To Do

Read Sections 4.6 - 4.7.

Do End-of-Chapter Problems 1-17 in preparation for Tutorial Test 2.

Today's Lecture

(1) The Likelihood Ratio Statistic and its Asymptotic Distribution

(2) Likelihood Intervals are Approximate Confidence Intervals

(3) Comparison of Likelihood Intervals and Approximate Confidence Intervals

(4) Confidence Interval for Gaussian mean μ when standard deviation σ is unknown.

Likelihood Intervals and Confidence Intervals

It turns out that a likelihood interval is an approximate confidence interval.

To show this we need the Chi-squared distribution ($\chi^2(1)$) with parameter *k* = 1.

Relationship Between Chi-squared(1) and G(0,1)

If $Z \sim G(0,1)$ then $W = Z^2 \sim \chi^2(1)$.

If $W \sim \chi^2(1)$ then

$$P(W \leq c) = 2P(Z \leq \sqrt{c}) - 1$$

and

$$P(W > c) = 2P\left(Z > \sqrt{c}\right) = 2[1 - P(Z \leq \sqrt{c})]$$

Likelihood Ratio Statistic

Let

$$\Lambda = -2\log\left[\frac{L(\theta)}{L(\tilde{\theta})}\right] = -2\log\left[\frac{L(\theta;Y)}{L(\tilde{\theta};Y)}\right]$$

where $\tilde{\theta} = \tilde{\theta}(Y)$ **is the maximum likelihood estimator of *θ*.**

Λ **is a random variable depending on the data *Y*.**

Λ **is called the likelihood ratio statistic.**

Approximate Distribution of the Likelihood Ratio Statistic

For large *n* it can be shown that Λ has approximately a $\chi^2(1)$ distribution.

This implies that Λ is an approximate pivotal quantity that can be used to obtain confidence intervals for θ.

Likelihood Based Confidence Interval

Find *c* such that

$$p = P\left(W \le c\right) = 2[1 - P(Z \le \sqrt{c}\,)]$$

where *W* ~ χ^2(1) and Z ~ G(0,1).

Then since

$$p = P(W \le c) \approx P\left\{-2\log\left[\frac{L(\theta)}{L(\tilde{\theta})}\right] \le c\right\}$$

an approximate 100p% confidence interval for *θ* is

$$\left\{\theta : -2\log\left[\frac{L(\theta)}{L(\hat{\theta})}\right] \le c\right\} = \{\theta : -2\log R(\theta) \le c\}$$

Likelihood Based Confidence Interval

But $$\{\theta : -2\log R(\theta) \le c\} = \left\{\theta : R(\theta) \ge e^{-c/2}\right\}$$

is just a likelihood interval.

For *c* = (1.96)^2

$$P\left(W \le (1.96)^2\right) = P(|Z| \le 1.96)] = 0.95$$

and $$\left\{\theta : R(\theta) \ge e^{-(1.96)^2/2}\right\} = \{\theta : R(\theta) \ge 0.147\}$$

A 14.7% or 15% likelihood interval is an approximate 95% confidence interval.

Example

What is the confidence coefficient of a 10% likelihood interval?

Approximate Confidence Intervals for Binomial

For data *y* from a Binomial(*n*,*θ*) distribution we have 2 methods for obtaining approximate 95% confidence intervals:

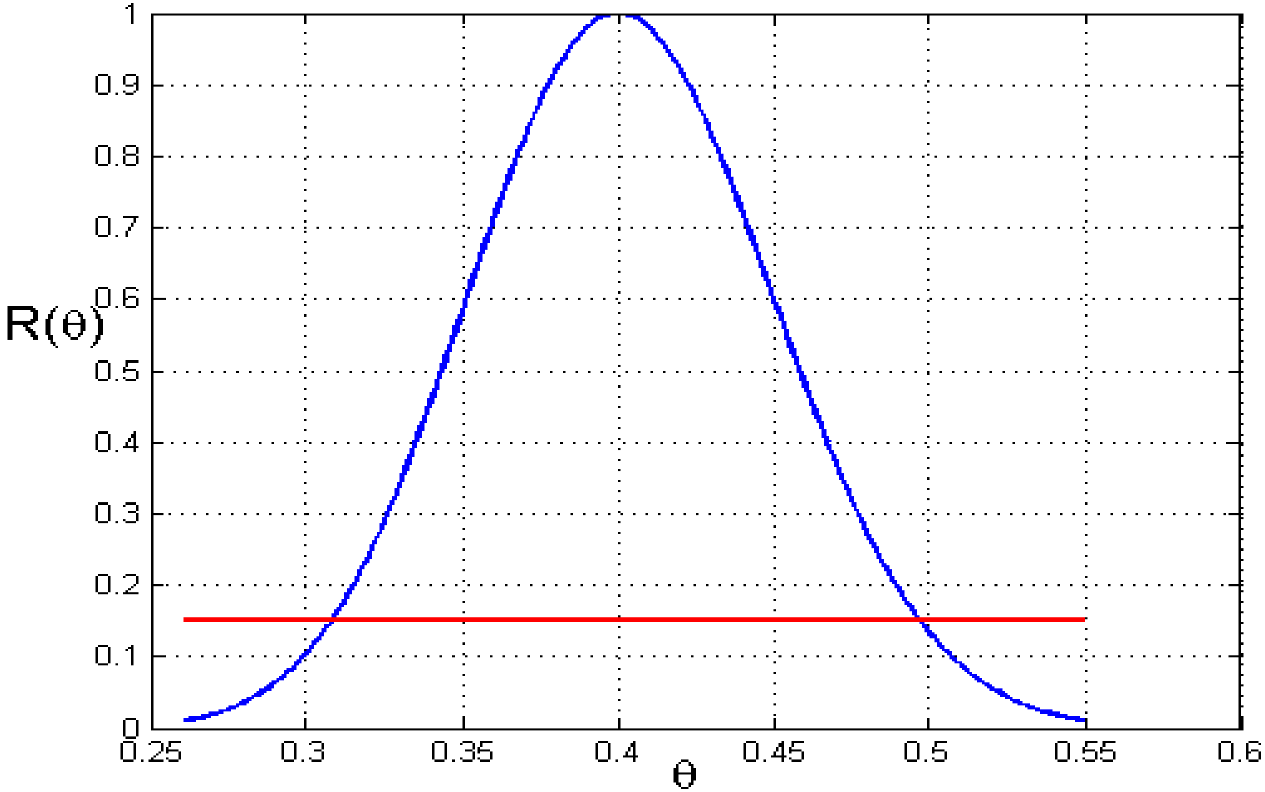
(1) a 15% likelihood interval

and

(2)

$$\hat{\theta} \pm 1.96\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}} \quad \text{where } \hat{\theta} = \frac{y}{n}$$

Example: n = 100, y = 40



15% likelihood interval: [0.31,0.50]

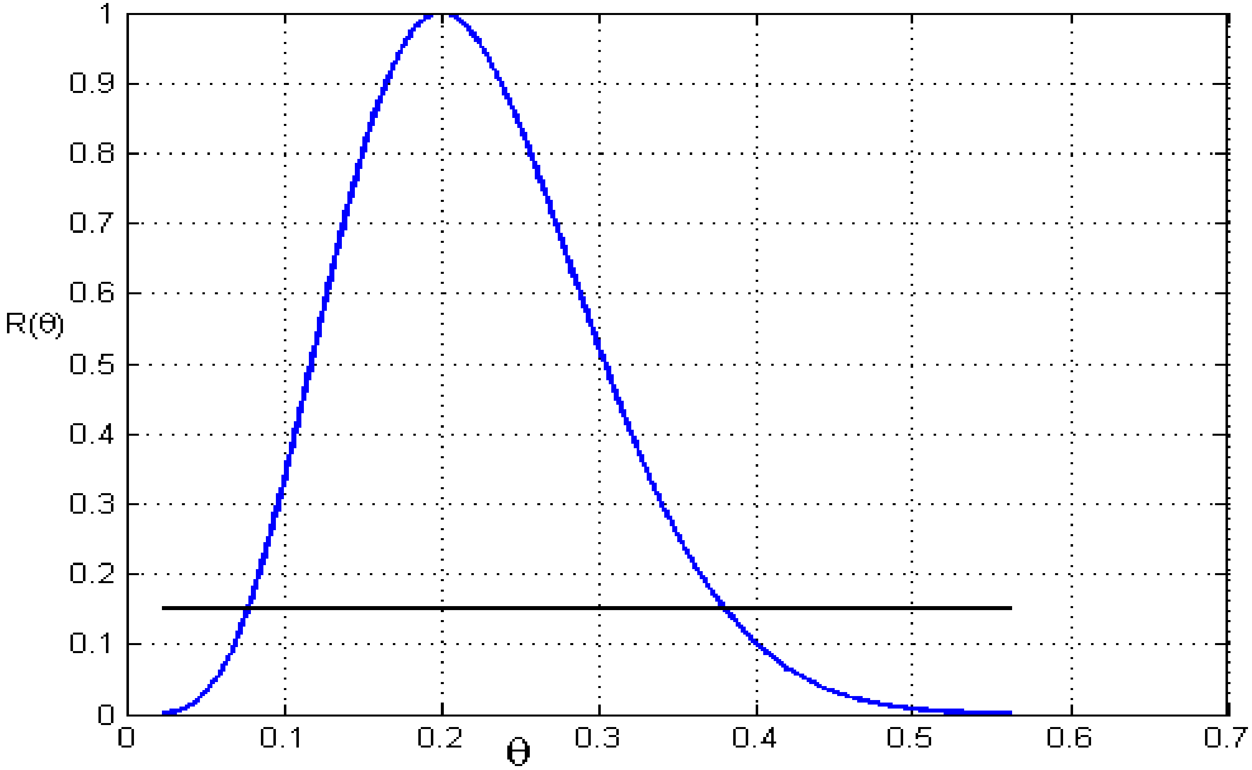
Example: n = 100, y = 40

Compare the 15% likelihood interval [0.31,0.50] with

$$\hat{\theta} \pm 1.96\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}} = 0.4 \pm 0.096 = [0.31, 0.50]$$

The two intervals are based on different approximations but they are the same to 2 decimal places.

Example: *n* = 25, *y* = 5



15% likelihood interval: [0.08,0.38]

Example: n = 100, y = 40

Compare the 15% likelihood interval [0.08,0.38] with

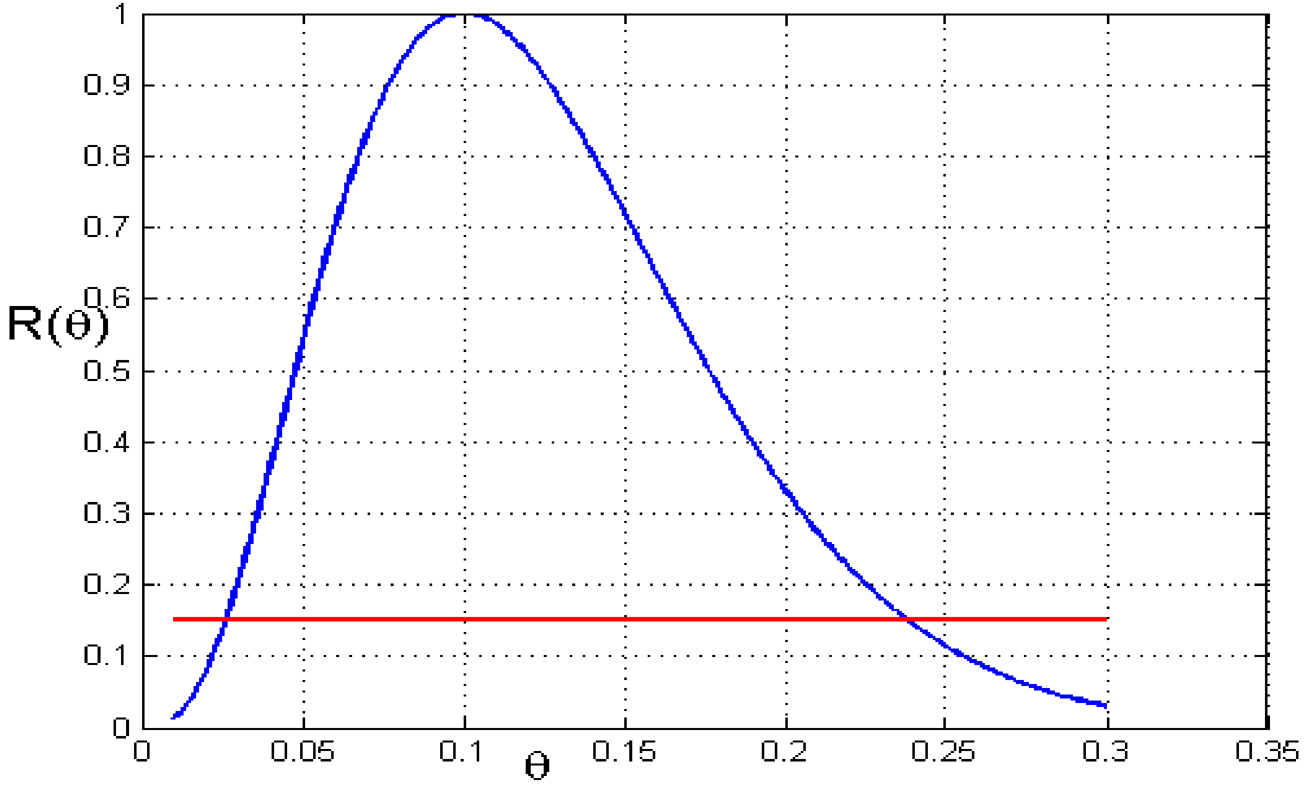
$$\hat{\theta} \pm 1.96\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}} = 0.2 \pm 0.157 = [0.04, 0.36]$$

The two confidence intervals are not as similar as the previous example. Why not?

Which interval gives a better summary of the values of θ which are reasonable given the observed data?

Which interval do you think is usually used? Why?

Example: *n* = 30, *y* = 3



15% likelihood interval: [0.03,0.24]

Example: n = 100, y = 40

Compare the 15% likelihood interval [0.08,0.38] with

$$\hat{\theta} \pm 1.96\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}} = 0.1 \pm 0.11 = [-0.01, 0.21]$$

The two confidence intervals are not very similar.

Do you notice anything unusual?

Which interval is better?

Exercise:

Suppose $y_1, y_2, \ldots, y_n$ is an observed random sample from a Poisson(θ) distribution.

A 95% confidence interval for θ is given by

(1) a 15% likelihood interval

and

(2) $$\hat{\theta} \pm 1.96\sqrt{\frac{\hat{\theta}}{n}} \quad \text{where } \hat{\theta} = \bar{y}$$

Exercise: Compare there two intervals for (i) n = 30 and $\bar{y}$ = 2 and (ii) n = 30 and $\bar{y}$ = 7.

Gaussian data with unknown mean μ and unknown standard deviation σ

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

A point estimator for μ is $\tilde{\mu} = \bar{Y}$ (the maximum likelihood estimator).

Point Estimator for σ^2

A point estimator for σ^2 is

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

(not the maximum likelihood estimate).

We prefer S^2 because $E(S^2) = \sigma^2$.
See Course Notes page 132.

RECALL: Confidence Interval for μ, when σ is known

If σ is known then a 100p% confidence for μ is

$$\bar{y} \pm a \frac{\sigma}{\sqrt{n}}$$

where P(-a ≤ Z ≤ a) = p and Z ~ G(0,1) or equivalently P(Z ≤ a) = (1+p)/2.

This interval was constructed using the pivotal quantity

$$\frac{\bar{Y} - \mu}{\sigma / \sqrt{n}} \sim G(0,1)$$

σ is unknown

If σ is unknown then we replace σ by the estimator *S* to obtain the random variable

$$\frac{\overline{Y} - \mu}{S / \sqrt{n}}$$

which turns out to also be a pivotal quantity.

This pivotal quantity has a Student *t* distribution – a new distribution.

Student *t* Distribution

Suppose *T* is a random variable with probability density function

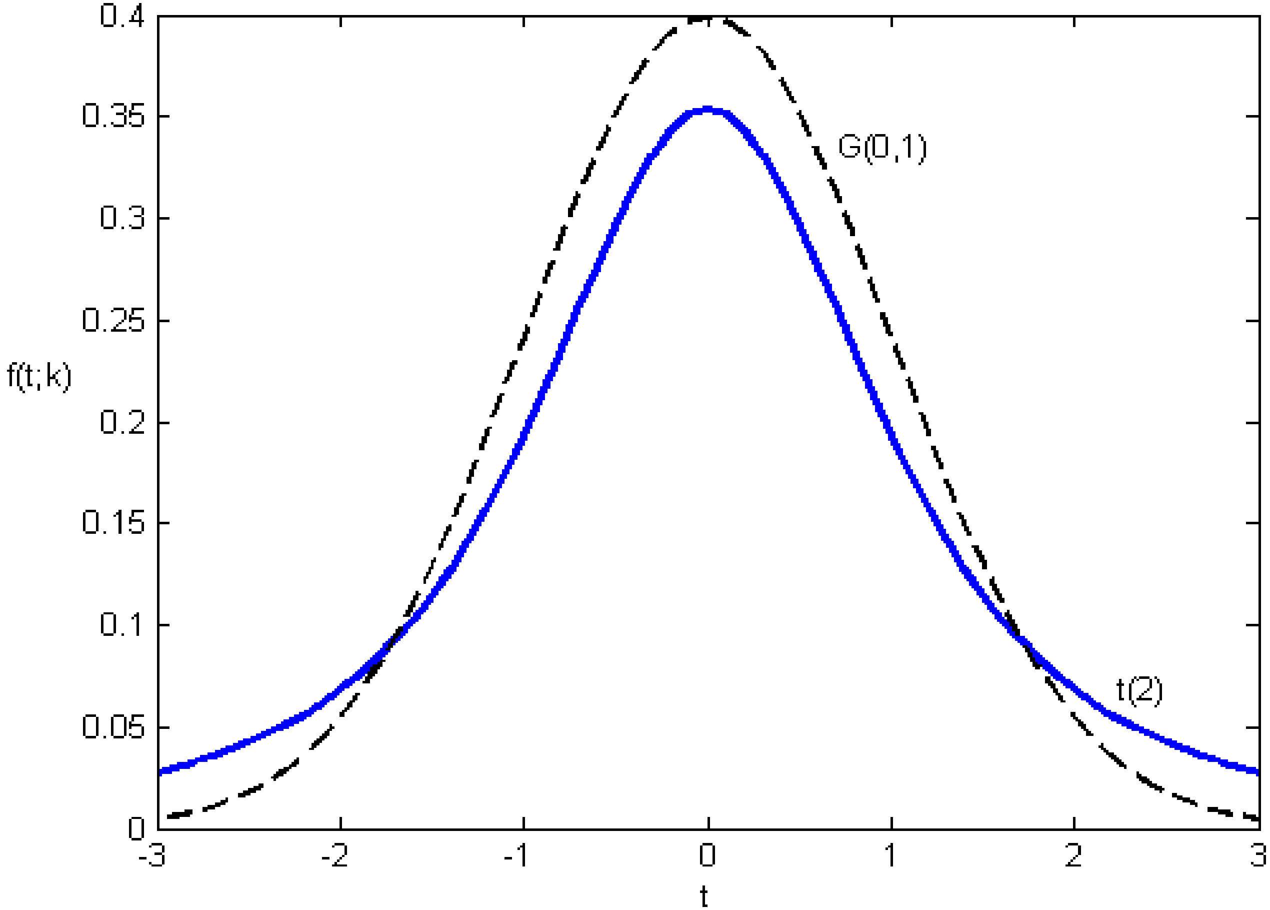
$$f(t;k)=\frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}\,\Gamma\left(\frac{k}{2}\right)}\left(1+\frac{t^2}{k}\right)^{-(k+1)/2},\quad t\in\Re$$

***T* is said to have a Student *t* distribution.**

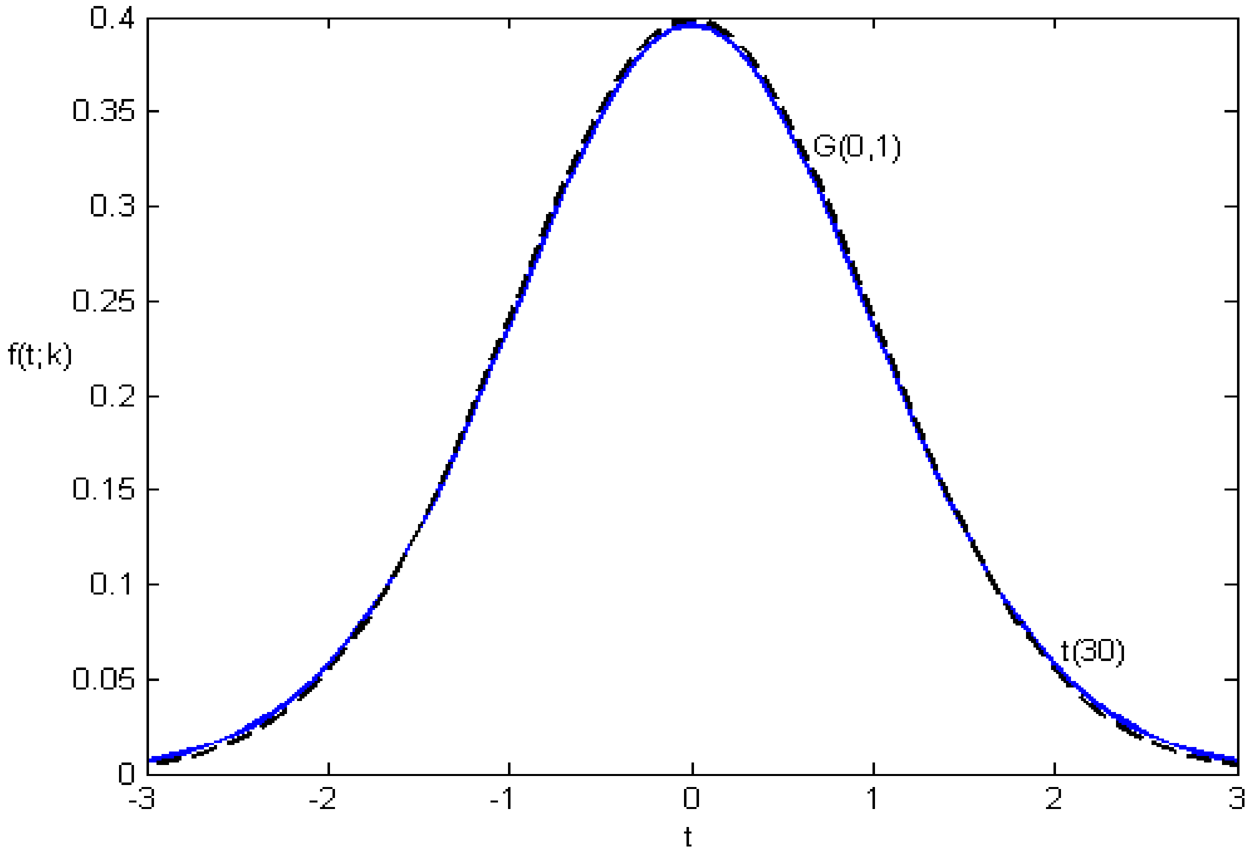
The parameter *k* is called the degrees of freedom.

We write *T* ~ *t*(*k*).

t(2) and G(0,1)



t(30) (blue) and G(0,1) (black)



Properties of the *t* Distribution

The *t* probability density function is similar to that of the G(0,1) distribution since it is unimodal and symmetric about the origin.

For small *k*, the *t* density has larger “tails” or more area in the extreme left and right tails.

For large *k*, the graph of the probability density function f(*t*;*k*) looks like the G(0,1) probability density function.

See Problem 18 at the end of Chapter 4 on reading *t* tables.

Theorem

Suppose $Y_1, Y_2, \ldots, Y_n$ is a random sample from a G(μ,σ) distribution.

Then

$$\frac{\bar{Y} - \mu}{S/\sqrt{n}} \sim t(n-1)$$