Unit 4

Testing of Hypothesis

Population:-

A group of Individuals consider under study is called population

Sampling:-

A finite subset of a population is called sample. And the process of such selection of such samples is called sampling.

Types of Sampling:-

- Purposive Sampling
- Random Sampling.
- Simple Sampling
- Systematic Sampling

Random Sampling:-

It is the one in which each unit of population has an equal of chance of being Included in it.

(X) **Parameters and Statistics:-**

The statistical constant of the population such as mean (μ) and standard deviation (σ) is called Parameters

The mean ($\bar{x}$) and standard deviation (s) of sample are known as statistics.

SMART
Date.
Page No.

* Standard Error:

The SD of the sampling distribution of the Statistics is known as standard error.

If,

n - Sample size

σ^2 - Population Varience

s^2 - Sample Varience.

P - Population Proportion

$Q = 1 - P$.

Statistics	Standard Error
$\bar{x}$	$\sigma / \sqrt{n}$
s	$\sqrt{\sigma^2 / 2n}$
s^2	$\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}$
p	$\sqrt{PQ/n}$
$P_1 - P_2$	$\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}$

2m Statistical hypothesis and Null hypothesis:-

↳ Null hypothesis is a hypothesis of no difference i.e., there is no significance difference between the observed value and the expected value.

↳ Denoted by H_0.

Alternate hypothesis:-

↳ Any hypothesis which is complementary to null hypothesis is called alternative hypothesis

↳ Denoted by H_1.

* **Critical Region:-**

A region corresponding to the statistics 't' in the sample space 'S' which amounts to rejection of the null hypothesis H_0 is called critical Region or Region of Rejection.

* **Critical Value (or) Significant Value:**

The Value of the test statistics which separates the critical region from the acceptance region is called critical Value or significant Value.

* **Type-1 and Type-2 Error:**

Type-1 Error:

↳ If H_0 is rejected while it should have been accepted is called Type-1 Error.

Type-2 Error:-

↳ If H_0 is accepted while it should have been rejected is called Type-2 Error.

Confidence ~~low~~ limit:

↳ If the sample statistics lies in the interval $\mu - 1.96\sigma$ and $\mu + 1.96\sigma$, We called 95% confident limit.

↳ If the sample statistics lies in the interval $\mu - 2.58\sigma$ and $\mu + 2.58\sigma$, We called 98% confident limit.

Critical Value	level of significance (α) 1%	5%	10%
2-tail	2.58	1.96	1.645
Right tail	2.33	1.645	1.28
Left tail	-2.33	-1.645	-1.28

Large sample test : (single mean):-

If $x_1, x_2 \ldots x_n$ is a random sample of size n from an normal population with mean μ and standard deviation σ and $\bar{x}$ is the mean of sample then the test.

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

(here we test single mean so standard deviation of sample and population is same $\sigma = s$)

Problem

1) A random sample of 200 tins coconut oil gave an average weight of 4.95 kgs. with a SD of 0.21 kg. i) To be accept the net weight 5kg per tin at 5% level.

ii) To be accept that net weight is 5 kg/tin at 1% level.

Given :-

$n = 200$; $\bar{x} = 4.95$; $s = \sigma = 0.21$

To check:-

$\mu_0 = 5$

Parameter of μ:

$H_0 : \mu = \mu_0 = 5$

$H_1 : \mu \neq \mu_0 = 5$

Level of significance (α): 5% and 1%.

test statistics: $z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{4.95 - 5}{0.21 / \sqrt{200}} = -3.37$

$|z| = 3.37$

Significant ~~test~~ level 5%.

$|z| = 3.37 > \underline{1.96} \rightarrow$ 5% in 2 tail

H_0 is rejected.

Significant level 1%.

$|z| = 3.37 > \underline{2.58} \rightarrow$ 1% in 2 tail

H_0 is rejected

2) A sample of 900 members has a mean of 3.4 cms and SD of 2.61 cms is the sample from large population of mean 3.25 cm and SD 2.61 cm. Find 95% confident limit.

Given:

$n = 900$; $\bar{x} = 3.4$; $S = \sigma = 2.61$

To check:

$\mu_0 = 3.25$

Parameter of μ:-

$H_0 \Rightarrow \mu = \mu_0$

$H_1 : \mu \neq \mu_0$

Level of significant 5% (95% confident limit)

$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{3.4 - 3.25}{\frac{2.61}{\sqrt{900}}} = \frac{0.15}{0.087}$

$|z| = 1.7241$

Significant level 5% $\Rightarrow |z| = 1.7241 > 1.96$

Ho is accepted.

95% Confidence limit

$\bar{x} \pm \alpha$(Standard error)

$\bar{x} \pm 1.96\ \sigma/\sqrt{n}$

$3.4 \pm 1.96\ (2.6/\sqrt{900})$

3.4 ± 0.1705

3.57 and 3.229.

3. An insurance agent has claimed average age of the policy holders who insure through him is less than the average of all agents which is 30.5 yrs. A random sample of 100 policy holders who had insured through him reveal that the mean and SD are 28.8 and 6.35 yrs respectively. test the claim at 5% level of Significance.

Soln:-

$n = 100,\ \bar{x} = 28.8,\ \sigma = s = 6.35,\ \mu_0 = 30.5$

Parameter μ

$H_0: \mu < 30.5$ (left tail test) $\mu = \mu_0$

$H_1: \mu \neq 30.5\ \ \mu < 30.5$

Level of significance $\alpha = 5\% = 0.05$

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{28.8 - 30.5}{6.35/\sqrt{100}}$$

$$= \frac{-1.7}{0.63} = -2.698$$

$|z| = 2.698$

Significant level 5% $\Rightarrow |z| = 2.698 > -1.645$

Ho is rejected.

4. A machine is expected to produce nails of length 2cm a random sample of 25 nails gave avg length of 2.1cm with SD 0.25 cm can it be said that the machine is producing nails as per specification?

Soln:

$n = 25$, $\bar{x} = 2.1$, $\sigma = S = 0.25$, $\mu_0 = 2$

Parameter μ.

$H_0 : \mu = \mu_0$

$H_1 = \mu \neq \mu_0$

$\alpha = 5\%$.

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{2.1 - 2}{0.25/\sqrt{25}} = \frac{0.1}{0.05} = 2.$$

$|Z| = 2$

Significant level 5% $\Rightarrow |Z| = 2 > 1.96$

H_0 is rejected, at 5% level (in 1% level H_0 is accepted $2 < 2.58$).

Test of significance of difference of mean:

If $\bar{x}_1$ and $\bar{x}_2$ are samples from the population mean μ_1 and μ_2 and σ_1 and σ_2 are the standard deviation of the 2 population, n_1 & n_2 are the size of 2 samples then the test statistic is

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

check $\bar{x}_1$ & $\bar{x}_2$ are taken from same population using $\mu_1 = \mu_2$

PROBLEM:-

1. The sales manager of the large company conducted a sample survey in state A and B, taking 400 samples in each case test whether avg sale is same at 1% level.

	State A	State B
Avg sale	2500	2200
SD	400	550

Soln:-

$\bar{x}_1 = 2500$, $\bar{x}_2 = 2200$, $n_1 = n_2 = 400$

$\sigma_1 = s_1 = 400$, $\sigma_2 = s_2 = 550$

The parameters are $\mu_1 \mu_2$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

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

$$= \frac{2500 - 2200}{\sqrt{\frac{400^2}{400} + \frac{550^2}{400}}}$$

$$= \frac{300}{\sqrt{\frac{462500}{400}}} = \frac{300}{34.00} = 8.823$$

Significant level 5% $\Rightarrow |z| = 8.823 > 1.96$

Ho is rejected.

Problem:-

A buyer wants to decide which of the 2 brands of electric bulbs. He buys 200 bulbs from brands A and B. By using these bulbs, he find that brand A has a meanlife 1400 hrs with SD 60hrs and brand B has meanlife 1250hrs. with SD of 50 hrs. Do these two brands differs significantly in quality? Test for 5% level of significant.

Soln:-

$n_1 = n_2 = 200$

$\sigma_1 = S_1 = 60$, $\sigma_2 = S_2 = 50$

$\bar{x}_1 = 1400$, $\bar{x}_2 = 1250$

Parameters are μ_1, μ_2

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{1400-1250}{\sqrt{\frac{60^2}{200} + \frac{50^2}{200}}} = \frac{150}{\sqrt{\frac{3600+2500}{200}}}$$

$$= \frac{150}{\sqrt{\frac{6100}{200}}} = \frac{150}{5.5226} = 27.17$$

$|z| = 27.17$

Significant level 5% $= \alpha = 0.05 \Rightarrow |z| = 27.17 > 1.96$

H_0 is rejected.

Date.
Page No.

2. The average hourly wages of a sample of 150 workers in a plant A was Rs. 2.56 its SD is Rs. 1.08. The avg wages of a sample of 200 workers in plant B was Rs 2.87 and SD is Rs 1.28. Can we assume that the wages paid by plant B is higher than that of plant A.
(level of significant not given so 5%)

Soln:-

$n_1 = 150$; $\bar{x}_1 = 2.56$; $\sigma_1 = 1.08$

$n_2 = 200$; $\bar{x}_2 = 2.87$; $\sigma_2 = 1.28$

$\alpha = 0.05$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 < \mu_2$

Test statistics $= \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

$= \dfrac{-0.31}{\sqrt{\dfrac{\cancel{1.664}\ 1.1664}{150} + \dfrac{1.63\cancel{9}84}{200}}}$

$\cancel{= \dfrac{-0.31}{\sqrt{0.0110 + 0.0081}}}$ $= \dfrac{-0.31}{\sqrt{0.00776 + 0.008192}}$

$\cancel{= \dfrac{-0.31}{0.1129}}$ $= \dfrac{-0.31}{0.12636}$

$= \cancel{2.7457}$

$= 2.46 > 1.645$

0.00776
0.008192
$\sqrt{0.01595}$
0.31

Ho is rejected

Scanned with OKEN Scanner

Page No

Test of Significance for proportions using

If 'x' is the no of success in 'n' independent trials with constant probability 'P' of Success of each trial. Then, $Q = 1-P$ is the probability of failure. Mean $= n(P)$, Variance $= n(P)(Q)$.

Then, the ~~st~~ test statistics,

$$Z = \frac{p-P}{\sqrt{\frac{PQ}{n}}}$$

Here, $p = S/n$.

PROBLEM

1) In a sample of 1000 people in Maharashtra 540 are Rice eaters and the rest are wheat eaters. Cane we assume that both Rice and wheat are equally popular in the state 1% of level of significance.

Soln:-

$n = 1000$, $x = 540$, $p = 0.54$ (x/n).

$P = 1/2$, $Q = 1/2$

H_0 : $P = \cancel{0.05}\ 0.5$ } → equally popular so (1/2)

H_1 : $P \neq \cancel{0.05}\ 0.5$

$\alpha = 0.01$

Test statistics $= \dfrac{p-P}{\sqrt{PQ/n}}$.

Scanned with OKEN Scanner

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$$

$$= \frac{0.54 - 0.5}{\sqrt{\frac{1/4}{1000}}}$$

$$= \frac{0.04}{\sqrt{\frac{1}{4000}}}$$

$$= \frac{0.04}{\sqrt{0.00025}}$$

$$= \frac{0.04}{\cancel{0.1581}} \quad 0.0158 = 0.\cancel{253}$$

$$= 2.53 < \boxed{2.58} \rightarrow \text{table Value}$$

Ho is accepted.

3) 20 People were attacked by diesease. 18 are Survived; Will you accept the hypothesis, that the survival rate if attacked by the diesase is 85% in favor of hypothesis at 5% level.

Soln:-

$n = 20, \quad x = 18, \quad p = 18/20 = 0.9,$

$P = 85, \quad Q = 15, \quad \alpha = 0.05.$

$H_0 : P = 85$

$H_1 : P \neq 85$

Test statistics, $Z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}}$

$$= \frac{0.9 - \cancel{85}\, 0.85}{\sqrt{\frac{85 \cdot \cancel{15}\, 3}{\cancel{20} \cdot 4}}}$$

$$= \frac{\cancel{.847}\ 0.05}{\sqrt{63.75}}$$

$$= \frac{\cancel{.847}}{7.9843}$$

$$= 10.533$$

$$= \frac{0.05}{\sqrt{\frac{0.85 \times 0.15}{20}}}$$

$$= \frac{0.05}{\sqrt{0.006375}}$$

$$= \frac{0.05}{0.07968}$$

$$\cancel{= 0.6273 > 2}$$

$$= 0.627$$

Significant level 5% $\Rightarrow |z| = 0.63 < 1.96$, H_0 is accepted

4) A coin is tossed 144 times and a person 80 heads. Can you say the coin is unbiased? (H_0 is unbiased, H_1 is biased).

Soln:-

$n = 144$, $x = 80$, $p = \frac{80}{144}$

$P = 1/2$, $Q = 1/2$, $\alpha = 0.05$.

H_0 : P is unbiased

H_1 : P is biased

Test statistics, $Z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}}$

$$= \frac{0.55 - 0.5}{\sqrt{\dfrac{1}{576}}}$$

$$= \frac{0.05}{\sqrt{0.0017}}$$

$$= \frac{0.05}{0.0412}$$

$$|z| = 1.2$$

Significant level 5% $\Rightarrow |z| = 1.32 < 1.96$

H_0 is accepted.

PROBLEM:-

1. During a countrywide investigation, the incidence of TB was found to be 1%. In a college of 400 students 5 were reported to be affected. In the other college of 1200 10 were reported to be affected. Thus, this indicate any significance difference.

Sdn:-

$p = 1\% = 0.01$, $x_1 = 5$, $x_2 = 10$, $Q = \cancel{99.99}$ 0.99.

$n_1 = 400$; $n_2 = 1200$

$P_1 = x_1/n_1$; $P_2 = x_2/n_2$

$P_1 = \cancel{5}^{1}/\cancel{400}^{80}$; $P_2 = \cancel{10/1200}$ $1\cancel{0}/120\cancel{0}$

$P_1 = 0.0125$; $P_2 = 0.0083$

$H_0: P_1 = P_2$

$H_1: P_1 \neq P_2$

$\alpha = 0.05$

Test statistics, $z = \dfrac{p_1 - p_2}{\sqrt{(0.01)(0.99)\left(\dfrac{1}{400} + \dfrac{1}{1200}\right)}}$

$= \dfrac{0.0125 - 0.0083}{\sqrt{0.0099\left(0.0033\right)}}$ 0.0025 + 0.0008

$= \dfrac{0.0125 - 0.0083}{\sqrt{0.00003267}}$

$= \dfrac{0.0042}{0.005715\,76766}$

$= 0.734 < 1.96$

H_0 accepted

2. Random sample of 400 men and 600 Women are asked if they would like to have a flyover near their residence. 200 mean and 325 Women were in favor of the proposal test the hypothesis that the proportion mean and women in favor of the proposal of sale against that they are not at 5% levels.

NOTE: 'P' not given, So, $\boxed{P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}}$

Soln:-

$n_1 = \cancel{200}\ 400$; $n_2 = \cancel{325}\ 600$;; $x_1 = 200$; $x_2 = 325$

$P_1 = x_1 / n_1$; $P_2 = x_2 / n_2$

$P_1 = 200/400$; $P_2 = 325/600$.

$P_1 = 0.5$; $P_2 = 0.5416$, $\alpha = 0.05$

3.

$$P = \boxed{\frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}} = \frac{400 \cdot (0.5) + 600(0.5416)}{1000}$$

$$= \frac{200 + 324.96}{1000}$$

$$= \frac{524.96}{1000}$$

$$P = 0.52496$$

$$Q = 1 - 0.52496$$

$$= 0.47504$$

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.5 - 0.5416}{\sqrt{(0.52496)(0.47504)\left(\frac{1}{400} + \frac{1}{600}\right)}}$$

$$= \frac{-0.0416}{\sqrt{0.2493769984\left(0.004166666667\right)}}$$

$$= \frac{0.0416}{\sqrt{0.0010390708 3}}$$

$$= \frac{0.0416}{0.0322346216}$$

$$= 1.29053 < 1.96$$

H_0 is accepted.

A machine produced 20 defective units in a sample of 400 after over handling it produced 10 defective in a batch of 300. Has the machine improved its production due to overhandling. test at 5% level.

Soln:-

$n_1 = 400,\ x_1 = 20,\quad n_2 = 300,\ x_2 = 10.$

$P_1 = x_1/n_1 \ ;\ P_2 = x_2/n_2$

$P_1 = 0.05 \ ;\ P_2 = 0.033$

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = \frac{400(0.05) + 300(0.033)}{700}$$

$$= \frac{\overset{20}{\cancel{8}} + 9.9}{700} = \cancel{0.0255}\ 0.04271$$

$Q = 1 - P = 1 - 0.0255 = \cancel{0.9745}\ 0.95729$

$$Z_t = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$= \frac{0.05 - 0.033}{\sqrt{0.040885855\left(\frac{1}{400}+\frac{1}{300}\right)}}$$

$$= \frac{0.017}{\sqrt{0.040885855\,(0.005833)}}$$

$$= \frac{0.017}{\sqrt{0.00023884879}}$$

$$= \frac{0.017}{0.01544305345}$$

$$= 1.1008 < 1.96$$

H_0 is accepted

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from a normal distribution with population mean μ and variance σ^2, then the test statistics,

$$t = \frac{\bar{x} - \mu}{s\sqrt{n}} \quad (\text{or}) \quad \frac{\bar{x} - \mu}{SD/\sqrt{n-1}}$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Degrees of Freedom: $\gamma = n-1$

PROBEM

1. A sample of 10 house owners is drawn and the following Values of their incomes are obtained mean is Rs. 6000 and SD is 650. Test the hypothesis that the average income of the house owners of the town is ~~5500~~ Rs. 5500.

Soln:-

$n = 10$ (sample size is small)

$\bar{x} = 6000$, $SD = 650 = \sigma$

$H_0 = \mu_0 = 5500$

$H_1 : \mu_1 \neq \mu_0$.

$\alpha = 0.05$

degrees of freedom ~~[illegible]~~

$t = n-1$

$= 9$.

Test statistics, $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n-1}}}$

$= \dfrac{6000 - 5500}{\frac{650}{3}}$

$= \dfrac{500}{216.6666}$

$= 2.307 > 2.262$

Ho is rejected at 5%. level.

SMART

Date.

Page No. ..../......./.......

The average strength of a steel rod is specified to 18.5, 1000 pounds. Test that A Sample of 14 rods are tested the mean and SD are obtained to be 17.85 and 1.955 respectively Is the result of the experiment significant with 95% Confidence.

Soln:-

$n = 14, \quad \bar{x} = 17.85, \quad \sigma = S.D = 1.955$

$\mu = 18.5$

degrees of freedom, $t = n-1$

$= 14-1 = 13.$

Test statistics, $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n-1}}}$

$= \dfrac{17.85 - \cancel{1.955}\ 18.5}{\frac{1.955}{\sqrt{13}}}$

$= \dfrac{-0.65}{0.5422}$

$|z| = 1.2 < 2.16$

Ho is ~~rejected~~ accepted

Date.
Page No.

A random sample of 16 values from a normal population showed a mean of 41.5 inches and the sum of squares of deviation from this mean is equal to 135 square inches. Show that the assumption of a mean of 43.5 inches for the population is not reasonable. Obtain 95% and 99% Fiducial Limit:-

Soln:-

$n=16$, $\bar{x}=41.5$

$\Sigma(x_i-\bar{x})^2=135$

$S=\sqrt{\frac{135}{15}}=\sqrt{9}=3$, $\mu_0=43.5$

$\alpha=1\%$ & 5%, $t=15$

$H_0 = \mu_0=43.5$

$H_1 : \mu_1 \neq 43.5$

Test Statistics,

$$Z=\frac{\bar{x}-\mu}{S}$$

$$=\frac{41.5-43.5}{3/\sqrt{16}}$$

$$=\frac{2}{3}\times 4$$

$$=2.66$$

$t=2.66<2.947$ at 1% level

$\therefore$ Ho is accepted at 1% level

$t=2.66>2.131$ at 5% level

$\therefore$ Ho is rejected at 5% level.

deviation
T Test for difference of mean

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1}+\left(\frac{1}{n_2}\right)\right)}}$$

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

Problem:-

1. The average no. of articles produced by 2 machines /day are 200 & 250 with SD 20 & 25 respectively on the basis of records of 25 days production. Can you regret both the machines equally efficient at 1% level equivalent.

Soln:-

$\bar{x}_1 = 200$, $\bar{x}_2 = 250$

$n_1 = 25$ $n_2 = 25$

$s_1 = 20$ $s_2 = 25$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$\alpha = 0.01$

degree of freedom $= n_1 + n_2 - 2$

$= 25 + 25 - 2$

$= 48$

$$s^2 = \frac{25(20)^2 + 25(25)^2}{25+25-2}$$

$$= \frac{10000 + 15625}{48} = 533.85$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$= \frac{200-250}{\sqrt{533.85\left(\frac{1}{25}+\frac{1}{25}\right)}}$$

$$= \frac{-50}{\sqrt{533.85(0.08)}}$$

$$= \frac{-50}{\sqrt{42.708}}$$

$$= \frac{-50}{6.535}$$

table Value
$= 2.58$

$t = -7.65$

2. A group of 10 rats are fed by diet A and another group of 8 rats are fed by diet B.

Diet A	5	6	8	1	12	4	3	9	6	10
Diet B	2	3	6	8	10	1	2	8.		

If there significant difference b/w diet A & diet B at 5% level?

$\boxed{\text{Soln}}$:-

x_1	x_2	x_1^2	x_2^2
5	2	25	4
6	3	36	9
8	6	64	~~12~~ 36
1	8	1	64
12	10	144	100
4	1	16	1
3	2	9	4
9	8	81	64
6		36	
10		100	
		512	

$$\bar{x}_1 = \sum_{i=1}^{n} = \frac{64}{10} = 6.4$$

$$\bar{x}_2 = \frac{\sum x_2}{n_2} = \frac{40}{8} = 5$$

$$s_1^2 = \frac{\sum x_1^2}{n_1} - (\bar{x}_1)^2$$

$$= \frac{512}{10} - (6.4)^2$$

$$= 10.24$$

$$s_2^2 = \frac{\sum x_2^2}{n_2} - (\bar{x}_2)^2$$

$$= \frac{282}{8} - 25$$

$$= 10.25$$

$n_1 = 10, \quad n_2 = 8$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$\alpha = 0.05$

$\gamma = n_1 + n_2 - 2$

$= 10 + 8 - 2$

$= 16$

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

$= 11.52$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{64 - 5}{\sqrt{11.5\left(\frac{1}{10} + \frac{1}{8}\right)}}$$

$$= \frac{1.4}{\sqrt{11.5\left(\frac{18}{80}\right)}}$$

$$= \frac{1.4}{\sqrt{11.52(0.225)}}$$

$$= \frac{1.4}{\sqrt{2.592}}$$

$$= \frac{1.4}{1.609}$$

$t = 0.870105655$

$0.86 < 1.746$

H_0 is accepted.

3. The following table gives the Values of ~~I~~ protein from cow milk and buffalo milk. Examine if these difference are significant.

Cow	1.9	1.95	2	2.02	1.85	1.8
Buffalo	2.12	2.00	2.20	2.45	2.20	2.10

Soln:

$n_1 = 6, \ n_2 = 6$

$\bar{x}_2 = \frac{13.07}{6} = 2.176$

$\bar{x}_1 = \frac{11.52}{6} = 1.92$

$s_1^2 = \frac{22.155}{6} - (1.92)^2$

$= 0.0061$

$s_2^2 = \frac{25.58}{6} - (2.17)^2$

$= \cancel{0.01936}\ 0.055$

$n_1 = 6, \ n_2 = 6$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$\alpha = 0.05$

$\gamma = (n_1 + n_2) - 2$

$= 10$

$S^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{10}$

$= \frac{6(0.0061) + 6(0.055)}{}$

$$= \frac{0.0366 + 0.33}{10}$$

$$= 0.03666$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{1.92 - 2.178}{\sqrt{0.036\left(\frac{1}{6} + \frac{1}{6}\right)}}$$

$$= \frac{-0.258}{\sqrt{0.036(0.333)}}$$

$$= \frac{-0.258}{\sqrt{0.011988}}$$

$$= \frac{-0.258}{0.1094897255}$$

$$t = +2.3563854846 > 2.228$$

H_0 is rejected.

SMART

Date.

Page No.

4. A random Sample of 10 boys had the following IQ's 70, 120, 110, 101, 88, 83, 95, 98, 107, 100. Do these data support the assumption of the population mean IQ 100? Find the reasonable range in which most of the mean IQ values of sample of 10 boys lie?

Soln:-

Date
Page No

Date.
Page No.

Assumptions made while using P-test:-

i) Parent Population from which samples have been drawn or normally distributed.

~~ii) Population Var~~

ii) Population Varience are equally unknown.

iii) Two samples are random and independent

F-test/distribution:-(Test for Varience):-

If s_1^2

~~F-test:-~~

Applications of F-distribution:-

↳ F-test is used to test if two independent samples have been drawn from the normal population with the same varience.

↳ Whether the 2 independent estimation of the population varience are homogenous or not.

Problem :-

1) Pumpkins were grown under 2 experimental conditions. Two samples of 11 and 9 Pumpkins show the sample standard deviation of their weight as 0.8 and 0.5 respectively. Assume, the weight distributions are normal. Test the hypothesis that the variance are equal at 10% level.

Soln:-

$n_1 = 11, \quad n_2 = 9, \quad s_1 = 0.8, \quad s_2 = 0.5$

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

$\alpha = 0.10$,

$\nu_1 = 11-1 = 10, \quad \nu_2 = 9-1 = 8$.

$$S_1^2 = \frac{n_1 s_1^2}{n_1 - 1}$$

$$= \frac{11 \times (0.8)^2}{10} = \frac{11 \times 0.64}{10} = \frac{7.04}{10} = 0.704$$

$$S_2^2 = \frac{n_2 (s_2)^2}{n_2 - 1}$$

$$= \frac{9 \times (0.5)^2}{8} = \frac{9 \times 0.25}{8} = 0.28125$$

$$F = \frac{S_1^2}{S_2^2}$$

$$= \frac{0.704}{0.28125}$$

$$= 2.5031 < 3.35$$

H_0 is accepted at 10% level.

SMART
Date.
Page No.

2. In 1 sample of 8 observations the sum of square of deviations of a sample values mean $\sum(x_1-\bar{x})^2$ is 84.4 and in the other sample of 10 observations it was ~~102.6~~ 102.6. Test the significance at 5% level by using F-test.

Soln:-

$n_1 = 8, \quad n_2 = 10,$

$\sum(x_1^i - \bar{x})^2 = 84.4$

$\sum(x_2^i - \bar{x})^2 = 102.6$

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

$\alpha = 0.05$

$\nu_1 = 7, \quad \nu_2 = 9$

$S_1^2 = \frac{84.4}{7}$

$= 12.057$

$S_2^2 = \frac{102.6}{9}$

$= 11.4.$

$F = \frac{S_1^2}{S_2^2} = \frac{12.057}{11.4} = 1.0576.$

Scanned with OKEN Scanner

2) Two independent sample of 9 and 7 ~~of~~ from a normal population have the following values.

x_1	18	13	12	15	12	14	16	14	15
x_2	16	19	13	16	18	13	15		

Do the estimation of the population variance differ significantly at 5% level.

x_1	x_2	x_1^2	x_2^2
18	16	324	256
13	19	169	361
12	13	144	169
15	16	225	256
12	18	144	324
14	13	196	169
16	15	256	225
14		196	
15		225	

$\Sigma x_1^2 = 1879$

$\Sigma x_2^2 = 1760$

$\Sigma x_1 = 129$

$\Sigma x_2 = 110$

$$\bar{x}_1 = \frac{\Sigma x_1^i}{n_1} = \frac{129}{9} = 14.33$$

$$\bar{x}_2 = \frac{\Sigma x_2^i}{n_2} = \frac{110}{7} = 15.714$$

$$S_1^2 = \frac{\Sigma x_1^2}{n_1} - (\bar{x}_1)^2$$

$$S_2^2 = \frac{\Sigma x_2^2}{n_1} - (\bar{x}_2)^2$$

$S_1^2 = 208.77 - 205.34$

$S_1^2 = 3.43$

$S_2^2 = 251.42 - 246.929$

$S_2^2 = 4.491$

$$~~F~~ \; F = \frac{3.43}{\cancel{4.4961}} \qquad \frac{4.491}{3.43}$$

$$\cancel{F = 0.7637} \qquad = 1.309.$$

χ^2 test:-

χ^2 test of goodness of fit $= \sum \frac{(O-E)^2}{E}$

O - Observed frequency
E - expected frequency.

Applications of χ^2 Test:-

i) To test the goodness of fit
ii) To test the independence of attributes
iii) To test if the hypothetical Value of the population Varience is σ^2
iv) " " the homogenity of independent estimates of the population Varience.

Conditions to apply χ^2 Test:-

↳ The sample observation should be Indopentent
↳ 'N' - total frequency should be atleast 50

Degrees of freedom = n-1

SMART
Date.
Page No.

PROBLEM:-

A company keeps records of accident during a recent safety review. A random sample of 60 accidents was selected and classified by the day of the week on which they occured.

Day	Mon	Tues	Wed	Thurs	Fri
No. of Acc.	8	12	9	14	17

Test if there is any evidence, that accidents are more likely on somedays than others.

Soln:-

$n=5, \quad \nu = n-1 = 4,$

Expected frequency $= 60/5 = 12.$

O	E	$(O-E)^2/E$	
8	12	1.33	$(16/12)$
12	12	0	0
9	12	0.75	$(9/12)$
14	12	0.33	$4/12$
17	12	2.08	$(25/12)$
		4.49	

$4.49 < 9.488$

H_0 is accepted at 5% level

2) A Sample analysis of examination results of ~~0~~ 500 Students were made. It was found that 200 students have failed 170 Scored 3^{rd} class 90's scored 2^{nd} class and the rest 1^{st} class. So, these figure support the general belief that the above category are in the ratio 4:3:2:1 ~~resp~~ respectively.

Soln:-

H_0 : The result is in the ratio 4:3:2:1

H_1 : The result is not in the ratio 4:3:2:1

$n=4$, $\nu=3$, $\alpha=0.05$

Expected frequency of ~~1st class~~ failure: $4/10 \times 500$
$=200$

" " of 3^{rd} class: $3/10 \times 500$
$=150$

" " of 2^{nd} class = $2/10 \times 500$
$=100$

" " of 1^{st} class = $1/10 \times 500$
$=50$

O	E	$(O-E)^2/E$	$(O-E)^2/E$
2~~1~~0	200	~~0~~ ~~2.66~~ 2	2
170	150	~~0.02~~ 2.66	2.66
90	100	1	1
20 ~~70~~	50	18	18
		~~1.32~~	23.66

$23.66 > 7.81$

H_0 is rejected at 5% level

χ^2 test to test Independence of Attributes

~~Let us consider~~

Let us consider 2 attributes A and B divided in 'r' classes $A_1, A_2 \ldots A_r$ are and B divided into 's' classes $B_1, B_2, \ldots, B_s$, this is $r \times s$ matrix. And is called $r \times s$ contingency table.

If the independence of attribute is a 2×2 contigency table, then $\chi^2 = \frac{N[ad-bc]^2}{(a+b)(a+c)(b+d)(c+d)}$

where,

Attribute	B_1	B_2
A_1	a	b
A_2	c	d

And, $N = a+b+c+d$.

Problem

1) Find if there is any association btw extra in wages in father and xtra wages in son from the following table.

	xtravages son	Miserly father
xtravagss son	a 327	741 b
Miserly son	c 545	234 d

Determine the co-efficient of xtra wages.

$\boxed{\text{Soln}}$:-

Given, H_0 : Extravages of father and son are equal

H_1 : Extravages of father and son aren't equal

$N = a+b+c+d$

$= 327+741+545+234 = 1847$

975

$$= \frac{1847(327 \times 234 - 741 \times 545)}{(327+741)(327+545)(741+234)(545+234)}$$

545
234
779

$$= \frac{1847(76518 - 403845)}{(1068)(872)(975)(779)}$$

$$= \frac{1847(327327)}{707342594400}$$

$= 604572969$

1000 students at College level are graded according to their IQ at their economic conditions what conclusion can you draw from the following data.

	High	Low
Rich	460	140
Poor	240	160.

Soln:

H_0: The given attributes are independent

H_1: The given attributes are not independent

$\alpha = 0.05$ $\quad$ $\nu = r-1(s-1)$

$= 2-1(2-1)$

$\nu = 1$

Scanned with OKEN Scanner

SMART
Date.
Page No. ..../......./.......

Test statistics, $|\psi^2| = \frac{\Sigma(O-E^2)}{E}$

Expected frequency = $\frac{\text{corresponding row total} \times \text{Column total}}{\text{grand total}}$:

Expected frequency 460 = $\frac{600 \times 700}{1000}$

$= \frac{420000}{1000}$

$= 420$.

Expected frequency 140 = $\frac{600 \times 300}{1000}$

$= \frac{180000}{1000}$

$= 180$.

Expected frequency 240 = $\frac{700 \times 400}{1000}$

$= 280$

Expected frequency, 160 = $\frac{300 \times 400}{1000}$.

$= 120$.

O	E	$(O-E)^2/E$
460	420	3.809
140	180	8.888
240	280	5.714
160	120.	13.333
		31.744

table Value = 3.841

31.744 > 3.841

Ho is rejected

3) ~~tt:~~ From the following information state if the condition of a child is associated with the condition of house. ~~From the~~

Condition of child	Condition of house		
	clean	dirty	total
clean	69	51	
fairly clean	81	20	
dirty	35	44	

Sdn ~~Ho~~

Ho : the given attributes are independent

H_1 : the given attributes are not independent

$\alpha = 0.05$

$\gamma = (r-1)(s-1)$

$= (3-1)(2-1)$

$= 2(1)$

$\gamma = 2$

Test statistics, $|\psi^2| = \Sigma(O-E)^2/E$

Expected frequency of 69 $= \frac{120 \times 185}{300} = \frac{222\cancel{00}}{3\cancel{00}} = 74$

Expected frequency of 51 $= \frac{120 \times 115}{300} = \frac{138\cancel{00}}{3\cancel{00}} = 46$

Expected frequency of 81 $= \frac{\overset{185}{\cancel{120}} \times 101}{300} = \cancel{40.4}\ 62.28$

Expected frequency of $\overset{20}{\cancel{35}} = \frac{115 \times 101}{300} = 38.716$.

Expected frequency of 35 $= \frac{185 \times 79}{300} = 48.716$.

SMART
Date.
Page No. ------/------/------

Expected frequency of 44 $= \frac{115 \times 79}{300} = 30.28$.

O	E	$(O-E)^2/E$
69	74	0.337
51	46	0.543
81	62.28	5.626.
20	38.716	9.047.
35	48.716	3.861
44	30.28	6.216
		25.626

5.991 - table Value.

$25.626 > 5.991$

Ho is rejected.