Binomial Distribution (263 problems 45-48, 286 problems 51-54)

The simplest of all distributions is the binomial distribution with one experiment. The binomial distribution is characterized by having two possible outcomes A and "not A" ($\overline{A}$). If the probability of A is p then the probability of "not A" must be (1-p). Clearly

$$p + (1 - p) = 1$$

which satisfies our definition of a probability distribution. To simplify equations it is common to let q = (1-p). We now have

$$P(A) = p \text{ and } P(\overline{A}) = q$$

Table 1 summarizes some of the large number of experiments that can be characterized with the binomial probability distribution. The classic example is flipping a coin. The sample space is S = {Heads, Tails}. Flipping the coin once is an experiment with two possible outcomes: H and T. In the case of a coin, p = q = 1/2. As Table 1 demonstrates there are many more examples where p and q are not necessarily equal. In fact, most of the time, they are not equal. In the case of reliability models, they are extremely different.

Table 1- Examples of binomial probability experiments

Flipping a coin	head or tails (.5, .5)
Product testing and reliability models	success or failure (.999, 001)
Batter coming to bat	base hit or strike out (.285, .715)
Choice of turn at a corner	left or right (.5, .5)
Generate a random number	less than 6, 6 or greater (.6, .4)
Roll a die	less than 3, 3 or more (1/3, 2/3)
Quality of professor	good or bad (.45, .55)
Attendance	show or no show
Voting	yes or no
Opinion survey	like or dislike
Weather	rain or shine
Disease	infected or not infected
Dropped object	breaks or not breaks
Baby born	boy/girl, male/female
Jury	guilty or innocent
Getting sea sick	windy or not windy
Stopping for light	red or green
Turning direction	left or right
Draw red or yellow balls from urn	draw red or draw yellow
Numbers	even/odd, <5 or >4
Size	large/small

The binomial distribution becomes much more interesting when an experiment or trial occurs repeatedly. A fundamental assumption of the binomial distribution is that each experiment is independent of all previous experiments. In the case of a coin, it is assumed that each flip of the coin is independent of all previous flips.

The random variable X for a binomial distribution is usually defined as the number of times that one of the binary conditions occurs in n experiments or trials. The terms "experiment" and "trial" are used synonymously. For example,

X = the number of successes in five flips
X = the number of girls born in 100 births
X = the number of heads in four flips

Independence means that if you flip a coin nine times and get heads each time, the probability of getting a head on the tenth flip is still ½. You might be getting suspicious at this point since the probability of getting nine heads in a row is very small, namely $(1/2)^9$. Nevertheless, that's what independence means.

Now lets look at the coin flipping example is detail.

One Trial

The simplest case is one flip with two possible outcomes: H or T. We can only have one head or one tail.

Trials	Possibilities	Probability
P(1 Head)	1	½
P (0 Heads	1	½

Two Trials

When the coin is flipped two times there are not just two possibilities but four possibilities:

Trials	Prob
HH	¼
HT	¼
TH	¼
TT	¼

Trials	Possibilities	Probability
P(2 Heads)	1	¼
P(1 Head)	2	½
P(0 Heads)	1	¼

Three Trials

Trials	Prob
HHH	1/8
HHT	1/8
HTH	1/8
THH	1/8
HTT	1/8
THT	1/8
TTH	1/8
TTT	1/8

Trials	Possibilities	Probability
P(3 Heads)	1	1/8
P(2 Heads)	3	3/8
P(1 Head)	3	3/8
P(0 Heads)	1	3/8

Four Trials- all equally likely

Trials	Prob
HHHH	1/16
HHHT	1/16
HHTH	1/16
HTHH	1/16
THHH	1/16
HHTT	1/16
HTHT	1/16
HTTH	1/16
THTH	1/16
THHT	1/16
TTHH	1/16
HTTT	1/16
THTT	1/16
TTHT	1/16
TTTH	1/16
TTTT	1/16

Since we have enumerated all the possible combinations, we can count the number of times that three heads occur, namely four, or we could be smarter and think about it as a combinatorial problem where we want to know the number of ways of arranging four things, three of which are the same and one is the same. That is, the number of ways of arranging four things when three are the same and one is the same is given by

$$\frac{4!}{3!1!} = \frac{4!}{3!(4-3)!} = \binom{4}{3} = {}_4C_3$$

(Note: Compare this to the problem of counting the number of ways that four dice can be rolled that total to 10. For example, 1126 can be rolled twelve ways.) Continuing for two heads, we

$$\frac{4!}{2!2!} = \frac{4!}{2!(4-2)!} = \binom{4}{2} = {}_4C_2$$

Note that having two heads implies having two tails. In general, if we have n heads, then we must have (n-k) tails, and the general formula for k heads out of n trials becomes

$$\frac{n!}{k!(n-k)!} = \binom{n}{k} = {}_nC_k$$

So we don't really have to enumerate all the possible outcomes. We can use the combinatorial formulas to do the counting for us. The probability distribution for the number of heads in four trials is summarized below.

Trials	Possibilities	Probability
P(4 Heads)	1	1/16
P(3 Heads)	4	4/16
P(2 Heads)	6	6/16
P(1 Head)	4	4/16
P(0 Heads)	1	1/16

Now let's stop and take a look at some patterns that are developing.

n	Coefficients	Sum
1	1 1	2
2	1 2 1	4
3	1 3 3 1	8
4	1 4 6 4 1	16

Notice how each number is the sum of the two numbers directly above it. It would be reasonable to guess that the next rows would be

n	Coefficients	Sum
5	1 5 10 10 5 1	32
6	1 6 15 20 15 6 1	64
7	1 7 21 35 35 21 7 1	128
8	1 8 28 56 70 56 28 8 1	256
9	1 9 36 84 126 126 84 36 9 1	512

This is known as **Pascal's triangle** and corresponds to frequencies of {0, 1, 2, 3, …, n} heads occurring in n trials. This famous triangle is named after the famous mathematician Blaise Pascal; however, it was discovered by others hundreds of years prior to his publication of Triangle Arithmetique in the 1600's. Notice that the sum of the numbers totals to the number in the rightmost column, which gives rise to another amazing fact. This sum is equal to 2^n, which can be written as

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

where

$$_nC_k = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

is known as the binomial coefficient. When we do counting problems, we will see that it also the count of the number of combinations that can be made by choosing k objects from n objects. In words, it is spoken as "n choose k." It is also the number of ways of arranging n things when k of them and (n-k) of them are the same.

Binomial Expansions

Now we will look at another amazing fact related to Pascal's Triangle. If you expand the binomials $(p+q)^n$ you will notice the astonishing coincidence that the coefficients are the same as the frequency counts generated above by the combinatorial formula. It is one of those pieces of mathematical magic like pi being equal to the circumference of a circle divided by its diameter. We don't know why it is that way, it just is.

$(p + q)^1$	$p+q$	1 1
$(p + q)^2$	$p^2+2pq+q^2$	1 2 1
$(p + q)^3$	$p^3+3p^2q+3pq^2+q^3$	1 3 3 1
$(p + q)^4$	$p^4+4p^3q+ 6p^2q^2+4pq^3+q^4$	1 4 6 4 1
$(p + q)^5$	$p^5+5p^4q+10p^3q^2+10p^2q^3+5pq^4+q^5$	1 5 10 10 5 1

We know these all add up to one since p+q = 1, so $(p+q)^n$ must equal one. The general expression is

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}$$

Binomial Probability Distribution

We can now state the binomial probability distribution in all its glory

P(k heads after n trials) = (n choose k) p^k $q^{(n-k)}$

Or more formally as

$$P(k) = \binom{n}{k} p^k q^{n-k}$$

Note that if you get k heads you must necessarily get (n-k) tails. In the case of a fair coin p = q = ½.

$$P(k) = \binom{n}{k} p^n$$

Mean and Standard Deviation

The mean and standard deviation of the binomial random variable are very easily computed from the formulas

Mean **= np**
Variance **= npq**
Standard Deviation **= √npq**

Binomial Coefficients as the Number of Subsets

Earlier we saw that

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

We can see this more clearly if we look at the formula for binomial expansion when p=1/2:

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{(n-k)} = \left(\frac{1}{2} + \frac{1}{2}\right) = 1 = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{2}\right)^n$$

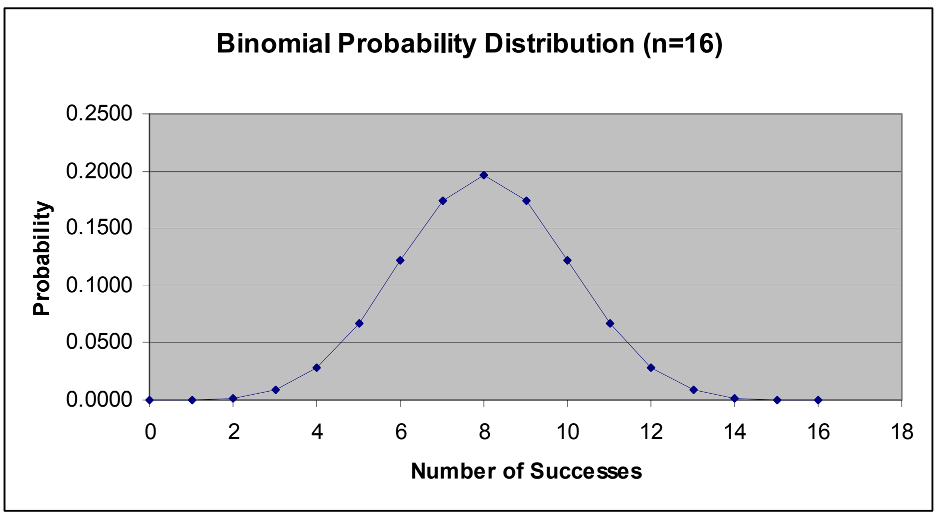
There is yet another fascinating fact related to the binomial coefficients. We have just seen that the sum of the binomial coefficients is equal to 2^n, but from that we can conclude that the number of subsets of n things is also equal to 2^n. That this is true can be deduced from a proper interpretation of each of the individual combinatorial terms. We will illustrate this with n = 4. Consider the set {a, b, c, d}. There are subsets with 0, 1, 2, 3, and 4 elements.

Subsets of 4 things = 2^n

Elements per subset	Subsets	Combinations	Count
0	∅	4C0	1
1	a.b.c.d	4C1	4
2	ab, ac, ad, bc, bd, cd	4C2	6
3	abc, abd, bcd, acd	4C3	4
4	abcd	4C4	1

Normal Approximation of Binomial Distribution

In the 1730's, the French mathematician De Moivre was searching for a more efficient method of computing the binomial coefficients. For large values of n, the computations, which had to be done by hand, were extremely laborious. De Moivre discovered an equation, which though is was an approximation, provided an accurate estimate of the binomial probabilities without having to do the detailed calculations. This approximation turned out to be what today is called the normal probability distribution or normal probability density function. What De Moivre discovered, in effect, was that for large values of n, the normal probability distribution is a good approximation of the binomial probability distribution. The normal shape is clearly evident in the following figure for n=16 and p=1/2. It should be noted that the approximation does not work well when p is significantly different from 1/2.



Binomial Distribution for n=16 and p=1/2