Lab 5: Reproducible Data Analysis and Recitation of Week 1-4 Models

Haohan Chen

February 9, 2018

Agenda

- Reproducible Research (some experience)
- Short paper I: Idea? Data? Concerns?
- Likelihood function and MLE

Reproducible Research

Reproducible Research (Workflow)

Idea – Theory – Empirical Analysis

- Idea, data, theory: Think about them at the same time

Reproducible Research (Workflow)

Idea – Theory – Empirical Analysis

- Idea, data, theory: Think about them at the same time
- Is it cheating to look at data when you think about theory? Depends

Reproducible Research (Workflow)

Idea – Theory – Empirical Analysis

- Idea, data, theory: Think about them at the same time
- Is it cheating to look at data when you think about theory? Depends
- Exploratory Data Analysis
    - Correlations
    - Shape of pattern

Reproducible Research (Workflow)

Idea – Theory – Empirical Analysis

- Idea, data, theory: Think about them at the same time
- Is it cheating to look at data when you think about theory? Depends
- Exploratory Data Analysis
  - Correlations
  - Shape of pattern
- Models: from simple to complex
  - Start with 'lm'
  - Use complex models to solve remaining problems
  - If you can find something ONLY with certain complex models. Don't trust it too much

Reproducible Research (Workflow)

Idea – Theory – Empirical Analysis

- Idea, data, theory: Think about them at the same time
- Is it cheating to look at data when you think about theory? Depends
- Exploratory Data Analysis
  - Correlations
  - Shape of pattern
- Models: from simple to complex
  - Start with ‘lm‘
  - Use complex models to solve remaining problems
  - If you can find something ONLY with certain complex models. Don't trust it too much
- Interpretation of your results: Be honest (to yourself), be confident (in front of the audience)

Reproducible Research (Pragmatics)

Data – researchers/ coauthors – readers

- Organize your folder
- Clean, extendible code
- Reproducible (but don't go to far)
- Allocate time for visualization

Reproducible Research (Example)

`(see Lab 4 material)`

Thoughts about your short paper?

Likelihood

What is likelihood?

The task: Data meets the Model

What is likelihood?

The task: Data meets the Model

Ex: Female representativeness in the supreme court We have:

- Data: Number of female judges in the supreme court $y = 2$
- Model: *assume* y is drawn from a ainomial Distribution

$$\underbrace{P(Y = y \mid \pi) = \binom{n}{y} \pi^y (1 - \pi)^{N-y}}$$

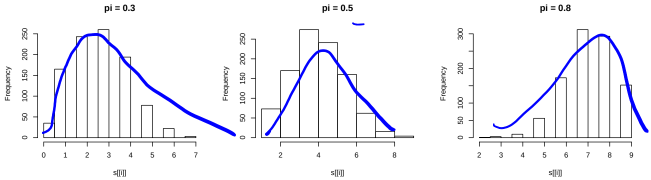
What is likelihood?

The task: Data meets the Model

Ex: Female representativeness in the supreme court We have:

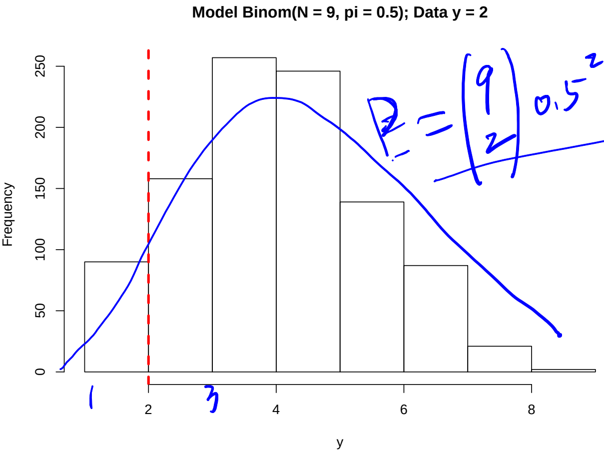
- Data: Number of female judges in the supreme court $y = 2$
- Model: *assume* y is drawn from a ainomial Distribution

$$P(Y = y \mid \pi) = \binom{n}{y} \pi^{y} (1 - \pi)^{N-y}$$



What is likelihood?

If, magically, we know the model! Say, $y \sim Binom(N = 9, \pi = 0.5)$, which means equal representativeness between male and female. What is the probability of observing our data?



$$P = \binom{9}{2} 0.5^2 (1-0.5)^{9-2}$$

What is likelihood?

But this is not a task we often do. Often, we have data, we have a sketch of a theoretical model (with parameterss). Want: Estimate the parameters. In our Supreme Court example, we want π.

What is likelihood?

But this is not a task we often do. Often, we have data, we have a sketch of a theoretical model (with parameterss). Want: Estimate the parameters. In our Supreme Court example, we want π.

π is a function of the data and other parameters. We define the "likelihood of π"

var parameter Var param $Y_i \overset{iid}{\sim} \text{Binom}(n, p)$

$$L(\pi \mid y) = P(Y = y \mid \pi) = \binom{n}{y} \pi^y (1 - \pi)^{N-y}$$

What is the nature of this likelihood function?

$L(\pi \mid Y) = P(Y \mid \pi)$ $Y = \{2, 1, 0, 0\}$.

$= \Pr(Y_1, Y_2, \ldots \mid \pi)$ $= \prod_{i=1}^{n} \Pr(Y_i \mid \pi)$

$= \prod_{i=1}^{n} \binom{n}{y_i} \pi^{y_i} (1-\pi)^{N-y_i}$

What is likelihood?

Can we call the likelihood $L(\pi \mid y)$ ''the probability density of π conditional on y''

What is likelihood?

Can we call the likelihood $L(\pi \mid y)$ ''the probability density of π conditional on y''

No. If $f(x)$ is a probability density function for a continuous random variable X then

$$(1) F(x) = Pr(X \leq x) = \int_{-\infty}^{x} \cancel{x} f(t)dt$$

$$(2) f(x) \geq 0 \text{ for any value of x}$$

$$(3) \int_{-\infty}^{\infty} f(x)dx = 1$$

What is likelihood?

Can we call the likelihood $L(\pi \mid y)$ ‘’the probability density of π conditional on y”

No. If $f(x)$ is a probability density function for a continuous random variable X then

$$(1)F(x) = Pr(X \leq x) = \int_{-\infty} xf(t)dt$$

$$(2)f(x) \geq 0 \text{ for any value of x}$$

$$(3)\int_{-\infty}^{\infty} f(x)dx = 1$$

A likelihood function $L(\pi \mid y)$ need not meet these criteria (e.g. $\int_{-\infty}^{\infty} L(\pi \mid y)d\pi \neq 1$).

It’s a function that leads us to some π of interest. Nothing more.

What is Maximum Likelihood Estimator?

What π do we want?

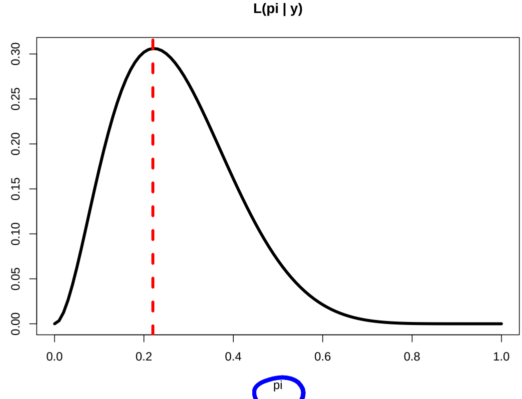
We want an estimated $\hat{\pi}$ that maximizes the likelihood $L(\pi \mid y)$

Why? (my tentative answer) Think of it as maximizing the **joint probability** of observing all your data points given the model you assume

The Shape of the $L(\pi \mid y)$ and MLE

$$L(\pi \mid y) = \binom{n}{y} \pi^y (1-\pi)^{N-y}$$

We can simulate this (see `.Rmd code`)

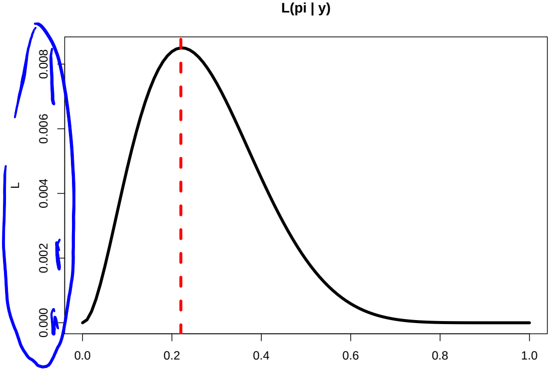


The Shape of the $L(\pi \mid y)$ and MLE

Actually, terms that do not include parameter π do not matter

$$L(\pi \mid y) = \boxed{\binom{n}{y}} \pi^y (1-\pi)^{N-y} \propto \pi^y (1-\pi)^{N-y}$$

We can simulate this (see `.Rmd` code)

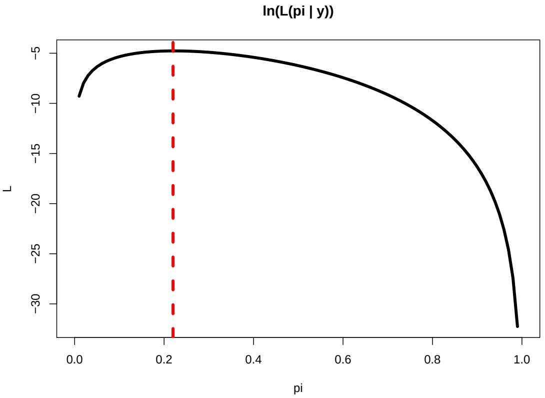


The Shape of the $L(\pi \mid y)$ and MLE

Then, taking the `logarithm` yields the same π_{MLE}

$$\ln L(\pi \mid y) \propto \ln \left\{ \pi^{y} (1-\pi)^{N-y} \right\} \propto \boxed{y \ln \pi + (1-y) \ln(1-\pi)}$$

We can simulate this `(see .Rmd code)`



The Shape of the $L(\pi \mid y)$ and MLE

Why take logarithm? (1) Likelihood can be very small. (2) A computational problem – Floating-Point Underflow. Everything goes to zero!

```
a <- 0.01^1000; b <- 0.02^1000
cat("a = ", a, "; b = ", b, "; a < b?", a < b)
```

```
## a =  0 ; b =  0 ; a < b? FALSE
```

```
log_a <- sum(rep(log(0.01), 1000))
log_b <- sum(rep(log(0.02), 1000))
cat("log(a) = ", log_a, "; log(b) = ", log_b)
```

```
## log(a) =  -4605.17 ; log(b) =  -3912.023
```

```
cat("log(a) < log(b)?", log_a < log_b)
```

```
## log(a) < log(b)? TRUE
```

The Shape of the $L(\pi \mid y)$ and MLE

We can derive MLE analytically

$$\ell \propto y \ln \pi + (1-y)\ln(1-\pi)$$

$$\frac{\partial}{\partial \pi}(\ell) = \frac{y}{\pi} - \frac{1-y}{1-\pi} = 0$$

$$\pi^*_{MLE} = \_\_\_\_\_\_$$

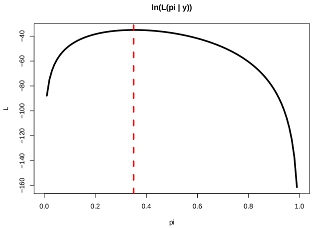
outcome. cat.

The Shape of the $L(\pi \mid y)$ and MLE

Many data points? $y = \{2, 1, 4, 4, 3, 5\}$

$$L(\pi \mid y) = \prod_i^n \binom{n}{y}_i \pi_i^y (1-\pi)^{N-y_i} \propto \prod_i n\pi_i^y (1-\pi)^{N-y_i}$$

$$\ln L(\pi \mid y) \propto \sum_i^n y_i \ln \pi + (1-y_i)\ln(1-\pi)$$



Regressions: MLE for Linear Models

Derive MLE for linear models (setup)

Regressions: MLE for Linear Models

Derive MLE for linear models (likelihood function)

Regressions: MLE for Linear Models

Derive MLE for linear models (get MLE)

Regressions: MLE for Linear Models

Derive MLE for linear models (get MLE)

Regressions: OLS for Linear Models

Derive OLS Estimator for linear models

Regressions: OLS for Linear Models

Derive OLS Estimator for linear models

Regressions: OLS for Linear Models

Derive OLS Estimator for linear models

Regressions: OLS for Linear Models

Derive OLS Estimator for linear models