The Binomial Distribution

Prerequisites

You should have studied discrete probability distributions.

Definition of a discrete random variable
Let *X* be a variable such that

(*a*) It is discrete, meaning it can only take *n* exact values $x_1, x_2, \ldots, x_n$. When *X* takes the value x_i we write $X = x_i$.

(*b*) It is random, meaning that with each value x_i that the variable takes, there is associated a probability p_i. We write this $P(X = x_i) = p_i$. The assignment of probabilities to each value that X can take is called a discrete probability distribution.

(*c*) Because it is random it obeys the law of total probability. The sum of all the probabilities for all *n* values is equal to 1.

You should also understand how to expand binomial products of the form $(a+b)^n$ by means of the binomial theorem.

Example (1)

Use Pascal's triangle to expand $(a+b)^5$.

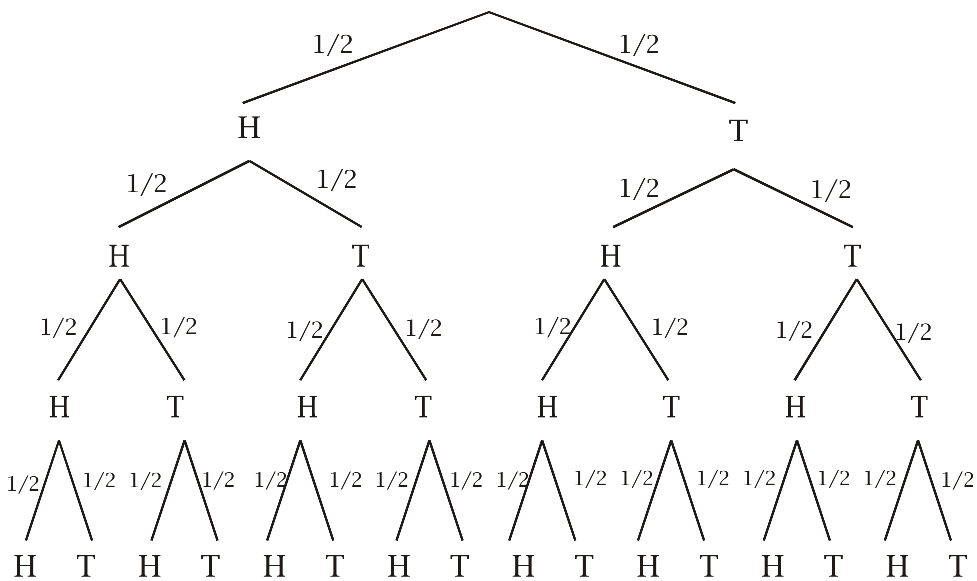
Solution

$(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$

You should be able to evaluate a binomial coefficient, from the definition: $\binom{n}{r} = \frac{n!}{r!(n-r)!}$.

The binomial distribution

The binomial distribution is an example of a discrete probability distribution. It arises as the result an experiment where at each trial or observation there are just two possible outcomes, one of which is called a "success" and the other a "failure". For example, the trial could be the spinning of a coin with a head as a "success" and a tail as a "failure". An experiment may consist of a single trial, or the trials may be repeated – in this example the coin may be spun several times. The outcome of each trial must be independent of each other, meaning the probability of a success does not vary from one trial to another. As trials are made there is a range of possible events, each of which nay be characterised in terms of the number of successes occurring. For example, if a coin is spun four times, there may be 0, 1, 2, 3, or 4 heads at the end. Defining a success in this case to be when the coin turns up heads and taking X to be a the discrete random variable standing for the number of successes obtained, and r to represent the values that X can take, then $P(X = r)$ represents the probability that X takes the value r. The assignment of probabilities to each value that X can take is a discrete probability distribution. One way to find the probability distribution for a particular experiment is by means of a probability tree. On the assumption that the coin is fair so that every outcome is equally likely, the probability of obtaining a head on any one trial is $\frac{1}{2}$. The whole probability tree is



	H	H	H	H	H	H	H	H	T	T	T	T	T	T	T	T
Outcome	H	H	H	H	T	T	T	T	H	H	H	H	T	T	T	T
	H	H	T	T	H	H	T	T	H	H	T	T	H	H	T	T
	H	T	H	T	H	T	H	T	H	T	H	T	H	T	H	T

The probability of each outcome is $\left(\frac{1}{2}\right)^4 = \frac{1}{16}$.

When a success is defined to be the number of times heads comes up the order in which the heads and tails are obtained is not important so several outcomes may contribute to an event.

Successes, x	0	1	2	3	4
no out outcomes	1	4	6	4	1

Thus we have the probability distribution

r	0	1	2	3	4
$P(X=x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

We have already indicated that in this example the probability of each outcome is $\left(\frac{1}{2}\right)^4 = \frac{1}{16}$. The probability of each event, (for example, the event $X=2$), is found by multiplying this by the number of outcomes belonging to that event. This number is the same as the binomial coefficient $\binom{4}{r} = \frac{4!}{r!(4-r)!}$. For example, the probability that the number of successes is 2 is given as $P(X=2) = \binom{4}{r} \times \frac{1}{16} = 4 \times \frac{1}{16}$.

These binomial coefficients are, of course, the same as those coefficients generated by Pascal's triangle.

$$
\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}
$$

What this means is that there no need to write out a probability tree for a binomial distribution. Knowledge of the binomial coefficients is all that is required to generate the probability distribution for any n in a binomial distribution.

Example (2)

A coin is spun n times. A success is defined to be when the coin turns up heads. X is defined to be a the discrete random variable standing for the number of successes obtained, and r to represent the values that X can take.

(*a*) Write down the probability distribution of X when

(*i*) $n = 5$

(*ii*) $n = 6$

(*b*) Expand the binomial products

(*i*) $(p+q)^5$

(*ii*) $(p+q)^6$

Solution

(*a*) (*i*)

r	0	1	2	3	4	5
$\binom{5}{r}$	1	5	10	10	5	1
$P(X=r)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

(*ii*)

r	0	1	2	3	4	5	6
$\binom{5}{r}$	1	6	15	20	15	6	1
$P(X=r)$	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$

(*b*)

(*i*)
$$\begin{aligned}(p+q)^5 &= \binom{5}{0}p^5q^0 + \binom{5}{1}p^4q^1 + \binom{5}{2}p^3q^2 + \binom{5}{3}p^2q^3 + \binom{5}{4}p^1q^4 + \binom{5}{5}p^0q^5 \\ &= p^5 + 5p^4q^1 + 10p^3q^2 + 10p^2q^3 + 5p^1q^4 + q^5\end{aligned}$$

(*ii*)
$$\begin{aligned}(p+q)^6 &= \binom{6}{0}p^6q^0 + \binom{6}{1}p^5q^1 + \binom{6}{2}p^4q^2 + \binom{6}{3}p^3q^3 + \binom{6}{4}p^2q^4 + \binom{6}{5}p^1q^5 + \binom{6}{6}p^0q^6 \\ &= p^6 + 6p^6q^1 + 15p^5q^2 + 20p^3q^3 + 15p^2q^4 + 6p^1q^5 + q^6\end{aligned}$$

It is quite clear that there is a very close relationship between the binomial theorem and the binomial distribution. To make this even clearer let us substitute

$$p = q = \frac{1}{2}$$

into the expansion

$$(p+q)^5 = p^5 + 5p^4q^1 + 10p^3q^2 + 10p^2q^3 + 5p^1q^4 + q^5$$

We obtain

$$\left(\frac{1}{2}+\frac{1}{2}\right)^5 = \frac{1}{32}+5\times\frac{1}{32}+10\times\frac{1}{32}+10\times\frac{1}{32}+5\times\frac{1}{32}+\frac{1}{32}$$
$$= \frac{1}{32}+\frac{5}{32}+\frac{10}{32}+\frac{10}{32}+\frac{5}{32}+\frac{1}{32}$$
$$= 1$$

We see that the fractions obtained in this expansion are precisely the fractions obtained in the probability distribution in example (2) where $n = 5$.

r	0	1	2	3	4	5
$\binom{5}{r}$	1	5	10	10	5	1
$P(X = r)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

From this we see that for this example, where the probability of a success is $p = \frac{1}{2}$, and there are n trials and r successes, then

$$P(X = r) = \binom{n}{r}\times\left(\frac{1}{2}\right)^n.$$

Example (3)

A coin is spun n times. A success is defined to be when the coin turns up heads. X is defined to be a the discrete random variable standing for the number of successes obtained, and r to represent the values that X can take. Without recourse to Pascal's triangle, evaluate the probability of 5 successes when

(i) $n = 7$
(ii) $n = 23$

Solution

$$P(X = r) = \binom{n}{r}\times\left(\frac{1}{2}\right)^n$$

(i) $n = 7, r = 5$ $\quad P(X = 5) = \binom{7}{5}\times\left(\frac{1}{2}\right)^7 = \frac{21}{128} = 0.164$ (3 s.f.)

(ii) $n = 23, r = 5$ $\quad P(X = 5) = \binom{23}{5}\times\left(\frac{1}{2}\right)^{23} = \frac{33649}{8388608} = 0.00401$ (3 s.f.)

In the examples we have been considering up to now the probability of a success ("heads") equals the probability of a failure ("tails") = $\frac{1}{2}$. Using the results we have obtained we could create a probability distribution for the spinning of a fair coin where there are any number of trials, *n*. However, there are other cases were the probability of a success is not $\frac{1}{2}$. For example let a fair cubical die be rolled, and let a success be defined as obtaining a 6, then the probability of a success is $p = \frac{1}{6}$, and the probability of a failure is $q = \frac{5}{6}$. In general, the probability of a success can be any value between 0 and 1. We denote

p = probability of a success in any one trial q = probability of a failure in any one trial

We also have $p + q = 1$, since at every trial the total probability is 1, and there are only two possible outcomes. If there are *n* trials and then the probability of *r* successes is

$$P(X = r) = \binom{n}{r} p^r q^{n-r}.$$

The binomial distribution

Suppose we have an experiment giving rise to a probability distribution where there are *n* trials and at every trial just two possible outcomes, a success with probability *p* and a failure with probability *q*. *X* is defined to be a the discrete random variable standing for the number of successes obtained, and *r* to represent the values that *X* can take. Then the probability of *r* successes is given by $P(X = r) = \binom{n}{r} p^r q^{n-r}$ and the resultant probability distribution is called a *binomial distribution.*

Example (4)

A fair cubical die is rolled 4 times. Find the probability of obtaining 3 sixes.

Solution

$$\mathrm{P}(X = 3) = \binom{4}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^1 = 4 \times \frac{5}{1296} = \frac{5}{324}$$

Bernoulli trials

An assumption underlying any binomial probability distribution is that the experiment comprises *n* trials each of which is a *Bernoulli trial.* A Bernoulli trial is defined to be a sequence of independent repetitions of an experiment with two possible outcomes, one of which is often called a success and the other a failure, whose probabilities do not vary between repetitions. All the examples we have considered so far have been based on Bernoulli trials. Jakob Bernoulli was a Swiss mathematician 1654 – 1705.

Example (5)

A Bernoulli trial with probability of success $\frac{1}{5}$ is repeated 6 times. Let X be the number of successes. Find the probability that

(a) $X = 2$

(b) $X = 5$

(c) $X > 0$

(d) $X \le 4$

(e) $0 \le X < 4$

Solution

The statement, "A Bernoulli trial with probability of success $\frac{1}{5}$ is repeated 6 times. Let X be the number of successes" means that X is binomially distributed with 6 trials and probability of success $p = \frac{1}{5}$. We write this information as

$X \sim B\left(6, \frac{1}{5}\right)$.

The probability of a failure is $q = 1 - p = 1 - \frac{1}{5} = \frac{4}{5}$.

The general formula is $P(X = r) = \binom{n}{r} p^r q^{n-r}$.

(a) $P(X = 2) = \binom{6}{2}\left(\frac{1}{5}\right)^2\left(\frac{4}{5}\right)^4 = 15 \times \frac{256}{15625} = \frac{768}{3125} = 0.248$ (3 s.f.)

(b) $P(X = 5) = \binom{6}{5}\left(\frac{1}{5}\right)^5\left(\frac{4}{5}\right)^1 = 6 \times \frac{4}{15625} = \frac{24}{15625} = 0.00154$ (3 s.f.)

(c) $P(X > 0) = 1 - P(X = 0) = 1 - \binom{6}{0}\left(\frac{1}{5}\right)^0\left(\frac{4}{5}\right)^6 = 1 - \frac{4096}{15625} = \frac{11529}{15625} = 0.738$ (3 s.f.)

(d) $P(X = 6) = \binom{6}{6}\left(\frac{1}{5}\right)^6\left(\frac{4}{5}\right)^0 = \frac{1}{12625} = 0.0000792$ (3 s.f.)

$P(X \le 4) = 1 - P(X = 5) - P(X = 6) = 1 - \frac{24}{15625} - \frac{1}{15625} = \frac{15600}{15625} = \frac{624}{625} = 0.998$ (3 s.f.)

(*e*) $P(0 < X \le 4) = P(X > 0) - P(X = 5) - P(X = 6) = \frac{11529}{15625} - \frac{25}{15625} = \frac{11504}{15625}$

The solution to example (5) should be studied carefully.

(**1**) In example (5) we introduced the expression $X \sim B\left(6, \frac{1}{5}\right)$. By means of this expression you demonstrate that the assumptions given in a question enable you to *model* the problem by means of a binomial distribution.

Identifying a binomial distribution

The experiment is a Bernoulli trial with *n* trials. As it is a Bernoulli trial at each trial there are two outcomes; one designated a success, the other a failure. The probability of success is *p*. The probability of failure is $q = 1 - p$.

Then $X \sim B(n,p)$.

Example (6)

In an experiment a drawing pin is thrown onto a table. It is known in advance that the probability that the pin will land "point upwards" is 0.25. In the experiment 60 pins are thrown onto the table in successive trials. Let *X* denote the number of drawing pins falling "point upwards". Identify the distribution of *X*. Find the probability $P(X = 25)$ giving your answer to 3 significant figures.

Solution

$X \sim B(60, 0.25)$

$$P(X = 25) = \binom{60}{25}(0.25)^{25}(0.75)^{35} = 0.00195 \ (3 \text{ s.f.})$$

Note – it is not explicitly stated in the question that each trial is independent but this may be assumed. The question states that the probability is (always) 0.25.

Example (7)

2000 tickets are printed for a lottery. 1500 are printed on blue paper and 500 are printed on red paper. A random sample of 50 of these lottery tickets is made. *X* denotes the number of red tickets in the sample. Explain

(*a*) Why $X \sim B(50, 0.25)$ is not an exact model for the distribution of *X*.

(*b*) Why this binomial model can, however, be used as an approximation to the distribution of *X*.

Solution

(*a*) This is a sample without replacement. The probability that the first ticket is red is 0.25. However, supposing that a red ticket is chosen, then the probability that the next ticket is red decreases to $\frac{499}{2000} = 0.2495$. So the probability at each successive trial is not the same. Thus, it is not a Bernoulli trial and the binomial distribution is not an *exact* model for the distribution of *X*.

(*b*) The changes in the probabilities from trial to trial are very small. Furthermore, we expect that there will be roughly 3 times the number of blue tickets in the sample as red tickets, so the probability of choosing a red ticket will not vary much from trial to trial. Hence the binomial distribution may be used as an approximation to the real distribution of *X*.

(2) In example (5) we have introduced techniques for evaluating expressions such as $P(X > 0)$ and $P(0 < X \le 4)$. Rather than use

$$P(X > 0) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6)$$

we used the short cut $P(X > 0) = 1 - P(X = 0)$. In order to make the solving of problems like this we could prepare in advance a table of values. In example (5) we had the probability of success $p = \frac{1}{5} = 0.2$ and the number of trials $n = 6$. Let us construct a table for this distribution, adopting also the rule that we will write the probabilities as decimals to 4 decimal places

r	0	1	2	3	4	5	6
$P(X = r)$	0.2621	0.3932	0.2458	0.0819	0.0154	0.0015	0.0001

From this table we can find the solution to $P(X \le 4)$ as

$$P(X \le 4) = 1 - P(X = 5) - P(X = 6) = 1 - 0.0015 - 0.0001 = 0.9984 .$$

Tables of binomial distributions for different values of n = number of trials and p = probability of a success would reduce the amount of time spent on calculations. In turns out that even more useful are tables of *cumulative probabilities.*

Example (7)

In the following table the values in the third row are found by summing cumulatively the values in the second row. For example

$$P(X \le 1) = P(X = 0) + P(X = 1)$$
$$= 0.2621 + 0.3932$$
$$= 0.6553$$

Complete the table

r	0	1	2	3	4	5	6
$P(X = r)$	0.2621	0.3932	0.2458	0.0819	0.0154	0.0015	0.0001
$P(X \le r)$	0.2621	0.6553					

Solution

r	0	1	2	3	4	5	6
$P(X = r)$	0.2621	0.3932	0.2458	0.0819	0.0154	0.0015	0.0001
$P(X \le r)$	0.2621	0.6553	0.9011	0.9830	0.9984	0.9999	1.0000

From the table of cumulative probabilities we can easily recover the probability of a single value or an interval.

Example (7) continued

Using the cumulative values in the table you have just constructed, find

(a) $P(X = 2)$
(b) $P(X > 3)$
(c) $P(1 \le X < 4)$

Solution

(a) $P(X = 2) = P(X \le 2) - P(X \le 1) = 0.9011 - 0.6553 = 0.2458$
(b) $P(X > 3) = 1 - P(X \le 3) = 1 - 0.9830 = 0.0017$
(c) $P(1 \le X < 4) = P(X \le 3) - P(X \le 0) = 0.9830 - 0.2621 = 0.7209$

Remark

When using tables of cumulative probabilities for the binomial distribution (or other) you should take care to understand what the cumulative value represents. Here each entry represents the cumulative probability $P(X \le r)$.

Tables of cumulative probabilities enable us to efficiently make evaluations in problems involving binomial distributions.

Example (8)

At a certain station the probability that a train will depart at the scheduled time is 0.75. A student interested in quality control observes 10 trains and records whether they depart on time or not. Let X denote the number of trains in the sample that depart on time.

(*a*) Assuming that this may be modelled as a binomial distribution use the following table of cumulative probabilities to find the probability that between 40% and 80% of the trains depart on time.

The WJEC use tables that give values of probabilities up to $p = 0.5$, so the probability that $p = 0.75$ is not covered by such tables. Nonetheless, the probability that the train will **not** depart on time is $q = 1 - p = 0.25$, so we will this fact to solve the problem.

		p						
	$n = 10$	0.20	0.25	0.30	0.35	0.40	0.45	0.50
$x =$	0	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
	1	0.3758	0.2440	0.1493	0.0860	0.0464	0.0233	0.0107
	2	0.6778	0.5256	0.3828	0.2616	0.1673	0.0996	0.0547
	3	0.8791	0.7759	0.6496	0.5138	0.3823	0.2660	0.1719
	4	0.9672	0.9219	0.8497	0.7515	0.6331	0.5044	0.3770
	5	0.9936	0.9803	0.9527	0.9051	0.8338	0.7384	0.6230
	6	0.9991	0.9965	0.9894	0.9740	0.9452	0.8980	0.8281
	7	0.9999	0.9996	0.9984	0.9952	0.9877	0.9726	0.9453
	8	1.0000	1.0000	0.9999	0.9995	0.9983	0.9955	0.9893
	9	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

(*b*) Explain why, in fact, it may be unsafe to model this as a binomial distribution.

Solution

(*a*) The probability of a success is $X \sim B(10, 0.75)$. 40% of 10 trains is 4 trains, and 80% of 10 trains is 8 trains. This question is asking for $P(4 \le X \le 8)$. The probability of a failure is $Y \sim B(10, 0.25)$. Hence

$$\begin{aligned} P(4 \le X \le 8) &= P(10-4 \le Y \le 10-8) \\ &= P(6 \le Y \le 2) \\ &= P(Y \le 6) - P(Y \le 1) \\ &= 0.9965 - 0.2440 \\ &= 0.7525 \end{aligned}$$

(*b*) The assumption in modelling this sample as a binomial distribution is that each observation (trial) of the variable *X* is independent of the other. Another way of putting this is that the sample is random. However, in practice when trains are late this may be for reason such as bad weather or signal failure. In such a case trains tend to be late in clusters. This means that if one train is late it is more likely that the next train is late. So it may be unsafe to model this sample as a binomial distribution.

Expectation and variance of a binomial distribution

The binomial distribution is a theoretical model of the distribution of probabilities that may be applied to real life situations on the basis of certain assumptions. We are interested in the mean and variance of a binomial distribution in a given application.

Let $X \sim B(n,p)$ be a binomial distribution where there are n trials of the independent variable X and the probability of a success is p. The mean and variance of X is given by

$$\mathrm{E}(X) = np$$

$$\mathrm{Var}(X) = npq = np(1-p)$$

where $q = 1 - p$ is the probability of a failure.

Example (8) continued

With *X* defined as in example (8) find the mean and the standard deviation of *X*.

Solution

$X \sim B(10, 0.75)$

That is $n = 10,\ p = 0.75,\ q = 0.25$.

Then

$\mathrm{E}(X) = np = 10 \times 0.75 = 7.5$

$\mathrm{Var}(X) = npq = np(1-p) = 10 \times 0.75 \times 0.25 = 1.875$

Standard deviation of $x = \sqrt{Var(X)} = 1.37$ (3 s.f.)