Important Ideas from Last Day

(1) The Sampling Distribution of the Sample Mean for Gaussian Data with Known Variance.

(3) Interval Estimation using Likelihood Intervals

(4) Interpretation of Likelihood Intervals

(5) Effect of Sample Size on Width of Likelihood Interval

To Do

Read Sections 4.3 - 4.4.

Start End-of-Chapter Problems 1-30.

Today's Lecture

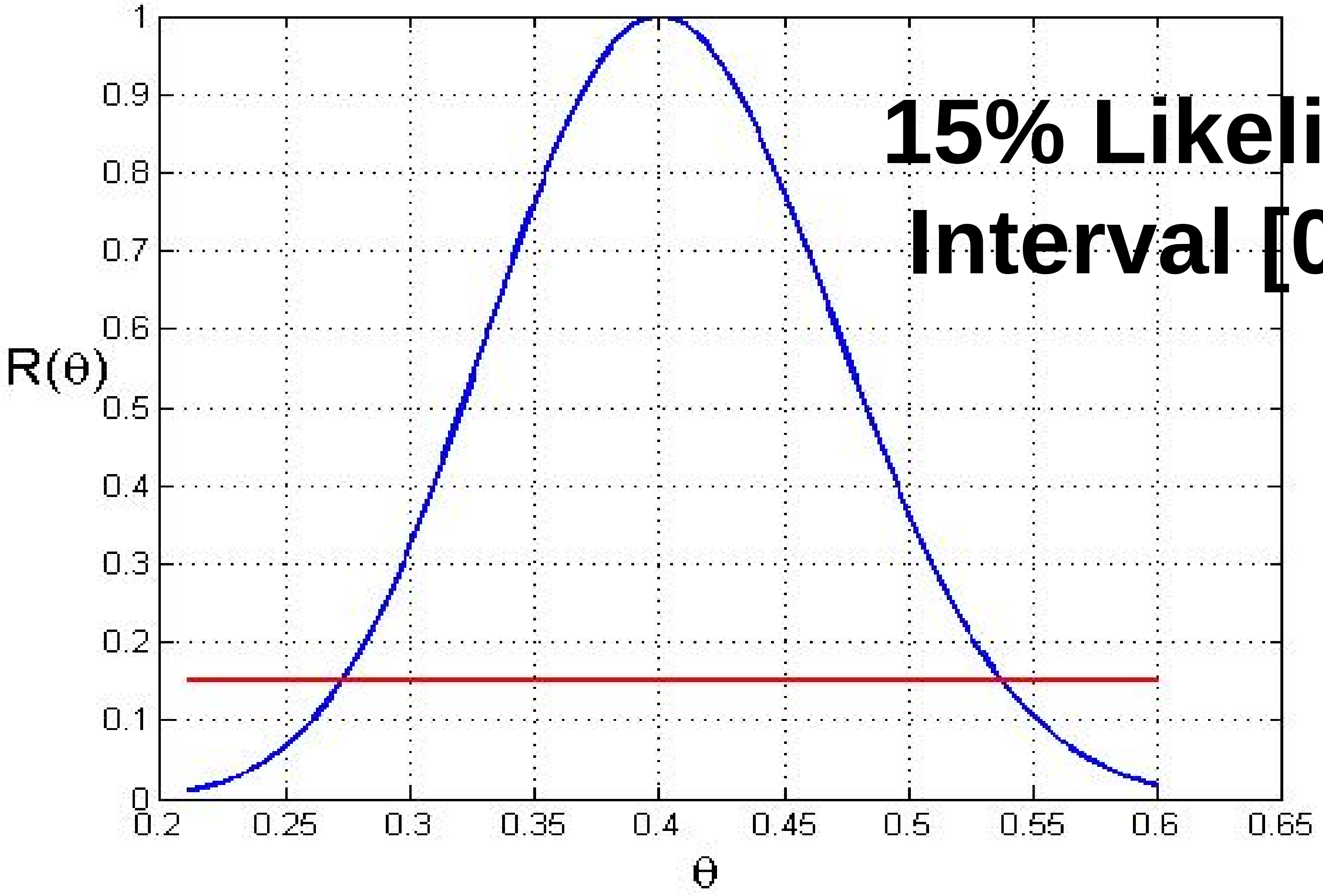
(1) Interval Estimation Using the Relative Likelihood Function (Review of Last Day)

(2) Interval Estimation Using the Log Relative Likelihood Function

(3) Definition of Coverage Probability

(4) Definition of a Confidence Interval

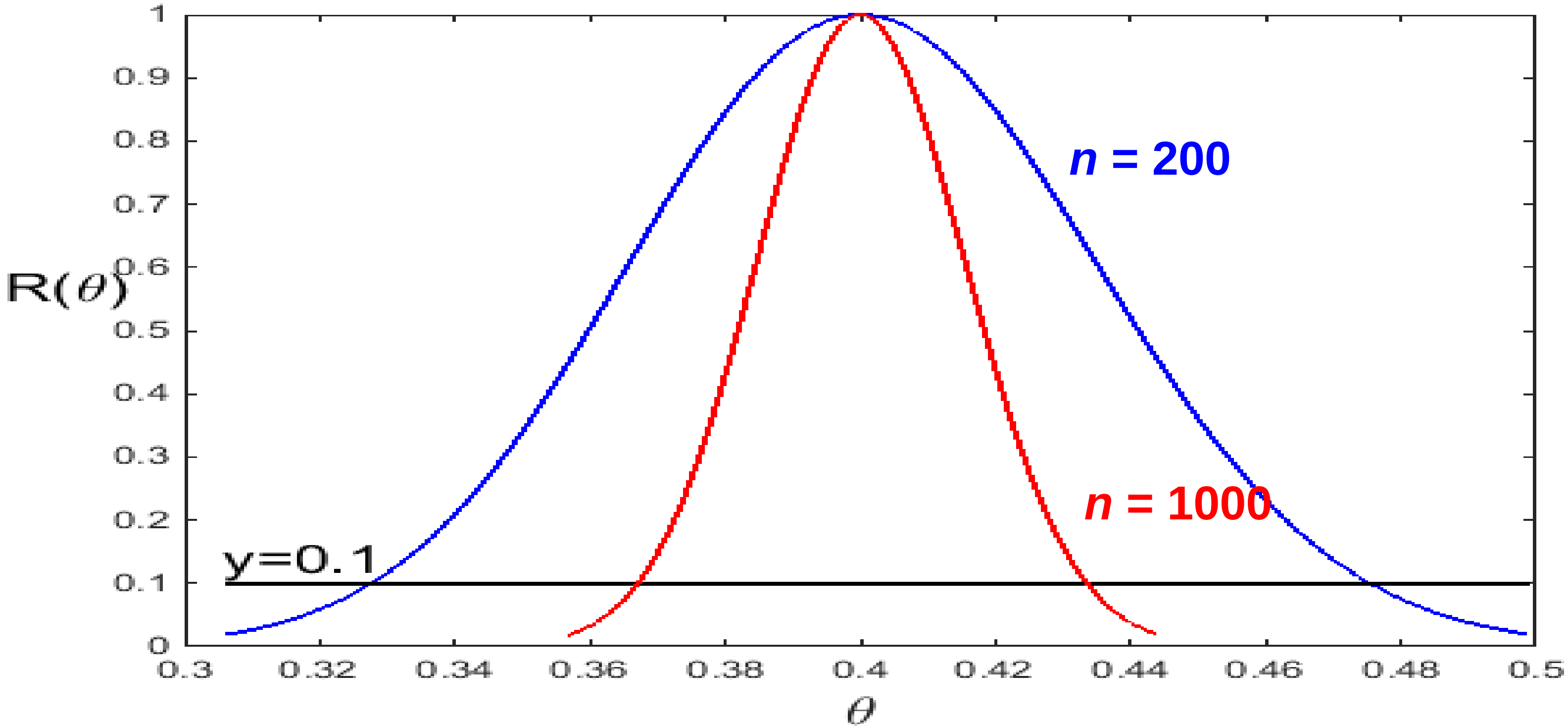
Binomial R.L. for n = 50 and y = 20 with 15% Likelihood Interval



15% Likelihood Interval [0.27,0.54].

Data set 1 : $n = 200$, $y = 80$

Data set 2 : $n = 1000$, $y = 400$



Data set 1 : $R(\theta) \geq 0.1$ for $\theta \in [0.33, 0.47]$

Data set 2 : $R(\theta) \geq 0.1$ for $\theta \in [0.37, 0.43]$

Sample Size and Likelihood Intervals

As the sample size n increases the graph of the relative likelihood function $R(\theta)$ becomes more "concentrated" around θ.

Consequently likelihood intervals become narrower as the sample size increases.

Both these statements reflect the fact that larger data sets contain more information about the unknown parameter θ.

Log Relative Likelihood Function

Definition:

The log relative likelihood function is given by

$$r(\theta) = \log[R(\theta)] = l(\theta) - l(\hat{\theta}) \quad \text{for } \theta \in \Omega$$

$$\text{where } l(\theta) = \log[L(\theta)].$$

***r*(*θ*) is often easier to compute.**

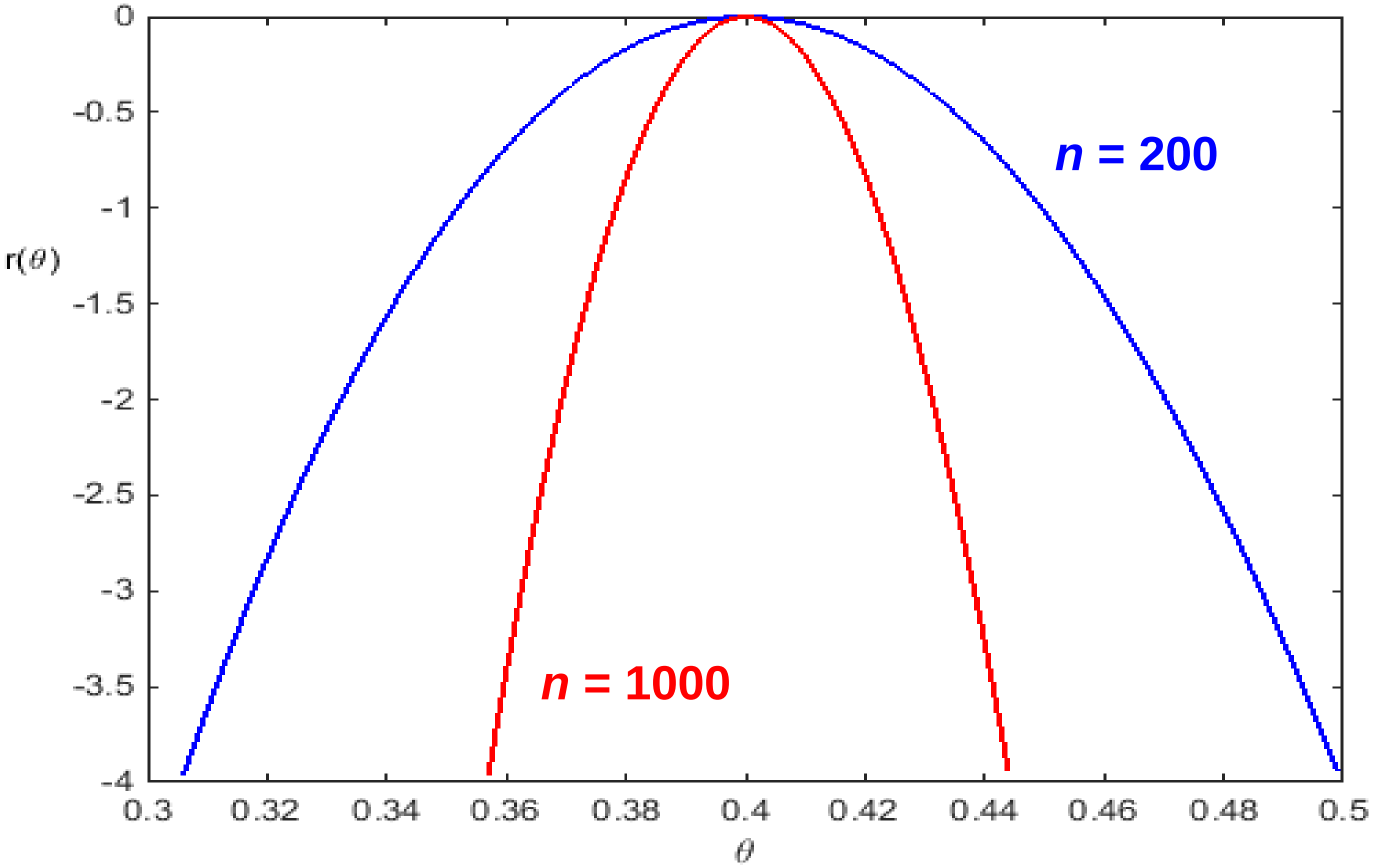
Log Relative Likelihood Function

If $R(\theta)$ is unimodal then $r(\theta)$ is unimodal and both graphs attain their maximum value at the maximum likelihood estimate of θ.

$R(\theta)$ and $r(\theta)$ differ in terms of their shape.

$R(\theta)$ often looks bell-shaped while the graph of $r(\theta)$ resembles a quadratic function of θ.

Binomial Log Relative Likelihood Function n = 200 versus n = 1000



Binomial Log Relative LF
n = 200 versus n = 1000

We see again that as the sample size n increases the graph of the log relative likelihood function $r(\theta)$ becomes more "concentrated" around $\hat{\theta}$.

Binomial Log Relative LF
n = 200 versus n = 1000

Since

$$R(\theta) \geq \text{p if and only if } r(\theta) \geq \log p$$

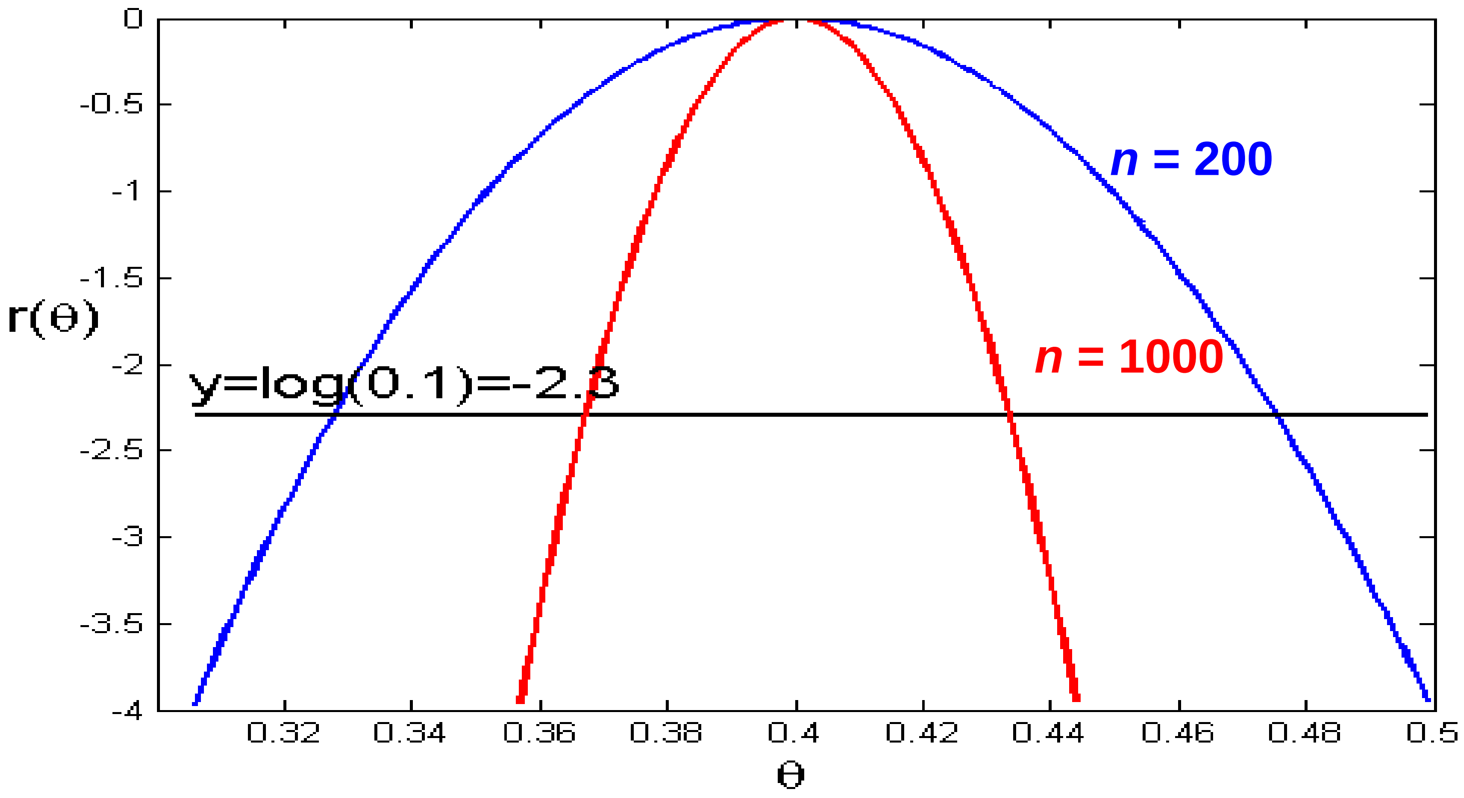
a 100p% likelihood interval for θ can also be easily obtained from the log relative likelihood function by plotting $r(\theta) = \log R(\theta)$ and drawing a line at $r(\theta) = \log(p)$.

Binomial Log Relative LF
n = 200 versus n = 1000

For example, for a 10% likelihood interval we would draw the line

$$y = \log(0.1) = -2.30.$$

Binomial Log Relative Likelihood Function $n = 200$ versus $n = 1000$



Data set 1 : $r(\theta) \geq \log(0.1) = -2.30$ for $\theta \in [0.33, 0.47]$

Data set 2 : $r(\theta) \geq \log(0.1) = -2.30$ for $\theta \in [0.37, 0.43]$

Interval Estimation

A 100p% likelihood interval for an unknown parameter θ gives us an interval of values for θ that are reasonable or plausible given the observed data.

How often does a 100p% likelihood interval contain the true but unknown value of θ?

To answer this question we need to talk about coverage probabilities.

Interval Estimation

Suppose we plan to collect data y or $y = (y_1, y_2, \ldots, y_n)$ in order to estimate an unknown parameter θ in a statistical model.

Let the random variable $Y = (Y_1, Y_2, \ldots, Y_n)$ represent the potential data that are to be collected.

Suppose [L(Y), U(Y)] is an interval estimator (a rule) which can be used to construct an interval of plausible values for the unknown parameter θ.

When we observe data y we construct the interval estimate [L(y), U(y)] for θ based on the observed data.

Definition of Coverage Probability

To determine how "good" our interval estimator or rule is we look at the coverage probability.

Definition

The value

$$P(\theta \in [L(Y), U(Y)]) = P[L(Y) \leq \theta \leq U(Y)]$$

is called the coverage probability for the interval estimator $[L(Y), U(Y)]$.

Interval Estimators and Coverage Probabilities

The coverage probability is the probability that the random interval [L(*Y*), U(*Y*)] contains the true (unknown) value of θ.

Remember that L(*Y*) and U(*Y*) are both random variables.

Interval Estimators and Coverage Probabilities

We try to choose interval estimators (rules) [L(Y), U(Y)] for which

(1) the coverage probability is large
(values 0.90, 0.95 and 0.99 are often used)

and

(2) the interval is as narrow as possible.

These 2 criteria are incompatible so we usually fix the coverage probability at some value and then try to find the narrowest interval.

Definition of a Confidence Interval

Definition

A 100*p*% confidence interval for a parameter is an interval estimate $[L(y), U(y)]$ for which

$$P(\theta \in [L(Y), U(Y)]) = P[L(Y) \le \theta \le U(Y)] = p.$$

The value *p* is called the confidence coefficient for the confidence interval.

Interpretation of a Confidence Interval – Be Careful!

The value θ is an unknown constant associated with the population. It is NOT a random variable and therefore does not have a distribution.

For an observed set of data, *y*, L(*y*) and U(*y*) are all numerical values.

It is not valid to say that the probability that θ lies in the interval [L(*y*), U(*y*)] is equal to *p* since θ is a constant. (This is the most common incorrect interpretation of a confidence interval.)

Interpretation of a Confidence Interval – Be Careful!

To give meaning to the interval [L(*y*), U(*y*)] which is calculated for ONE data set *y* we need to put the idea of calculating such an interval into the larger context of what happens if we repeat the experiment of interest over and over.

Interpretation of a Confidence Interval

Suppose

$$0.95 = P(\theta \in [L(Y), U(Y)]) = P[L(Y) \le \theta \le U(Y)]. \quad (1)$$

Suppose also that we draw repeated independent random samples from the same population and each time we construct the interval $[L(y), U(y)]$ based on the observed data y.

Then (1) tells us that we should expect 95% of these constructed intervals to contain the true but unknown value of θ.

We also expect that 5% of these constructed intervals will not contain the true but unknown value of θ.

Interpretation of a Confidence Interval

Of course we usually only construct one such interval based on one set of data.

We hope that we are one of the lucky 95%!

For a particular sample we say that we are 95% confident that the true value of θ is contained in the interval we have constructed.

Confidence Interval for the Mean μ of a Gaussian Population, Known Standard Deviation σ

Suppose $Y_1, Y_2, \ldots, Y_n$ is a random sample from a G(μ,1) distribution where E(Y_i) = μ is unknown but sd(Y_i) = 1 is known.

For the random interval

$$\left[\bar{Y} - \frac{1.96}{\sqrt{n}}, \bar{Y} + \frac{1.96}{\sqrt{n}} \right]$$

we have

$$P\left(\mu \in \left[\bar{Y} - \frac{1.96}{\sqrt{n}}, \bar{Y} + \frac{1.96}{\sqrt{n}} \right] \right) = 0.95$$

Confidence Interval for the Mean μ of a Gaussian Population, Known σ

The interval

$$\left[\bar{y} - \frac{1.96}{\sqrt{n}}, \bar{y} + \frac{1.96}{\sqrt{n}}\right] \text{ or } \bar{y} \pm \frac{1.96}{\sqrt{n}}$$

is a **95% confidence interval** **for the unknown mean** μ**.**

Note that the width of the confidence interval equals $2\left(\frac{1.96}{\sqrt{n}}\right)$ **which decreases as** n **increases.**

Why does this make sense?

Confidence Interval for the Mean μ of a Gaussian Population, Known σ

Suppose for a particular sample of size n = 16 we observed a sample mean of $\bar{y}$ =10.4.

A 95% confidence interval for μ is

$$\left[10.4-\frac{1.96}{\sqrt{16}},10.4+\frac{1.96}{\sqrt{16}}\right]=[9.91,\ 10.89].$$

We CANNOT say P($\mu \in$ [9.91, 10.89]) = 0.95.

We can only say we are 95% confident that the interval [9.91, 10.89] contains the true but unknown value of μ.