my team using the outcomes together with their associated probabilities. This is the discrete probability distribution, with X as the random variable taking values x:

x:	0	1	2	3	4
P(X = x):	0.1	0.2	0.3	0.3	0.1

Using this model, we can find the mean the variance, make predictions and even calculate betting odds for future matches.

The first thing to notice about a probability distribution is that the sum of the probabilities is equal to 1

$\sum p(x) = \mathbf{1}$

If the distribution had been written

x:	0	1	2	3	4
P(X = x):	0.1	0.2	0.3	**a**	0.1

, then **a** would be found by summing to 1.

a = 1 – 0.7

THE MEAN - This is known as the expected value of X written E(X), and just as we work out the expected number of pupils in a class of **30** wearing glasses when **p = 0.4**, by multiplying **30x0.4** to give 12, we multiply each x-value by its probability and sum:

$$\mathbf{E(X) = \sum xp(x) = 0x0.1 + 1x0.2 + 2x0.3 + 3x0.3 + 4x0.1 = 2.1}$$

So the expected number of goals for my team is **2.1** and I can expect them to have accumulated 21 goals after 10 games.

E(X) = μ = $\sum xp(x)$

CODING - We can code the random variable and find that the mean will be adjusted accordingly:

Let **Y = X + 2**, add 2 onto each value:

y:	2	3	4	5	6
P(Y = y):	0.1	0.2	0.3	0.3	0.1

The probabilities will be the same as before:

$$\mathbf{E(X+2) = \sum (x+2)p(x) = 2x0.1 + 3x0.2 + 4x0.3 + 5x0.3 + 6x0.1 = 4.1 = (2.1 + 2)}$$

E(a) = a

E(X + a) = E(x) + a

E(X+2) may be written as E(X) + E(2) and since the expected value of a constant must be the constant itself,
E(2) = 2 **E(X) + E(2) = 2.1 + 2 = 4.1**

Further properties of Expectation.

Let Y = 3X, multiply each value of the RV by 3: : y: 0 3 6 9 12
The probabilities will be the same as before: P(Y = y): 0.1 0.2 0.3 0.3 0.1

$$\mathbf{E(3X) = \sum 3xp(x) = 0x0.1 + 3x0.2 + 6x0.3 + 9x0.3 + 12x0.1 = 6.3 = (3x2.1)}$$

Capital **E** behaves naturally even with a linear coding of the random variable…….
E(2x + 3) = 2E(X) + 3, which in this case, will be (2 x 2.1 + 3) = 7.2

If we decide to square the values of the random variable and look at $E(X^2)$:

Let $Y = X^2$, square each value of the RV: : y: 0 1 4 9 16
The probabilities will be the same as before: P(Y = y): 0.1 0.2 0.3 0.3 0.1

$$\mathbf{E(X^2) = \sum x^2 p(x) = 0x0.1 + 1x0.2 + 4x0.3 + 9x0.3 + 16x0.1 = 5.7 \quad but \quad [E(X)]^2 = 2.1^2 = 4.41}$$

So there is no simple way of working out $E(X^2)$ and this is where capital E stops behaving like a constant.

The Variance

$$\mathbf{Var(X) = E(x - \mu)^2 = \sum (x-\mu)^2 p(x)}$$

The variance of a probability distribution is defined as the expected value of $(x - \mu)^2$, **where μ is the mean.**

We could set up a table to evaluate this:

x:	0	1	2	3	4
x – μ:	**-2.1**	**-1.1**	**-0.1**	**0.9**	**1.9**
$(x - \mu)^2$:	**4.41**	**1.21**	**0.01**	**0.81**	**3.61**
***p*:**	0.1	0.2	0.3	0.3	0.1

where **μ = 2.1**

$$\mathbf{Var(X) = E(x - \mu)^2 = \sum (x-\mu)^2 p(x) = 4.41x0.1 + 1.21x0.2 + 0.01x0.3 + 0.81x0.3 + 3.61x0.1 = 1.29}$$

We could however use the properties of E: $\mathbf{E(x - \mu)^2 = E(X^2 - 2\mu X + \mu^2) = E(X^2) - 2\mu E(X) + E(\mu^2)}$

$$\mathbf{= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - [E(X)]^2 = 5.7 - 4.41 = 1.29}$$

E(aX) = a E(x)

E(aX + b) = E(aX) + E(b)
= aE(X) + b

Beware!

$E(X^2) \neq [E(X)]^2$

E(X) = μ = $\sum xp(x)$

$Var(X) = E(X^2) - [E(X)]^2$

$E(X^2) = \sum x^2 p(x)$

The Cumulative distribution function (c.d.f)

Just like a cumulative frequency curve, the c.d.f gives the cumulative probabilities up to and including the value of the random variable.

With a probability density function:

x:	0	1	2	3	4
P(X = x):	0.1	0.2	0.3	0.3	0.1

, the c. d. f. would be written

$F(x) =$

x:	0	1	2	3	4
P(X ≤ x):	0.1	0.3	0.6	0.9	**1,**

with the final value being **1**.

F(x) is not only confined to integer values, even though we have a discrete probability distribution.

F(2.6) means the sum of all probabilities up to the value 2.6 and since this is not a recognised value, we stop at x = 2.
F(2.6) = P(X ≤ 2.6) = P(X ≤ 2) = 0.6

THE EFFECT OF CODING ON THE VARIANCE - We can code the random variable and find that the **variance** will be adjusted accordingly:

Let **Y = X + 2**, add 2 onto each value:

y:	2	3	4	5	6
P(Y = y):	0.1	0.2	0.3	0.3	0.1

The probabilities will be the same as before:

E(X+2) may be written as E(X) + E(2) and
Var(X + 2) may be written as Var(X) + Var(2) = Var(X) + 2, since the variance of a constant is the constant itself.

E(aX) = a E(x) but Var(aX) = a^2Var(X). This may be verified by looking at the distribution of 3X in the example above.

Var(aX + b) = Var(aX) + Var(b) = a^2Var(X).

The discrete Uniform or Rectangular distribution. **All probabilities are equally likely, like the numbers on a die.**

x:	1	2	3	4	n
P(X = x):	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$

Learn: $E(X) = \frac{n+1}{2}$ and $Var(X) = \frac{(n+1)(n-1)}{12}$