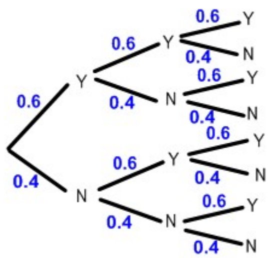
Bernoulli Trials and the Binomial Probability Model

At a college, 60% of the students voted in the latest election. If you select 3 students at random, what is the probability that exactly 2 voted in the election?



AND (multiply)

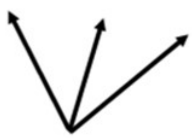
YYN (.6)(.6)(.4)
YNY (.6)(.4)(.6)
OR (add)
NYY (.4)(.6)(.6)

$$P(exactly\ 2\ voted) = P(YYN) + P(YNY) + P(NYY)$$

$$P(exactly\ 2\ voted) = (.6)(.6)(.4) + (.6)(.4)(.6) + (.4)(.6)(.6)$$

$$P(exactly\ 2\ voted) = (.6)^2(.4)^1 + (.6)^2(.4)^1 + (.6)^2(.4)^1$$

$$P(exactly\ 2\ voted) = 3(.6)^2(.4)^1$$



$$P(exactly\ 2\ voted) = {}_3C_2\left(P(voted)\right)^2\left(P(didn't\ vote)\right)^1$$

number of ways the 2 students who voted can be chosen

multiple trials, all independent
probability is constant
each trial only two options: "success" (voted), "failure" (did not vote)

Situations like this are called Bernoulli Trials

If we are running a **fixed number of trials** (here, 3 students) the probability answer is always in the form...

$$P(exactly\ 2\ voted) = {}_3C_2\left(P(voted)\right)^2\left(P(didn't\ vote)\right)^1$$

...which looks like a term from a binomial theorem expansion, so this model is called the **Binomial probability model**. More generally:

$$P(exactly\ k\ successes) = {}_nC_k\,(p)^k\,(q)^{n-k}$$

where n = *number of trials*
k = *number of successes*
p = *probability of success for a single trial*
q = *probability of failure for a single trial*

At a college, 60% of the students voted in the latest election. If you select 3 students at random, what is the probability that exactly 2 voted in the election?

We use this model so often that there is a calculator function dedicated to it:

$$binompdf(n,\ p,\ k)$$

Slightly different problem:
At a college, 60% of the students voted in the latest election. If you select 3 students at random, what is the probability that up to 2 voted in the election?

What are all the possibilities of number of students of 3 who voted?

0 voted | 1 voted | 2 voted | 3 voted

$$binompdf(3, .6, 0) + binompdf(3, .6, 1) + binompdf(3, .6, 2)$$

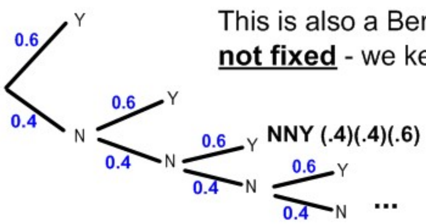
$$binomcdf(3, .6, 2)$$

c for cumulative

Another variation:
At a college, 60% of the students voted in the latest election. If you select 3 students at random, what is the probability that at least 2 voted in the election?

Bernoulli Trials and the Geometric Probability Model

At a college, 60% of the students voted in the latest election. If you keep asking students until you find one that voted, what is the probability that you have to ask exactly 3 students in order to find one that voted?



This is also a Bernoulli trial case, but **the number of trials is not fixed** - we keep asking *until* we have a 'success' case.

The results is in the form $P(\text{exactly } k \text{ successes}) = (q)^{k-1}(p)$
and this is called the **Geometric probability model**.

Like the binomial model, there are calculator functions for computation:

$geometpdf(p, k)$ probability that 1st success is on the k^{th} trial

$geometcdf(p, k)$ probability that 1st success is on or before the k^{th} trial

At a college, 60% of the students voted in the latest election. If you keep asking students until you find one that voted, what is the probability that you have to ask more than 3 students in order to find one that voted?

Bernoulli Trials

Compare these three scenarios...

A) If you draw two cards from a deck, what is the probability they are kings?

B) If you flip a fair coin twice, what is the probability you get tails both times?

C) If the ball pit in a MacDonald's play area contains 10,000 balls, and 10% of them are green, what is the probability that if you randomly choose 2 balls from the pit they will both be green?

The shape, mean, standard deviation of the Geometric Distribution

The Geometric distribution is a distribution so it has a shape which has a mean and standard deviation.

$$G(\mu, \sigma)$$

$p = 0.2$

$geometpdf(0.2,1) = .200$

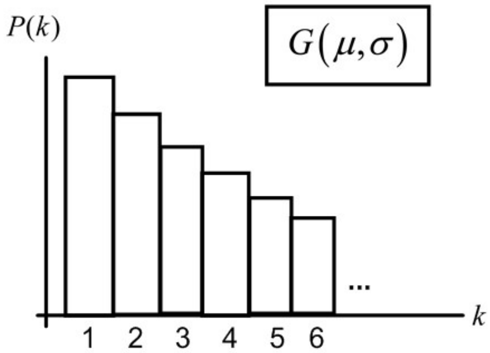
$geometpdf(0.2,2) = .160$

$geometpdf(0.2,3) = .128$

$geometpdf(0.2,4) = .102$

$geometpdf(0.2,5) = .082$

$geometpdf(0.2,6) = .066$



(Math box (p.388) has an interesting derivation for the mean.)

$$\mu = \frac{1}{p} \qquad \sigma = \sqrt{\frac{q}{p^2}}$$

The shape, mean, standard deviation of the Binomial Distribution

The Binomial distribution is also a distribution and has a shape, a mean, and standard deviation. The shape depends upon the probability of success and the number of trials.

$$B(\mu, \sigma)$$

$p = 0.1, n = 5$

$binompdf(5,0.1,0) = .590$

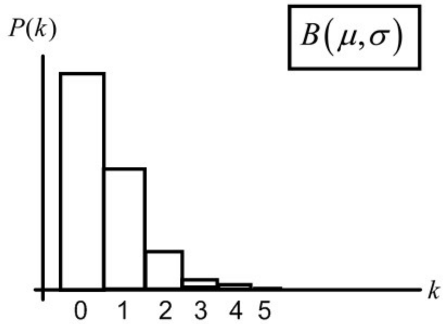
$binompdf(5,0.1,1) = .320$

$binompdf(5,0.1,2) = .073$

$binompdf(5,0.1,3) = .008$

$binompdf(5,0.1,4) = .00045$

$binompdf(5,0.1,5) = .00001$



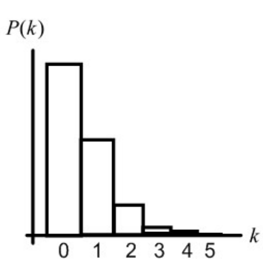
(Math box (p.392) has an interesting derivation for the mean and standard deviation.)

$$\mu = np \qquad \sigma = \sqrt{npq}$$

The shape of the Binomial Distribution - varying p

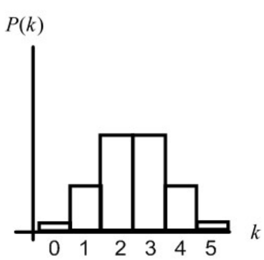
Changing the probability of success, p, changes the shape...

$p = 0.1, n = 5$



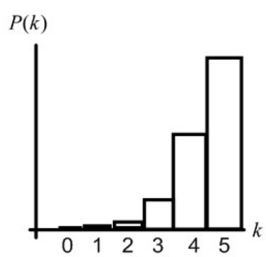
p close to 0
skewed right

$p = 0.5, n = 5$



p close to 0.5
symmetric

$p = 0.9, n = 5$

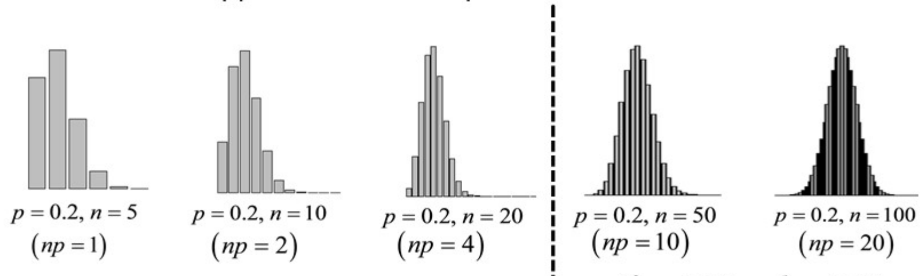


p close to 1
skewed left

(imagine flipping a coin 5 times with different probability of 'heads')

The shape of the Binomial Distribution - varying n

As the number of trials, n, increases, the shape of the Binomial distribution approaches the shape of a Normal distribution:



The "Success/Failure Condition"
If we expect at least 10 successes and 10 failures, we can use a Normal model to approximate the Binomial model. → *If $np \geq 10$ and $nq \geq 10$ the Binomial model is approximately Normal.*

(Math box (p.395) explains this condition in more detail.)

(see practice packet for examples of approximating a Binomial distribution with a Normal model)

Summary of probability models

<u>Binomial Model (Bernoulli trials) Setting</u> $\mu = np$ $\sigma = \sqrt{npq}$

1) Only 2 outcomes (success/failure)
2) p does not change
3) Trials are independent
4) ***Fixed*** number of trials

$$binompdf(n, p, k) = \text{exactly } k \text{ successes out of } n \text{ trials}$$

$$binomcdf(n, p, k) = \text{at most } k \text{ successes out of } n \text{ trials}$$

Can approximate Binomial $B(\mu, \sigma)$ with Normal $N(\mu, \sigma)$ if $np \geq 10$ and $nq \geq 10$
$(np = \#\,successes,\ \ nq = \#\,failures)$

<u>Geometric Model (Bernoulli trials) Setting</u> $\mu = \frac{1}{p}$ $\sigma = \sqrt{\frac{q}{p^2}}$

1) Only 2 outcomes (success/failure)
2) p does not change
3) Trials are independent
4) ***Number of trials not fixed*** (continue until first success)

$$geometpdf(p, k) = 1^{st} \text{ success is on } k^{th} \text{ trial}$$

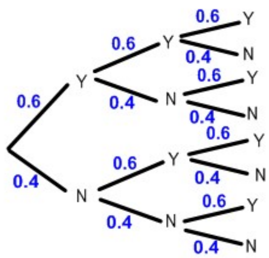
$$geometcdf(p, k) = 1^{st} \text{ success is on or before } k^{th} \text{ trial}$$

where $n = \text{number of trials}$
$k = \text{number of successes}$
$p = \text{probability of success for a single trial}$
$q = \text{probability of failure for a single trial}$
$q = 1 - p$

Bernoulli Trials *can be considered independent if sample is <10% of the entire population.*

Bernoulli Trials and the Binomial Probability Model

At a college, 60% of the students voted in the latest election. If you select 3 students at random, what is the probability that exactly 2 voted in the election?



AND (multiply)

YYN (.6)(.6)(.4)
YNY (.6)(.4)(.6)
OR (add)
NYY (.4)(.6)(.6)

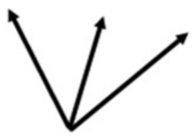
$$P(exactly\ 2\ voted) = P(YYN) + P(YNY) + P(NYY)$$

$$P(exactly\ 2\ voted) = (.6)(.6)(.4) + (.6)(.4)(.6) + (.4)(.6)(.6)$$

$$P(exactly\ 2\ voted) = (.6)^2(.4)^1 + (.6)^2(.4)^1 + (.6)^2(.4)^1$$

$$P(exactly\ 2\ voted) = 3(.6)^2(.4)^1$$

$$P(exactly\ 2\ voted) = {}_3C_2(P(voted))^2(P(didn't\ vote))^1$$



number of ways the 2 students who voted can be chosen

Situations like this are called Bernoulli Trials

multiple trials, all independent
probability is constant
each trial only two options: "success" (voted), "failure" (did not vote)

If we are running a **fixed number of trials** (here, 3 students) the probability answer is always in the form...

$$P(exactly\ 2\ voted) = {}_3C_2(P(voted))^2(P(didn't\ vote))^1$$

...which looks like a term from a binomial theorem expansion, so this model is called the **Binomial probability model**. More generally:

$$P(exactly\ k\ successes) = {}_nC_k(p)^k(q)^{n-k}$$

$$\begin{aligned} where\ n &= number\ of\ trials \\ k &= number\ of\ successes \\ p &= probability\ of\ success\ for\ a\ single\ trial \\ q &= probability\ of\ failure\ for\ a\ single\ trial \end{aligned}$$

At a college, 60% of the students voted in the latest election. If you select 3 students at random, what is the probability that exactly 2 voted in the election?

We use this model so often that there is a calculator function dedicated to it:

$$binompdf(n, p, k)$$

Slightly different problem:
At a college, 60% of the students voted in the latest election. If you select 3 students at random, what is the probability that up to 2 voted in the election?

What are all the possibilities of number of students of 3 who voted?

0 voted | 1 voted | 2 voted | 3 voted

$$binompdf(3, .6, 0) + binompdf(3, .6, 1) + binompdf(3, .6, 2)$$

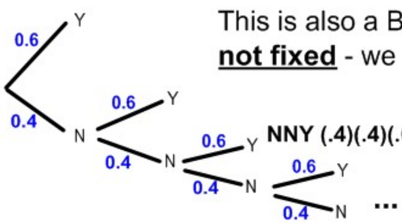
$$binomcdf(3, .6, 2)$$

c for cumulative

Another variation:
At a college, 60% of the students voted in the latest election. If you select 3 students at random, what is the probability that at least 2 voted in the election?

Bernoulli Trials and the Geometric Probability Model

At a college, 60% of the students voted in the latest election. If you keep asking students until you find one that voted, what is the probability that you have to ask exactly 3 students in order to find one that voted?



This is also a Bernoulli trial case, but **the number of trials is not fixed** - we keep asking *until* we have a 'success' case.

The results is in the form $P(exactly\ k\ successes) = (q)^{k-1}(p)$
and this is called the **Geometric probability model**.

Like the binomial model, there are calculator functions for computation:

$geometpdf(p,k)$ probability that 1st success is on the k^{th} trial

$geometcdf(p,k)$ probability that 1st success is on or before the k^{th} trial

At a college, 60% of the students voted in the latest election. If you keep asking students until you find one that voted, what is the probability that you have to ask more than 3 students in order to find one that voted?

Bernoulli Trials

Compare these three scenarios...

A) If you draw two cards from a deck, what is the probability they are kings?

B) If you flip a fair coin twice, what is the probability you get tails both times?

C) If the ball pit in a MacDonald's play area contains 10,000 balls, and 10% of them are green, what is the probability that if you randomly choose 2 balls from the pit they will both be green?

The shape, mean, standard deviation of the Geometric Distribution

The Geometric distribution is a distribution so it has a shape which has a mean and standard deviation.

$G(\mu,\sigma)$

$p = 0.2$

$geometpdf(0.2,1) = .200$

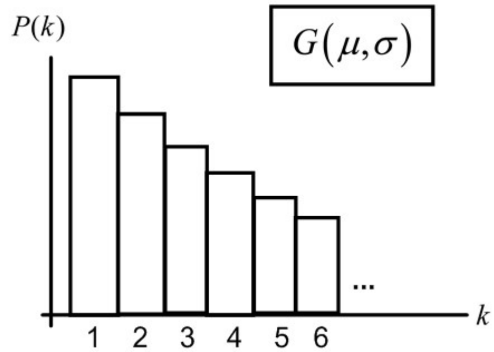
$geometpdf(0.2,2) = .160$

$geometpdf(0.2,3) = .128$

$geometpdf(0.2,4) = .102$

$geometpdf(0.2,5) = .082$

$geometpdf(0.2,6) = .066$



(Math box (p.388) has an interesting derivation for the mean.)

$\mu = \frac{1}{p}$ $\sigma = \sqrt{\frac{q}{p^2}}$

The shape, mean, standard deviation of the Binomial Distribution

The Binomial distribution is also a distribution and has a shape, a mean, and standard deviation. The shape depends upon the probability of success and the number of trials.

$B(\mu,\sigma)$

$p = 0.1, n = 5$

$binompdf(5,0.1,0) = .590$

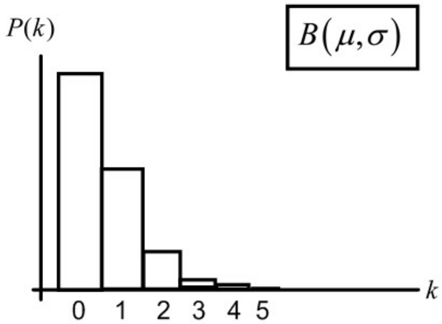
$binompdf(5,0.1,1) = .320$

$binompdf(5,0.1,2) = .073$

$binompdf(5,0.1,3) = .008$

$binompdf(5,0.1,4) = .00045$

$binompdf(5,0.1,5) = .00001$



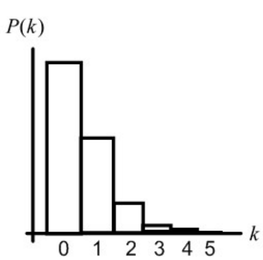
(Math box (p.392) has an interesting derivation for the mean and standard deviation.)

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

The shape of the Binomial Distribution - varying p

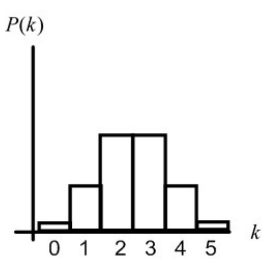
Changing the probability of success, p, changes the shape...

$p = 0.1, n = 5$



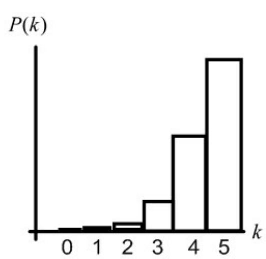
p close to 0
skewed right

$p = 0.5, n = 5$



p close to 0.5
symmetric

$p = 0.9, n = 5$

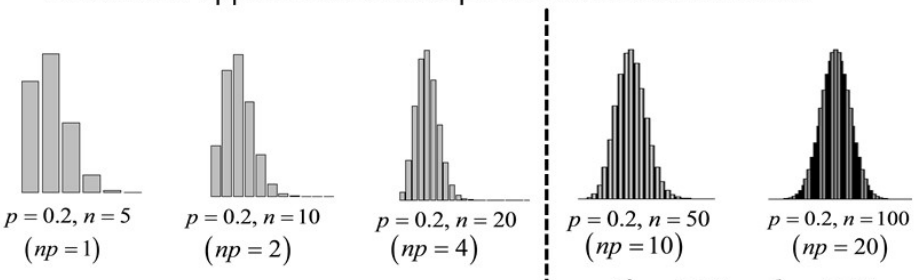


p close to 1
skewed left

(imagine flipping a coin 5 times with different probability of 'heads')

The shape of the Binomial Distribution - varying n

As the number of trials, n, increases, the shape of the Binomial distribution approaches the shape of a Normal distribution:



The "Success/Failure Condition"
If we expect at least 10 successes and 10 failures, we can use a Normal model to approximate the Binomial model. → *If $np \geq 10$ and $nq \geq 10$ the Binomial model is approximately Normal.*

(Math box (p.395) explains this condition in more detail.)

(see practice packet for examples of approximating a Binomial distribution with a Normal model)

Summary of probability models

<u>Binomial Model (Bernoulli trials) Setting</u> $\mu = np$ $\sigma = \sqrt{npq}$

1) Only 2 outcomes (success/failure)
2) p does not change
3) Trials are independent
4) ***Fixed*** number of trials

$$binompdf(n, p, k) = exactly\ k\ successes\ out\ of\ n\ trials$$

$$binomcdf(n, p, k) = at\ most\ k\ successes\ out\ of\ n\ trials$$

Can approximate Binomial $B(\mu, \sigma)$ with Normal $N(\mu, \sigma)$ if $np \geq 10$ and $nq \geq 10$
$(np = \#\,successes,\ \ nq = \#\,failures)$

<u>Geometric Model (Bernoulli trials) Setting</u> $\mu = \frac{1}{p}$ $\sigma = \sqrt{\frac{q}{p^2}}$

1) Only 2 outcomes (success/failure)
2) p does not change
3) Trials are independent
4) ***Number of trials not fixed***
   (continue until first success)

$$geometpdf(p, k) = 1^{st}\ success\ is\ on\ k^{th}\ trial$$

$$geometcdf(p, k) = 1^{st}\ success\ is\ on\ or\ before\ k^{th}\ trial$$

where $n = number\ of\ trials$

$k = number\ of\ successes$

$p = probability\ of\ success\ for\ a\ single\ trial$

$q = probability\ of\ failure\ for\ a\ single\ trial$

$q = 1 - p$

Bernoulli Trials *can be considered independent if sample is <10% of the entire population.*