PROBABILITY AND STATISTICS (UCS401)

Lecture-31

(Interval of Estimation and confidence interval with illustrations)

Sampling Distributions and Theory of Estimation (Unit –V & VI)

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Dr. Rajanish Rai
Assistant Professor
School of Mathematics
Thapar Institute of Engineering and Technology, Patiala

Priyanish Rai

Interval of Estimation and Confidence limits

The probability that we associated with an interval is called the confidence level. Higher the probability, the more is the confidence.

* If t is the sample statistics used to estimate the corresponding population parameter θ, then $(1-\alpha)$ 100% confidence limits for θ are $t \pm S.E.(t)\, t_\alpha$, where α is the level of significance.

$S.E.(t)$ is the standard error of t and t_α is the significant or critical value of t at the significant level α.

$t - S.E(t)\, t_\alpha$ and $t + S.E(t)\, t_\alpha$ are the limits of confidence interval.

By central limit theorem, if t is the sample statistics, then $Z = \frac{t - E(t)}{S.E(t)}$ is standard normal variate with mean zero and variance $= 1$ as n tends to ∞.

Confidence interval of the mean :

(i) Consider a population. Let μ be the population mean and $\bar{x}$ be the sample mean of sampling distribution. Assume that the sample is large. When the standard deviation σ of the normal population is known, the $S.E.(\bar{x})$ is $\frac{\sigma}{\sqrt{n}}$ (n is large sample size).

The confidence interval

$$[\bar{x} - S.E(\bar{x}) Z_\alpha, \ \bar{x} + S.E.(\bar{x}) Z_\alpha]$$

or $$\left[\bar{x} - \frac{\sigma}{\sqrt{n}} Z_\alpha \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}} Z_\alpha\right]$$

* Confidence interval for μ, when σ^2 is unknown :

Let $X_1, X_2, X_3, \dots X_n$ are random sample from $N(\mu, \sigma^2)$ then we know that

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

and

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

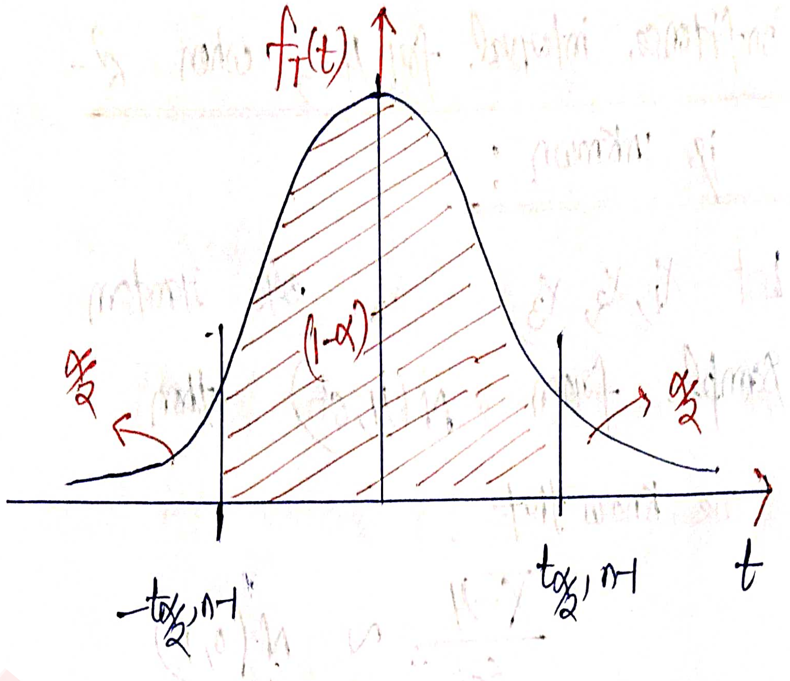
$$\Rightarrow \quad T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{(n-1)}$$

$$\Rightarrow \quad 1-\alpha = P\left(-t_{\alpha/2, n-1} < T < t_{\alpha/2, n-1}\right)$$

$$\Rightarrow \quad 1-\alpha = P\left(-t_{\alpha/2, n-1} < \frac{\bar{X}-\mu}{s/\sqrt{n}} < t_{\alpha/2, n-1}\right)$$

$$= P\left(-t_{\alpha/2, n-1}\frac{s}{\sqrt{n}} - \bar{X} < -\mu < \frac{s}{\sqrt{n}} t_{\alpha/2, n-1} - \bar{X}\right)$$

$$1-\alpha = P\left(\bar{X} - \frac{s}{\sqrt{n}} t_{\alpha/2, n-1} < \mu < \bar{X} + \frac{s}{\sqrt{n}} t_{\alpha/2, n-1}\right)$$



* $\bar{x} = 53.92$

$s = 10.07$

$n = 24$

find a 99% Confidence interval for μ.

$\Rightarrow \quad (1-\alpha)\,100\% = 99\%$

$\Rightarrow \quad \alpha = 0.01$

$\Rightarrow \quad \alpha/2 = 0.005$

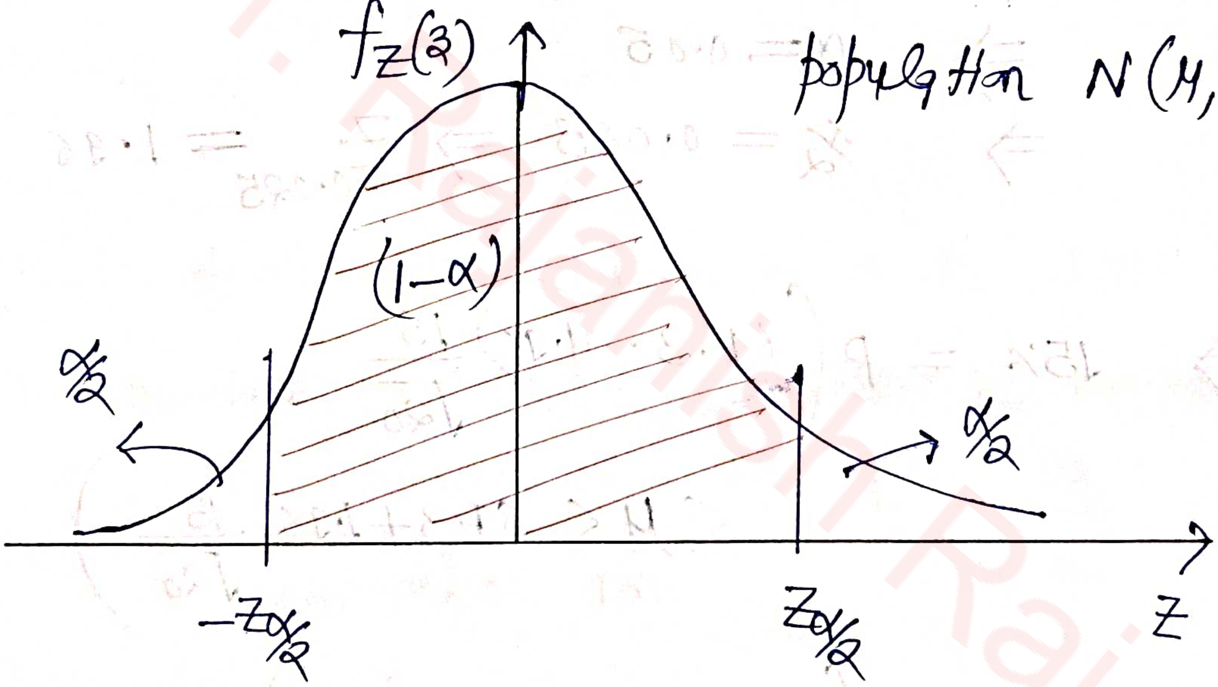
$$1 - 0.01 = P\left(53.92 - t_{0.005, 23}\frac{10.07}{\sqrt{24}} < \mu < 53.92 + t_{0.005}\frac{10.07}{\sqrt{24}}\right)$$

$$= P\left(52.92 - 2.807 \times \frac{10.07}{\sqrt{24}} < \mu < 53.92 + 2.807 \times \frac{10.07}{\sqrt{24}}\right)$$

$$\Rightarrow \quad 0.99 = P\left(48.15 < \mu < 59.69\right)$$

Ans

Case-II Confidence interval for μ if σ^2 is known from a Normal population $N(\mu, \sigma^2)$.



$$1-\alpha = P\left(-Z_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < Z_{\alpha/2}\right)$$

$$1-\alpha = P\left(-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{X}-\mu < \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}\right)$$

$$1-\alpha = P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}\right).$$

* Given

$n = 20$

$\sigma^2 = 225$

$\bar{x} = 64.3$

Construct a 95% Confidence interval for μ.

$1-\alpha = 95\%$

$(1-\alpha)\,100\% = 95\%$

$\Rightarrow \alpha = 0.05$

$\Rightarrow \alpha/2 = 0.025 \Rightarrow Z_{0.025} = 1.96$

$$\Rightarrow 95\% = P\left(64.3 - (1.96)\frac{15}{\sqrt{20}} < \mu < 64.3 + 1.96\,\frac{15}{\sqrt{20}}\right)$$

$$\Rightarrow 95\% = P\left(57.7 < \mu < 70.9\right)$$

Question : [2] The average zinc concentration recovered from a sample of measurements taken in 36 different locations in a river is found to be 2.6 grams per ml. Find the 95% and 99% confidence intervals for the mean zinc concentration in the river. Assuming that the population standard deviation is 0.3 gram per ml.

Given $n = 36$

$\sigma = 0.3$

$\bar{x} = 2.6$

To construct 95% confidence interval for μ

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

$(1-\alpha)100\% = 95\%$

$\Rightarrow \alpha = 0.05$

$\Rightarrow \alpha/2 = 0.025 \Rightarrow Z_{0.025} = 1.96$

$$95\% = P\left(2.6 - 1.96\frac{0.3}{\sqrt{36}} < \mu < 2.6 + 1.96\frac{0.3}{\sqrt{36}}\right)$$

$$95\% = P\left(2.50 < \mu < 2.70\right),$$

To Construct 99% Confidence interval

$$(1-\alpha)\,100\% = 99\%$$

$$\Rightarrow \quad \alpha = 0.01$$

$$\Rightarrow \quad \alpha/2 = 0.005 \quad \Rightarrow Z_{0.005} = 2.575$$

$$99\% = P\left(2.6 - 2.575\frac{0.3}{\sqrt{36}} < \mu < 2.6 + 2.575\frac{0.3}{\sqrt{36}}\right)$$

$$99\% = P\left(2.47 < \mu < 2.73\right)$$

Ans

Theorem: If $\bar{x}$ is used as an estimate of μ, we can be $100(1-\alpha)\%$ confident that the error will not exceed a specified amount e, when sample size is

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

Question ③: How large a sample is required if we want to be 95% confident that our estimate of μ in Ques ② is off by less than 0.05.

Solution: The population standard deviation is

$\sigma = 0.3$ then we know that

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

$$Z_{\alpha/2} = Z_{0.025} = 1.96$$

$$n = \left[\frac{(1.96)(0.3)}{0.05}\right]^2$$

$$n = 138.3$$

Therefore, we can be 95% confident that a random sample of size 139 will provide an estimate $\bar{x}$ differing from μ by an amount less than 0.05.

The Case of σ unknown:

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

Solution-:

	x	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$
	9.8	-0.2	0.04
	10.2	0.2	0.04
	10.4	0.4	0.16
	9.8	-0.2	0.04
	10.0	0	0
	10.2	0.2	0.04
	9.6	-0.4	0.16
sum	70		0.48

$$n = 7$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{70}{7} = 10$$

$$s^2 = \frac{\sum (x_i-\bar{x})^2}{n-1} = \frac{0.48}{6} = 0.08$$

$$s = \sqrt{0.08} = 0.283$$

and $(1-\alpha)100\% = 95\%$

$$\Rightarrow \alpha = 0.05$$

$$\Rightarrow \alpha/2 = 0.025 \qquad t_{0.025,6} = 2.447.$$

Hence, 95% Confidence interval for μ

$$95\% = P\left(\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right)$$

$$95\% = P\left(10 - 2.447\frac{0.283}{\sqrt{7}} < \mu < 10 + 2.447\frac{0.283}{\sqrt{7}}\right)$$

$$95\% = P\left(9.74 < \mu < 10.26\right)$$ Ans