L- 6: Inferential statistics

Agenda

- Quick Review of the topics covered in previous class
- Testing of Hypothesis

Statistical Inferences

Theory of statistical inference is divided into two major areas

- Estimation
- Tests of hypothesis

Estimation

Point Estimation

Interval Estimation

Interval Estimation

Sampling $\longrightarrow$ $\bar{x}$: Mean

Mean of the population

$$\left(\bar{x} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \;,\; \bar{x} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

with $(1-\alpha)$ level of significance

How?

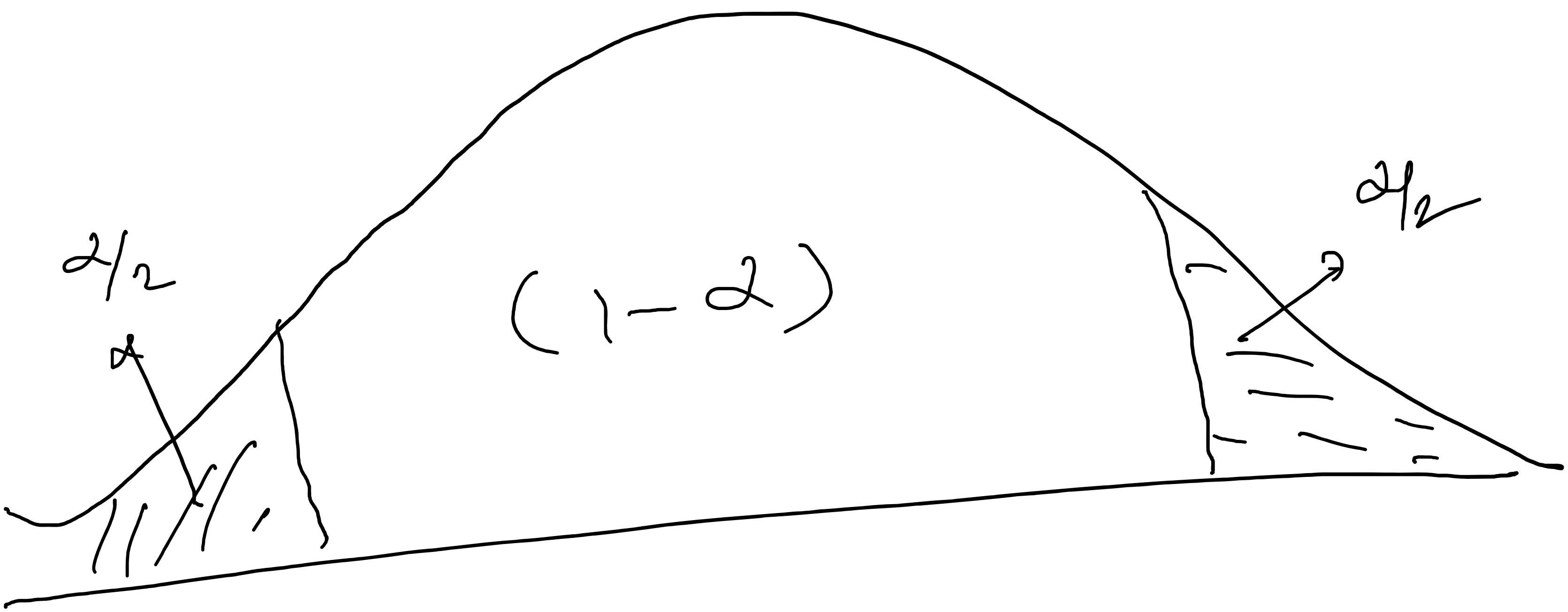
$$P\left(-z_{\alpha/2} < z < z_{\alpha/2}\right) = 1-\alpha$$

$$\downarrow$$

$$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$

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

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


α/2
α
(1-α)
α/2

Estimate for μ

large

$$\left(\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

Replace by S

small

$$\left(\bar{x} - t_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}},\ \bar{x} + t_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

Replace by S

Example:

A company wants to estimate the average life of the product. The S.D is known to be 100 hours. A random sample of 50 gave a sample average life of 350 hours.

Estimate the confidence interval for the mean.

Example:

σ

A company wants to estimate the average life of the product. The S.D is known to be 100 hours. A random sample of 50 gave a sample average life of 350 hours.

n

$\bar{x}$

Estimate the confidence interval for the mean.

95% is LOS suppose

Example:

σ

$$\left(\bar{x} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}},\ \bar{x} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

$$\left(350 \pm 1.96 \cdot \frac{100}{\sqrt{50}}\right)$$

100

$$(350 \pm 27.72)$$

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

σ 50

n

350

confidence interval

$\bar{x}$

$$(322.28, 377.72)$$

95% is LoS suppose

Sample Size

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \rightarrow E$$

i.e.

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

Hypothesis Testing Steps

- Null and alternative hypotheses
- Test statistic
- P-value and interpretation
- Significance level (optional)

→ H_0 and H_1

↓ Type of test

one sided
- left
- right

Two sided

→ Level of significance

→ Critical region

⟹ Decision

Example

Null hypothesis $H_{0:}\ \mu = 170$

The **alternative hypothesis** can be either $H_{1:}\ \mu > 170$ (**one-sided test**)

or

$H_{1:}\ \mu \neq 170$ (**two-sided test**)

Example

A. Hypotheses:

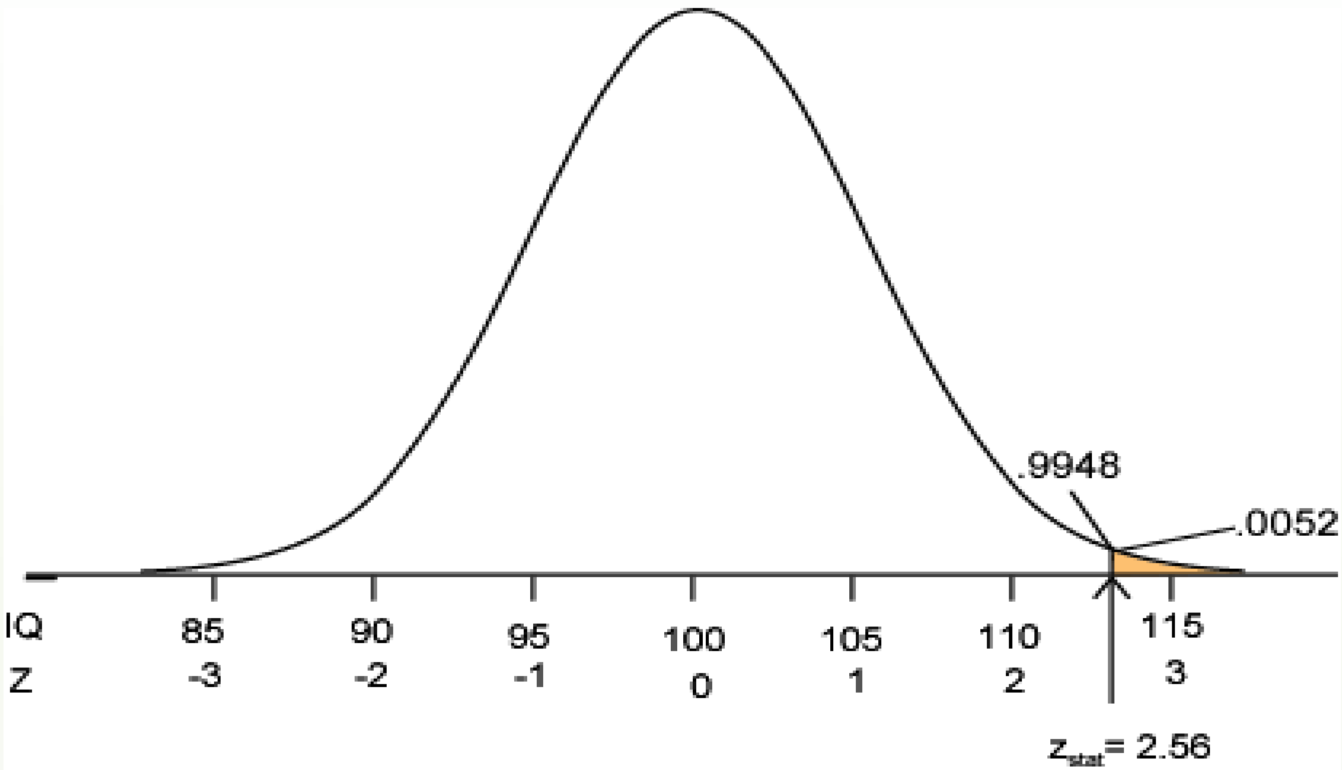
H_0: μ = 100 versus
H_a: μ > 100 (one-sided)
H_a: μ ≠ 100 (two-sided)

B. Test statistic:

$$SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{9}} = 5$$

$$z_{\text{stat}} = \frac{\bar{x} - \mu_0}{SE_{\bar{x}}} = \frac{112.8 - 100}{5} = 2.56$$

C. *P*-value: $P = \Pr(Z \geq 2.56) = 0.0052$



$P = .0052 \Rightarrow$ it is unlikely the sample came from this null distribution $\Rightarrow$ strong evidence against H_0

Hypothesis Testing

Test Result – True State	H_0 True	H_0 False
H_0 True	Correct Decision	Type I Error
H_0 False	Type II Error	Correct Decision

$$\alpha = P(Type\ I\ Error) \quad \beta = P(Type\ II\ Error)$$

- Goal: Keep α, β reasonably small

Problem

It is claimed that a random sample 49 tyres has a mean life of 15200 kms. This sample was drawn from a population whose mean is 15150 kms and a standard deviation of 1200kms. Test the significance at 0.05 level.

Solution:

1. Null hypothesis $H_0 : \mu = 15200$

2. Alternate hypothesis $H_1 : \mu \neq 15200$

3. Level of significance $\alpha = 0.05$

4. critical region :- This is a two tailed test (large sample). So reject H_0 if $(Z_{cal} = Z) < -\ Z_{\frac{\alpha}{2}}$ or $(Z = Z_{cal}) > Z_{\frac{\alpha}{2}}$

Here $\alpha = 0.05$

$$\frac{\alpha}{2} = \frac{0.05}{2}$$

$$= 0.025$$

From table we get

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

i.e; if

$Z_{cal} = Z < -1.96$ or $Z_{cal} > 1.96$ we reject null hypothesis.

6. Computation :

Test statistic

$$Z_{cal} = Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{15200 - 15150}{\frac{1200}{\sqrt{49}}}$$

$$= 0.2916$$

7. Decesion:

Since $Z_{cal} = 0.2916 < 1.96$ we accept the mull hypothesis.

Problem

A trucking firm is suspicious of the claim that the average life time of certain tyres is at least 28,000 miles. To check the claim, the firm puts 40 of these tyres on its trucks and get a mean life of 27,463miles with a standard deviation of 1,348 miles. What can it conclude if the probability of Type I error is to be at most 0.01

Solution

1. Null hypothesis : H_0 : $\mu \geq 28{,}000$ miles

2. Alternate hypothesis: H_1: $\mu < 28{,}000$ miles

3. Level of significance: $\alpha = 0.01$

4. Critical region

This is a left tailed test (large sample)

If $Z = Z_{cal} < -Z_{\alpha}$ we reject null hypothesis

If $Z = Z_{cal} < -Z_{\alpha} = -Z_{0.01} = -2.33$ we reject null hypothesis

5.Computation

Test statistic

$$Z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{27,463-28,000}{\frac{1,348}{\sqrt{40}}} = -2.52$$

6.Conclusion

Since Z = Z_{cal} = -2.52 < -2.33 , we reject null hypothesis at level of significance 0.01. In other words the trucking firm's suspicion that μ< 28,000 miles is confirmed.

Hypothesis concerning one mean (small sample)

Procedure

1. Null hypothesis $H_0 : \mu = \mu_0$

2. Alternate Hypothesis $H_1 : \mu \neq \mu_0$ (Two tailed test)

Or

$H_1 : \mu > \mu_0$ (Right tailed test)

Or

$H_1 : \mu < \mu_0$ (left tailed test)

3. Level of significance : α

4. Critical region

For two tailed test $H_1 : \mu \neq \mu_0$

Reject H_0 if $t < -t_{\frac{\alpha}{2}}$ or

$t > t_{\frac{\alpha}{2}}$ with (n-1) degrees of freedom

For right tailed test $H_1 : \mu > \mu_0$

Reject H_0 if $t > t_{\alpha}$ with (n-1) degrees of freedom

For left tailed test $H_1 : \mu < \mu_0$

Reject H_0 if $t < -t_{\alpha}$ (n-1) degrees of freedom

5. Test statistic

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$ with (n-1) degrees of freedom

6. Calculation

7. Decision

Example:

A random sample of 6 steel beams has a mean compressive strength of 58,392 p.s.i (pounds per square inch) with a standard deviation of 648 p.s.i . use this information at the level of significance $\alpha = 0.05$ to test

whether the true average compressive strength of steel from which the sample came is 58,000 p.s.i

n $\bar{x}$ s

A random sample of 6 steel beams has a mean compressive strength of 58,392 p.s.i (pounds per square inch) with a standard deviation of 648 p.s.i . use this information at the level of significance $\alpha =$ 0.05 to test

whether the true average compressive strength of steel from which the sample came is 58,000 p.s.i μ

$\mu = 58{,}000$

$\mu \neq 58{,}000$ → Two tailed test

$n = 6$ → small sample t - test.

Thanks