THE BINOMIAL DISTRIBUTION

In Statistics, we use many distributions to model uncertain events in the real world. And just like in Mechanics, if we apply the wrong model the whole system will fall apart:- Bridges will fall down and our predictions of events will be nowhere near accurate.

Simple statistical models could include:

A model for a die

outcomes	1	2	3	4	5	6
probabilities	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

A model for the goals scored per match for a football team

goals	0	1	2	3	4	>4
probabilities	$\frac{5}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

The payout on a simple fairground stall based on throwing two coins.
Pay 50p Throw two coins:

2 Heads – win £1.50, 1 Head – win 20p, No Heads – lose

outcomes	2 Heads	1 Head	No Heads
probabilities	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
Profit	100p	-30p	-50p

My expected profit would be E(X) $= \frac{1}{4} \times 100 - \frac{1}{2} \times 30 - \frac{1}{4} \times 50 = -2.5$pence.

Using these models and the probabilities associated with them, we can predict how many sixes we would expect in 60 tosses of a die, how long it should take for our team to accumulate 20 goals and how much we would expect to make (or lose) after 10 goes at the stall. These models are known as probability distributions since they contain the outcomes and the associated probabilities.

On this page, we will consider the **Binomial distribution** as a model for all situations where there are two outcomes and the trials are repeated a finite number of times.

The probability of success is denoted by **p** and the probability of failure by **q** and the number of trials by **n**. **p + q = 1**

The probability that I win a game of tennis is $\frac{3}{8}$ Pr(lose) = $\frac{5}{8}$. If I play two games:

First Game	Second Game	Code for outcomes	Probability Multiply along branches
$\frac{3}{8}$ win	$\frac{3}{8}$ win	win, win	Pr(two wins) $\frac{3}{8} \times \frac{3}{8} = \frac{9}{64}$
	$\frac{5}{8}$ lose	win, lose	Pr(one win) $\frac{3}{8} \times \frac{5}{8} = \frac{15}{64}$
	$\frac{3}{8}$ win	lose, win	Pr(one win) $\frac{5}{8} \times \frac{3}{8} = \frac{15}{64}$
$\frac{5}{8}$ lose	$\frac{5}{8}$ lose	lose, lose	Pr(no wins) $\frac{5}{8} \times \frac{5}{8} = \frac{25}{64}$

This game is modelled by the Binomial Distribution.

outcomes	No Wins	1 Win	2 Wins
probabilities	$\frac{25}{64}$	$\frac{30}{64}$	$\frac{9}{64}$

Note there are **two** ways of winning one out of two games.

If I played 5 such games:

Number of wins	0	1	2	3	4	5
Code	LLLLL	WLLLL	WWLLL	WWWLL	WWWWL	WWWWW
probabilities	$\left(\frac{5}{8}\right)^5$	$\left(\frac{3}{8}\right)^1 \left(\frac{5}{8}\right)^4$	$\left(\frac{3}{8}\right)^2 \left(\frac{5}{8}\right)^3$	$\left(\frac{3}{8}\right)^3 \left(\frac{5}{8}\right)^2$	$\left(\frac{3}{8}\right)^4 \left(\frac{5}{8}\right)^1$	$\left(\frac{3}{8}\right)^5$

But there is more than one way of getting, for instance, 3 wins

WWWLL
WWLWL
WWLLW ……………..etc

And, using combination theory, there are $^5C_3$ or $\binom{5}{3}$ ways of choosing 3 from 5 $\boxed{\frac{5!}{3!2!}}$

I.E. There are **10** ways of getting 3 wins out of 5 games.

This coefficient needs to go in front of the probability. $\mathbf{10}\left(\frac{3}{8}\right)^3 \left(\frac{5}{8}\right)^2$

and this is the true probability of 3 wins.

Using combination theory (see the factsheet on the Binomial Distribution in Pure Maths) we can obtain all the coefficients and hence the probabilities.

Number of wins	0	1	2	3	4	5
probabilities	$\left(\frac{5}{8}\right)^5$	$5\left(\frac{3}{8}\right)^1 \left(\frac{5}{8}\right)^4$	$10\left(\frac{3}{8}\right)^2 \left(\frac{5}{8}\right)^3$	$10\left(\frac{3}{8}\right)^3 \left(\frac{5}{8}\right)^2$	$5\left(\frac{3}{8}\right)^4 \left(\frac{5}{8}\right)^1$	$\left(\frac{3}{8}\right)^5$

So the probability of obtaining r successes from n trials is $^nC_r\, p^r q^{(n-r)}$. The formula for binomial probabilities.

Example. The probability of a new "cheapo" calculator being faulty is $\frac{1}{5}$. I buy a dozen calculators. Find the probabilities of:

Three faulty calculators: 3 faulty, 9 good ……… $^{12}C_3 \left(\frac{1}{5}\right)^3 \left(\frac{4}{5}\right)^9$

More than 10 faulty calculators: $^{12}C_{11} \left(\frac{1}{5}\right)^{11} \left(\frac{4}{5}\right)^1 + {}^{12}C_{12} \left(\frac{1}{5}\right)^{12}$ 11 or 12 faulty

At least one faulty calculator: 1 – Pr(none faulty)……… $1 - \left(\frac{4}{5}\right)^{12}$

Use of tables

Binomial tables give the probability of **r or less** successes in n trials. They give **cumulative** probabilities.

n = 20, p = 0.30 in a section of the table, looks like this:

From tables:

Pr ($X \le 5$) 5 or less successes = 0.4164

Pr ($X < 2$) less than 2 successes = 0.0076

Pr ($X = 6$) exactly 6 successes (≤ 6) - (≤ 5)

0.6080 – 0.4164 = 0.1916

Let us check this value using the Binomial Formula:

Pr ($X = 6$) = $^{20}C_6\,(0.3)^6\,(0.7)^{14}$ = **0.1916**

n = 20

p =	0.3
x=0	0.0008
x=1	0.0076
x=2	0.0355
x=3	0.1071
x=4	0.2375
x=5	0.4164
x=6	0.6080
x=7	0.7723

Etc………