Binomial Example.

An accounting firm audits eight files per day. Twenty five percent of all files audited have some accounting error. If the files are independent, which implies that one file with errors does not affect the probability of other files having errors, then the number of files with errors out of eight files audited is a Binomial random variable. The variable is

X = # of files with errors out of eight files.

The values of a Binomial random variable are always 0, 1, 2, ... up to the n value or number of trials in the problem, so in this problem the number of files with errors out of eight has possible values of

$$X = 0, 1, 2, \ldots, 8.$$

The distribution of this variable is described as $X \sim Bin(n=8, p=.25)$. That is read as, "the variable X is distributed according to the Binomial distribution with parameters $n=8$ and $p=.25$." Notice, n is the number of independent trials and p is the probability of the outcome of interest, called success. In this example, the outcome of interest is a file having an accounting error. Success is not what you want to happen, but rather the event of interest.

The Binomial Mass Function is

$$P(X=x) = \binom{n}{x} p^x q^{n-x} \quad \text{where } q = 1-p, x = 0, 1, \ldots, n$$

so for $X \sim \text{Binomial}(n=8, p=.25)$ the Binomial Mass Function is

$$P(X=x) = \binom{8}{x} .25^x .75^{8-x} \quad , x = 0, 1, 2, \ldots, 8.$$

The part of the mass function that appears $\binom{8}{x}$ is called a Combination. This is a math function that calculates how many combinations of x can be made from 8. For example, how many ways are there for one file out of 8 to have an error? There are eight ways, the first could, or the second could, or the third could, and so forth, down to the eighth one. Those events would appear as, when E represents error and N represents no error, the following:

(E, N, N, N, N, N, N, N), (N, E, N, N, N, N, N, N), (N, N, E, N, N, N, N, N),

... (N, N, N, N, N, N, E, N), (N, N, N, N, N, N, N, E).

Consider writing out all the ways there are for two files to have an error and six NOT to have an error. It is much easier to calculate $\binom{8}{2} = \frac{8!}{2!(6!)} = \frac{8 \cdot 7 \cdot 6!}{(2 \cdot 1)\, 6!} = 28$ ways.

When eight files are sampled and each file is investigated for an error how many possible outcomes are there? There would only be one point with no errors, that is, when each of the eight is error free. There are eight points with one error. Those are the ones represented previously and there are twenty-eight points with two errors. But, how many ways total?

$$\binom{8}{0}+\binom{8}{1}+\binom{8}{2}+\binom{8}{3}+\binom{8}{4}+\binom{8}{5}+\binom{8}{6}+\binom{8}{7}+\binom{8}{8}=$$

$$=1+8+28+56+70+56+28+8+1=256$$

They could be represented as

$$S=\{(N,N,N,N,N,N,N,N),(E,N,N,N,N,N,N,N),\ldots,$$
$$(E,E,E,E,E,E,E,N),(E,E,E,E,E,E,E,E)\}.$$

Notice, that the total number of points or outcomes is $2^8=256$, just the same as the sum of the combinations used in the probability calculations.

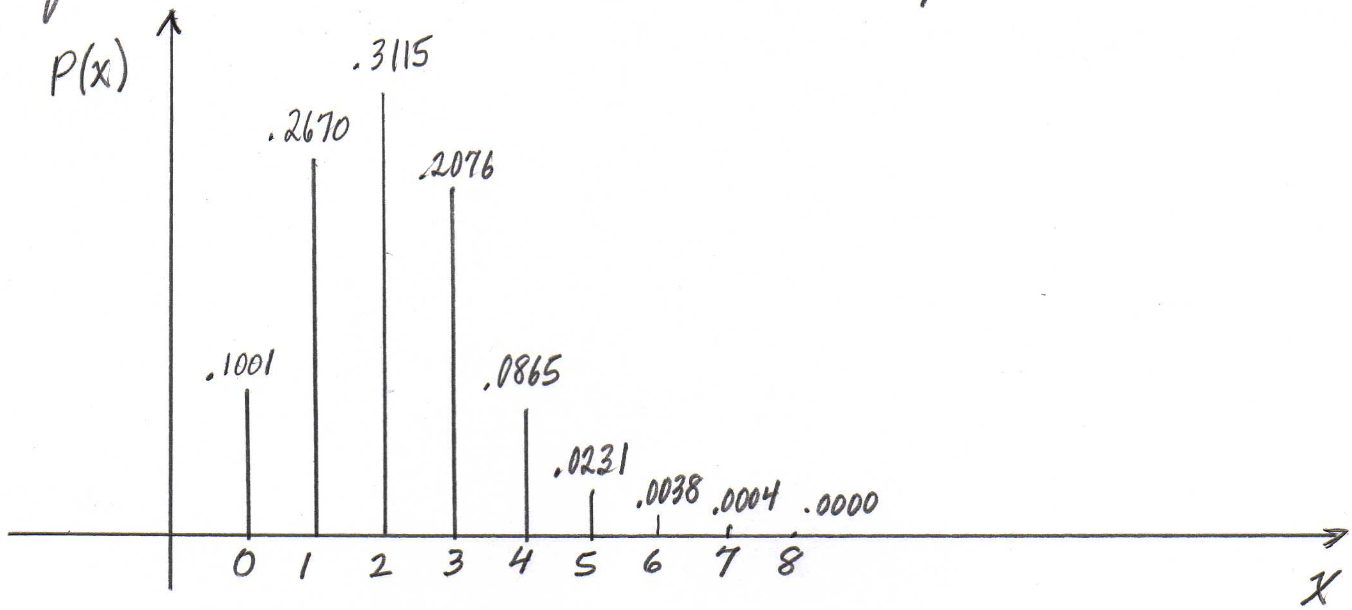
The combination part of the mass function tells how many of the points have zero files with errors, how many points have one file with errors, how many points have two files with errors, how many points have three files with errors, and so forth up to how many points have eight files with errors.

The Binomial Mass Function, $P(X=x)=\binom{8}{x}.25^{x}.75^{8-x}$ is used to calculate the probability of each possible value of the variable.

Values, X	$P(X=x)=P(x)$.
0	$P(X=0)=\binom{8}{0}.25^{0}.75^{8}=\frac{8!}{0!\,8!}.25^{0}.75^{8}=1.25^{0}.75^{8}=.1001$
1	$P(X=1)=\binom{8}{1}.25^{1}.75^{7}=\frac{8!}{1!\,7!}.25^{1}.75^{7}=8.25^{1}.75^{7}=.2670$
2	$P(X=2)=\binom{8}{2}.25^{2}.75^{6}=\frac{8!}{2!\,6!}.25^{2}.75^{6}=28.25^{2}.75^{6}=.3115$
3	$P(X=3)=\binom{8}{3}.25^{3}.75^{5}=\frac{8!}{3!\,5!}.25^{3}.75^{5}=56.25^{3}.75^{5}=.2076$
4	$P(X=4)=\binom{8}{4}.25^{4}.75^{4}=\frac{8!}{4!\,4!}.25^{4}.75^{4}=70.25^{4}.75^{4}=.0865$
5	$P(X=5)=\binom{8}{5}.25^{5}.75^{3}=\frac{8!}{5!\,3!}.25^{5}.75^{3}=56.25^{5}.75^{3}=.0231$
6	$P(X=6)=\binom{8}{6}.25^{6}.75^{2}=\frac{8!}{6!\,2!}.25^{6}.75^{2}=28.25^{6}.75^{2}=.0038$
7	$P(X=7)=\binom{8}{7}.25^{7}.75^{1}=\frac{8!}{7!\,1!}.25^{7}.75^{1}=8.25^{7}.75^{1}=.0004$
8	$P(X=8)=\binom{8}{8}.25^{8}.75^{0}=\frac{8!}{8!\,0!}.25^{8}.75^{0}=1.25^{8}.75^{0}=.0000$

Note, $0!$ is defined to be 1.

Consider a graph of the Binomial mass function for the variable, $X \sim Bi(N=8, p=.25)$.



Notice, this graph is right skewed; the majority of the unit of probability is on values on the left side of this graph. That is because $p=.25$ is closer to zero than to one. What value of p would cause the Binomial mass function to be symmetric?

What would the graph look like if $p=1$? What would the graph look like if $p=0$? What would it mean in terms of the accounting files for $p=1$? For $p=0$?

What is the expected number of files with accounting errors out of eight files if twenty-five percent of all files have errors? That is the expected value or mean of the variable, $X \sim Bin(8, .25)$.

Expected value or the Mean.

$EX = \mu = \Sigma x \cdot p(x)$, but if the variable is Binomial then this equation reduces to np. That is,

$$EX = \mu = \Sigma x \cdot p(x) = np \quad \text{if } X \sim Bi(n, p).$$

So, $EX = \mu = np = 8(.25) = 2$.

The variance of $X \sim Bin(n, p)$ is

$\sigma^2 = \Sigma x^2 p(x) - \mu^2$ reduces to npq.

So, $\sigma^2 = \Sigma x^2 \cdot p(x) - \mu^2 = npq = 8(.25)(.75)$

$= 1.5$

The standard deviation, $\sigma = \sqrt{\sigma^2} = \sqrt{1.5}$.

What is the probability that this variable, $X \sim Bin(n=8, p=.25)$ is within one standard deviation of the mean?

$$P(\mu-\sigma < X < \mu+\sigma) =$$

$$= P(2-\sqrt{1.5} < X < 2+\sqrt{1.5}) =$$

$$= P(.775 < X < 3.225) =$$

Ask yourself, what values of the variable are contained in this interval? 1, 2, 3

$$= P(1 \leq X \leq 3) =$$

$$= P(X=1) + P(X=2) + P(X=3) =$$

$$= .2670 + .3115 + .2076 = .7861.$$

Probability within two standard deviations of μ?

$$P(\mu - 2\sigma < X < \mu + 2\sigma) =$$
$$= P(2 - 2\sqrt{1.5} < X < \mu + 2\sqrt{1.5}) =$$
$$= P(-.449 < X < 4.449) =$$

$$= P(0 \leq X \leq 4) =$$

$$= P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4) =$$
$$= .1001 + .2670 + .3115 + .2076 + .0865 = .9727.$$

What is the probability that the number of files with accounting errors out of eight files is at least six?

$$P(X \geq 6) = P(X=6) + P(X=7) + P(X=8)$$

$$= .0038 + .0004 + .0000$$

$$= .0042.$$

What is the probability that fewer than two files have accounting errors out of eight files, if the probability of any one file having an error is twenty-five percent?

$$X \sim Bin(n=8,\ p=.25)$$

$$P(X < 2) = P(X \leq 1)$$

$$= P(X=0) + P(X=1)$$

$$= .1001 + .2670$$

$$= .3671.$$