Distributions and Intro to Likelihood

Gov 2001 Section

February 4, 2010

Outline

Why should we become familiar with these distributions?

Why should we become familiar with these distributions?

- Part of the point of this class is to get you to fit models to a variety of data.

Why should we become familiar with these distributions?

- Part of the point of this class is to get you to fit models to a variety of data.
- But the first step is recognizing what kind of data you are working with.

Why should we become familiar with these distributions?

- Part of the point of this class is to get you to fit models to a variety of data.
- But the first step is recognizing what kind of data you are working with.
- If you see that your data are Poisson, Binomial, Normal, etc., then you can analyze the data using a model (likelihood or Bayesian) appropriate for that data.

- So learning about the distributions is a bit like eating your spinach! It's not pleasant, but it's really useful.

- So learning about the distributions is a bit like eating your spinach! It's not pleasant, but it's really useful.
- It's a lot better for you than forcing on the data a distribution assumption that doesn't make sense (e.g., assuming the data are normal and then using OLS).

- So learning about the distributions is a bit like eating your spinach! It's not pleasant, but it's really useful.
- It's a lot better for you than forcing on the data a distribution assumption that doesn't make sense (e.g., assuming the data are normal and then using OLS).
- What's the best way to learn the distributions? Learn the "stories" behind them.

- So learning about the distributions is a bit like eating your spinach! It's not pleasant, but it's really useful.
- It's a lot better for you than forcing on the data a distribution assumption that doesn't make sense (e.g., assuming the data are normal and then using OLS).
- What's the best way to learn the distributions? Learn the "stories" behind them.
- Remember that you can always look up the "specs" of the distributions later – just focus on trying to identify them for now.

Outline

Outline

The Bernoulli Distribution

- Takes value 1 with success probability π and value 0 with failure probability $1 - \pi$.
- Ideal for modelling **one-time** yes/no (or success/failure) events.
- The best example is **one** coin flip – if your data resemble a single coin flip, then you have a Bernoulli distribution.

The Bernoulli Distribution

- Takes value 1 with success probability π and value 0 with failure probability $1 - \pi$.
- Ideal for modelling **one-time** yes/no (or success/failure) events.
- The best example is **one** coin flip – if your data resemble a single coin flip, then you have a Bernoulli distribution.
- ex) one voter voting yes/no

The Bernoulli Distribution

- Takes value 1 with success probability π and value 0 with failure probability $1 - \pi$.
- Ideal for modelling **one-time** yes/no (or success/failure) events.
- The best example is **one** coin flip – if your data resemble a single coin flip, then you have a Bernoulli distribution.
- ex) one voter voting yes/no
- ex) one person being either a man/woman

The Bernoulli Distribution

- Takes value 1 with success probability π and value 0 with failure probability $1 - \pi$.
- Ideal for modelling **one-time** yes/no (or success/failure) events.
- The best example is **one** coin flip – if your data resemble a single coin flip, then you have a Bernoulli distribution.
- ex) one voter voting yes/no
- ex) one person being either a man/woman
- ex) the New Orleans Saints winning/losing the Super Bowl

The Bernoulli Distribution

$Y \sim \text{Bernoulli}(\pi)$

$y = 0, 1$

probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \pi^y (1 - \pi)^{(1-y)}$

$E(Y) = \pi$

$\text{Var}(Y) = \pi(1 - \pi)$

The Bernoulli Distribution

$Y \sim \text{Bernoulli}(\pi)$

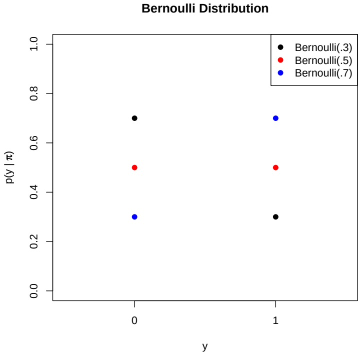
$y = 0, 1$

probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \pi^y (1 - \pi)^{(1-y)}$

$E(Y) = \pi$

$\text{Var}(Y) = \pi(1 - \pi)$



The Binomial Distribution

- Let's say you run a bunch of Bernoulli trials, and, instead of seeing the result of each trial separately, you just see the grand total.

The Binomial Distribution

- Let's say you run a bunch of Bernoulli trials, and, instead of seeing the result of each trial separately, you just see the grand total.
- So, for example, you flip a coin three times and count the total number of heads you got. (The order doesn't matter.)

The Binomial Distribution

- Let's say you run a bunch of Bernoulli trials, and, instead of seeing the result of each trial separately, you just see the grand total.
- So, for example, you flip a coin three times and count the total number of heads you got. (The order doesn't matter.)
- This is the Binomial. It's ideal for modelling **repeated** yes/no (or success/failure) events.

The Binomial Distribution

- Let's say you run a bunch of Bernoulli trials, and, instead of seeing the result of each trial separately, you just see the grand total.
- So, for example, you flip a coin three times and count the total number of heads you got. (The order doesn't matter.)
- This is the Binomial. It's ideal for modelling **repeated** yes/no (or success/failure) events.
- ex) the number of women in a group of 10 Harvard students

The Binomial Distribution

- Let's say you run a bunch of Bernoulli trials, and, instead of seeing the result of each trial separately, you just see the grand total.
- So, for example, you flip a coin three times and count the total number of heads you got. (The order doesn't matter.)
- This is the Binomial. It's ideal for modelling **repeated** yes/no (or success/failure) events.
- ex) the number of women in a group of 10 Harvard students
- ex) the number of rainy days in the seven week

The Binomial Distribution

$Y \sim \text{Binomial}(n, \pi)$

$y = 0, 1, \ldots, n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \binom{n}{y} \pi^y (1 - \pi)^{(n-y)}$

$E(Y) = n\pi$

$\text{Var}(Y) = n\pi(1 - \pi)$

The Binomial Distribution

$Y \sim \text{Binomial}(n, \pi)$

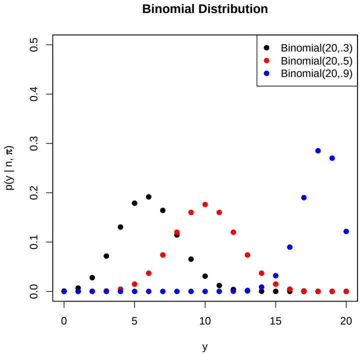
$y = 0, 1, \ldots, n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \binom{n}{y} \pi^y (1-\pi)^{(n-y)}$

$E(Y) = n\pi$

$\text{Var}(Y) = n\pi(1-\pi)$



The Multinomial Distribution

- Suppose you had more than just two outcomes – e.g., vote for Republican, Democrat, or Independent. Can you use a binomial?

The Multinomial Distribution

- Suppose you had more than just two outcomes – e.g., vote for Republican, Democrat, or Independent. Can you use a binomial?
- We can't use a binomial, because a binomial requires two outcomes(yes/no, 1/0, etc.). Instead, we use the multinomial.
- Multinomial lets you work with several mutually exclusive outcomes.

The Multinomial Distribution

- Suppose you had more than just two outcomes – e.g., vote for Republican, Democrat, or Independent. Can you use a binomial?
- We can't use a binomial, because a binomial requires two outcomes(yes/no, 1/0, etc.). Instead, we use the multinomial.
- Multinomial lets you work with several mutually exclusive outcomes.
- ex) you toss a die 15 times and get outcomes 1-6

The Multinomial Distribution

- Suppose you had more than just two outcomes – e.g., vote for Republican, Democrat, or Independent. Can you use a binomial?
- We can't use a binomial, because a binomial requires two outcomes(yes/no, 1/0, etc.). Instead, we use the multinomial.
- Multinomial lets you work with several mutually exclusive outcomes.
- ex) you toss a die 15 times and get outcomes 1-6
- ex) ten undergraduate students are classified freshmen, sophomores, juniors, or seniors

The Multinomial Distribution

- Suppose you had more than just two outcomes – e.g., vote for Republican, Democrat, or Independent. Can you use a binomial?
- We can't use a binomial, because a binomial requires two outcomes(yes/no, 1/0, etc.). Instead, we use the multinomial.
- Multinomial lets you work with several mutually exclusive outcomes.
- ex) you toss a die 15 times and get outcomes 1-6
- ex) ten undergraduate students are classified freshmen, sophomores, juniors, or seniors
- ex) Gov graduate students divided into either American, Comparative, Theory, or IR

The Multinomial Distribution

$Y \sim \text{Multinomial}(n, \pi_1, \ldots, \pi_k)$

$y_j = 0, 1, \ldots, n; \;\; \sum_{j=1}^{k} y_j = n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success for j: $\pi_j \in [0, 1]; \;\; \sum_{j=1}^{k} \pi_j = 1$

$p(\mathbf{y}|n, \boldsymbol{\pi}) = \frac{n!}{y_1! y_2! \ldots y_k!} \pi_1^{y_1} \pi_2^{y_2} \ldots \pi_k^{y_k}$

$E(Y_j) = n\pi_j$

$\text{Var}(Y_j) = n\pi_j(1 - \pi_j)$

$\text{Cov}(Y_i, Y_j) = -n\pi_i\pi_j$

The Poisson Distribution

- Represents the number of events occurring in a fixed period of time.
- Can also be used for the number of events in other specified intervals such as distance, area, or volume.
- Can never be negative – so, good for modeling events.

The Poisson Distribution

- Represents the number of events occurring in a fixed period of time.
- Can also be used for the number of events in other specified intervals such as distance, area, or volume.
- Can never be negative – so, good for modeling events.
- ex) the number Prussian solders who died each year by being kicked in the head by a horse (Bortkiewicz, 1898)

The Poisson Distribution

- Represents the number of events occurring in a fixed period of time.
- Can also be used for the number of events in other specified intervals such as distance, area, or volume.
- Can never be negative – so, good for modeling events.
- ex) the number Prussian solders who died each year by being kicked in the head by a horse (Bortkiewicz, 1898)
- ex) the of number shark attacks in Australia per month

The Poisson Distribution

- Represents the number of events occurring in a fixed period of time.
- Can also be used for the number of events in other specified intervals such as distance, area, or volume.
- Can never be negative – so, good for modeling events.
- ex) the number Prussian solders who died each year by being kicked in the head by a horse (Bortkiewicz, 1898)
- ex) the of number shark attacks in Australia per month
- ex) the number of search warrant requests a federal judge hears in one year

The Poisson Distribution

$Y \sim \text{Poisson}(\lambda)$

$y = 0, 1, \ldots$

expected number of occurrences: $\lambda > 0$

$p(y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}$

$E(Y) = \lambda$

$\text{Var}(Y) = \lambda$

The Poisson Distribution

$Y \sim \text{Poisson}(\lambda)$

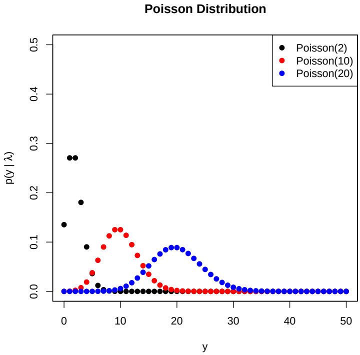
$y = 0, 1, \ldots$

expected number of occurrences: $\lambda > 0$

$p(y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}$

$E(Y) = \lambda$

$\text{Var}(Y) = \lambda$



Outline

The Univariate Normal Distribution

- Probably the one distribution you are already familiar with – describes data that cluster in a bell curve around the mean.

The Univariate Normal Distribution

- Probably the one distribution you are already familiar with – describes data that cluster in a bell curve around the mean.
- A lot of naturally occurring processes are normally distributed.

The Univariate Normal Distribution

- Probably the one distribution you are already familiar with – describes data that cluster in a bell curve around the mean.
- A lot of naturally occurring processes are normally distributed.
- ex) the weights of male students in our class

The Univariate Normal Distribution

- Probably the one distribution you are already familiar with – describes data that cluster in a bell curve around the mean.
- A lot of naturally occurring processes are normally distributed.
- ex) the weights of male students in our class
- ex) high school students' SAT scores

The Univariate Normal Distribution

$Y \sim \text{Normal}(\mu, \sigma^2)$

$y \in \mathbb{R}$

mean: $\mu \in \mathbb{R}$
variance: $\sigma^2 > 0$

$$p(y|\mu, \sigma^2) = \frac{\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}}$$

$E(Y) = \mu$

$\text{Var}(Y) = \sigma^2$

The Univariate Normal Distribution

$$Y \sim \text{Normal}(\mu, \sigma^2)$$

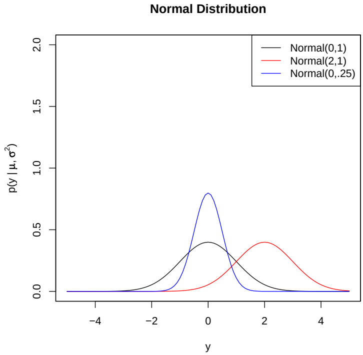
$$y \in \mathbb{R}$$

mean: $\mu \in \mathbb{R}$
variance: $\sigma^2 > 0$

$$p(y|\mu, \sigma^2) = \frac{\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}}$$

$$E(Y) = \mu$$

$$\text{Var}(Y) = \sigma^2$$



The Uniform Distribution

- Any number in the interval you chose is equally probable.

The Uniform Distribution

- Any number in the interval you chose is equally probable.
- Intuitively easy to understand, but hard to come up with examples. (Easier to think of discrete uniform examples.)

The Uniform Distribution

- Any number in the interval you chose is equally probable.
- Intuitively easy to understand, but hard to come up with examples. (Easier to think of discrete uniform examples.)
- ex) the numbers that come out of random number generators

The Uniform Distribution

- Any number in the interval you chose is equally probable.
- Intuitively easy to understand, but hard to come up with examples. (Easier to think of discrete uniform examples.)
- ex) the numbers that come out of random number generators
- ex) rolling 1-6 in a die roll (discrete)

The Uniform Distribution

- Any number in the interval you chose is equally probable.
- Intuitively easy to understand, but hard to come up with examples. (Easier to think of discrete uniform examples.)
- ex) the numbers that come out of random number generators
- ex) rolling 1-6 in a die roll (discrete)
- ex) the lottery tumblers out of which a person draws one ball with a number on it (also discrete)

The Uniform Distribution

$Y \sim \text{Uniform}(\alpha, \beta)$

$y \in [\alpha, \beta]$

Interval: $[\alpha, \beta]$; $\beta > \alpha$

$p(y|\alpha, \beta) = \frac{1}{\beta - \alpha}$

$E(Y) = \frac{\alpha + \beta}{2}$

$\text{Var}(Y) = \frac{(\beta - \alpha)^2}{12}$

Quiz: Test Your Knowledge of Discrete Distributions

Are the following Bernoulli (coin flip), Binomial(several coin flips), Multinomial (Rep, Dem, Indep), Poisson (Prussian soldier deaths), Normal (SAT scores), or Uniform (die)?

Quiz: Test Your Knowledge of Discrete Distributions

Are the following Bernoulli (coin flip), Binomial(several coin flips), Multinomial (Rep, Dem, Indep), Poisson (Prussian soldier deaths), Normal (SAT scores), or Uniform (die)?

- The heights of trees on campus?

Quiz: Test Your Knowledge of Discrete Distributions

Are the following Bernoulli (coin flip), Binomial(several coin flips), Multinomial (Rep, Dem, Indep), Poisson (Prussian soldier deaths), Normal (SAT scores), or Uniform (die)?

- The heights of trees on campus?
- The number of airplane crashes in one year?

Quiz: Test Your Knowledge of Discrete Distributions

Are the following Bernoulli (coin flip), Binomial(several coin flips), Multinomial (Rep, Dem, Indep), Poisson (Prussian soldier deaths), Normal (SAT scores), or Uniform (die)?

- The heights of trees on campus?
- The number of airplane crashes in one year?
- A yes or no vote cast by Senator Brown?

Quiz: Test Your Knowledge of Discrete Distributions

Are the following Bernoulli (coin flip), Binomial(several coin flips), Multinomial (Rep, Dem, Indep), Poisson (Prussian soldier deaths), Normal (SAT scores), or Uniform (die)?

- The heights of trees on campus?
- The number of airplane crashes in one year?
- A yes or no vote cast by Senator Brown?
- The number of parking tickets Cambridge PD gives out in one month?

Quiz: Test Your Knowledge of Discrete Distributions

Are the following Bernoulli (coin flip), Binomial(several coin flips), Multinomial (Rep, Dem, Indep), Poisson (Prussian soldier deaths), Normal (SAT scores), or Uniform (die)?

- The heights of trees on campus?
- The number of airplane crashes in one year?
- A yes or no vote cast by Senator Brown?
- The number of parking tickets Cambridge PD gives out in one month?
- The poll your Facebook friends took to choose their favorite sport out of football, basketball, and soccer

Outline

Meet the Distributions!

- Discrete Distributions
- Continuous Distributions

Basic Likelihood

Likelihood

The whole point of likelihood is to leverage information about the data generating process into our inferences.

Likelihood

The whole point of likelihood is to leverage information about the data generating process into our inferences.

Here are the basic steps:

Likelihood

The whole point of likelihood is to leverage information about the data generating process into our inferences.

Here are the basic steps:

- Think about your data generating process. (What do the data look like? Use your substantive knowledge.)

Likelihood

The whole point of likelihood is to leverage information about the data generating process into our inferences.

Here are the basic steps:

- Think about your data generating process. (What do the data look like? Use your substantive knowledge.)
- Find a distribution that you think explains the data. (Poisson, Binomial, Normal? Something else?)

Likelihood

The whole point of likelihood is to leverage information about the data generating process into our inferences.

Here are the basic steps:

- Think about your data generating process. (What do the data look like? Use your substantive knowledge.)
- Find a distribution that you think explains the data. (Poisson, Binomial, Normal? Something else?)
- Derive the likelihood.

Likelihood

The whole point of likelihood is to leverage information about the data generating process into our inferences.

Here are the basic steps:

- Think about your data generating process. (What do the data look like? Use your substantive knowledge.)
- Find a distribution that you think explains the data. (Poisson, Binomial, Normal? Something else?)
- Derive the likelihood.
- If you want, maximize the likelihood to get the MLE.

Likelihood

The whole point of likelihood is to leverage information about the data generating process into our inferences.

Here are the basic steps:

- Think about your data generating process. (What do the data look like? Use your substantive knowledge.)
- Find a distribution that you think explains the data. (Poisson, Binomial, Normal? Something else?)
- Derive the likelihood.
- If you want, maximize the likelihood to get the MLE.

Note: This is the case in the univariate context. We'll be introducting covariates later on in the term.

Likelihood: An Example

Let's walk through an example.

Likelihood: An Example

Let's walk through an example.

Suppose I am a lawyer and I want to study the rate of convictions in Massachusetts. Here is my data:

Likelihood: An Example

Let's walk through an example.

Suppose I am a lawyer and I want to study the rate of convictions in Massachusetts. Here is my data:

- There are 100 cases in my data set. In each a defendant is either found innocent or guilty.

Likelihood: An Example

Let's walk through an example.

Suppose I am a lawyer and I want to study the rate of convictions in Massachusetts. Here is my data:

- There are 100 cases in my data set. In each a defendant is either found innocent or guilty.
- I observe that defendants are found innocent in 65 of them.

Likelihood: An Example

Let's walk through an example.

Suppose I am a lawyer and I want to study the rate of convictions in Massachusetts. Here is my data:

- There are 100 cases in my data set. In each a defendant is either found innocent or guilty.
- I observe that defendants are found innocent in 65 of them.
- And they are found guilty in 35 of them.

Likelihood: An Example

Let's walk through an example.

Suppose I am a lawyer and I want to study the rate of convictions in Massachusetts. Here is my data:

- There are 100 cases in my data set. In each a defendant is either found innocent or guilty.
- I observe that defendants are found innocent in 65 of them.
- And they are found guilty in 35 of them.

These data follow what distribution?

Likelihood: An Example (ctd)

So our data are binomial.
Now what do we do?

Likelihood: An Example (ctd)

So our data are binomial.
Now what do we do?

- Look up the appropriate PDF. (If you are unsure, talk to your friends, look at Wikipedia, look at a probability textbook.)

Likelihood: An Example (ctd)

So our data are binomial.
Now what do we do?

- Look up the appropriate PDF. (If you are unsure, talk to your friends, look at Wikipedia, look at a probability textbook.)
- We know from lecture that the binomial PDF is $p(y|\pi) = \binom{n}{y}\pi^y(1-\pi)^{(n-y)}$

Likelihood: An Example (ctd)

So our data are binomial.
Now what do we do?

- Look up the appropriate PDF. (If you are unsure, talk to your friends, look at Wikipedia, look at a probability textbook.)
- We know from lecture that the binomial PDF is $p(y|\pi) = \binom{n}{y}\pi^y(1-\pi)^{(n-y)}$
- Here, $n = 100$ and $y = 35$. Note that y is the number of "successes," here the number of guilty defendants.

Likelihood: An Example (ctd)

So our data are binomial.
Now what do we do?

- Look up the appropriate PDF. (If you are unsure, talk to your friends, look at Wikipedia, look at a probability textbook.)
- We know from lecture that the binomial PDF is $p(y|\pi) = \binom{n}{y}\pi^{y}(1-\pi)^{(n-y)}$
- Here, $n = 100$ and $y = 35$. Note that y is the number of "successes," here the number of guilty defendants.
- Plugging in this info gives us $p(y|\pi) = \binom{100}{35}\pi^{35}(1-\pi)^{(100-35)}$

Likelihood: An Example (ctd)

Next, let's calculate the likelihood function. Where does the likelihood come from?

Likelihood: An Example (ctd)

Next, let's calculate the likelihood function. Where does the likelihood come from?

- From Bayes' Rule, we get $p(\pi|y) = \frac{p(y|\pi)p(\pi)}{p(y)}$

Likelihood: An Example (ctd)

Next, let's calculate the likelihood function. Where does the likelihood come from?

- From Bayes' Rule, we get $p(\pi|y) = \frac{p(y|\pi)p(\pi)}{p(y)}$
- Let $k(y) = \frac{p(\pi)}{p(y)}$

Likelihood: An Example (ctd)

Next, let's calculate the likelihood function. Where does the likelihood come from?

- From Bayes' Rule, we get $p(\pi|y) = \frac{p(y|\pi)p(\pi)}{p(y)}$
- Let $k(y) = \frac{p(\pi)}{p(y)}$
- Note that the π in $k(y)$ is the true π, a constant that doesn't vary. So $k(y)$ is just a function of y.

Likelihood: An Example (ctd)

Next, let's calculate the likelihood function. Where does the likelihood come from?

- From Bayes' Rule, we get $p(\pi|y) = \frac{p(y|\pi)p(\pi)}{p(y)}$
- Let $k(y) = \frac{p(\pi)}{p(y)}$
- Note that the π in $k(y)$ is the true π, a constant that doesn't vary. So $k(y)$ is just a function of y.
- Define $L(\pi|y) = p(y|\pi)k(y)$

Likelihood: An Example (ctd)

Next, let's calculate the likelihood function. Where does the likelihood come from?

- From Bayes' Rule, we get $p(\pi|y) = \frac{p(y|\pi)p(\pi)}{p(y)}$
- Let $k(y) = \frac{p(\pi)}{p(y)}$
- Note that the π in $k(y)$ is the true π, a constant that doesn't vary. So $k(y)$ is just a function of y.
- Define $L(\pi|y) = p(y|\pi)k(y)$
- $\rightarrow L(\pi|y) \propto p(y|\pi)$

Likelihood: An Example (ctd)

So here are our steps:

Likelihood: An Example (ctd)

So here are our steps:

- First, we got $p(y|\pi)$ from the Biomial PDF: $\binom{100}{35}\pi^{35}(1-\pi)^{(100-35)}$

Likelihood: An Example (ctd)

So here are our steps:

- First, we got $p(y|\pi)$ from the Biomial PDF: $\binom{100}{35}\pi^{35}(1-\pi)^{(100-35)}$
- Second, we derived the likelihood, $L(\pi|y) \propto p(y|\pi)$

Likelihood: An Example (ctd)

So here are our steps:

- First, we got $p(y|\pi)$ from the Biomial PDF: $\binom{100}{35}\pi^{35}(1-\pi)^{(100-35)}$
- Second, we derived the likelihood, $L(\pi|y) \propto p(y|\pi)$
- Third, we can pull this all together, $L(\pi|y) \propto \binom{100}{35}\pi^{35}(1-\pi)^{(100-35)}$

Likelihood: An Example (ctd)

So here are our steps:

- First, we got $p(y|\pi)$ from the Biomial PDF: $\binom{100}{35}\pi^{35}(1-\pi)^{(100-35)}$
- Second, we derived the likelihood, $L(\pi|y) \propto p(y|\pi)$
- Third, we can pull this all together, $L(\pi|y) \propto \binom{100}{35}\pi^{35}(1-\pi)^{(100-35)}$

That's it!

Likelihood: An Example (ctd)

So now we have our likelihood function,
$L(\pi|y) \propto \binom{100}{35}\pi^{35}(1-\pi)^{(100-35)}$.

- The interpretation is that it's the likelihood of our model having generated the data.
- The likelihood doesn't make much sense in the abstract. How to make sense?
- (1) It's a good idea to plot it to get a sense of what's going onf
- (2) deriving (analytically or via simulation) the maximum of the likelihood, which is the maximum likelihood estimate (MLE)

Plotting the example

First, note that we can take advantage of a lot of pre-packged R functions

- `rbinom`, `rpoisson`, `rnorm`, `runif` → gives random values from that distribution
- `pbinom`, `ppoisson`, `pnorm`, `punif`→ gives the cumulative distribution (the probability of that value or less)
- `dbinom`, `dpoisson`, `dnorm`, `dunif`→ gives the density (i.e., height of the PDF – useful for drawing)
- `qbinom`, `qpoisson`, `qnorm`, `qunif`→ gives the quantile function (given quantile, tells you the value)

We can also write our own function using the `plot` command.

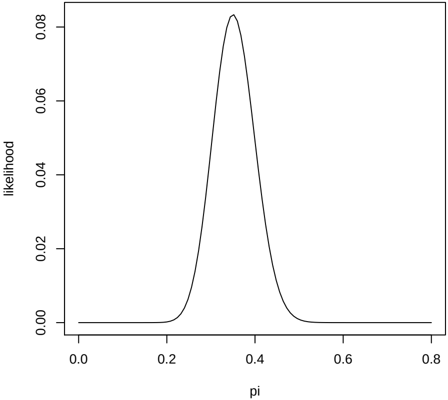
Plotting the example

We want to plot $L(\pi|y) \propto \binom{100}{35}\pi^{35}(1-\pi)^{(100-35)}$

```
> ## example using the dbinom
> dbinom(35, size = 100, prob =.35)
[1] 0.0834047
> ## prob of getting 35 successes given that prob = .35
> ## it's actually kind of low

> curve(dbinom(35, size = 100, prob = x),
  xlim =c(0,.8), xlab ="pi", ylab = "likelihood")
```

Plotting the example



Can we eyeball what the maximum likelihood estimate will be?

Other things to keep in mind

Other things to keep in mind

- What if we have two or more data points that we believe come from the same model?

Other things to keep in mind

- What if we have two or more data points that we believe come from the same model?
- We can derive a likelihood for the combined data by multiplying the independent likelihoods together.

Other things to keep in mind

- What if we have two or more data points that we believe come from the same model?
- We can derive a likelihood for the combined data by multiplying the independent likelihoods together.
- Taking the log of the likelihood (the “log-likelihood”) someimes makes this easier.

Other things to keep in mind

- What if we have two or more data points that we believe come from the same model?
- We can derive a likelihood for the combined data by multiplying the independent likelihoods together.
- Taking the log of the likelihood (the "log-likelihood") someimes makes this easier.
- But we will address this – as well as finding the MLE – in the weeks to come.