To Do List

Read Chapter 2, Sections 2.1-2.3

Do Problems 1-20 in Chapter 1

Tutorial Test 1 is on Wednesday September 28 – see detailed instructions posted on Learn

Today's Class: Parameter Estimation

1) Definition of a (Point) Estimate of an Unknown Parameter

2) Method of Maximum Likelihood

i) Definition of the Likelihood Function

ii) Definition of the Maximum Likelihood Estimate

iii) Definition of the Relative Likelihood Function

iv) Definition of the Log Likelihood Function

Definition 7, page 47

A (point) estimate of a parameter $\boldsymbol{\theta}$ is the value of a function of the observed data *y*.

The estimate is denoted by

$$\hat{\theta} \text{ where } \hat{\theta} = \hat{\theta}(\mathbf{y}).$$

Note: most often the data are of the form *y* = $(y_1, y_2, \ldots, y_n)$.

How do we estimate the unknown parameters when we have data that we assume have come from a:

(1) Binomial model
(2) Poisson model
(3) Exponential model
(4) Gaussian model

How do we estimate an unknown parameter for other models?

We need a method of estimation which has a mathematical justification and which can be used when a reasonable estimate is not obvious.

An Important Idea

Values of θ which make the observed data, $y = 10$ heads in 25 tosses, more probable seem more reasonable or plausible than values of θ which make the observed data, $y = 10$ heads in 25 tosses, improbable.

What is the value of θ which make the observed data, $y = 10$ heads in 25 tosses, most probable?

Binomial Example

Let *Y* = number of heads in 25 tosses of the coin.

Then *Y* ~ Binomial(25, *θ*),

$$P(Y = y;\theta) = \binom{25}{y}\theta^{y}(1-\theta)^{25-y} \quad y = 0,1,\ldots,25;\ 0 < \theta < 1.$$

and P(observing 10 heads in 25 tosses ; *θ*)

$$= P(Y = 10;\theta) = \binom{25}{10}\theta^{10}(1-\theta)^{15} \quad \text{for}\ \ 0 < \theta < 1.$$

Binomial Example Continued

$$P(Y = 10; \theta) = \binom{25}{10} \theta^{10} (1-\theta)^{15} \quad \text{for } 0 < \theta < 1$$

The value of $\boldsymbol{\theta}$ which maximizes this function of $\boldsymbol{\theta}$ is $\boldsymbol{\theta}$ = 10/25 = 0.4. Can you show this?

Note: We can see from the graph on the next slide that the value of

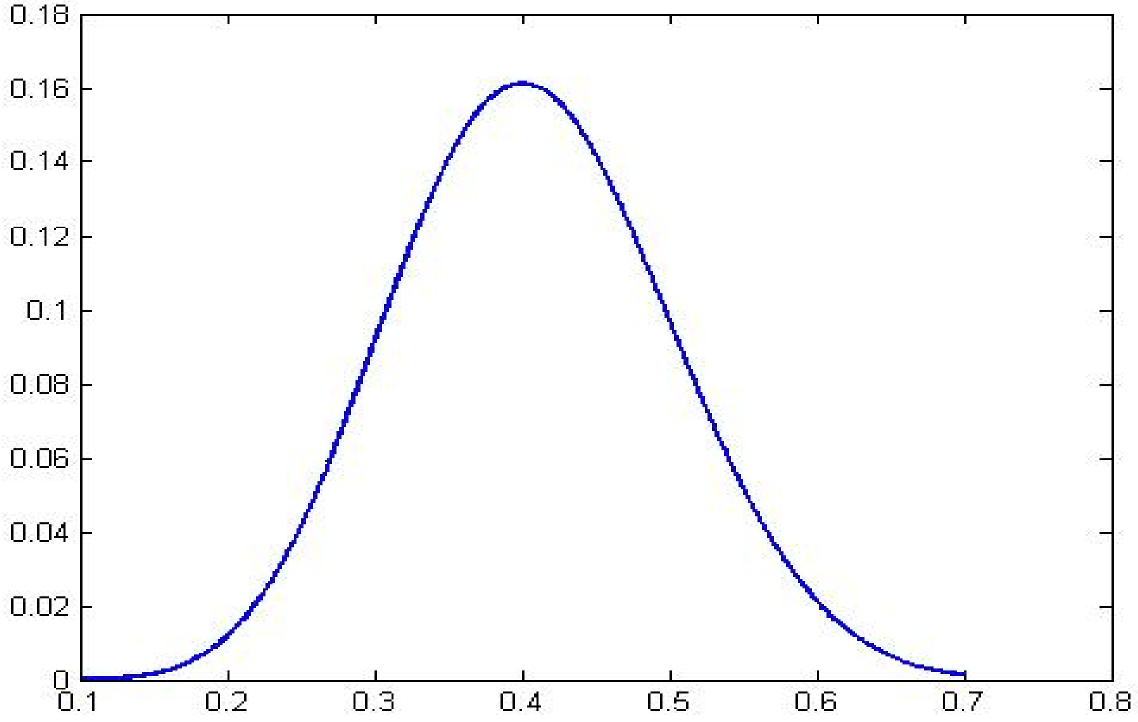
$$P(Y = 10; \theta = 0.42)$$

is bigger than the value of

$$P(Y = 10; \theta = 0.1).$$

Probability of Observing y=10

$P(Y = 10; \theta)$



θ

$P(Y = 10; \theta)$ is maximized at $\theta = 0.4$

Method of Maximum Likelihood

In this simple example we have used what is called the Method of Maximum Likelihood to estimate an unknown parameter $\boldsymbol{\theta}$ in an assumed model for observed data $\boldsymbol{y}$.

The method of maximum likelihood is the most widely used method of estimation.

Method of Maximum Likelihood

Let the random variable *Y* represent potential data that will be used to estimate *θ* and let *y* represent the actual observed data.

We begin by supposing *Y* is a discrete random variable.

Note: most often *Y* = ($Y_1, Y_2, \ldots, Y_n$ **).**

Definition of the Likelihood Function

The likelihood function for $\boldsymbol{\theta}$ is defined as

$$L(\theta) = L(\theta; \boldsymbol{y}) = P(\boldsymbol{Y} = \boldsymbol{y}; \theta) \text{ for } \theta \in \Omega$$

$$= \text{the probability that we observe the data } \boldsymbol{y} \text{ as a function of } \theta$$

where Ω = parameter space

= set of possible values of $\boldsymbol{\theta}$.

Note: The likelihood is a function of both $\boldsymbol{\theta}$ and the data $\boldsymbol{y}$ but usually we write just $\boldsymbol{L(\theta)}$.

Important Idea

We decide on how plausible (reasonable) a value of θ is by looking at how probable it makes the observed data.

Values of θ which make the observed data probable are considered to be more plausible than values of θ which make the observed data improbable.

The Maximum Likelihood Estimate

In other words, values of θ which correspond to large values of L(θ) are more consistent with the observed data y.

The value of θ that maximizes L(θ) for given data y is called the maximum likelihood estimate (m.l. estimate) of θ

and is denoted by $\hat{\theta} = \hat{\theta}(y)$.

Likelihood Function and Maximum Likelihood Estimate for Binomial Data - Summary

Let *Y* = number of successes in *n* Bernoulli trials with P(Success) = *θ*. Then *Y* ~ Binomial(*n*, *θ*).

Suppose a Binomial experiment is conducted and *y* successes are observed. The likelihood function for *θ* based on the observed data is

$$L(\theta) = P(Y = y; \theta)$$

$$= \binom{n}{y} \theta^{y} (1-\theta)^{n-y} \quad for \;\; 0 < \theta < 1.$$

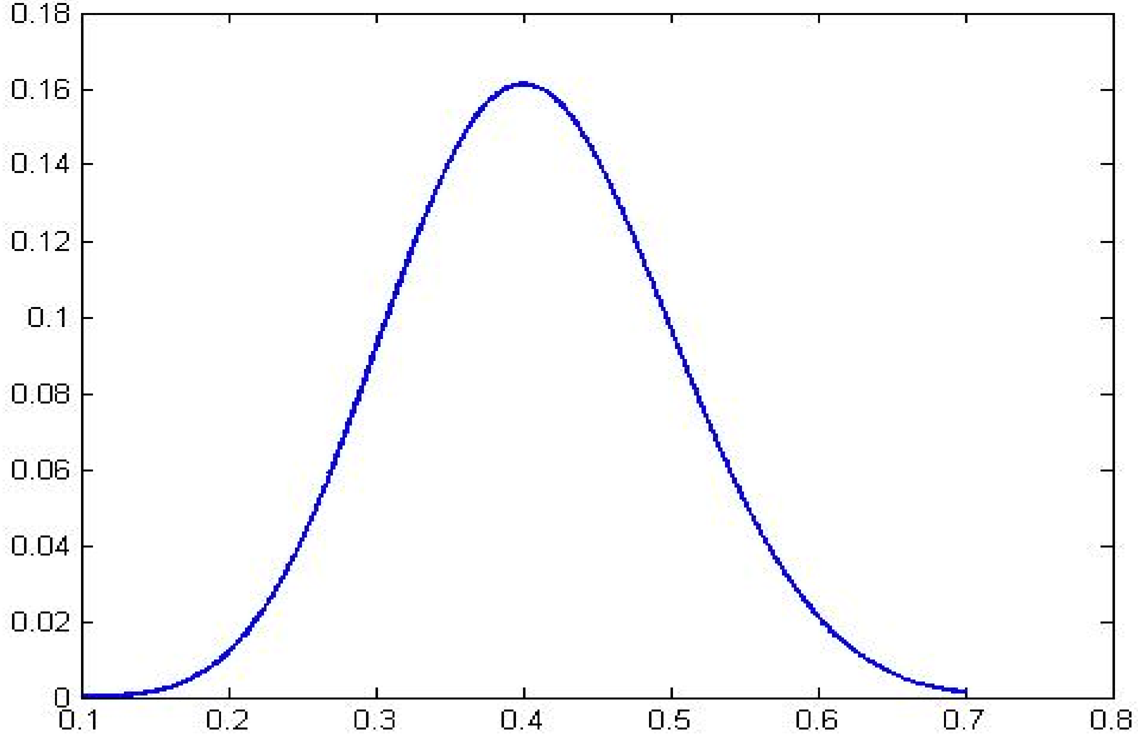
and the maximum likelihood estimate of *θ* is

$$\hat{\theta} = \frac{y}{n}$$

$L(\theta)$ for Coin Example

Graph of $L(\theta)=\binom{25}{10}\theta^{10}(1-\theta)^{15}$

$L(\theta)$



θ

$L(\theta)$ is maximized at θ = 10/25 = 0.4