(1) Sonia Marjla

Testing of Hypothesis

Standard Error:- (S.E.) The standard deviation of the sampling distribution of a statistic is known as the Standard Error (S.E.)

If t is any statistic, for large sample

$$\boxed{Z = \frac{t - E(t)}{S.E(t)}}$$

For large sample, the standard error of some of the well known statistic are listed below.

n = sample size

σ^2 = Population variance

s^2 = Sample variance

P = Population proportion

Q = $1 - P$

n_1, n_2 are sizes of two independent random samples

No.	Statistic	Standard Error
1	$\bar{x}$	$\sigma/\sqrt{n}$
2.	s	$\sqrt{\sigma^2/2n}$
3.	Difference of two sample means $\bar{x}_1 - \bar{x}_2$	$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
4.	Difference of two standard deviations $s_1 - s_2$	$\sqrt{\frac{\sigma_1^2}{2n_1} + \frac{\sigma_2^2}{2n_2}}$
5.	Difference of sample proportions $(P_1 - P_2)$	$\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}$
6.	Observed sample proportion p	$\sqrt{\frac{PQ}{n}}$

Introduction:- The main objective of the sampling theory is the study of the testing of Hypothesis or Tests of Significance

In many circumstances, we are to make decisions about the Population on the basis of only sample information.
eg:- on the basis of sample data
(i) a quality control manager is to determine whether a process is working properly
(ii) a drug chemist is to decide whether a new drug is really effective in curing a disease.
(iii) a statistician has to decide whether a given coin is unbiased etc.

Basic Concepts of Hypothesis testing:-

Null Hypothesis:- (H_0) In tests of hypothesis we always begin with an assumption or hypothesis
hypothesis is a definite statement about the Population parameter called Null hypothesis denoted by H_0.
Null hypothesis asserts that there is no (significant) difference between the sample statistic and population Parameter and whatever the observed difference is there, is merely due to fluctuations in sampling from the same population]

Alternative hypothesis:- (H_1) Any hypothesis which is complementary to the null hypothesis (H_0) is called an alternative hypothesis denoted by (H_1)

If we want to test the null hypothesis that the population has a specified mean μ_0, then we have

$$H_0: \mu = \mu_0$$

Alternative hypothesis will be

(i) $H_1 : \mu \neq \mu_0 \; (\mu > \mu_0$ or $\mu < \mu_0)$ (two tailed alternative hypothesis)

(ii) $H_1 : \mu > \mu_0$ [~~righted~~ right tailed alternative hypothesis (single tailed)]

(iii) $H_1 : \mu < \mu_0$ [left tailed alternative hypothesis (single tailed)]

Level of significance:- This refers to the degree of significance with which we accept or reject a particular hypothesis. Since 100% accuracy is not possible in taking a decision over the acceptance or rejection of a hypothesis. This is usually denoted by α

Critical Region or rejection region:- The critical region or rejection region is the region of the standard normal curve corresponding to a pre-determined level of significance.

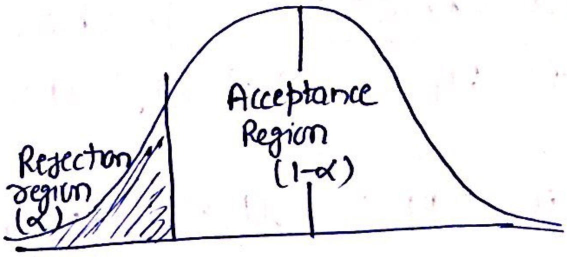
The statistic which leads to the rejection of null hypothesis H_0 gives us a region known as rejection region or critical region, while those which lead to acceptance of H_0 give us a region called as acceptance region.

One tailed test and two tailed test:- A test of any statistical hypothesis where the alternative hypothesis expressed by the symbol ($<$) or the symbol ($>$) called a one tailed test

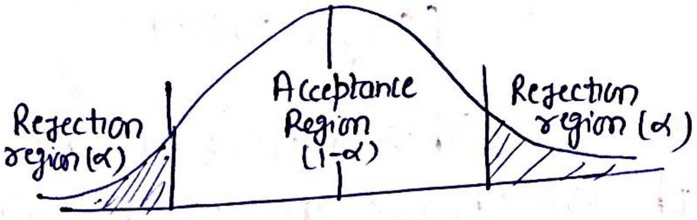
(4) (Sonia maryla)

Right tailed test

Left tailed test



Two tailed test



Critical value:- The critical values of standard normal variate (Z) for both the two-tailed and one tailed tests at different level of significance

Level of Significance (α)	α=0.10	α=0.05 (α=5%)	α=0.01 (α=1%)	α=0.005
Critical value of Z (for one tailed test)	−1.28 or +1.28	−1.645 or +1.645	−2.33 or +2.33	−2.58 or +2.58
Critical value of Z (for two tailed test)	−1.645 or +1.645	−1.96 or +1.96	−2.58 or +2.58	−2.81 or +2.81

Procedure for testing of hypothesis:-

(i) **Setup a null hypothesis:-** It is denoted by H_0, Null hypothesis assumes that difference between any values to be compared is not significant.

~~ii) Set up a suitable level of significance:-~~

ii) **Set up alternative hypothesis** → set up H_1, so that we could decide whether we should used one-tailed test or two tailed test.

iii) **Set up suitable level of significance**

(iv) Test statistic:- calculate the test statistic (Sonia marla) (5)

$$Z = \frac{t - E(t)}{S.E.(t)}$$ under the null hypothesis

V) Conclusion:- Compared the calculated value of Z with critical value Z_α at level of significance α.

if $|Z| > Z_\alpha$, we reject the H_0 and conclude that there is significant difference.

If $|Z| < Z_\alpha$ we accept H_0 and conclude that there is no significant difference.

Tests of significance for large samples

If the sample size $n > 30$ the sample is taken as large sample. For such sample ~~we~~ we apply following tests

1. Testing of significance for single proportion
2. Testing of significance for difference of proportions
3. Testing of significance for single mean.
4. Testing of significance for difference of means

(I) Testing of significance for single proportion:-

This test is used to find the significant difference between proportion of the sample and the population.

test statistic $(Z) = \frac{p - P}{\sqrt{PQ/n}}$

p = sample proportion
P = Population Proportion
$Q = 1 - P$

eg:- A coin was tossed 400 times and the head turned up 216 times. Test the hypothesis that the coin is unbiased.

Soln:- H_0: The coin is unbiased i.e. $P = 0.5$
H_1: The " " not " i.e. $P \neq 0.5$

Here $n = 400$, X = no of successes $= 216$

p = proportion of success in the sample $= \frac{216}{400} = 0.54$

Population Proportion $P = 0.5$, $Q = 1 - P = 0.5$

Test statistic $Z = \frac{p - P}{\sqrt{PQ/n}} = \left| \frac{0.54 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{400}}} \right| = 1.6$

at 5% level of significance Z_α for two tailed test is (1.96)

Since $|Z| = 1.6 < 1.96$

i.e. $|Z| < Z_\alpha$, value of test statistic is less that significant value of Z at 5% level of significance

$\therefore$ we accept the null hypothesis

$\therefore$ coin is unbiased

---

eg: A cubical die is thrown 9000 times and a throw of 3 or 4 is observed 3240 times. Show that die can not be regarded as an unbiased one and find the extreme limits between which the probability of a throw of 3 or 4 lies.

Solution $n = 9000$

Probability of success (i.e getting 3 or 4 on die)

$P = \frac{2}{6} = \frac{1}{3}$, $Q = 1 - \frac{1}{3} = 2/3$

$p = \frac{3240}{9000} = 0.36$

H_0: die is unbiased $P = 1/3$

H_1: $P \neq 1/3$ (two tailed test)

Test statistic $Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.36 - 0.33}{\sqrt{\frac{1}{3} \times \frac{2}{3} \times \frac{1}{9000}}} = 0.03496$

Scanned with CamScanner