CI IN STATISTICS 11/2/2024

11 February 2024 12:11

Advanced Statistical Analysis

Zone

Smaller Data set

Larger Data Set

How to choose ??

Poisson → Parameter

$e^{-\lambda}$

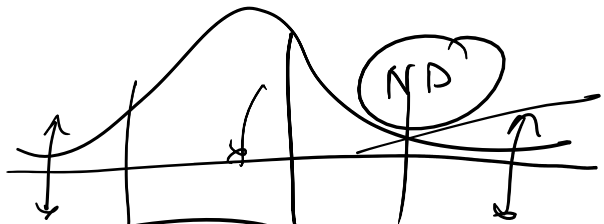
λ → Parameter

Parameter | Known facts | Errors

95% 5%

99% 1%

100% ??



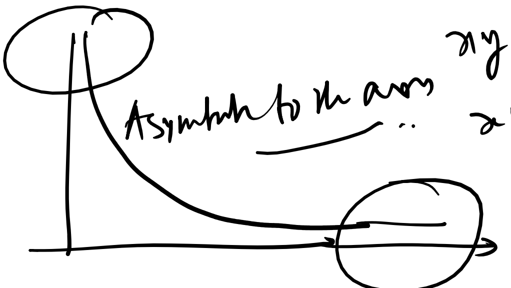
Asymptotic

$y = \frac{1}{x}$

$xy = \text{Const}$

$xy = 7063.21$

Asymptote to the axis ..



Econometrics

$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots$

3 ✓

Case I $y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

n ✓

Case II $y = \alpha_2 + \beta_1' x_1' + \beta_2 x_2' + \dots + \beta_n' x_n'$ $\boxed{n > 3}$

Number of independent variables may not be the deciders.

Number of Independent variables

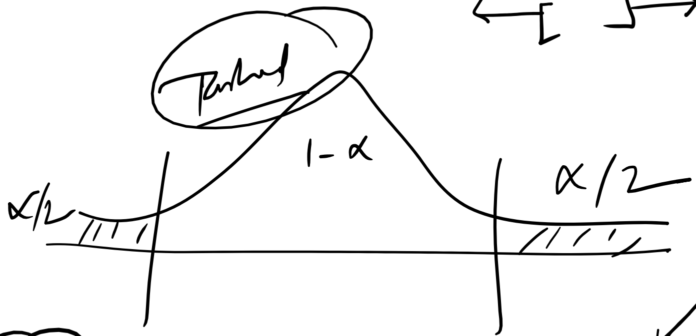
PRJ | Professor | Govt Employee | Unemployed

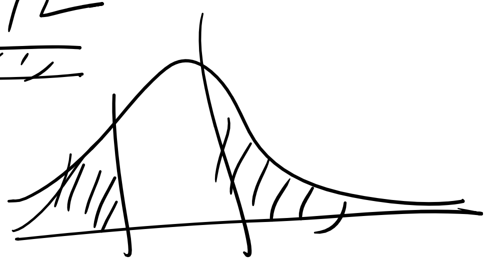
Time x

Already ✓

Current value ← On What Basis??

4 yr | 1 | 18th Jan 2023





CLT

ND

$$\mu \in \left(\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}\right)$$

Sample Size → Larger is better X

→ More Relevant is Better ..

#

Boxing →

Person's Strength ??

Weight ?? → Weight dynamics

42
50
60

weight ?? → weight avg...

Arm wrestling — log

50
60
80
100
120
140

Warm up

Q. $x_1, x_2 \dots x_n$ → r.v. x

$N(\mu, 16)$ .. Population

If a 95% CI for μ is $[\bar{X} \pm 0.98]$

$n = ?$

ISI 2012

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

width → ± 0.98 — (i)

width → $\pm Z_{0.025} \frac{\sigma}{\sqrt{n}}$ — (ii)

Expanding (i) & (ii)

$$\pm 0.98 = \pm 1.96 \frac{4}{\sqrt{n}}$$

$$\sqrt{n} = \frac{1.96}{0.98} \times 4 = 8$$

$$n = \boxed{64}$$

Z | t | x | Z^2 → χ^2

$n \geq 30$ | $n \cancel{<} 30$

\# $n = 9$, $\mu = 5$,

observed statistics $\frac{1}{8} \sum X_i^2 = 39.125$ → $\sum x_i^2 = 313$

$\sum X_i = 45$

$$\sum x_i = 45$$

95% CI for σ^2 ?

Ans

$$\sum (x_i-\mu)^2 = \sum x_i^2 - 2\mu \sum x_i + 9\mu^2$$

$$= 313 - 450 + 225$$

$$= 88$$

Left to be given on the CI

95% CI

$$\sigma^2 \in \left[\frac{\sum (x_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}, n-1}}, \frac{\sum (x_i-\mu)^2}{\chi^2_{1-\alpha/2, n-1}} \right]$$

$$\sigma^2 \in \left(\frac{88}{17.53}, \frac{88}{2.180} \right) \ldots$$

$\rightarrow$ 5 to 41

ISS 2018

$x_1, x_2 \ldots x_n \rightarrow$ PDF

$$f(x|\theta) = \begin{cases} \theta e^{-\theta x} & x>0, \theta>0 \\ 0 & \text{otherwise} \end{cases}$$

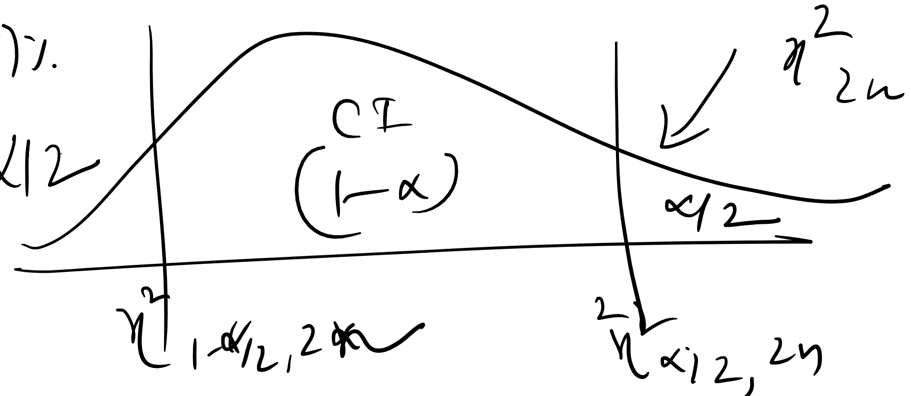
Which of the following is/are $100(1-\alpha)\%$ CI for θ ?

Ans: P.Q. : $-2\sum \ln\left(1-F(x_i,\theta)\right) \sim \chi^2_{2n}$

Here : $-2\sum \ln\left[1-F(x_i,\theta)\right] = 2\theta\sum x_i \sim \chi^2_{2n}$

So, formally, P.Q. : $2\theta \sum_{i=1}^{n} x_i \sim \chi^2_{2n}$

Then CI $(1-\alpha)\%$.



~~Now~~ find

$$(1-\alpha) = P\left[0 < 2\theta \sum x_i < \chi^2_{2n,\alpha}\right]$$

$$= P\left[0 < \theta < \frac{\chi^2_{2n,\alpha}}{2\sum x_i}\right]$$

$$\Rightarrow \theta \in \left[0, \frac{\chi^2_{2n,\alpha}}{2\sum x_i}\right]$$

Case of 2 variables

Question Let. $X_1, X_2 \rightarrow N(0, \theta)$ $\boxed{\theta > 0}$

Then the value of K, for which

$$\left(0, \frac{X_1^2 + X_2^2}{K}\right)$$ is a 95% CI for θ

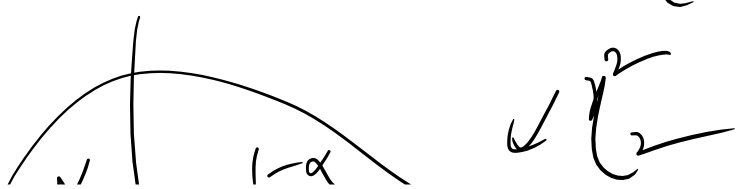
Approach

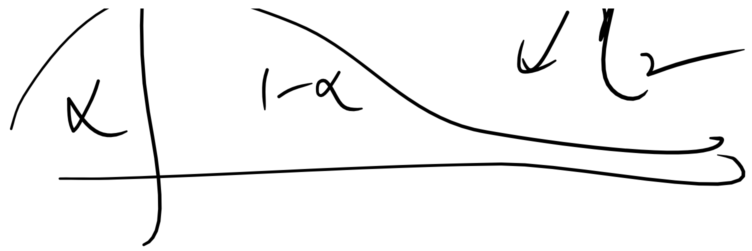
$X_1 \sim N(0, \theta) \Rightarrow \frac{X_1}{\sqrt{\theta}} \sim N(0,1) \Rightarrow \frac{X_1^2}{\theta} \sim \chi^2_1$

X_2 - - - - - - . $\frac{X_2^2}{\theta} \sim \chi^2_1$

Let, $Y = \frac{X_1^2 + X_2^2}{\theta} \sim \chi^2_2$

95% CI $\rightarrow \theta \rightarrow \left(0, \frac{X_1^2 + X_2^2}{\cancel{\theta}\, K}\right)$





Table

$1-\alpha = 95\%$ $\alpha = 5\%$ $\chi^2_{0.95,2} = 0.1026$

So, $0.95 = P(0.1026 < \chi^2 < \infty)$

$= P\left(0.1026 < \frac{X_1^2 + X_2^2}{\theta} < \infty\right)$

$= P\left(\frac{1}{\infty} < \frac{\theta}{X_1^2 + X_2^2} < \frac{1}{0.1026}\right)$

$= P\left(0 < \theta < \frac{X_1^2 + X_2^2}{0.1026}\right)$

$\Rightarrow \theta \in \left(0, \frac{X_1^2 + X_2^2}{0.1026}\right)$

$K = 0.1026$

Self-study

$\cancel{2 \log(0.95)}$

$\cancel{n \log}$

Q Type 5

Approximation Types .. (Normal app. to Binomial Distn)

A poll is taken to UNI students before election.

78 → 33 voted for ... Population 2200

95% CI for the proportion of votes in favour of

95% CI for the proportion of votes in favour of Mr Smith??

* Cambridge Analytica

$\boxed{\text{Ans:}}$ The sample proportion $\hat{p} = \frac{x}{n} = \frac{37}{78} = 0.4231$

$(1-\alpha) = 95\%$

$$p \in \left[\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

$$p \in \left[0.4231 \pm 1.96\sqrt{\frac{0.4231(1-0.4231)}{78}}\right]$$

$$p \in (0.3135,\ 0.5327)$$

Alternative Solution..

$\hat{p} = 37/78 = 0.4231$

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.4231 \times 0.5769}{78}}$$

$= 0.559$

The 2.5th Percentile of N.D is

$z_{0.025} = 1.96$ (table)

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$= 0.4231 - (1.96)(0.559)$

$= 0.4231 - 0.1096$

$= 0.4231 - (1.96)(0.5\dots)$

$= 0.4231 - 0.1096$

$= \underline{0.3135}$

& $\hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ ✓

$= \underline{0.5327}$ ..