

THE BINOMIAL DISTRIBUTION



The *binomial probability distribution* generally occurs when a procedure has two outcomes that may be categorized as either *success* or *failure*.

More specifically, the procedure must meet the following conditions.

More specifically, the procedure must meet the following conditions.

- The procedure has a fixed number of trials n .

More specifically, the procedure must meet the following conditions.

- There procedure has a fixed number of trials n .
- The outcome of each trial is independent of the others.

More specifically, the procedure must meet the following conditions.

- There procedure has a fixed number of trials n .
- The outcome of each trial is independent of the others.
- The outcome of each trial can be classified as either *success* (S) or *failure* (F).

More specifically, the procedure must meet the following conditions.

- There procedure has a fixed number of trials n .
- The outcome of each trial is independent of the others.
- The outcome of each trial can be classified as either *success* (S) or *failure* (F).
- The probability of *success* is the same for each trial, $P(S)=p$, and the probability of *failure* is $P(F)=1-p=q$.

Suppose we take a 5-question multiple choice test and that there are 4 choices for each problem.

Even though each problem has 4 choices, we can still think of each answer as resulting in one of two outcomes. Our answer is either a *success* or a *failure*.

Now suppose that we guess on each question. Then the probability of *success* is $p = 1/4 = 0.25$, and the probability of *failure* is $q = 3/4 = 0.75$.

Now suppose that we guess on each question. Then the probability of *success* is $p = 1/4 = 0.25$, and the probability of *failure* is $q = 3/4 = 0.75$.

Also, we can assume that each “guess” is independent of the others, and so this procedure is a *binomial experiment*. It leads to a binomial distribution.

Now suppose that we guess on each question. Then the probability of *success* is $p = 1/4 = 0.25$, and the probability of *failure* is $q = 3/4 = 0.75$.

Also, we can assume that each “guess” is independent of the others, and so this procedure is a *binomial experiment*. It leads to a binomial distribution.

Furthermore, our random variable is $X = \text{the number of successes}$.

We now just need to figure out how to determine the probability of getting x successes.

Suppose that we guess the first three questions correctly and then miss the next two. The probability of this outcome is shown below.

$$P(\text{correct \& correct \& correct \& wrong \& wrong})$$

$$= (.25)(.25)(.25)(.75)(.75) = .25^3 .75^2 \approx 0.0088$$

Notice, though, that any result that involves three correct and two wrong will have exactly the same probability.

$P(\text{correct \& wrong \& correct \& wrong \& correct})$

$$= (.25)(.75)(.25)(.75)(.25) = .25^3 .75^2 \approx 0.0088$$

Now the question is just how many ways can we complete the test so that three of the answers are correct and two are wrong?

$P(\text{correct \& wrong \& correct \& wrong \& correct})$

$$= (.25)(.75)(.25)(.75)(.25) = .25^3 .75^2 \approx 0.0088$$

The answer is given by one of our counting formulas. Think of it in terms of having five answers in front of us and we want to pick exactly three of them to be *correct*. The number of ways we can do this is equal to the number of combinations of five objects choose three.

$${}_5C_3 = \frac{5!}{(5-3)!3!} = 10$$

So we now know that there are 10 outcomes that contain three successes and two failures, and each individual outcome has a probability of approximately 0.0088.

$${}_5C_3 = \frac{5!}{(5-3)!3!} = 10$$

$$.25^3 .75^2 \approx 0.0088$$

Furthermore, since these outcomes are *mutually exclusive* from one another, we can find the overall probability of getting one or the other by just adding the individual probabilities together.

$${}_5C_3 = \frac{5!}{(5-3)!3!} = 10$$

$$.25^3 .75^2 \approx 0.0088$$

In other words, the probability of getting exactly three successes is given by the formula below.

$$P(3) = {}_5C_3 (.25^3)(.75^2) \approx 0.088$$

More generally, in a binomial experiment with n trials, the probability of getting x successes is:

$$P(x) = {}_n C_x (p^x)(q^{n-x})$$

We can now complete the probability distribution for this experiment.

$$P(x) = {}_n C_x (p^x)(q^{n-x})$$

$$P(0) = {}_5 C_0 (.25^0)(.75^5) \approx .23730$$

$$P(1) = {}_5 C_1 (.25^1)(.75^4) \approx .39551$$

$$P(2) = {}_5 C_2 (.25^2)(.75^3) \approx .26367$$

$$P(3) = {}_5 C_3 (.25^3)(.75^2) \approx .08789$$

$$P(4) = {}_5 C_4 (.25^4)(.75^1) \approx .01465$$

$$P(5) = {}_5 C_5 (.25^5)(.75^0) \approx .00098$$

x	$P(x)$
0	0.23730
1	0.39551
2	0.26367
3	0.08789
4	0.01465
5	0.00098

As usual, however, there is also a way to do this on our calculator.

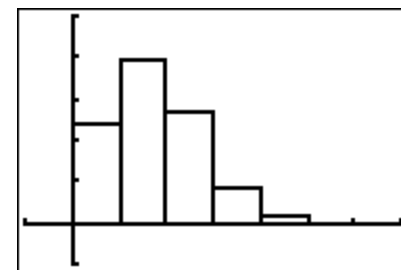
```
binompdf(5,.25,3)
)
.087890625
```

```
binompdf(5,.25)→
L2
(.2373046875 .3...
```

L1	L2	L3	2
0	.23730	-----	
1	.39551		
2	.26367		
3	.08789		
4	.01465		
5	9.8E-4		
-----	-----		
L2(1)=.2373046875			

```
Plot1 Plot2 Plot3
Off Off
Type: [Bar] [Line] [Pie]
[Box] [Dot] [None]
Xlist:L1
Freq:L2
```

```
WINDOW
Xmin=-1
Xmax=7
Xscl=1
Ymin=-.1
Ymax=.5
Yscl=.1
Xres=1
```



There's an easy way to figure the *mean* or *expected value* of this distribution.

```
binompdf(5,.25,3
)
      .087890625
```

```
binompdf(5,.25)→
L2
(.2373046875 .3...
```

L1	L2	L3	2
0	.2373		
1	.39551		
2	.26367		
3	.08789		
4	.01465		
5	9.8E-4		
-----	-----		
L2(1)=.2373046875			

```
Plot1 Plot2 Plot3
Off Off
Type: [Bar] [Line] [Pie]
      [Box] [Dot] [None]
Xlist:L1
Freq:L2
```

```
WINDOW
Xmin=-1
Xmax=7
Xscl=1
Ymin=-.1
Ymax=.5
Yscl=.1
Xres=1
```



If we took several 5 problem tests of this sort and guessed on each question with probability of success equal to $\frac{1}{4}$, then in the long run we would expect to get about $\frac{1}{4}$ of the questions correct on each test. In other words, the *mean* of a binomial distributions is:

$$\mu = np$$

$$(5)(.25) = 1.25$$

The *standard deviation* of a binomial distribution is a simple formula, but we will give it without proof.

$$\mu = np$$

$$(5)(.25) = 1.25$$

$$\sigma = \sqrt{npq}$$

$$\sqrt{5 \cdot .25 \cdot .75} \approx .96825$$

And of course, this can also be done just using the calculator.

L1	L2	L3	Z
0	.2373	-----	
1	.39551		
2	.26367		
3	.08789		
4	.01465		
5	9.8E-4		

L2(1)=.2373046875			

```
1-Var Stats L1,L
2
```

```
1-Var Stats
x̄=1.25
Σx=1.25
Σx²=2.5
Sx=
σx=.9682458366
↓n=1
```