07/23

Unit-3 Hypothesis Testing

*) Definations

a) Population: An aggregate of objects under study. It is also called universe.

b) Sample: A finite subset of population. It is small part of universe.

c) Sample Size: The no individuals in the sample

d) Sampling: The process of selecting the sample from universe.

e) Parameter: Statistics which is associated with population. Statististical concept of population such as mean, variance etc.

f) Statistics: The statistical concept from the sample from the members of the sample.

*) Population mean, variance is given by μ and σ^2.

Sample mean and variance is given by $\bar{x}$ and s^2

Test of Significance: It enables us to decide on the basis of the results of the sample whether

a) Deviation between observed sample statistics & hypothetical parameter value Or

b) Deviation between 2 sample statistic

is significant or might be attributed to fluncuations in sampling.

For applying the test of significance

1) Null hypothesis H_0 and alternate hypothesis H_1 are set up.

Null hypothesis H_0 is definite statement and alternate hypothesis H_1 complimentary of null hypothesis

Egs: Null hypothesis is that the population has a specific mean.

$H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$ (two tailed test)

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

$\mu < \mu_0$ (Left tailed test)

Critical Region: A region corresponding to statistics in a sample space which amounts to the rejection of Null hypothesis is called critical region or region of rejection. The region of sample space which amounts to the acceptance of null hypothesis is called acceptance region.

* Level of Significance: The probablity of the value of a variant falling in the critical region is known as level of significance.

* Errors:

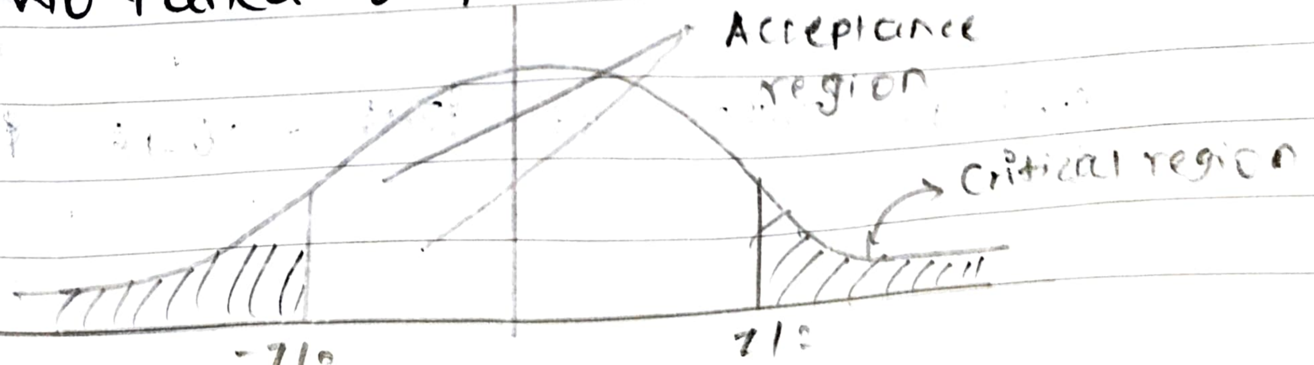
Type 1: When null hypothesis is true it is rejected

② When Null hypothesis is false it's accepted

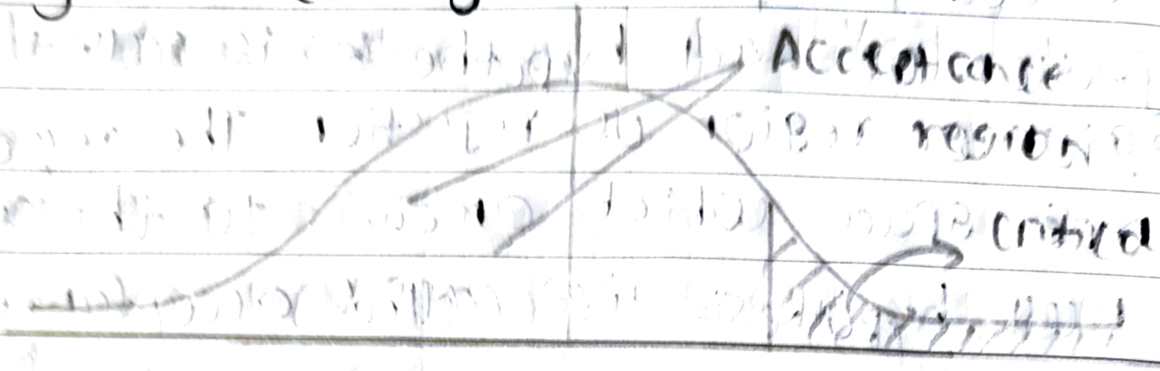
* Critical Values / Significant Values: (z_α)

The values of the test statistics which separates the ~~vat~~ critical region & acceptance region.

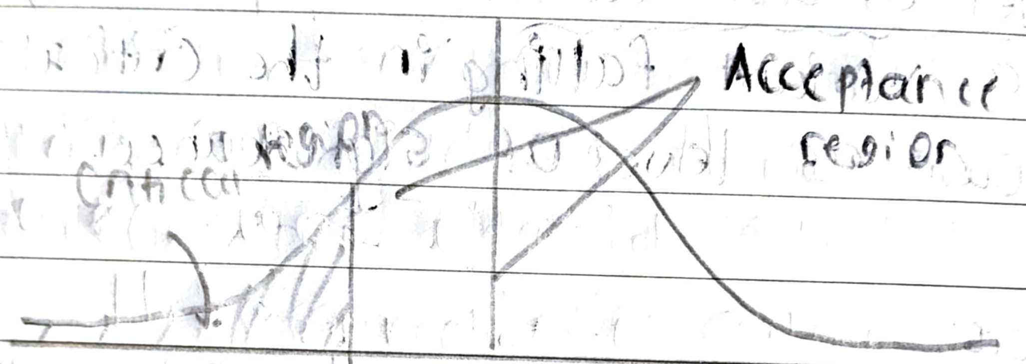
Two tailed test Diagram:



Right tailed region



Left tailed Region



Level Of Significance

	1% (0.01)	5% (0.05)	10% (0.1)
Two Tailed test	$\lvert Z\alpha \rvert = 2.58$	$\lvert Z\alpha \rvert = 1.966$	$\lvert Z\alpha \rvert = 0.645$
Right Tailed test	$\{Z\alpha\} = +2.33$	$\{Z\alpha\} = +1.645$	$\{Z\alpha\} = +1.28$
Left tailed test	$\{Z\alpha\} = -2.33$	$\{Z\alpha\} = -1.645$	$\{Z\alpha\} = -1.28$

* P is probablity of success for one trial ie for one head = 0.5

Problems

1) Test of significance for large samples: If the sample size ~~n<30~~ n > 30 the sample is taken as large sample.

a) Testing of significance for single proportion

$$Z = \frac{p - P}{\sqrt{\frac{P \times Q}{n}}}$$ where $p = \frac{X}{n}$ is the observed proportion of success ie X is the no of successors in n independent trials with constant probablity P of success of each trial and $Q = 1 - P$, z is called test statistic which is used to test significant difference of a sample in population proportion.

a) A coin wass tossed 400 times & head turned up 216 times. Test the hypothesis that the coin is unbaised

Given: $n = 400$, $X = 216$.

$$p = \frac{X}{n} = \frac{216}{400} = 0.54$$

$P = 0.5$ $\quad Q = 1 - P = 1 - 0.5 = 0.5$

$$Z = \frac{p - P}{\sqrt{PQ/n}} = \frac{0.54 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{400}}} = 1.6.$$

... baised [$P \neq 1/2$] { two Tailed Test }.

$Z = 1.6 < 2.58$, Significant value Of 1% level Of Significance $\therefore$ H_0 is accepted ~~at~~ unbiased.

2) ~~(b)~~ A certain cubical die is thrown 9000 times. ~~f~~ 5 or ~~a~~6 was Obtained ~~3104~~ 3240 times. On the assumption Of certain throwing do the data indicate the unbaised die.

Ans:

Given: $n = 9000$, $X = 3240$.

$P = 1/3$ $Q = 2/3$.

$$p = \frac{X}{n} = \frac{3240}{9000} = 0.36.$$

Hypothesis:

H_0: Die is unbaised [$P = 1/3$]

H_1: Die is baised [$P \neq 1/3$]

$$Z = \frac{p - P}{\sqrt{PQ/n}} = \frac{0.36 - 0.3333}{\sqrt{\frac{1/3 \times 2/3}{9000}}} = 5.3732.$$

$Z = 5.3732 > 2.58$ is significant test of 1% level of significance ∴ die is baised.

20 people where attacked by a disease and only 18 survived will you reject the hypothesis that the survival rate if attacked by this disease is 85% in favour of the hypothesis that it's more. Test 5% level of significance. using (large sample test.)

$n = 20$, $X = 18$, $P = 0.85$, $Q = 0.15$

$$p = \frac{X}{n} = \frac{18}{20} = 0.9$$

Hypothesis:

H_0: Survival ratie is attacked is 0.85

H_1: Survival rate is attacked is > 0.85

Right tailed test

$$Z = \frac{p - P}{\sqrt{PQ/n}} = \frac{0.9 - 0.85}{\sqrt{\frac{0.85 \times 0.15}{20}}} = 0.6262$$

~~Z = 0.6262~~ $Z = 0.6262 < 1.645$ significant value is 5% level of signitcance ∴ H_0 is accepted ∴ people will survive by 0.85

NOTE: Probable limits for observed proportion of success is given by $p \pm Z_\alpha \sqrt{\frac{pq}{n}}$ and $q = 1 - p$.

A random sample of 500 apples was taken from a large consignment and 60 where found bad. Obtain 99% confidence limits for the percentage of bad apples in consignment.

$n = 500$, $x = 60$ ~~$Z_\alpha = 0.99$~~

$$p = \frac{x}{n} = \frac{60}{500} = 0.12$$

$q = 1 - p = 0.88$

$\therefore$ 99% for given confidence limits

$\therefore$ It will 1% level of significance

$\therefore Z_\alpha = 2.58$

$$0.12 + 2.58 \sqrt{\frac{0.12 \times 0.88}{500}} = \cancel{0.1343}\ 0.157 = 15 = 16\%$$

$$0.12 - 2.58 \sqrt{\frac{0.12 \times 0.88}{500}} = 0.08 = 8\%$$

Hence 99% confidence limits for percentage of bad apples in the consignment is (8%, 16%)

5) In a sample of 11000 people in Karnataka 540 are rice eaters and the rest are wheat eaters. Can we assume that both rice and wheat are equally popular in this state at 1% level of significance.

$N = 1000$ $x = 540$ $P = 0.5$ $Q = 0.5$

$$p = \frac{x}{n} = 0.54$$

H_0: Person is rice eater ($P = 0.5$)

H_1: Person is not rice eater ($P \neq 0.5$) (Two tailed test)

$$Z = \frac{p - P}{\sqrt{PQ/n}} = \frac{0.54 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{1000}}} = 2.53.$$

$Z = 2.53 < 2.58$. Significant ~~level~~ ~~of~~ value is 1% level of significance. H_0 is accepted.

∴ People will eat rice.

* Test of difference between proportion

Consider 2 samples of sizes n_1 & n_2 respectively taken from 2 different populations. The test of significance of difference between the sample proportion P_1 and P_2 the test statistics under the null hypothesis that there is no significant difference betw^n 2 sample proportions. is

$$\boxed{Z = \frac{P_1 - P_2}{\sqrt{\hat{P}\hat{Q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}}$$

where $\hat{P} = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$ and $\hat{Q} = 1 - \hat{P}_1$

Problems

Before an increase in tax on Tea 800 people out of 1000 people where found to be tea drinkers. After an increase in tax 800 people where known to be tea drinkers in a sample of 1200. Do you think there has been a significant decrease in consumption of tea after the increase in tax.

$n_1 = 1000$ $n_2 = 1200$ $x_1 = 800$ $x_2 = 800$

ns

Ho: People who drink tea after increase tax reno
H1: People who drank less tea after tax.

Ho $\boxed{P_1 = P_2}$ $H_1 := P_1 > P_2$

$P_1 = \frac{x_1}{n_1} = \frac{800}{1000} = 0.8$ $P_2 = \frac{x_2}{n_2} = \frac{800}{1200} = 0.67$.

$\hat{P} = \frac{n_1 P_1 + P_2 n_2}{n_1 + n_2} = \frac{1000(0.8) + 0.67(1200)}{1000+1200}$

$\boxed{\hat{P} = 0.729}$

$$\boxed{\hat{q} = 0.271}$$

$$Z = \frac{P_1 - P_2}{\sqrt{\hat{P}\hat{Q}\left(1/n_1 + 1/n_2\right)}} = \frac{0.8 - 0.67}{\sqrt{0.729(0.271)\left[\frac{1}{1000} + \frac{1}{1200}\right]}}$$

$$\boxed{Z = 6.83}$$

$Z = 6.83 > 2.58$ at significant value of 1% level of significance. H_0 is not accepted ∴ People will drink less tea after tax.

) Random samples of 400 men and 600 women where asked whether they would like to have flyover near there residenee 200 men and 325 women where in favour of the proposal. Test the hypothesis that the propotion of men and women in favour of the proposal are same against that they are not. Test thi at 5% level of significance

$n_1 = 400 \quad n_2 = 600 \quad x_1 = 200 \quad x_2 = 325.$

$$P_1 = \frac{x_1}{n_1} = \frac{200}{400} = 0.5. \quad P_2 = \frac{x_2}{n_2} = \frac{325}{600} = 0.5416.$$

$$\hat{P} = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = \frac{400(0.5) + 600(0.5416)}{400 + 600}$$

$$\hat{P} = 0.525$$

$q = 1 - 0.525$

$\hat{q} = 0.475$

$H_0: P_1 = P_2$ (They are in favour)
$H_1: P_1 \neq P_2$ (They aren't in favour)

{two tailed test}

$$Z = \frac{P_1 - P_2}{\sqrt{\hat{p}\hat{q}\left(1/n_1 + 1/n_2\right)}}$$

$$Z = \frac{0.5 - 0.5416}{\sqrt{0.525 \times 0.475\left[1/400 + 1/600\right]}}$$

$|Z| = |-1.29| = 1.29 < 1.966$, is the significant value at 5% level of significance. ∴ H_0 is accepted at 5% level of sig

∴ People will accept the proposal.

3) 500 articles from a factory are examined and found to be 2% defective. 800 similiar articles from a second factory are found to have 1.5%. Can it be reasonably concluded that the product of the 1st factory is inferior to those of 2nd.

$P_1 = 0.02 \quad P_2 = 0.015.$

$$\hat{p} = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = \frac{500(0.02) + 800(0.015)}{500 + 800}$$

$\hat{P} = 0.017$

$\hat{q} = 0.983$

H_0: $P_1 = P_2$ (Products of both factory equal)

H_1: $P_1 < P_2$ (Products of factory 2 is greater than factory 1) {Left tailed test}

$$Z = \frac{P_1 - P_2}{\sqrt{\hat{p}\hat{q}(1/n_1 + 1/n_2)}} = \frac{0.02 - 0.015}{\sqrt{0.017 \times 0.983\left(\frac{1}{500} + \frac{1}{800}\right)}}$$

$Z = \cancel{2.445 \; 000}$ $\boxed{Z = 0.6784}$

$Z = 0.6784 > -2.33$ the test of significance value at 1% $\therefore$ H_0 is accepted (as value should be greater for left tailed test).

On basis of total scores 200 candidates of UPSC examinations are divided into 2 groups. The upper 30% and remaining 70%. Consider 1st question of the examination among the 1^{st} group 40 had correct answer whereas among the 2^{nd} group 80 had correct answer. On the basis of this results can I conclude that the 1^{st} question is no good at discrimination ability of the type being examined.

Ans: $n_1 = 60 \quad x_1 = 40 \quad n_2 = 140 \quad x_2 = 80.$

$$p_1 = \frac{40}{60} = \cancel{1.333} = 0.6666 \qquad p_2 = \frac{80}{140} = \cancel{1.1428} = 0.5714$$

$$\hat{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{40 + 80}{60 + 140} = \frac{120}{200}$$

$$\hat{p} = 0.6.$$

$$\hat{q} = 0.4$$

H_0: It is not discriminating

H_1: $P_1 \neq P_2$ It is discriminating

(Two tailed test)

$$Z = \frac{P_1 - P_2}{\sqrt{\hat{P}\hat{Q}\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} = \frac{0.6666 - 0.5714}{\sqrt{0.6 \times 0.4\left[\frac{1}{60} + \frac{1}{140}\right]}}$$

$Z = 1.2598 < 1.966$ is the significance value of 5%. $\therefore$ H_0 is accepted $\therefore$ There will be no discriminating ability of the type being examined.

A company has a the head office at Bengaluru and a branch at Mumbai. The personal director wanted to know if the workers at 2 places would like the introduction at a new plan of work & a survey was conducted for this purpose. Out of the sample of 500 workers are Bengaluru 62% favoured new plant, at Mumbai out of a sample of 400 workers 41% where against the plan. Is there significant difference betw^n the 2 grps in there attitude at 5% level.

$n_1 = 500$ $n_2 = 400$ $P_1 = 0.62$ $P_2 = 0.59.$ (41% against

$\therefore P_2 = 1 - 0.41 = 0.59$)

{two tailed test}.

$$\hat{P} = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = \frac{500(0.62) + 400(0.59)}{500 + 400}$$

$\hat{P} = 0.6066$

$\hat{q} = 0.3934.$

Ho: There is no difference in significant level of 2 grps in attitute. ($P_1 = P_2$)

H_1: There is a difference in significant level of 2 grps in attitude ($P_1 \neq P_2$)

{two tailed test}.

$$Z = \frac{P_1 - P_2}{\sqrt{\hat{P}\hat{Q}\left[1/n_1 + 1/n_2\right]}}$$

$$Z = \frac{0.62 - 0.59}{\sqrt{0.6066 \times 0.3934\left[1/500 + \frac{1}{400}\right]}}$$

$Z = 0.9154 < 1.966$ ∴ ~~There~~ It is significance ~~difference in~~ lev value of 5%. ∴ H_0 is accepted ∴ People will not have change in attitude.

Test of Significance of Single Mean

Here the difference between the sample mean and population mean is significant or not is tested.

→ Under the H_0 there is no difference between sample mean and population mean

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

$\bar{x}$ = Sample Mean, σ = S.D of pop

μ = Population Mean, n = sample

Note: ① If σ is not known then $Z = \frac{\bar{x} - \mu}{(s/\sqrt{n})}$

② here s = S.D of sample.

The limit of population mean μ are given by $\bar{x} - z_\alpha\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{x} + z_\alpha\left(\frac{\sigma}{\sqrt{n}}\right)$

Problems:

A normal population has mean of 6.8 and S.D of 1.5. A sample of 400 members gave a mean of 6.75 is the difference significant?

$\mu = 6.8 \quad \sigma = 1.5 \quad \bar{x} = 6.75 \quad n = 400$

Ho: No significant difference between sample of population mean $(\bar{x} = \mu)$

H1: There is a difference between sample of population mean i.e $(\bar{x} \neq \mu)$

$$|z| = \left|\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}\right| = \left|\frac{6.75 - 6.8}{1.5/20}\right|$$

$$|z| = |0.6667| < 1.966$$

$$|0.6667| < 2.58$$

=> Ho is accepted at 1% and 5% level of significance

=> There is no significant difference between sample mean and populated mean.

2) The mean obtained from a random sample of size 100 and 64g. The S.D of the weight distribution of the population is 3g. Test the statement that the mean weight of population is 67g at 5% level of significance. Also set up 99% confidence limit of the mean weight of the population.

Ans. $\bar{x} = 64$ $n = 100$ $\mu = 67$ $\sigma = 3$

Ho: There is no significant difference betwⁿ population mean and sample mean.

H1: There is significant difference betwⁿ population mean and sample mean.

$$|z| = \left| \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \right| = \left| \frac{64 - 67}{3/10} \right| = 10.$$

$10 > 1.966$

∴ Ho is rejected at 5% level of significance

∴ There is no significant difference betwⁿ mean

$$\bar{x} - z_\alpha \left(\sigma/\sqrt{n}\right) < \mu < \bar{x} + z_\alpha \left(\sigma/\sqrt{n}\right)$$

$$64 - 2.58\left(3/10\right) < \mu < 64 + 2.58\left(3/10\right)$$

$$63.227 < \mu < 64.772$$

3) The avg marks in mathematics of sample of 100 students is 51 with S.D 6 marks could this have been from random population with avg marks 50.

$\mu = 50 \quad \bar{x} = 51 \quad s = 6 \quad n = 100$

$H_0 : \Rightarrow \mu = 50$

$H_1 : \Rightarrow \mu \neq 50$ {two tailed test}

$$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{51 - 50}{6/10} = \frac{10}{6}$$

$$Z = 1.6667 < 1.966$$

$$1.6667 < 2.58$$

$\therefore$ H_0 is accepted at both 1% and 5% level of significance

4) A sample of 900 members has a mean 3.4 cm and S.D 2.61 cm is the sample from the large population of mean 3.25cm and S.D is 2.61cm ; If the population is normal and it's mean is not known find the 95% and 99% confidence limits of true mean

$\bar{x} = 3.4 \quad n = 900 \quad \sigma = 2.61 \quad s = 2.61$

$\mu = 3.25$

H_0: There is no significant difference in $\mu = 3.25$

H_1: There is significant difference ie $\mu \neq 3.25$.

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{3.4 - 3.25}{2.61/30}$$

$$Z = 1.724$$

$\therefore Z = 1.724 < 1.966$

$1.724 < 2.58$

H_0 is accepted at both 1% and 5% level of significance.

$$\bar{x} - Z\alpha\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{x} + Z\alpha\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$3.4 - 2.58\left(\frac{2.61}{30}\right) < \mu < 3.4 + 2.58\left(\frac{2.61}{30}\right)$$

$$3.17 < \mu < 3.624$$

$$\bar{x} - Z\alpha\left(\sigma/\sqrt{n}\right) < \mu < \bar{x} + Z\alpha\left(\sigma/\sqrt{n}\right)$$

$$3.4 - 1.966\left(\frac{2.61}{30}\right) < \mu < 3.4 + 1.966\left(\frac{2.61}{30}\right)$$

$$3.22 < \mu < 3.57$$

~~Test Of Signif.~~ The guaranted average life of a certain type of electric bulbs is 1000 hrs and standard deviation 125 hrs. It is decided to sample the o/p so as to ensure that 90% of the bulbs don't fall short of the garantied average by more than 2.5%. what must be the minimum size of sample.

$\mu = 1000 \quad \sigma = 125$

Since we don't want the sample mean to be less than guaranted average mean by more than 2.5% we have $\bar{x} > 1000 - \frac{2.5(1000)}{100}$

ie $\bar{x} > 975$.

Let n be the size of the sample and $z = \frac{\bar{x} - \mu}{(\sigma/\sqrt{n})}$

$$z = \frac{\bar{x} - \mu}{(\sigma/\sqrt{n})} = \frac{975 - 1000}{125/\sqrt{n}} = \frac{-\sqrt{n}}{5}$$

$$P(z > -\sqrt{n}/5) = 0.9.$$

$$0.5 + A\left(-\frac{\sqrt{n}}{5}\right) = 0.9$$

$$A(-\sqrt{n}/5) = 0.4$$

$$A(1.28) = 0.4$$

$$\frac{\sqrt{n}}{5} = 1.28 \qquad \sqrt{n} = 6.4$$

$$\boxed{n = 40.96} \approx 41$$

Test of significance for difference of Mean

Let $\bar{x}_1$ be the mean of a sample of size n_1 and $\bar{x}_2$ be the mean of the sample

Of size n_2. with the population mean μ
and variance $(\sigma_1)^2$ and population mean
μ_2 and variance $(\sigma_2)^2$ respectively. Then

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

Under the null hypothesis that the sam
are drawn from same population ie $\mu_1 =$
and $\sigma_1 = \sigma_2 = \sigma$ then $Z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

NOTE: 1) If the σ_1 and σ_2 are not know
and $\sigma_1 \neq \sigma_2$ $Z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ //

2) If $\sigma_1 = \sigma_2 = \sigma$ and σ is not known

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}\right) \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$ //

The means of 2 single large samples of 10
and 2000 members are 67.5 inches and
inches respectively. Can the sample be regar

$n_1 = 1000$ $n_2 = 2000$ $\bar{x}_1 = 67.5$ $\bar{x}_2 = 68$ $\sigma = 2.5$

Ho: Samples are from same population

H1: Samples are not from same population.

{two tailed test}

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{67.5 - 68}{2.5\sqrt{\frac{1}{1000} + \frac{1}{2000}}}$$

$$|Z_k| = |-5.163| > 1.966$$

∴ Ho is rejected at 5% level of significance

Intelligence test where given 2 groups of boys and girls.

	Mean	S.D	Sample size
Girls	75	8	60
Boys	73	10	100.

Examine if there is difference betwn the mean score which is significant!

$\bar{x}_1 = 75$ $\bar{x}_2 = 73$ $S_1 = 8$ $S_2 = 10$ $n_1 = 60$ $n_2 = 100$.

Ho: There is a ~~diff~~ no difference betwn the mean score

H1: There is a significant difference betwn the mean score

{two tailed test}

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{75 - 73}{\sqrt{\frac{64}{60} + \frac{100}{100}}}$$

$Z = 1.39 < 1.966$ there is no significant difference between the mean scores.

$\therefore$ H_0 is accepted.

Q) The mean height of 50 male students who showed above average participation in college atheletics was 68.2 inches with the standard deviation of 2.5 inches, while 50 male students who showed no interest in such participation has a mean height of 67.5 inches with the std deviation of 2.8 inches. Test the hypothesis that male students who participate in college atheletics are taller than other male students. ii) By how much should the sample size of each of 2 groups be increased in order that the observed difference of 0.75 inches in the mean height be significant at 5% level of significance

10m

Ans:

Ans:

$n_1 = 50$ $n_2 = 50$ $\bar{x}_1 = 68.2$ $\bar{x}_2 = 67.5$

$s_1 = 2.5$ $s_2 = 2.8$

H_0: There is no differente between the height.

H_1: There is significant difference between the height ($\bar{x}_1 > \bar{x}_2$) (Right tailed test)

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{68.2 - 67.5}{\sqrt{\frac{6.25}{50} + \frac{7.84}{50}}}$$

$$z = 1.32$$

$z = 1.32 < 1.645$, ∴ The H_0 is accepted at 5% level of significance. ∴ There is no significant difference in heights.

$$z > 1.645$$

$$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2 + s_2^2}{n}}} > 1.645$$

(The difference betw^n the mean heights of 2 groups each of size n will be significant if $z \geq 1.645$)

$$\therefore \frac{68.2 - 67.5}{\sqrt{\frac{6.25 + 7.84}{n}}} > 1.645$$

$$\frac{0.7\sqrt{n}}{3.753} > 1.645$$

$$\sqrt{n} > \frac{1.645 \times 3.753}{0.7}$$

$$\sqrt{n} > 8.81955 \quad \therefore n > 77.78$$

$$\boxed{n = 78}$$

~~Difference between rejection~~ Hence sample size of 2 groups should be increased by ~~78~~ $-50 = \boxed{28}$ in order that the difference betwn mean heights of 2 groups is significant.

Test of significance for difference of std deviation.

If s_1 and s_2 are standard deviation of 2 independent samples then under the null hypothesis that sample S.D don't differ signif

$$Z = \frac{s_1 - s_2}{\sqrt{\frac{\sigma_1^2}{2n_1} + \frac{\sigma_2^2}{2n_2}}}$$

where σ_1 = S.D of population

σ_2 = S.D of population

NOTE: If population of S.D. is not known then

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

Random samples drawn from 2 countries gave the following data relating to the height of adults.

	A	B
Mean Height (inches)	67.42	67.25
S.D	2.58	2.50
No of members	1000	1200.

Is the difference between the means are significant.

Is the difference between the S.D significant

i) Ho: There is no significant difference in the mean

H1: There is significant difference in mean

{Two tailed test}

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

$$= \frac{67.42 - 67.25}{\sqrt{\frac{(2.58)^2}{1000} + \frac{(2.50)^2}{1200}}}$$

$$Z = 1.56$$

$Z = 1.56 < 1.966$ $\therefore$ Ho is accepted at the 5% level of significance.

ii) Ho: There is no significant difference in S.D

H1: There is a significant difference in S.D.

$$Z = \frac{s_1 - s_2}{\sqrt{\frac{s_1^2}{2n_1} + \frac{s_2^2}{2n_2}}} = \frac{2.58 - 2.5}{\sqrt{\frac{(2.58)^2}{2000} + \frac{(2.5)^2}{2400}}}$$

$Z = \cancel{0.73}$. 1.088

$\therefore$ $Z = 1.038 < 1.966$ $\therefore$ H0 is accepted at 5% level of significance.

—x—x—x—x—x—x—x—