Lecture 35

15 November 2022 13:18

We saw that for population parameter

$\mu \longrightarrow$ $\bar{X}$ estimator point $\longrightarrow$ (when population variance is known) $\left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$ confidence interval

This interval is s.t.

$P(\mu \in \text{interval}) = (1-\alpha)$

$\therefore$ it is $(1-\alpha)100\%$ confidence interval.

Que The average zinc concentration recovered from a sample of measurements taken in 36 different locations in a river is found to be 2.6 gm/ml. Find the 95% and 99% confidence intervals for mean zinc concentration in the river. Assume that the population s.d. = 0.3 gm/ml.

Sol^n 95% confidence interval

$$1-\alpha = 0.95$$

$$\alpha = 1 - 0.95 = 0.05$$

interval $\left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$

$$\left[2.6 - z_{0.025}\frac{0.3}{\sqrt{36}},\ \bar{x} + z_{0.025}\frac{0.3}{\sqrt{36}}\right]$$

Value from table = 1.96

95% confidence interval $[2.50,\ 2.70]$

99% $1-\alpha = 0.99 \Rightarrow \alpha = 1 - 0.99 = 0.01$

$\dfrac{\bar{X}-\mu}{\sigma/n} \rightarrow$ z distribution

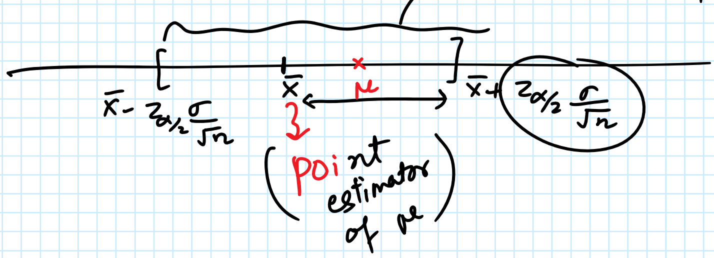
$\dfrac{\bar{X}-\mu}{s/\sqrt{n}} \rightarrow$ t distribution

$$\left[2.6 - z_{0.005}\frac{0.3}{\sqrt{36}},\ 2.6 + z_{0.005}\frac{0.3}{\sqrt{36}}\right]$$

Table 2.575

99% $[2.47,\ 2.73]$ $\rightarrow$ $[2\quad 3]$

Now we have the following situation $\rightarrow$ interval estimation of μ.



what is max. value of

$|\bar{x} - \mu| = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Result If e is the error then

$$e = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \Rightarrow \boxed{n = \left(\frac{z_{\alpha/2}\,\sigma}{e}\right)^2}$$

minimum no. of samples to be used.

Que How large a sample is required if we want to be 95% confident that our estimate of μ in the last example is off by less than 0.05?

error

Soln

$$n = \left(\frac{z_{\alpha/2}\,\sigma}{0.05}\right)^2 = \left(\frac{z_{0.025}\times 0.3}{0.05}\right)^2 = 138.3 \qquad (z_{0.025} = 1.96)$$

at least 139 samples.

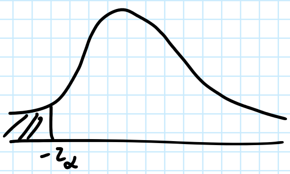
---

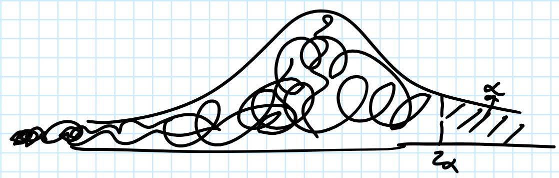
interval - $\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \;,\; \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$

Sometimes we are interested in only one side confidence bounds.

$$P\left(\mu < \hat{\Theta}_U\right) = 1-\alpha$$

upper one sided bound.





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

$P(Z > -z_\alpha) = 1-\alpha$

$P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > -z_\alpha\right) = 1-\alpha$

$P\left(\mu < \bar{X} + z_\alpha\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$

$P(Z < z_\alpha) = 1-\alpha$

$P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z_\alpha\right) = 1-\alpha \Rightarrow P\left(\bar{X} - z_\alpha\frac{\sigma}{\sqrt{n}} < \mu\right) = 1-\alpha$

One side confidence bounds

Upper one sided bound $\quad \bar{x} + z_\alpha\frac{\sigma}{\sqrt{n}}$

lower " " " $\quad \bar{x} - z_\alpha\frac{\sigma}{\sqrt{n}}$

In the last example

for 95% confidence

upper one sided bound $\quad \hat{\Theta}_U = \bar{x} + z_{0.05}\frac{\sigma}{\sqrt{n}} \rightarrow P(\mu \le \hat{\Theta}_U) = 0.95$

lower " " " $\quad \bar{x} - z_{0.05}\frac{\sigma}{\sqrt{n}} \rightarrow P(\hat{\Theta}_L \le \mu) = 0.95$

Next when σ is not known

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \quad t \text{ distribution with } (n-1) \text{ dof}$$

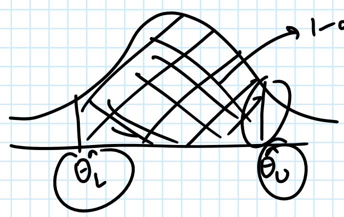
$P(\_ < \mu < \_)$

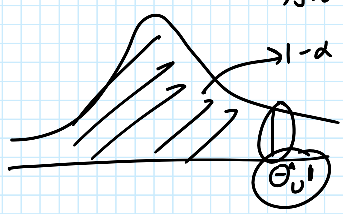
$(1-\alpha)100\%$ confidence interval $= \left[\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}},\ \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right]$

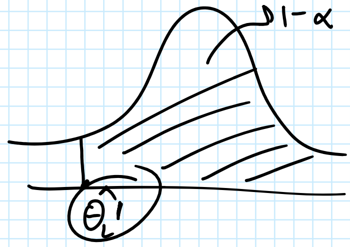
$P(\mu < \_)$

One sided upper bound $= \bar{x} + t_{\alpha} s/\sqrt{n}$ ✓

One " lower " $= \bar{x} - t_{\alpha} s/\sqrt{n}$ ✓



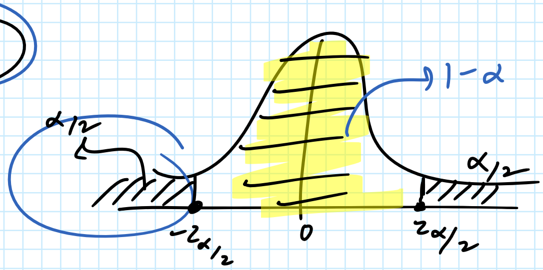




Interval

$$P(\theta_L < \mu < \theta_U) = 1-\alpha$$

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

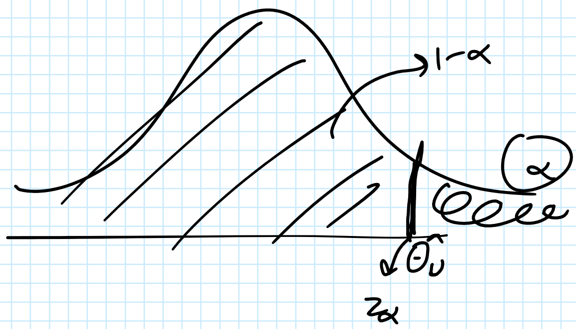


$$z_{\alpha/2} < Z < z_{\alpha/2}$$

$$\left[z_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right]$$

$$P(\mu < \hat{\theta}_U) = 1-\alpha$$

$$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$



Que The contents of seven similar containers of sulphuric acid are 9.8, 10.2, 10.4, 9.8, 10.0, 10.2 and 9.6 litres. Find a 95% confidence interval for the mean content of all such containers, assuming an approximately normal distribution.

Sol^n Since σ is not given so we get

$$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t \text{ distributed with dof } n-1 = 7-1 = 6$$

95% confidence interval $\left[\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}},\ \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right]$

$\Sigma(x_i-10)^2$

$[10.0 - t \ldots 0.283,\ 10 + t_{0.025}\ 0.283]$

95% confidence interval $\left[\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}}, \ \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right]$

$\frac{\sum(x_i - 10)^2}{n-1}$

$$\left[10.0 - t_{0.025}\frac{0.283}{\sqrt{7}}, \ 10 + t_{0.025}\frac{0.283}{\sqrt{7}}\right]$$

Table 2.447

95% $[9.74 \quad 10.26]$

Last note

When your sample is very big $(n \geq 30)$, even if σ is not given, you can approximate $s \sim \sigma$

$$\left[\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}}, \ \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}}\right]$$

(concept of a large sample confidence interval)

Que

SAT scores of a random sample of 500 students has sample mean $= 501$ & S.D. $= 112$. Find a 99% confidence interval.

Sol^n

$n = 500$ big

$\therefore \sigma \approx s = 112$

$\therefore$ 99% $\left[\bar{x} - z_{\alpha/2}\frac{112}{\sqrt{500}}, \ \bar{x} + z_{\alpha/2}\frac{112}{\sqrt{500}}\right]$

$$\left[501 - z_{0.005}\frac{112}{\sqrt{500}}, \ \bar{x} + z_{0.005}\frac{112}{\sqrt{500}}\right]$$

$$[488.1, \ 513.9]$$