STAT 231

November 2, 2016

Roadmap

- Wrap up Interval estimation using Confidence Intervals.
- Relationship between Likelihood Intervals and C.I for large samples
- Hypothesis Testing

Case I : C I for μ; Gaussian (μ, σ).

σ known.

Case II : Binomial problem. $Y \sim Bin(n, \theta)$

C.I for θ

Case III : $Y_1, \ldots Y_n \sim G(\mu, \sigma)$

undep. μ, σ unknown.

C.I for μ ?

PIVOTAL QUANTITY : $\dfrac{\bar{Y} - \mu}{\frac{S}{\sqrt{n}}}$ ✓

PIVOTAL DISTRIBUTION

Student's T distribution:

with $n-1$ df

COVERAGE INTERVAL

$$\bar{Y} \pm t^* \frac{S}{\sqrt{n}}.$$

where t^* = from the t-table

Row = $n-1$

Column = depends on the level of confidence.

CONFIDENCE INTERVAL : $\left[\bar{y} \pm t^* \frac{s}{\sqrt{n}} \right]$

$\bar{y}$: sample mean

$$s : \sqrt{\frac{1}{n-1} \sum (y_i - \bar{y})^2}$$

<u>Case IV</u> $\quad Y_1, \ldots Y_n \sim G(\mu, \sigma)$

μ, σ unknown.

Objective : To find the C.I for σ

PIVOTAL QUANTITY = $\dfrac{(n-1)S^2}{\sigma^2}$

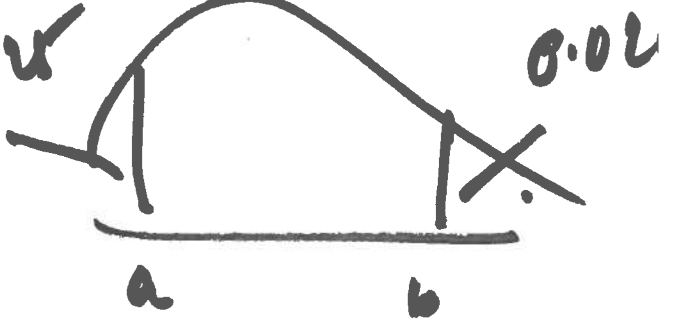
PIVOTAL DISTRIBUTION = χ^2_{n-1}

$df = n-1$

COVERAGE INTERVAL

$$\left[\frac{(n-1)S^2}{b}, \frac{(n-1)S^2}{a}\right].$$

where b, a are computed from the χ^2 table



Confidence Interval for σ^2 = $\left[\frac{(n-1)s^2}{b}, \frac{(n-1)s^2}{a}\right]$

C.I for σ : $\left[\sqrt{\frac{(n-1)s^2}{b}}, \sqrt{\frac{(n-1)s^2}{a}}\right]$

Equal-tailed Confidence Interval

but that need not be the case.

Case V Other distributions with large sample sizes

(a) **Poisson** : $Y_1, \ldots Y_n \sim Poi(\psi)$

P.Q : $\dfrac{\bar{Y} - \psi}{\sqrt{\bar{Y}/n}}$

PIVOTAL DISTRIBUTION = Z

Coverage : $\bar{Y} \pm z^{\alpha} \sqrt{\bar{Y}/n}$.

C.I = $\boxed{\bar{y} \pm z^{\alpha} \sqrt{\bar{y}/n}}$.

(b) <u>Exponential distribution</u>

$Y_1, \ldots Y_n \sim Exp(\mu)$

Objective : To find the C.I for μ.

(The following result is true for all n)

<u>Result</u>: If $Y \sim Exp(\mu)$, then

$$\frac{2Y}{\mu} \sim Exp(2) \quad \text{——} \ (1)$$

$$= \chi^2(2)$$

From (1),

$$\underline{\underline{\sum_{i=1}^{n} \frac{2Y_i}{\mu}}} \sim \underline{\underline{\chi^2_{2n}}}$$

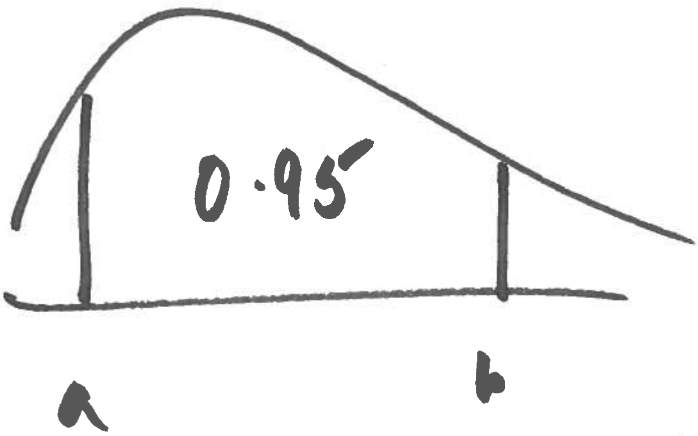
PIVOTAL QUANTITY

PIVOTAL DISTRIBUTIO

Example: $n = 10$

Go to the χ^2_{20} table and find the two end points a and b

95% C.I



$$P\left(a \le \chi^2_{20} \le b\right) = 0.95$$

$$P\left(a \le \frac{2\sum Y_i}{\mu} \le b\right) = 0.95$$

$$\mu \le \frac{2\sum Y_i}{a} \qquad \mu \geq \frac{2\sum Y_i}{b}$$

Coverage Interval is

$$\left(\frac{2\sum Y_i}{b} , \frac{2\sum Y_i}{a} \right)$$

Confidence Interval

$$\left(\frac{2\sum y_i}{b} , \frac{2\sum y_i}{a} \right)$$

$$\left(\frac{2\bar{y}n}{b} , \frac{2n\bar{y}}{a} \right)$$

There are two approaches to interval estimation

- Through likelihood function.
- Through sampling distributions.

~~Re~~

<u>Question</u>: How are these two intervals related?

For large samples, we call the following theorem.

Theorem: If $L(\theta)$ is based on $(Y_1, \ldots, Y_n)$

θ = unknown parameter, then

$$\cancel{\Lambda(\theta) = \frac{L(\theta)}{L(\tilde{\theta})} \sim \chi^2_1}$$

if n is large.

$\tilde{\theta}$ = Estimator corresponds

$$\boxed{\Lambda(\theta) = -2\log\frac{L(\theta)}{L(\tilde{\theta})} \sim \chi^2_1}$$

$\Lambda(\theta)$ = LIKELIHOOD RATIO TEST STATISTIC.

Theorem: $p = 0.1$

a $100p\%$ likelihood interval is an approximate $100q\%$ C.I.

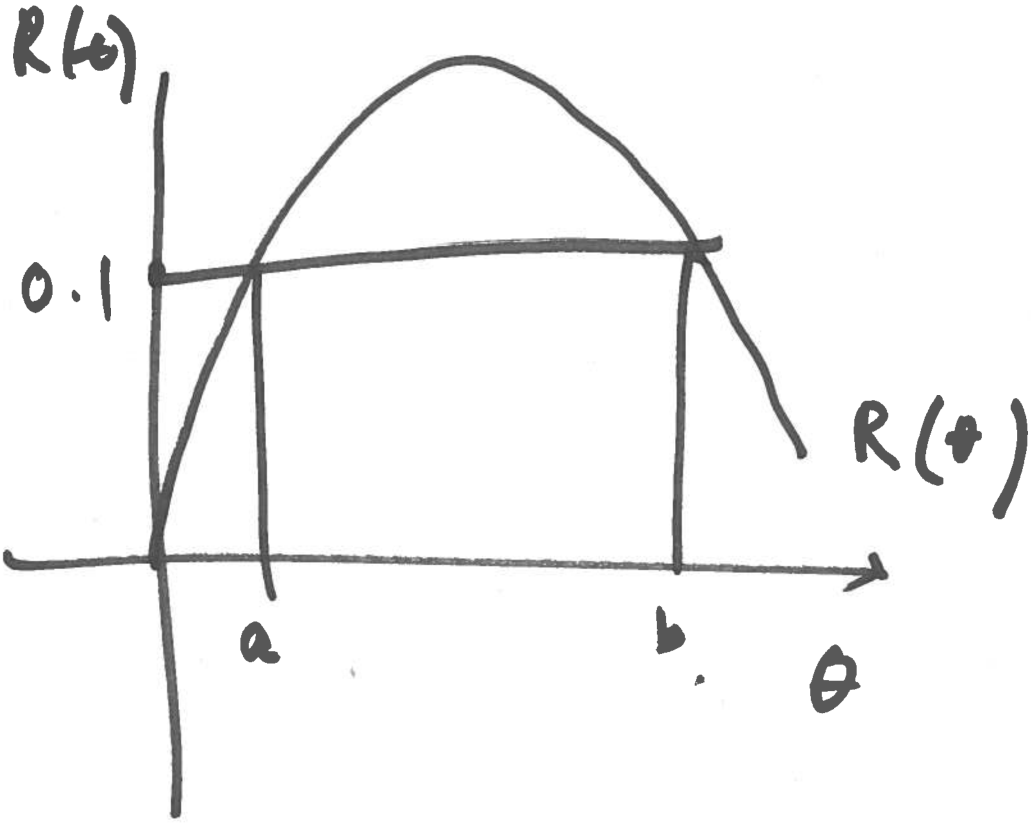
where $q = 2P(Z \leq \sqrt{-2\log p}) - 1$

Application

10% Likelihood Interval.

$2P(Z \leq \sqrt{-2\log 0.1}) - 1$

$= 96.8\%$



10%

Outline of the proof

100 p% l.i

$$= \{\theta : R(\theta) \geq p\}$$

$$= \{\theta : -2\log R(\theta) \leq -2\log p\}$$

$$\underbrace{-2\log \frac{L(\theta)}{L(\hat{\theta})}}$$

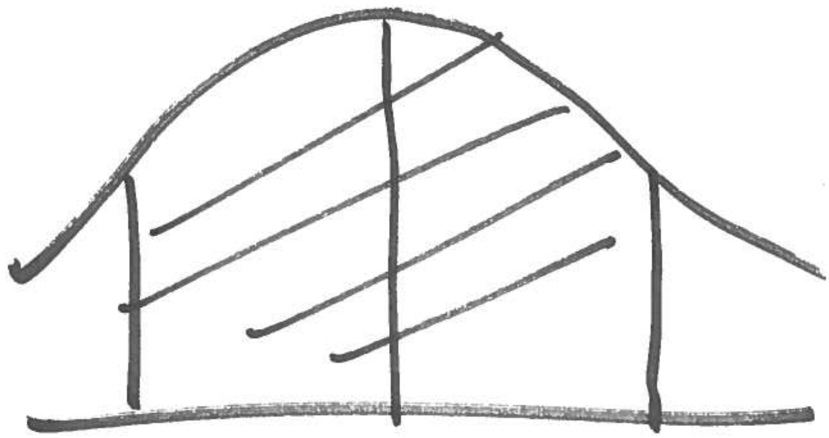
Corresponding coverage probability

$$P\left(\underbrace{-2\log \frac{L(\theta)}{L(\hat{\theta})}} \le -2\log p \right)$$

$$P\left(\Lambda(\theta) \le -2\log p \right)$$

$$P\left(Z^2 \le -2\log p \right)$$

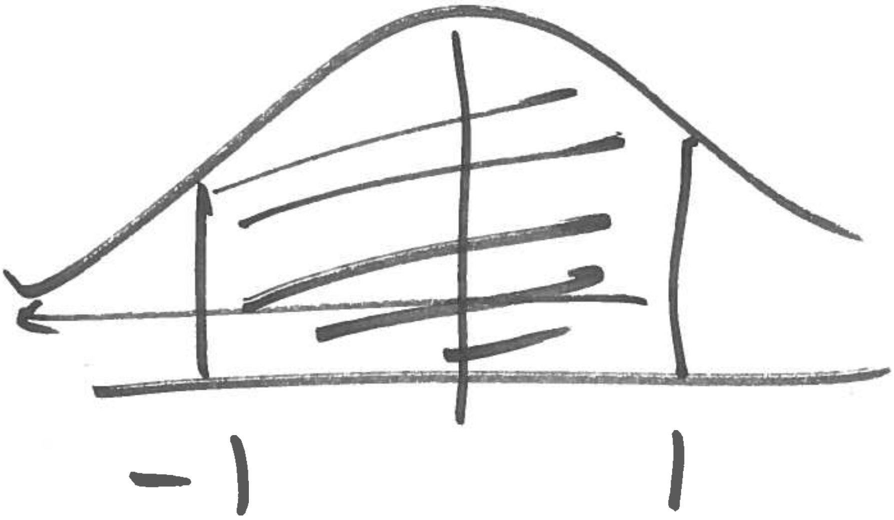
$$= P\left(|Z| \le \sqrt{-2\log p} \right)$$



$-\sqrt{-2\log p}$ $\sqrt{-2\log p}$

$$P(Z \le \sqrt{-2\log p}) - (1 - P(Z \le \sqrt{-2\log p}))$$

$$\boxed{2P(Z \le \sqrt{-2\log p}) - 1}$$



$$P(Z \leq 1) - \left(P(Z \leq -1)\right)$$

$$\therefore P(Z \geq 1)$$

$$= 1 - P(Z \leq 1)$$

$p = 0.1 \Rightarrow$ Confidence Interval
96.8%

$p = 0.01 \Rightarrow$

Example: Suppose we have a 95% C.I. What likelihood interval would this correspond to?

$$\left\{ \theta : R(\theta) \geq \underbrace{e^{-z^{*2}/2}}_{p} \right\}$$

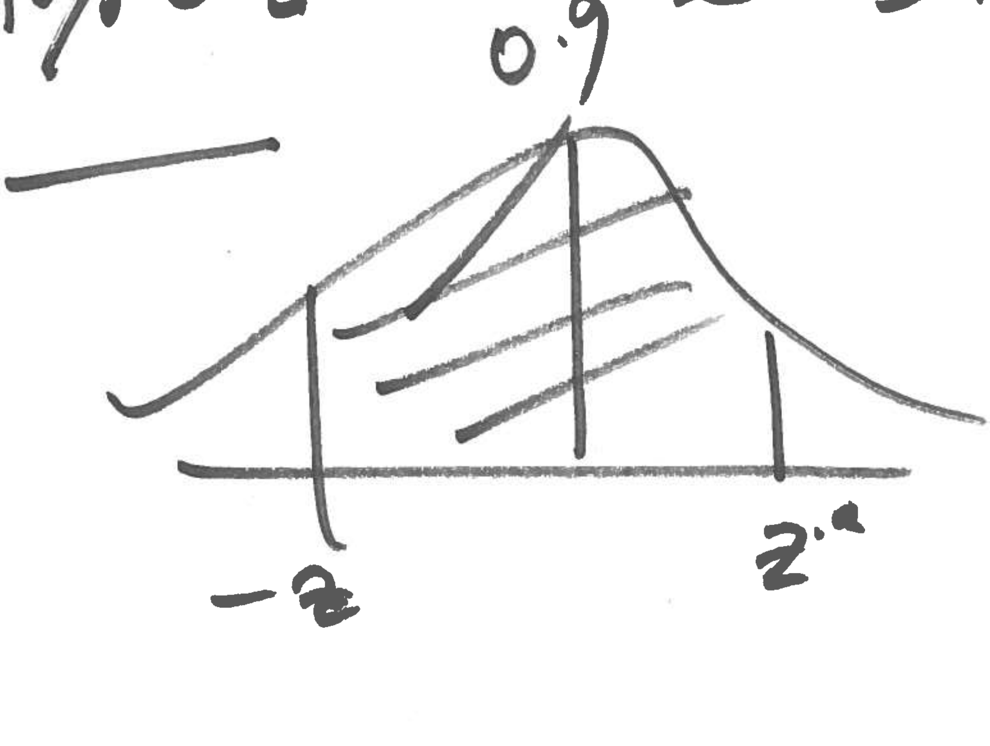
where z^{*} = value from the Z-table

For a 95% C.I, the p would be

$$e^{-1.96^2/2} = 14.65\%$$

For large samples, a 95% C.I is equivalent to a 15% l.i approximately

90% C.I $\quad z^{\alpha} = 1.64$



95% C.I

$$P(-1.96 \leq Z \leq 1.96) = 0.95$$

$$\Rightarrow P(Z^2 \leq 1.96^2) = 0.95$$

$$P(\chi^2_1 \leq 1.96^2) = 0.95$$

$$P(\Delta(\theta) \leq 1.96^2) = 0.95$$

$$P\left(-2\log\frac{L(\theta)}{L(\hat{\theta})} \leq 1.96^2\right)$$

$$P\left(\frac{L(\theta)}{L(\hat{\theta})} \geq e^{-1.96^2/2}\right)$$

$$R(\theta) \geq p$$

C.I

90% → 26% l.c