Chapter 8 Confidence Interval:

$\bar{x}$, S

An interval estimate a population parameter is Confidence Interval.

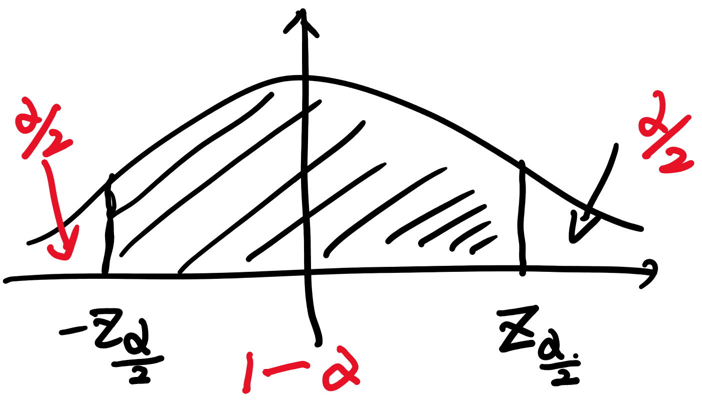
Suppose $X_1, \dots, X_n$ is a random sample from a normal distribution with unknown mean μ and known variance σ^2.

Central Limit Theorem.

$\bar{X}$ is normally distributed with mean μ and variance $\frac{\sigma^2}{n}$.

$$Z_{\frac{\alpha}{2}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

standard normal
mean 0, variance 1.



$0 \le \alpha \le 1$

$$P(-Z_{\frac{\alpha}{2}} \le Z \le Z_{\frac{\alpha}{2}}) = 1 - \alpha.$$

$$p\left(-Z_{\frac{\alpha}{2}} \le Z \le Z_{\frac{\alpha}{2}}\right) = 1-\alpha.$$

Example. $\alpha = 5\%$.

$$Z_{\frac{\alpha}{2}} = Z_{0.025}$$

$$Z_{\frac{\alpha}{2}} = 1.96 \qquad -Z_{\frac{\alpha}{2}} = -1.96$$

$$p(-1.96 \le Z \le 1.96) = 95\%$$

sample mean

population parameter

$$p\left[-Z_{\frac{\alpha}{2}} \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le Z_{\frac{\alpha}{2}}\right] = 1-\alpha.$$

σ known. $\quad$ n: sample size.

$$p\left[-Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \le \bar{X}-\mu \le Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right] = 1-\alpha.$$

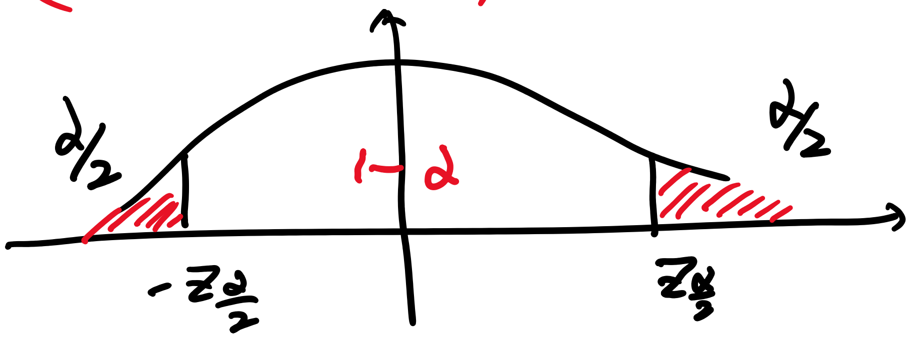
$$p\left[\bar{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right] = 1-\alpha.$$

If $\bar{x}$ is the sample mean of a random sample of size n from a normal population with known variance.

Then, $100(1-\alpha)\%$ C.I. on μ is given by.

$$\bar{x} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

$$\left(\bar{x} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right)$$



Example:

Assume we have heights of students.

$n=10$, $\bar{x} = 164.46$, $\sigma = 1$

Find the 95% CI of the heights:

$100(1-\alpha)\%$. $\alpha = 5\%$

$$\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}.$$

$$164.46 - 1.96 \frac{1}{\sqrt{10}} \le \mu \le 164.46 + 1.96 \cdot \frac{1}{\sqrt{10}}$$

$-z_{\frac{\alpha}{2}}$ $\qquad$ $z_{\frac{\alpha}{2}}$.

Interpreting a Confidence Interval.

Lower $\qquad$ Upper

R.V $\qquad$ R.V.

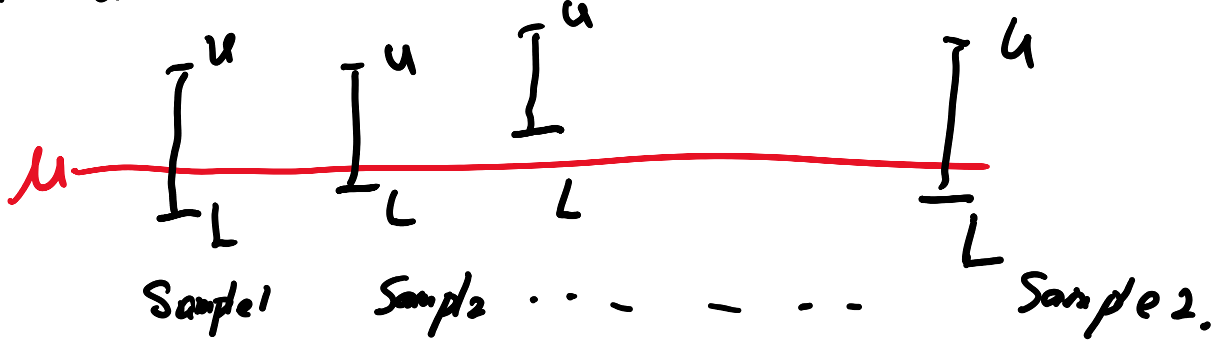
95% confidence interval.

repeat doing sampling n times.

repeat doing sampling n times.

$n \to$ large.

95% of these n C.I will contain the true mean.



We are 95% certain / confident that μ lies in the interval.

Marginal Error:

$$E = Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$ Marginal Error (Error).

the larger the sample size, the less the error is.

$$\bar{x} - E \le \mu \le \bar{x} + E .$$

$\underline{E = 0.01}$

Rearrange.

$$n = \left(\frac{Z_{\frac{\alpha}{2}} \sigma}{E}\right)^2$$

If $\bar{x}$ is used as an estimate of μ, we can be $100(1-\alpha)\%$. confident that we can be $100(1-\alpha)\%$ confident ...

the error $|\bar{x}-\mu|$ will not exceed a specified amount E when the sample size is

$$n = \left(\frac{Z_{\frac{\alpha}{2}}\,\sigma}{E}\right)^2$$

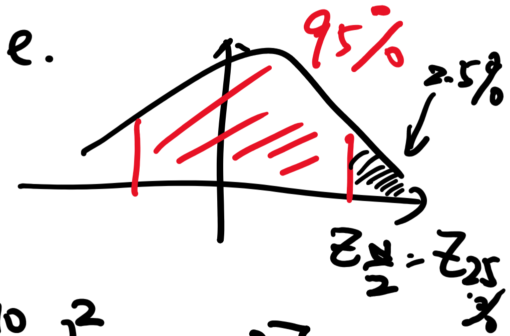
Given any 3, you could find the last one

Example:

knowning that the population variance for # of customes is 100.

Find 95% CI with error of 0.5.

what's the minimum sample size.



$$E \le 0.5$$

$$n = \left(\frac{Z_{\frac{5\%}{2}}\,\sigma}{E}\right)^2 = \left(\frac{1.96 \cdot 10}{0.5}\right)^2 = \lceil 153 \rceil.$$

97.5%

Narrow Confidence Interal:

$$\bar{x} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

$100(1-\alpha)\%$

1. Lower Conifidence Level (Larger α)
2. Lower standard devation.

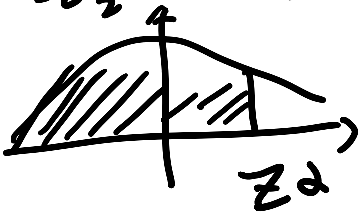
$(1-3,\ 1+3)$ $\quad$ $(1-2,\ 1+2)$

$(-2,\ 4)$ 6 $\quad$ $(-1,\ 3)$ 4

3. Larger sample size n.

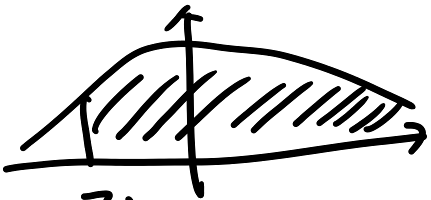
③ Larger sample size n.

two sided C.I.



$(-\infty, U)$

one sided C.I.



$(L, +\infty)$

$L \geq \bar{x} - Z_\alpha \frac{\sigma}{\sqrt{n}}$ $\quad (L, +\infty)$

$U \leq \bar{x} + Z_\alpha \frac{\sigma}{\sqrt{n}}$ $\quad (-\infty, U)$.

General method to derive a C.I.

Pivot: A distibution does not depend on any unknown parameter.

For example: $Z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

Large Sample Confidence Interval:

Let $X_1, \dots, X_n$ a random sample from a population with unknown mean μ and unknown variance σ^2.

IID

$\bar{x}, S^2$

variance σ^2. $\bar{x}$, s^2

When n is Large enough. ($n \geq 30$)

$$Z \approx \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim N(0,1).$$

Approximately Standard Normal.

Sampe

- σ known $\Rightarrow Z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$
- σ unknown $\Rightarrow$
  - $n \geq 30$ → $Z = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$
  - $n < 30$? t distribution → $T = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

T distribution:

Gosset. (student t distribution).

Let $X_1 \dots X_n$ be a random sample from a normal distribution with unknown μ and unknown variance σ^2.

Then

$$T = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

has a t distribution with $n-1$ degree of freedom.

freedom.

$\underbrace{X_1 + \cdots X_n = 1000}$ $n-1$ degree of freedom.

P.D.F. for t:

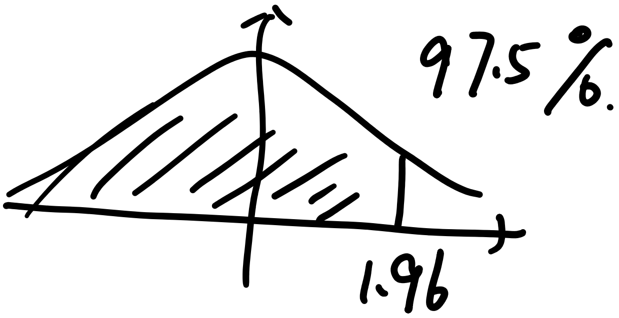
$$f(x) = \frac{\Gamma[(k+1)/2]}{\sqrt{\pi k}\,\Gamma(k/2)} \frac{1}{\left[\frac{x^2}{k}+1\right]^{(k+1)/2}}, \quad -\infty < x < \infty.$$

where k is the number of degree of freedom.

$E(X) = 0$

$Var(X) = \frac{k}{k-2}$.

T-table. Z table



Example: $n=22$. $\bar{x} = 13.71$, $S = 3.55$.

Find 95% two sided
one sided upper / lower } C.I.

(1) two sided:

$n-1 = 21$ two sided $\alpha = 5\% = 0.05$

$n-1=21$ two sided $\alpha = 5\%$ $= 0.05$

$t_{21,\text{two side}}\ \alpha=5\%$

$= 2.08$

$$\bar{x} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}$$

$$13.71 - 2.08\cdot\frac{3.55}{\sqrt{22}} \le \mu \le 13.71 + 2.08\cdot\frac{3.55}{\sqrt{22}}$$

(2) One sided:

$t_{21},\ \text{one side}\ \alpha = 5\%$

$= 1.721$

Upper : $\mu \le 13.71 + 1.721\,\frac{3.55}{\sqrt{22}}$

Lower : $\mu \ge 13.71 - 1.721\,\frac{3.55}{\sqrt{22}}$

Two sided

$$P\left[\bar{X} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}} \le \mu \le \bar{X} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

$$\left(\bar{X} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}},\ \bar{X} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\right)$$

One sided

$\left(\bar{X} - t_{\alpha,n-1}\frac{S}{\sqrt{n}},\ +\infty\right)$ Lower

$\left(-\infty,\ \bar{X} + t_{\alpha,n-1}\frac{S}{\sqrt{n}}\right)$ Upper

25. Sample Test 2.

$n=15$. $95\% \Rightarrow (83.9219, 90.6781)$

What's 99% C I.

$$\bar{x} = \frac{83.9219 + 90.6781}{2} = 87.3$$

95%

$ME = 87.3 - 83.9219 = 3.3781.$

$\bar{x} - \text{Lower}$

$n-1=14$

$ME = t_{\frac{\alpha}{2},n-1} \cdot \frac{s}{\sqrt{n}}$

$t_{2.5\%,14} = 2.145$

$3.3781 = 2.145 \frac{s}{\sqrt{n}}$

$\frac{s}{\sqrt{n}} = 1.57487$

99%

$ME = t_{0.5\%,14} \cdot \frac{s}{\sqrt{n}}$

$t_{0.5\%,14} = 2.977$

$= 2.977 \cdot 1.57487$

$= 4.6883933$

99% C.I.: $(87.3 - 4.6883933, 87.3 + 4.6883933)$

$= (82.61, 91.99)$

Large Sample C.I. for population proportion.

$$Z = \frac{X - np}{\quad}$$

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \quad \text{where } \hat{p} = \frac{X}{n}$$

$$P\left(\hat{p} - Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} \le p \le \hat{p} + Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right)$$

$$= 1 - \alpha.$$

Example: $X = 10$ out of $n = 85$.

$\hat{p} = \frac{10}{85} = 0.12$.

What's 95% CI for p (two-sided)

$n = 85$, $\hat{p} = \frac{10}{85} = 0.12$.

$-Z_{\frac{\alpha}{2}} = -1.96$

$$\left(0.12 - 1.96 \cdot \sqrt{\frac{0.12 \cdot 0.88}{85}},\ 0.12 + 1.96\sqrt{\frac{0.12 \cdot 0.88}{85}}\right)$$

95% C.I. $(0.0509, 0.2\cancel{24}3)$

Choice of sample size:

$$ME = Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

$$\therefore \quad n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 p(1-p)$$

if p is not specified $p = 0.5$.

$$n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 \cdot 0.5^2$$

One-sided p C.I.: $(100(1-\alpha)\%)$

Upper: $p \leq \hat{p} + Z_{\alpha} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Lower: $p \geq \hat{p} - Z_{\alpha} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Test 2.