Unit 2: Testing Of Hypothesis:

Hypothesis: A statement related to population parameter.

Two types

- **NULL:** No difference statement
  Ex: $H_0: \mu = 3$, $H_0: \mu_1 = \mu_2$, $H_0: \sigma^2 = 2$
- **Alternate:** Complementary to NULL hypothesis.
  Ex: $H_1: \mu \neq 3$, $H_1: \mu_1 \neq \mu_2$, $H_1: \sigma^2 \neq 2$

Alternate → Two types

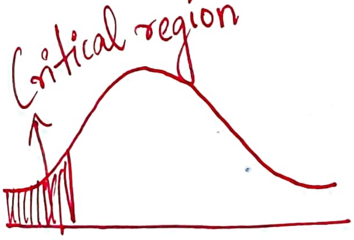
Single tailed (one tailed)

Tests the possibility of relationship in one direction and completely ignores the possibility of a relationship in another direction. The values defer from reference only in one direction.

Left Tailed

A hypothesis test where the rejection region is located to the extreme left of the distribution.

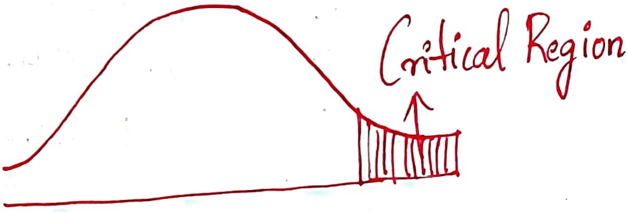
$H_A < \alpha$



Right tailed.

A hypothesis test where the rejection region is located to the extreme right of the distribution

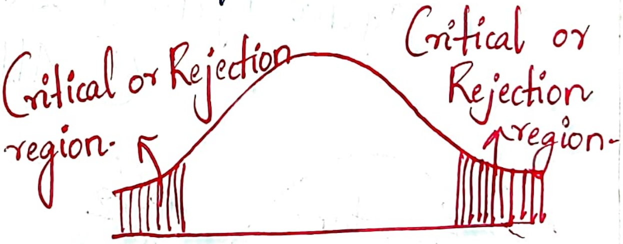
$H_A > \alpha$



Two tailed

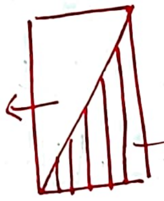
Tests whether a sample is greater than or less than a certain range of values. It is non-directional because the effects are tested for in both directions.

Ex: $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$ | $H_0: \sigma^2 = 3$, $H_1: \sigma^2 \neq 3$



Critical Region (Rejection Region)

The set of all values of the test static that would cause us to reject the NULL hypothesis. This area is defined by Critical value(s).



Level of Significance:

The probability of rejecting the null hypothesis when it is true

$$\boxed{\begin{aligned} &P(t \in w \mid H_0) = \alpha \\ &P(t \in \bar{w} \mid H_1) = \beta \\ &w \cup \bar{w} = S \\ &w \cap \bar{w} = \phi \end{aligned}}$$

Critical Value or Significant Value

↓

depends upon: ① Level of significance
② Alternate Hypothes

Two tailed alternative

Right-tailed Alternative

Left tailed Alternative

Types of Error:

Type I: probability of rejecting a good lot (producer's risk) denoted by α

Type II: probability of accepting a bad lot (consumer's risk) denoted by β

Level of Significance

	1%	5%	10%
Two-tailed	$Z_{\alpha/2} = 2.58$	$Z_{\alpha/2} = 1.96$	$Z_{\alpha/2} = 1.65$
Right-tailed	$Z_\alpha = 2.33$	$Z_\alpha = 1.65$	$Z_\alpha = 1.29$
left-tailed	$Z_\alpha = -2.33$	$Z_\alpha = -1.65$	$Z_\alpha = -1.29$

power of the test $= 1-\beta$

For 2-tailed & right-tailed
Cal Value < table value → accept H_0
Cal Value > table value → reject H_0

For left-tailed:
Calculated value < table value → reject H_0
Calculated value > table value → accept H_0

Test for single proportion:

COIN TOSS PROBLEMS

$$Z = \frac{x - np}{\sqrt{npq}} \sim N(0,1)$$

n → no. of times coin is tossed
p → prob of success
q → " " failure

Test for single proportion Cont'd:

$$Z = \frac{p - P}{\sqrt{PQ/n}} \sim N(0,1)$$

$p \rightarrow$ sample prop portion

$P \rightarrow$ population proportion

$Q \rightarrow 1-P \quad p = x/n$.

$n \rightarrow$ sample size

Confidence interval for true proportion:

$$p - z_{\alpha/2}\sqrt{pq/n} < P < p + z_{\alpha/2}\sqrt{pq/n}$$

NOTE: When level of significance not given, take default 5% (0.05 LOS)

Test for single proportion means: One sample and a claim about a proportion of the sample is tested.

Test for difference of proportions:

① $Z = \dfrac{p_1 - p_2}{\sqrt{PQ(1/n_1 + 1/n_2)}}$

$p_1 =$ I sample proportion

$p_2 =$ II " "

$Q = 1-P$

$P = \dfrac{X_1 + X_2}{n_1 + n_2} = \dfrac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$

$n_1 \rightarrow$ I sample size

$n_2 \rightarrow$ II " "

② $Z = \dfrac{P_1 - P_2}{\sqrt{\dfrac{P_1 Q_1}{n_1} + \dfrac{P_2 Q_2}{n_2}}} \sim N(0,1)$

→ Used when two different populations are discussed. Men, Women / City A, City B.

$P_1 \rightarrow$ I population proportion

$P_2 \rightarrow$ II " "

$Q_1 \rightarrow 1 - P_1$, $Q_2 \rightarrow 1 - P_2$

$n_1 \rightarrow$ I sample size

$n_2 \rightarrow$ II " "

(3) Confidence interval:

$$(p_1 - p_2) \pm z_{\alpha/2}\sqrt{pq\left(1/n_1 + 1/n_2\right)}$$

Both (1) & (2) are used when two different categories of populations are given.

(1) used used when the number of proportion is known.

(2) is " " " percentage " " " ".

TEST FOR SINGLE MEAN:

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

$\bar{x} \rightarrow$ Sample mean

$\mu \rightarrow$ population mean

$\sigma \rightarrow$ S.D of popⁿ. If not given, use S.D of sample (s).

$n \rightarrow$ ~~no~~ sample size.

Confidence Interval:

$$(\bar{x} - \mu) \pm z_{\alpha/2}\left(s/\sqrt{n}\right)$$

TEST FOR DIFFERENCE OF MEANS

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$x_1 \rightarrow$ 1st Sample mean

$x_2 \rightarrow$ 2nd " "

$s_1 \rightarrow$ 1st " S.D

$s_2 \rightarrow$ 2nd " "

$n_1 \rightarrow$ 1st " size

$n_2 \rightarrow$ 2nd " "

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$\sigma \rightarrow$ population S.D

Confidence interval formula:

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Confidence Interval Problem for True Proportions.

Note points:

* Confidence interval problems are always Two-tailed.
* Confidence interval is the same as confidence limits.

Both have upper and lower bounds.

Problems:

① In a random sample of 160 workers exposed to a certain amount of radiation, 24 experienced some ill effects. Construct a 99% confidence interval for the corresponding true proportion.

Soln: Given that:

$n = 160$

$X = 24$

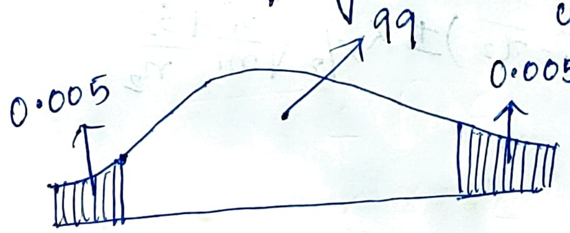
$p = X/n = \frac{24}{160} = 0.15$

$q = 1 - p = 1 - 0.15 = 0.85$

Given confidence region $\rightarrow$ 99

Critical region $\rightarrow$ $100 - 99 = 1$

$\Rightarrow \alpha = 1\%$ level of significance. Since it is two-tailed, we consider $\alpha/2 = \frac{0.01}{2} = 0.005$



[Why are we doing this? It is two tailed. The fraction part is considered as 0.005 → we take the absolute value disregarding the sign. Knowing the value for any one of the extreme ends is sufficient. Hence, we redefine the confidence level).

$1 - 0.005 = 0.995$ i.e. $\boxed{0.995}$

Search this value in the table (Normal Distribution table).

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	$\boxed{0.08}$	0.09
0.0										
0.1										
0.2										
⋮										
$\boxed{2.5}$									0.995	

$\Rightarrow Z_{\alpha/2} = 2.58.$

$$p - Z_{\alpha/2}\sqrt{pq/n} < P < p + Z_{\alpha/2}\sqrt{pq/n}$$

$$0.07717 < P < 0.2228.$$

(2) A random sample of 500 apples was taken from a large consignment and 60 were found to be bad. Obtain 98% confidence limits.

Soln: Given:

$n = 500$

$x = 60$

$p = 0.12$

$q = 1 - p = 0.88$

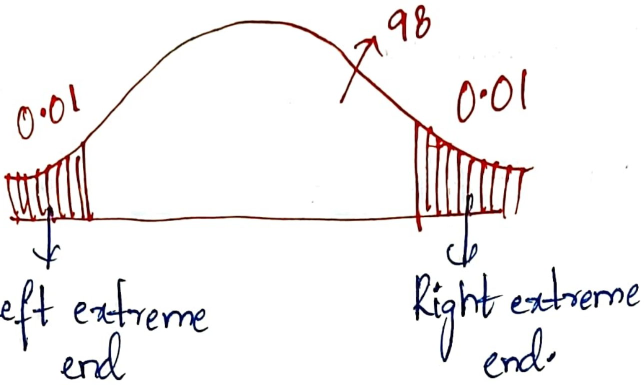
Given confidence region → 98

Critical region → 100 − 98

= 2

Two-tailed → $z_{\alpha/2}$ ⇒ $\alpha = 0.02$

$$\frac{\alpha}{2} = \frac{0.02}{2} = 0.01.$$



It is sufficient that we find the z-score for one extreme end. So we redefine the confidence area as follows:

$1 - 0.01$

$= 0.99$

~~fraction part~~ → $\boxed{0.99}$

Search it in the table.

⇒

	0.01	0.02	0.03	0.04	0.05	0.06	0.07
0.0							
0.1							
0.2							
⋮							
2.3			0.9901				

$$\Rightarrow p - z_{\alpha/2}\sqrt{pq/n} < P < p + z_{\alpha/2}\sqrt{pq/n}$$

$$9 < P < 16$$

Suppose we are asked for, 96 % Confidence interval

Confidence region $\rightarrow$ 96

Critical " $\rightarrow$ 100 - 96

= 4

4 % LOS $\Rightarrow$ $Z_{\alpha/2} = \frac{0.04}{2} = 0.02$

New confidence region $\rightarrow$ ~~100 - 0.02~~ 1 - 0.02

= ~~99.98~~ = 0.98

~~[illegible]~~ $\rightarrow$ 0.98

Search in table

$Z_{\alpha/2} = 2.05.$

SM Model Question Paper

UNIT - 2

3a) Given:

$H_0 : P_1 = 0.85$ [Two-tailed]

$H_1 : P_1 \neq 0.85$

$X = 18$

$n = 30$

$\alpha = 0.05$

$p = \frac{18}{30}$

$= 0.6$

$P = 0.85$

$Q = 1 - 0.85$

$= 0.15$

Formula to be used: $\frac{p - P}{\sqrt{PQ/n}}$

$$Z = \frac{0.6 - 0.85}{\sqrt{(0.85 \times 0.15)/30}}$$

$$= \frac{0.6 - 0.85}{0.06519}$$

$$= \frac{-0.25}{0.06519}$$

$$= |-3.834|$$

$$= 3.834.$$

Z at 5% LOS for 2 → tailed is 1.96

Cal Val > Tab Val.

Reject H_0

Inference: The survival rate cannot be taken as 85%.

3) b) Given:

$H_0: p_1 = 0.5$ $\left[\begin{array}{l}0.5 \text{ because} \\ \text{equally populated}\end{array}\right]$

$p_2 \neq 0.5$

$\alpha = 0.01$

$n = 1000$

$x = 540$

$p = \frac{x}{n}$

$p = 0.54$

$P = 0.5$

$Q = 1 - 0.5$

$= 0.5$

Formula to be used $\Rightarrow$ $Z = \frac{p - P}{\sqrt{PQ/n}}$

$= \frac{0.54 - 0.5}{\sqrt{(0.5 \times 0.5)/1000}}$

$Z = 2.531$

Z at 1% LOS for 2-tailed $\rightarrow$ 2.58.

Cal Val < Tab Val

Accept H_0

Inference: Rice & Wheat eaters are equally populated.

4) a) Given:

POPULATION PERCENTAGES.

$n_1 \rightarrow 1200$ $P_1 \rightarrow 0.3$

$n_2 \rightarrow 900$ $P_2 \rightarrow 0.25$

$Q_1 \rightarrow 1 - P_1$

$= 1 - 0.3$

$= 0.7$

$Q_2 \rightarrow 1 - P_2$

$= 1 - 0.25$

$= 0.75$

H_0: Differently likely hidden

H_1: Not likely to be hidden.

$\alpha = 5\%$ (default)

Formula to be used: $\dfrac{P_1 - P_2}{\sqrt{\dfrac{P_1 Q_1}{n_1} + \dfrac{P_2 Q_2}{n_2}}}$

$$Z = \frac{0.3 - 0.25}{\sqrt{\dfrac{0.3 \times 0.7}{1200} + \dfrac{0.25 \times 0.75}{900}}}$$

$Z = 2.55$

Z at 5% LOS for two-tailed $\rightarrow 1.96$

Cal Val > Tab Val

Reject H_0

Inference: Not likely to be hidden.

4)b) Given: H_0: P_1 → Australians height = Englishmen
H_1: P_1 → " " > "

Test for difference of Means problem:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$\bar{x}_1 = 67.85$ $\quad$ $\bar{x}_2 = 68.55$

$s_1 = 2.56$ $\quad$ $s_2 = 2.52$

$n_1 = 6400$ $\quad$ $n_2 = 1600$

$$Z = \frac{68.55 - 67.85}{\sqrt{\frac{(2.52)^2}{1600} + \frac{(2.56)^2}{6400}}}$$

$$Z = 9.9$$

Z at 5% Level of Significance for right tailed is 1.65.

Cal Val > Tab Val

Reject H_0.

Inference: Australians are on avg taller than Englishmen.