UNIT 4

Population : set of collection of objects. Population consists of numbers, measurements or observations of interest.

$N \rightarrow$ size of population

Sampling is the process of drawing sample from a population.

large sampling if $n \geq 30$

small sampling if $n < 30$.

Parameters : statistical measure or constants obtained from the population.

eg: mean, variance

Statistics : statistical quantities computed from sample

eg: sample mean, sample variance.

Populatⁿ :	mean (μ)	SD (σ)	proportⁿ (p)
Sample :	" ($\bar{x}$)	" (s)	" (P)

Statistical Inference :

Deals with methods of drawing valid or logical generalizations and predictⁿs about popⁿ using informatⁿ contained in sample alone, with an indication of accuracy of such inferences.

Inference : Process of drawing conclusion about popⁿ parameters based on sample taken from popⁿ.

Hypothesis: Claim or statement about a population parameter that we want to test.

Null Hypothesis H_0: currently accepted value for parameter.

Alternative Hypothesis H_1: involves the claim to be tested.

Significance level: In a hypothesis test, α is the prob. of making wrong decision when the null hypothesis is true.

Confidence level: $C = 1-\alpha$. Prob. that if poll / test were repeated over and over again, the result obtained would be the same.

20/12/21

Test of Hypothesis concerning single populatn mean (μ) for large sample:

Let μ and σ^2 be mean and variance of a populatn from which a random sample of size n is drawn. Let $\bar{x}$ be mean of sample. Then for large samples from central limit theorem, it follows that the sampling distributn of $\bar{x}$ is approx. normally distributed with mean $\mu_{\bar{x}} = \mu$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.

The test statistic for single mean is $Z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$

Q) The length of life x of certain computers is approx. normally distributed with mean 800 hrs and S.D 40 hrs. If a random sample of 30 computers has an avg. life of 788 hrs, test whether there is diff. in avg. value. Test whether $\mu \neq 800$ hrs at $\alpha = 5\%$.

→ x → length of life of computer

$n = 30$

$\bar{x} = 788$ hrs.

$\mu = 800$ hrs

S.D $= 40$ hrs
(σ)

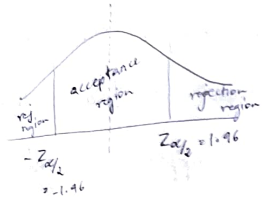
$H_0 : \mu = 800$ hrs

$H_1 : \mu \neq 800$ hrs

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

Decision rule :

At $\alpha = 5\%$ accept H_0 when $|z| < Z_{\alpha/2}$



$0.025 = 0.5 - \phi(Z_{\alpha/2})$

$\phi(Z_{\alpha/2}) = 0.5 - 0.025$

$= 0.475$

$Z_{\alpha/2} = 1.96$

Numerical computat^{n}:

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{788 - 800}{40/\sqrt{30}} = \frac{12}{40/\sqrt{30}} = 1.6432$$

Conclusion : $Z = -1.6432$ $\Rightarrow$ Accept H_0

$|z| = 1.6432 < Z_{\alpha/2} = 1.96$

---

⑤ Mice with an avg. life span of 32 months will live upto 40 months when fed by certain nutritious food. If 64 mice fed on this diet have an avg. life span of 38 months and SD of 5.8 months, is there any reason to believe that life span is less than 40 ?

$\rightarrow$ $x \rightarrow$ lifespan of mice

$n = 64$ $\mu = 40$

$\sigma =$

$\bar{x} = 38$ months

$s = 5.8$ months

(SD of sample)

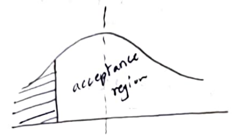
$H_0 : \mu = 40$

$H_1 : \mu < 40$

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

Decision rule :

at $\alpha = 5\%$ accept H_0 when $z > Z_\alpha$



$0.05 = 0.5 - \phi(Z_\alpha)$

$\phi(Z_\alpha) = 0.5 - 0.05$

$\phi(Z_\alpha) = 0.45$

$Z_\alpha = -1.65$

Numerical Computation :

(when S.D of pop^{n} is not given, we take SD of sample)

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{38 - 40}{5.8/\sqrt{64}}$$

$= -2.7586$

Conclusion :

$Z = -2.758 \not> -1.65$ $\Rightarrow$ Reject H_0

---

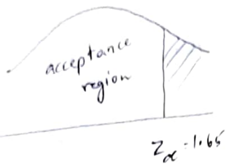
⑥ A machine runs on a avg. of 125 hrs/yr. A random sample of 49 machines has an annual avg. use of 126.9 hrs. with SD 8.4 hrs. Does this suggest to believe that machines are used on the avg., more than 125 hrs annually at 0.05 level of significance.

$\rightarrow$ $X \rightarrow$ run time in hrs/yr of machine

$n = 49$ $\mu = 125$ hr/yr.

$\bar{x} = 126.9$ lbs

$s = 8.4$



$H_0 : \mu = 125$ lbs/yr

$H_1 : \mu > 125$

$\alpha = 0.05$

Decision rule :

at $\alpha = 0.05$, accept H_0 when $Z < Z_\alpha$

$0.05 = 0.5 - \phi(Z_\alpha)$

$\phi(Z_\alpha) = 0.5 - 0.05$

$= 0.45$

$Z_\alpha = 1.65$

Numerical Computation :

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

$Z = \frac{126.9 - 125}{8.4/\sqrt{49}}$

$= 1.5833$

Conclusion :

$1.5833 < 1.65 \Rightarrow$ Accept H_0

---

9) To determine whether the mean breaking strength of synthetic fibre produced by certain company is 8 kilo or not. A random sample of 50 fibres were tested yielding a mean breaking strength of 7.8 kilo. If SD is 0.5 kilo, test at 0.01 level of significance.

→ $n = 50$

$x \rightarrow$ breaking strength of fibre

$\mu = 8$ kilo, $\sigma = 0.5$ kilo

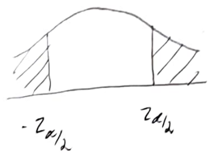
$\bar{x} = 7.8$, $\alpha = 0.01$

$s = \sigma$

$H_0 : \mu = 8$ kilo

$H_1 : \mu \neq 8$ kilo

Decision rule : at $\alpha = 1\%$ accept H_0 if $|Z| < Z_{\alpha/2}$



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

$0.005 = 0.5 - \phi(Z_{\alpha/2})$

$\phi(Z_{\alpha/2}) = 0.5 - 0.005$

$= 0.495$

$Z_{\alpha/2} = 2.58$

Numerical Computat^n :

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

$Z = \frac{7.8 - 8}{0.5/\sqrt{50}} = -2.828$

Conclusion : $|Z| = 2.828 \nless Z_{\alpha/2} = 2.58 \Rightarrow$ reject H_0

---

10) A manufacturer of tyre guarantees that avg. lifetime of its tyre is more than 28000 miles. If 40 tyres of this company tested, yields a mean lifetime of 27463 miles, with SD = 1348 miles, can guarantee be accepted at 0.01 LoS?

→

$x \rightarrow$ lifetime of tyre

$H_0 : \mu \geq 28000$ $\bar{x} = 27463$ $s = 1348$

$H_1 : \mu \ngtr 28000$ $n = 40$ $\alpha = 0.01$

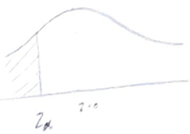
Decision rule:

at $\alpha = 0.01$ accept H_0 when $Z \geq Z_\alpha$

$0.01 = 0.5 - \phi(Z_\alpha)$

$\phi(Z_\alpha) = 0.49$

$Z_\alpha = -2.33$



Numerical Computa^n:

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{27463 - 28000}{1348/\sqrt{40}} = -2.519$$

Conclusion:

$Z = -2.519 \ngtr Z_\alpha = -2.33 \Rightarrow$ reject H_0.

Type-2:

Test of significance for diff. b/w 2 means (for large sample):

x_1 and x_2 follows normal distribut^n.

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$\mu_1 < \mu_2$

$\mu_1 > \mu_2$

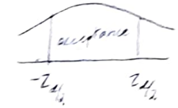
Test statistic $$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

at α % LOS,

accept H_0

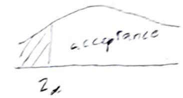
case 1: $H_1 : \mu_1 \neq \mu_2$

accept H_0 when $|Z| < Z_{\alpha/2}$



case 2: $H_1 : \mu_1 < \mu_2$

accept H_0 when $Z > Z_\alpha$



case 3: $H_1 : \mu_1 > \mu_2$ accept H_0 when $Z < Z_\alpha$

---

② In a random sample of 100 tubelights produced by company A, the mean lifetime of tubelight is 1190 hrs with SD of 90 hrs. Also in a random sample of 75 tubelights from company B, the mean lifetime is 1230 hrs with SD of 120 hrs. Is there a diff. b/w mean lifetime of 2 brands of tubelight at significance level of 0.05

→ x_1 → lifetime of tubelight by company A

x_2 → " " " " B.

$n_1 = 100$, $n_2 = 75$

$x_1 = 1190$ hrs, $x_2 = 1230$ hrs

$SD_1 = 90$ hrs, $s_2 = 120$ hrs

$\alpha = 0.05$

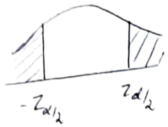
$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

Decision rule:

at $\alpha = 0.05$, accept H_0 when $|z| < z_{\alpha/2}$

$0.025 = 0.5 - \phi(z_{\alpha/2})$



$\phi(z_{\alpha/2}) = 0.475$

$z_{\alpha/2} = 1.96$

Numerical computatn: $Z = \dfrac{(\bar{x}_1 - \bar{x}_2) - \cancelto{0}{(\mu_1 - \mu_2)}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \dfrac{-40}{\sqrt{\dfrac{8100}{100} + \dfrac{(120)^2}{75}}}$

$= \cancel{21.789}$

$= -2.4209$

Conclusion: $|z| = 2.4209 \not< z_{\alpha/2} = 1.96$

$\Rightarrow$ reject H_0

---

5) To test the effects of new pesticide on rice product, a farmland was divided into 60 units of equal areas. All ~~proportions~~ portions have identical qualities as to soil, exposure to sunlight etc. The new pesticide is applied to 30 units while old pesticide to remaining 30. Is there reason to believe that new pesticide is better than old pesticide if mean no. of kilos of rice harvested per unit using new pesticide is 496.31 with SD of 17.18. Test at 0.01 Los.

For old pesticide mean = 485.41 kilo with SD = 14.73 kilo

$\rightarrow$ $x_1 \rightarrow$ new pesticide usage to unit

$x_2 \rightarrow$ old pesticide usage

$\bar{x}_1 = 496.31$ $\quad s_1 = 17.18$ $\quad \alpha = 0.01$

$\bar{x}_2 = 485.41$ $\quad s_2 = 14.73$

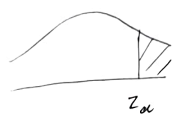
$n_1 = 30$, $n_2 = 30$

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 > \mu_2$

Decision rule: accept H_0 when $z < z_\alpha$

$0.5 - \phi(z_\alpha) = 0.01$



$z_\alpha = 2.33$

Numerical computatn: $Z = \dfrac{(\bar{x}_1 - \bar{x}_2) - \cancelto{0}{(\mu_1 - \mu_2)}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

$= 2.638$

Conclusion: $z \not< z_\alpha$ $\Rightarrow$ reject H_0

---

6) If a random sample data shows that 42 men earn on the avg. 744.85 with SD = 397.7 while 32 women earn on the avg. 516.78 with SD = 162.523. Test at $\alpha = 0.05$, whether avg. income for men and women is same or not.

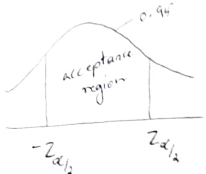
$\rightarrow$ $n_1 = 42$ $\quad \bar{x}_1 = 744.85$, $s_1 = 397.7$, $\alpha = 0.05$

$n_2 = 32$ $\quad \bar{x}_2 = 516.78$, $s_2 = 162.523$

$H_0: \mu_1 = \mu_2$ (avg. income for men & women are same)

$H_1: \mu_1 \neq \mu_2$ (" not same)

Decision rule: at $\alpha = 0.05$, accept H_0 if $|z| < z_{\alpha/2}$



$0.025 = 0.5 - \phi(z_{\alpha/2})$

$\phi(z_{\alpha/2}) = 0.475$

$z_{\alpha/2} = 1.96$

Numerical Computation: $z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

$= 3.3659$

Conclusion:

As $|z| = 3.365 \nless z_{\alpha/2} = 1.96$, reject H_0.

---

⑤ Test at 0.05 LOS, A manufacturer's claim that mean tensile strength (m) of thread A exceeds the metres of thread B by atleast 12 kilo. If 50 pieces of each type of thread are tested under similar condit^ns yielding the following data,

	Sample size	Mts	SD
Thread A	50	86.7	6.28
Thread B	50	77.8	5.61

→ $n_1 = 50$ $\quad \bar{x}_1 = 86.7$ $\quad s_1 = 6.28 = \sigma_1$, $\alpha = 0.05$

$n_2 =$ $\quad \bar{x}_2 = 77.8$ $\quad s_2 = 5.61 = \sigma_2$

$H_0: \mu_1 - \mu_2 \geq 12$

$H_1: \mu_1 - \mu_2 < 12$

Decision rule: at $\alpha = 0.05$, accept H_0 if $z > z_\alpha$

~~Num~~ $\quad 0.05 = 0.5 - \phi(z_\alpha)$

$\phi(z_\alpha) = 0.45$

$z_\alpha = -1.65$

Numerical Computation: $z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

$= -2.6031$

Conclusion: at $\alpha = 0.05$,

$z = -2.603 \ngtr z_\alpha = -1.65 \Rightarrow$ reject H_0.

---

⑧ A random sample of 40 geysers produced by company A have a mean lifetime (MLT) of 647 hrs. of continuous use with SD of 27 hrs, while a sample of 40 geysers produced by another company B has MLT of 638 hrs. with SD 31 hrs. Does this substantiate the claim of company A that their geysers are superior to those produced by company B at

(i) $\alpha = 0.05$

(ii) $\alpha = 0.01$

→ $n_1 = 40$ $\bar{x}_1 = 647$ hrs $s_1 = 27$

$n_2 = 40$ $\bar{x}_2 = 638$ hrs $s_2 = 31$

$H_0 : \cancel{\mu_1 \neq \mu_2}$ $\mu_1 \le \mu_2$

$H_1 : \mu_1 > \mu_2$

Numerical computatⁿ:

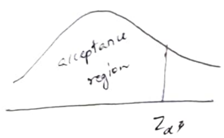
$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$Z = 1.3846$

~~i) at α = 0.05~~

Decision rule:

accept H_0 if $Z < Z_\alpha$



i) at $\alpha = 0.05$

$\phi(Z_\alpha) = 0.45$

$Z_\alpha = 1.65 > Z = 1.38 \Rightarrow$ accept H_0

ii) at $\alpha = 0.01 \Rightarrow 0.5 + \phi(Z_\alpha) = 0.99$

$\phi(Z_\alpha) = 0.49$

$Z_\alpha = 2.33 > Z = 1.38 \Rightarrow$ accept H_0

Note: Looking for a particular characteristic in a popⁿ ⇒ proportion.

Populatⁿ → binomial distⁿ

sample → normal distⁿ

(though proportⁿ of sample, proportⁿ of popⁿ is estimated ⇒ p is the expectatⁿ of $\hat{p}$)

Test - 3

Test of significance concerning single proportion.

Suppose we wish to estimate the proportion of individuals in a popⁿ who has certain attribute, then we shall use this test.

$x \to$ no. of items possessing the attribute / characteristics

$x \sim B(n, p)$ where p is popⁿ proportion.

As $n \to \infty$ $x \sim N(\mu = np, \sigma^2 = npq)$

H_0 is $P = P_0$

H_1 : $P \neq P_0$ (accept when $|z| < Z_{\alpha/2}$)

$P < P_0$ (" when $z > Z_\alpha$)

$P > P_0$ (" when $z < Z_\alpha$)

Test statistic $Z = \frac{\hat{p} - P}{\sqrt{\frac{PQ}{n}}}$ where $\hat{p} = \frac{x}{n}$ is the sample proportion.

Q) If in a random sample of 600 cars, making a right turn at certain traffic junctⁿ, 157 drove into the wrong lane. Test whether actually 30% of all the drivers make this mistake or not at this given junction. Use a) 0.05 LOS b) 0.01 LOS

→ $n = 600$, $\hat{p} = \frac{157}{600}$ (proportⁿ in sample)

→ X : No. of cars in wrong lane

H_0 : $P = 0.3$

H_1 : $P \neq 0.3$

Decision rule :

i) at $\alpha = 0.05$ accept H_0 when $|Z| < Z_{\alpha/2}$

$Z_{\alpha/2} = 1.96$

$-Z_{\alpha/2} = -1.96$ $\quad Z_{\alpha/2} = 1.96$

Numerical Computatⁿ : $Z = \dfrac{p - P}{\sqrt{\frac{PQ}{n}}} = \dfrac{0.2616 - 0.3}{\sqrt{\frac{0.3 \times 0.7}{600}}}$

$p = 0.2616$

$= -2.0493$

Conclusion : at $\alpha = 0.05$, $|Z| = +2.0493 \nless Z_\alpha = 1.65$

$\Rightarrow$ reject H_0

ii) Decision rule :

at $\alpha = 0.01$, accept H_0 when $|Z| < Z_{\alpha/2}$

$0.005 = 0.5 - \phi(Z_{\alpha/2})$

$(Z_{\alpha/2} = \cancel{2.33})$

$Z_{\alpha/2} = 2.58$

$-Z_{\alpha/2} = -2.33$ $\quad Z_{\alpha/2} = 2.33$

Conclusion : at $\alpha = 0.01$,

$|Z| = +2.049 < 2.58 = Z_{\alpha/2}$

$\Rightarrow$ accept H_0

⑧ Test the claim of manufacturer that 95% of his stabilisers confirm to ISI specificatⁿˢ. If out of random sample of 200 stabilisers, produced, 18 were faulty use 0.01 LOS

→ $n = 200$ $\quad$ x : No. faulty stabilisers (that)

H_0 : $P \leq 0.05$: (Manufacturers claim of 95% of stabilisers are not faulty)

H_1 : $P > 0.05$

$p = \dfrac{18}{200} = 0.09$

Decision rule : at $\alpha = 0.01$, accept H_0 when $Z < Z_\alpha$

$\phi(Z_\alpha) = 0.49$

$Z_\alpha = 2.33$

$Z_\alpha = 2.33$

Numerical Computatⁿ : $Z = \dfrac{p - P}{\sqrt{\frac{PQ}{n}}} = \dfrac{0.09 - 0.05}{\sqrt{\frac{(0.05) \times 0.95}{200}}}$

$= \cancel{(H-3+)}\ 2.5955$

Conclusion : at $\alpha = 0.01$,

$Z = 2.59 \nless Z_\alpha = 2.33$ $\Rightarrow$ reject H_0

⑨ If in a random sample of 200 persons suffering with headache, 160 people got cured by a drug, can we accept the claim of manufacturer that his drug cures 90% of sufferers ? Use 0.01 LOS.

(90% get cured)

→ H_0 : $P \geq 0.9$ $\quad n = 200$

H_1 : $P < 0.9$ $\quad p = \dfrac{160}{200} = 0.8$

Decision rule : at $\alpha = 0.01$, accept H_0 when $Z > Z_\alpha$

$\phi(Z_\alpha) = 0.49$

$Z_\alpha = 2.33$

$-Z_\alpha = 2.33$

Numerical computatⁿ $\quad z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.8 - 0.9}{\sqrt{\frac{0.9 \times 0.1}{200}}}$

$= -4.714$

Conclusion : at $\alpha = 0.01$,

$z = -4.714 \nless z_\alpha = -2.33 \Rightarrow$ reject H_0.

---

6) In a random sample of 400 people, from a large population 120 are female. Can it be said that males and females are in the ratio 5:3 in the popⁿ. Use 1% LOS.

→ $X \to$ No. of female in the popⁿ.

$n = 400, \quad x = 120, \quad p = \frac{120}{400} = \frac{3}{10} = 0.3$

$H_0 : \quad P = 3/8 = 0.375$ [M & F are in ratio 5:3]

$H_1 : \quad P \neq 3/8 \neq 0.375$ [M & F in popⁿ are not in ratio 5:3]

Decision rule : at $\alpha = 0.01$, accept H_0 if $-z_{\alpha/2} < z < z_{\alpha/2}$

or $|z| < z_{\alpha/2}$

$\phi(z_{\alpha/2}) = 0.495$

$-z_{\alpha/2} \qquad z_{\alpha/2}$

$z_{\alpha/2} = 2.58$

Numerical computation :

$z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.3 - (0.375)}{\sqrt{\frac{0.375 \times (1 - 0.375)}{400}}} = -3.098$

Conclusion : at $\alpha = 0.01$, $|z| = 3.098 \nless z_{\alpha/2}$

$\Rightarrow$ Reject H_0.

---

7) It is observed that 174 out of a random sample of 200 truck drivers on highway are drunk. Is it valid to state that atleast 90% of truck drivers are drunk? Use 5% LOS.

→ X : No. of truck drivers that are drunk.

$n = 200, \quad x = 174, \quad p = ~~0.9~~ \; 0.87$

$H_0 : \quad P \geq 0.9$ [90% of truck drivers are drunk]

$H_1 : \quad P < 0.9$

Decision rule : at $\alpha = 0.05$, accept H_0 if $z > z_\alpha$

z_α

$\phi(z_\alpha) = 0.5 - 0.05$

$= 0.45$

$z_\alpha = -1.65$

Numerical Computation :

$z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.87 - 0.9}{\sqrt{\frac{0.9 \times 0.1}{200}}} = -1.414$

Conclusion : at $\alpha = 0.05$, $z_\alpha = -1.65 < z = -1.414$

$\Rightarrow$ Accept H_0

Type - 4

Test of hypothesis - Difference b/w proportⁿs with large sample :

$H_0 : P_1 = P_2$

$H_1 : \quad P_1 \neq P_2 \quad$ accept H_0 when $|z| < z_{\alpha/2}$

$P_1 < P_2 \quad$ accept H_0 when $z > z_\alpha$

$P_1 > P_2 \quad$ accept H_0 when $z < z_\alpha$

Test Statistic:

$$Z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$ where $P = \frac{x_1 + x_2}{n_1 + n_2}$

& $Q = 1 - P$

Q) If 48 out of 400 people in rural area posess cell phone, while 120 out of 500 in urban area, can it be accepted that proportion of cellphones in rural and urban area is same? Use $\alpha = 5\%$.

→ x_1 : No. of people posessing cellphone in rural area

x_2 : " " " in urban ".

$n_1 = 400$, $p_1 = \frac{48}{400} = 0.12$, $n_2 = 500$, $p_2 = \frac{120}{500} = 0.24$

$x_1 = 48$, $x_2 = 120$

$H_0 : P_1 = P_2$ [Proport^n of cellphone users in rural is same as that of in urban]

$H_1 : P_1 \neq P_2$

Decision rule: at $\alpha = 0.05$, accept H_0 when $|z| < z_{\alpha/2}$

$\phi(z_{\alpha/2}) = 0.5 - 0.025 = 0.475$

$z_{\alpha/2} = 1.96$

Numerical Computat^n:

$P = \frac{x_1 + x_2}{n_1 + n_2} = \frac{168}{900} = 0.1867$, $Q = 1 - P = 0.8133$

$$Z = \frac{(0.12 - 0.24) - \cancel{0(0.8133)(\tfrac{1}{400} + \tfrac{1}{500})}}{\sqrt{(0.8133)(0.1867)\left(\frac{1}{400} + \frac{1}{500}\right)}}$$

$= \cancel{-0.12} = -4.5906$

Conclusion: at $\alpha = 0.05$, $|z| = 4.59 \nless z_{\alpha/2} = 1.96$

⇒ reject H_0

Q) If 57 out of 150 patients suffering with certain disease are cured by allopathy and 33 out of 100 are cured by homeopathy, patients with same disease are cured by homeopathy. is there any reason to believe that allopathy is better than homeopathy at 5% LOS?

→ $n_1 = 150$, $n_2 = 100$, $p_1 = \frac{57}{150} = 0.38$

$x_1 = 57$, $x_2 = 33$, $p_2 = \frac{33}{100} = 0.33$

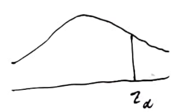
$H_0 : P_1 \leq P_2$

$H_1 : P_1 > P_2$

Decision rule: at $\alpha = 0.05$, accept H_0 if $z < z_\alpha$

$\phi(z_\alpha) = 0.45$

$z_\alpha = 1.65$



Numerical Computat^n:

$P = \frac{x_1 + x_2}{n_1 + n_2} = 0.36$ $Q = 1 - 0.36 = 0.64$

$$Z = \frac{p_1 - p_2}{\sqrt{0.36 \times 0.64 \left(\frac{1}{150} + \frac{1}{100}\right)}}$$

$$= \frac{1/20}{0.0619} = 0.8068$$

Conclusion: at $\alpha = 0.05$, $Z = 0.8068 < Z_\alpha = 1.65$

$\Rightarrow$ Accept H_0

$\Rightarrow$ reject H_1 $\Rightarrow$ Allopathy is not better than homeopathy.

---

8) A study of TV viewers was conducted to find opinion about megaserial Ramayana. If 56% of a sample of 300 viewers from South and 48% of 200 viewers from north preferred the serial, test the claim at 5% LOS that

i) There is a difference of opinion b/w south & north

ii) Ramayana is preferred in the south.

$\rightarrow$ x_1 : no. viewers from south , x_2 : no. viewers from north

$n_1 = 300$, $x_1 = 168$, $p_1 = 0.56$

$n_2 = 200$ $x_2 = 96$ $p_2 = 96/200 = 0.48$

i) $H_0 : P_1 = P_2$

$H_1 : P_1 \neq P_2$

Decision rule : at $\alpha = 0.05$ accept H_0 if $|Z| < Z_{\alpha/2}$

$\phi(Z_{\alpha/2}) = 0.495$

$Z_{\alpha/2} = 2.58$

Numerical computation :

$$P_1 = \frac{168 + 96}{300 + 200} = 0.528 \quad , \quad Q = 0.472$$

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.08}{(21.9433)^{-1}} = \cancel{\frac{45.64}{3.651}} \; 1.755$$

Conclusion : at $\alpha = 0.05$, $|Z| = \cancel{3.651}$ 1.755 $\cancel{\not<}$ $< Z_{\alpha/2} = 2.58$

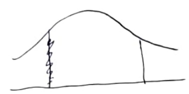
$\Rightarrow$ ~~reject~~ accept H_0

---

ii) $H_0 : P_1 \leq P_2$

$H_1 : P_1 > P_2$ [Ramayana is preferred in south]

Decision rule : at $\alpha = 0.05$, accept H_0 if $Z \leq Z_\alpha$


Z_α Z_α

$\phi(Z_\alpha) = 0.45$

$Z_\alpha = +1.65$

Numerical computation :

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.08}{(21.9433)^{-1}} = 1.755$$

Conclusion : at $\alpha = 0.05$, $Z = 1.755 \nleq Z_\alpha = 1.65$

$\Rightarrow$ reject H_0

$\Rightarrow$ accept H_1 $\Rightarrow$ Ramayana is preferred in south

9) A question in true-false quiz is considered to be smart if it discriminates b/w intelligent person (IP) and avg person (AP). Suppose 205 of 250 IP's and 137 of 250 AP's answer a quiz question correctly, test at 0.01 LOS whether for the given question, the proport^{n} of correct answers can be expected to be atleast 50% higher among IP's than among AP's

→

Type - 3 :

Small sample test concerning single mean : t - distribut^{n}

Here x does not follow normal distribut^{n} but follows t-dist^{n}. As $n \to \infty$, t-dist^{n} becomes normal dist^{n}

$H_0 : \mu = \mu_0$

$H_1 : \mu \neq \mu_0$ accept H_0 when $|t| < t_{\alpha/2, n-1}$

$\mu < \mu_0$ accept H_0 when $t > t_{\alpha, n-1}$

$\mu > \mu_0$ accept H_0 when $t < t_{\alpha, n-1}$

Test statistic :

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$ where $s^2 = \frac{1}{n-1} \Sigma (x_i - \bar{x})^2$

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

(or)

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}}$$

1) The mean weekly sales of a chocolate bar in a candy store was 140.3 bars / store. After an advertisement campaign, the mean weekly sales in 22 stores for a typical week increased to 153.7 & showed SD of 17.2. Was the ad campaign successful ?

→ x → weekly sales of chocolate box

$\mu = 140.3$ $\bar{x} = 153.7$ $n = 22$, $s = 17.2$

$H_0 : \mu \leq 140.3$

$H_1 : \mu \cancel{<} > 140.3$ [Sales increases after advertisement]

Decision rule:

At $\alpha = 5\%$ accept H_0 when $t < t_{\alpha, n-1}$

$t_{\alpha, n-1} = t_{0.05, 21} = 1.721$

N.C

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} \qquad s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$

$$t = \frac{153.7 - 140.3}{\frac{17.2}{\sqrt{21}}}$$

$t = 3.57$

Conclusion: At $\alpha = 5\%$, $t = 3.57 \nless t_{\alpha, n-1} = 1.721$

⇒ Reject H_0

⇒ The advertisement campaign was successful.

---

2) An ambulance service company claims that, on an avg. it takes 20 min. b/w a call for an ambulance and patients arrival at hospital. If in 6 calls, the time taken b/w a call & arrival at hospital are 27, 18, 26, 15, 20, 32, can the company's claim be accepted at $\alpha = 5\%$.

→ x → time taken b/w phonecall & arrival at hospital (in min)

$n = 6$

$\mu = 20$ min

$H_0 : \mu \leq 20$

$H_1 : \mu > 20$

Decision rule: at $\alpha = 0.05$, accept H_0 when $t < t_{\alpha, n-1}$

$t_{\alpha, n-1} = t_{0.05, 5}$

$t_{0.05, 5} = 2.015$

N.C

$$\bar{x} = \frac{27 + 18 + 26 + 15 + 20 + 32}{6} = 23$$

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}} = \sqrt{\frac{(4)^2 + 5^2 + 3^2 + 8^2 + 3^2 + 9^2}{6}}$$

$= \sqrt{34} = 5.83$

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} = \frac{23 - 20}{\frac{5.83}{\sqrt{5}}} = 1.1506$$

conclusion: at $\alpha = 0.05$, $t = 1.15 < t_{0.05, 5} = 2.015$

⇒ accept H_0

⇒ ~~reject~~ ⇒ company's claim can be accepted.

3 Q Following are the systolic BP of 12 patients under-going drug therapy for hypertension : 183, 152, 178, 157, 194, 163, 144, 114, 178, 152, 118, 158. Can we conclude on the basis of this data that popn mean is less than 165? Use $\alpha = 5\%$.

→ $n = 12$

$x \to$ systolic BP of patients.

$\mu = 165$

$H_0 : \mu \geq 165$

$H_1 : \mu < 165$

$\bar{x} = 157.58$

$s = \sqrt{\frac{\Sigma (x-\bar{x})^2}{n}} = \sqrt{(25.41)^2 +}$

Decision rule : at $\alpha = 0.05$, accept H_0 when $t > t_{\alpha, n-1} = t_{0.05, 11}$

$-t_{0.05, 11} = -1.796$

N.C

$$s = \sqrt{\frac{646.006 + 31.17 + 416.84 + 0.3402 + 1326.17 + 29.34 + 184.506 + \cancel{43.58} + 416.840 + 31.173 + 1566.84 + 1899.506 + 0.173611}{12}}$$

$= \cancel{16.5403}\ 20.7$

$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} = -1.2411$

Conclusion : $t = -1.24 > t_{\alpha, n-1} = -1.796$

$\Rightarrow$ accept H_0.

4 Q In a random sample of 10 bolts produced by a machine the mean length of bolt is ~~0.0~~ 0.53 mm and SD of 0.03 mm. Can we claim from this that the machine is in proper working order if in the past, it produced bolts of length 0.5 mm? use 0.01 LOS.

→ $x \to$ length of bolt

$n = 10$, $\bar{x} = 0.53$, $s = 0.03$

$\mu = 0.5$

$H_0 : \mu = 0.5$ (Machine works properly)

$H_1 : \mu \neq 0.5$

Decision rule : At $\alpha = 1\%$, accept H_0 when $|t| < t_{\alpha/2, n-1}$

$= t_{0.005, 9}$

$= 3.250$

$-t_{\alpha/2, n-1}$ $t_{\alpha/2, n-1}$

N.C :

$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} = \frac{0.53 - 0.5}{\frac{0.03}{\sqrt{9}}} = 3$

Conclusion : At $\alpha = 0.01$, $t = 3 < t_{\alpha/2, n-1} = 3.25$

$\Rightarrow$ accept H_0

31/12/21

③ (σ is given) 31/12/21

Mean lifetime of computers manufactured by a company is 1120 hrs, with SD 125 hrs

i) Test the hypothesis that mean lifetime of computer has not changed if a sample of 8 computers has a mean lifetime of 1070 hrs.

ii) Is there a decrease in MLT? Use 1% LOS.

→ x → lifetime of computers

$\mu = 1120$ hrs $n = 8$ computers

$\sigma = 125$ hrs

$\bar{x} = 1070$ hrs

N.C

$$t = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = -1.1313$$

i) $H_0 : \mu = 1120$

$H_1 : \mu \neq 1120$

Decision rule: at $\alpha = 1\%$ accept H_0 when $|t| < t_{\alpha/2, n-1}$

$t_{0.005, 7}$

3.499

$-t_{\alpha/2, n-1}$ $t_{\alpha/2, n-1}$

Conclusion: at $\alpha = 1\%$ $|t| = 1.1313 < 3.499 = t_{\alpha/2, n-1}$

∴ accept H_0

ii) $H_0 : \mu \geq 1120$

$H_1 : \mu < 1120$

Decision rule: At $\alpha = 1\%$ accept H_0 if $t > t_{\alpha, n-1}$

$t_{0.01, 7}$

$= 2.998$

$t_{\alpha, n-1}$ t

-2.998

Conclusion: at $\alpha = 1\%$,

$t = -1.1313 > \cancel{\ngtr}\ t_{\alpha, n-1} = -2.998$

⇒ accept H_0

---

③ If 5 pieces of certain ribbon selected at random have mean breaking strength of 169.5 pounds with SD 5.7 pounds. Do they confirm to specificatⁿ mean breaking strength of 180 pounds? LOS ⇒ 5%.

→ $n = 5$

x → breaking strength of ribbon

$\bar{x}$ ~~μ~~ $= 169.5$

& ~~σ~~ $s = 5.7$

μ ~~$\bar{x}$~~ $= 180$

$H_0 : \mu = 180$

$H_1 : \mu \neq 180$

Decision rule: at $\alpha = 0.05$,

accept H_0 if $|t_{cal}| \ngtr t_{\alpha/2, n-1}$

2.776

$-t_{\alpha/2, n-1}$ $t_{\alpha/2, n-1}$

NC

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}}$$

$= -3.6842$

Conclusion: at $\alpha = 0.05$

$|t| = 3.684 \nless t_{\alpha/2, n-1} = 2.776$

⇒ reject H_0

Type - 6 :

Small Sample test concerning difference b/w means (for small sample) :

Test Statistic :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where $s^2 = \frac{\Sigma(x_{1i} - \bar{x}_1)^2 + \Sigma(x_{2i} - \bar{x}_2)^2}{n_1 + n_2 - 2}$

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

Hypothesis testing :

$H_0 : \mu_1 - \mu_2 = \delta$

$H_1 :$ $\mu_1 - \mu_2 \neq \delta$, accept H_0 when $|t| < t_{\alpha/2,\ n_1+n_2-2}$

$\mu_1 - \mu_2 < \delta$, accept H_0 when $t > t_{\alpha, n_1+n_2-2}$

$\mu_1 - \mu_2 > \delta$, accept H_0 when $t < t_{\alpha, n_1+n_2-2}$

---

(5) In mathematics examination, 9 students of class ~~B~~A & 6 students of class B obtained the following marks. Test at 1% LOS whether performance in maths are same or not for the 2 classes. Assume that the samples are drawn from normal popula^n having equal variance.

A	44	71	63	59	68	46	69	54	48
B	52	70	41	62	36	50			

$n_1 = 9, \quad n_2 = 6$

$x_1 \to$ marks of students in class A

$x_2 \to$ " " " " B

$H_0 : \mu_1 = \mu_2$ [performance of class A and class B are same]

$H_1 : \mu_1 \neq \mu_2$ [not same]

Decision rule : At $\alpha = 1\%$ accept H_0 when $|t| < t_{\alpha/2, n_1+n_2-2}$

$|t| < t_{0.005, 13}$

$= 3.012$

$-t_{\alpha/2, n_1+n_2-2}$ $\quad t_{\alpha/2, n_1+n_2-2} = 3.012$

N.C

$$\bar{x}_1 = \frac{44+71+63+59+68+46+69+54+48}{9} = 58$$

$$\bar{x}_2 = \frac{52+70+41+62+36+50}{6} = 51.83$$

3/1/22

Q. Out of random sample of 9 mice, suffering with a disease 5 mice were treated with a new serum while the remaining were not treated. From the time of commencement of experiment, the following are the survival time. Test whether the serum treatment is effective in curing the disease at 5% los, assuming that the 2 distributions are normally distributed with equal variance.

→

Treatment	2.1	5.3	1.4	4.6	0.9
No treatment	1.9	0.5	2.8	3.1	

→ $x_1 \to$ ~~No. of mice~~ ~~treated with new serum~~ Survival rate with treatment.

$x_2 \to$ ~~No. of mice~~ Survival ~~time~~ rate without treatment

$n_1 = 5, \quad n_2 = 4$

$H_0 : \mu_1 \leq \mu_2$,

$H_1 : \mu_1 > \mu_2$. [Serum treatment is effective]

$$\sigma^2 = \frac{\sum (x_{1i} - \bar{x}_1)^2 + \sum (x_{2i} - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

$$\bar{x}_1 = \frac{2.1 + 5.3 + 1.4 + 4.6 + 0.9}{5} = 2.86$$

$\bar{x}_2 = 2.075$

$$s^2 = \frac{(2.1-2.86)^2 + (5.3-2.86)^2 + \dots + (1.9-2.075)^2 + \dots}{9-2}$$

=

$s^2 = 0.5776 + 5.9536 + 2.1316 + 3.0276 + 3.8416 + 0.030625 + 2.4806 + 0.725 + 1.0506$

$s^2 = 2.801$

$s = 1.6874$

Decision rule :

$t_{\alpha, n_1+n_2-2} = 1.895$

N.C $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

$= \dfrac{(2.86 - 2.075)}{1.678\sqrt{\frac{1}{5} + \frac{1}{4}}}$

$t = 0.697$

Conclusion : at $\alpha = 5\%$,

$t = 0.697 < 1.895$ $\Rightarrow$ Accept H_0

---

⑧ ~~To test the claim~~ A study is conducted to determine whether the wear of material A exceeds that of B by more than 2 units. If test of 12 pieces of material A is yields a mean wear of 85 units and S.D of 4, while test of 10 pieces of material B yields a mean of 81 and SD 5, what conclusion can be drawn at 5% LOS assume that popl^{ns} are normally distributed with equal variance.

→ $\bar{x}_1 = 85$, $s_1 = 4$, $n_1 = 12$

$\bar{x}_2 = 81$, $s_2 = 5$, $n_2 = 10$

$x_1 \rightarrow$ wear of material A

$x_2 \rightarrow$ wear of material B

$H_1: \mu_1 - \mu_2 > 2$ [wear of A exceeds that of B by more than 2]

$H_0: \mu_1 - \mu_2 \leq 2$

Decision rule : at $\alpha = 5\%$, accept H_0 when $t < t_{\alpha, n_1+n_2-2}$

$= t_{0.05, 20}$

$= 1.725$

N.C

$s^2 = \dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} = \dfrac{12(16) + 10(25)}{20} = 22.1$

$s = 4.7010$

$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \dfrac{(85 - 81) - (2)}{4.701\sqrt{\frac{1}{12} + \frac{1}{10}}} = 0.99$

Conclusion: at $\alpha = 0.05$, $t = 0.99 < t_{\alpha, n_1+n_2-2} = 1.725$

$\Rightarrow$ accept H_0

---

⑨ To determine whether vegetarian & non-vegetarian diet affects significantly on increase in weight, a study was conducted yielding the following data of gain in weight.

veg	34	24	14	32	25	32	30	24	30	31	35	25			
non veg	22	10	47	31	44	34	22	40	30	32	35	18	21	35	29

Can we claim that 2 diets differ pertaining to weight gain assuming that samples are drawn wit from normal pop^n with same variance ? $\alpha = 1\%$

→ x_1 → gain in weight by veg diet

x_2 → " " non-veg diet.

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

Decision rule : at $\alpha = 0.01$, accept H_0 when $|t| < t_{\alpha/2, n_1+n_2-2}$

$= t_{0.005, 25}$

$= 2.787$

N.C

$\bar{x}_1 = 28$ $\quad \bar{x}_2 = 30$

$$s = \frac{\sum (x_{1i} - \bar{x}_1)^2 + \sum (x_{2i} - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

$$s_1^2 = \frac{36 + 16 + 196 + 16 + 9 + 16 + 4 + 16 + 4 + 9 + 49 + 9}{12}$$

$= 31.67$

$$s_2^2 = \frac{64 + 400 + 289 + 1 + 196 + 16 + 64 + 100 + 0 + 4 + 25 + 144 + 25 + 1}{15}$$

~~$s^2 = 68.36 \Rightarrow s = 8.26$~~ $\quad s^2 = 71.6 \Rightarrow s = 8.46$

N.C

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{-2}{8.46\sqrt{\frac{1}{12} + \frac{1}{15}}} = -0.6103$$

Conclusion : at $\alpha = 1\%$, $|t| = 0.6103 < t_{\alpha/2, n_1+n_2-2}$

$\Rightarrow$ accept H_0

Type - 7

Paired - t test :

Hypothesis : $H_0 : \mu_d = \mu_d$

$H_1 : \mu \neq \mu_d$, accept H_0 when $|t| < t_{\alpha/2, n-1}$

$\mu < \mu_d$, accept H_0 when $t > t_{\alpha, n-1}$

$\mu > \mu_d$, accept H_0 when $t < t_{\alpha, n-1}$

Test Statistic :

$$t = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n}}} \quad \text{(or)} \quad t = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n-1}}}, \text{ where } s_d^2 = \frac{\sum (d_i - \bar{d})^2}{n}$$

where $S_d^2 = \frac{\sum (d_i - \bar{d})^2}{n-1}$, $d \Rightarrow$ difference between 2 values

---

Q) The BP of 5 women before and after intake of certain drug are given below. Test at 1% LOS whether there is significant change in BP.

Before	110	120	125	132	125
After	120	118	125	136	121

$x_1 \rightarrow$ BP of women before taking drug

$x_2 \rightarrow$ " " after " ".

$H_0 : \mu_1 = \mu_2$, $\mu_d = 0$

$H_1 : \mu_1 \neq \mu_2$, $\mu_d \neq 0$

Decision rule: at $\alpha = 0.01$, accept H_0 when $|t| < t_{\alpha/2, n-1}$

$= t_{0.005, 4}$

$= 4.604$

$t_{\alpha/2, n-1} = 4.604$

N.C.

d_{1-2} : -10 2 0 -4 4 . $\bar{d} = -1.6$

$$s_d^2 = \frac{\sum (d_i - \bar{d})^2}{n-1} = \frac{70.56 + 12.96 + 2.56 + 5.76 + \cancel{5.76}\ 31.36}{4}$$

$= 30.8$ $\qquad s_d = 5.54$

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{-1.6\sqrt{5}}{\cancel{30.8}\ 5.54} = -0.644$$

conclusion: at $\alpha = 0.01$, $|t| = 0.644 < t_{\alpha/2, n-1} = 4.604$

$\Rightarrow$ accept H_0

(5) Marks obtained in mathematics by 11 students before & after intensive coaching are given below. At 5% los, check whether intensive coaching is useful?

Before	24	17	18	20	19	23	16	18	21	20	19
After	24	20	22	20	17	24	20	20	18	19	22

$x_1 \rightarrow$ marks before coaching

$x_2 \rightarrow$ " after coaching.

$H_0 : \mu_2 \leq \mu_1$

$H_1 : \mu_2 > \mu_1$

Decision rule: at $\alpha = 0.05$, accept H_0 when $t \ngtr t_{\alpha, n-1}$

$t_{\alpha, n-1} =$

5/1/22

Q) Pulsality index of 11 patients before & after contacting a disease are given below. Test at 0.05 LOS, whether there is increase of mean of PI values.

Before	.9	.45	.44	.54	.48	.62	.48	.6	.45	.46	.35
After	.5	.6	.57	.65	.63	.78	.63	.8	.69	.62	.68

→ x_1: PI of patient before disease
x_2: " " " after disease

H_0: $\mu_1 \geq \mu_2$
H_1: $\mu_1 < \mu_2$ [PI increases after disease]

Decision rule: at $\alpha = 0.05$, accept H_0 when $t > t_{\alpha, n-1}$
" $t_{0.05, 10} = 1.812$

N.C

d: -.1 -.15 .13 .11 .215 -.16 -.15 -.2 -.24 .16 -.33

$\bar{d} = -0.1709 = -0.171$

$s_d^2 = \frac{\Sigma (d_i - \bar{d})^2}{n-1}$

$= (5.02 \times 10^{-3} + 4.3681 \times 10^{-4} + 1.67281 \times 10^{-3} + 3.70881 \times 10^{-3}$
$+ 4.3681 \times 10^{-4} + 1.1881 \times 10^{-4} + 4.3681 \times 10^{-4} + 8.4681 \times 10^{-4}$
$+ 4.77481 \times 10^{-3} + 2.456 \times 10^{-4} + 0.02381) / 10$

$= 0.665$

$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \frac{-0.17}{\frac{0.665}{\sqrt{11}}} = -8.62$

Conclusion: $t \ngtr t_{\alpha, n-1}$ ⇒ reject H_0.

Q) Avg weekly losses of man-hours due to strike in an Institute before & after a disciplinary program was implemented are as follows.

Is there a reason to believe that the disciplinary program was effective at 5% LOS?

Before	45	73	46	124	33	57	83	34	26	17
After	36	60	44	119	35	51	77	29	24	11

→

H_0: $\mu_1 \leq \mu_2$
H_1: $\mu_1 > \mu_2$ [Disciplinary program is effective]

Decision rule: at $\alpha = 0.05$ accept H_0 when $t < t_{\alpha, n-1} = t_{0.05, 9}$
$= 1.833$

$t_{\alpha, n-1} = 1.833$

N.C
d: 9, 13, 2, 5, -2, 6, 6, 5, 2, 6

$\bar{d} = \frac{52}{10} = 5.2$

$$S_d^2 = \frac{\sum (d_i - \bar{d})^2}{n-1} = 16.617$$

$$S_d = 4.076$$

$$t = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n}}} = \frac{5.2\sqrt{10}}{4.076} = 4.034 .$$

Conclusion: at $\alpha = 5\%$, $t = 4.034 \nless t_{\alpha, n-1} = 1.833$

$\Rightarrow$ reject H_0

Type - 8

Ratio of Variance : F - Distribution :

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

$H_1 : \sigma_1^2 \neq \sigma_2^2$, accept H_0 when $F < F_{n_1-1, n_2-1, \alpha}$

$\cancel{\sigma_1^2 < \sigma_2^2}$,

$\cancel{\sigma_1^2 > \sigma_2^2}$), x

Test Statistic:

$F = \frac{s_1^2}{s_2^2}$ if $s_1^2 > s_2^2$ (or) $F = \frac{s_2^2}{s_1^2}$ if $s_2^2 > s_1^2$

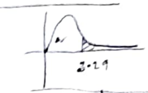
where $s_1^2 = \frac{\sum (x_{1i} - \bar{x}_1)^2}{n_1 - 1}$, $s_2^2 = \frac{\sum (x_{2i} - \bar{x}_2)^2}{n_2 - 1}$

① In one sample of 8 observatⁿs, the sum of squares of deviations of the sample values from the sample mean was 84.4 & in the other sample of 10 observatⁿ it was 102.6. Test whether this diff is significant at 5% LOS

→ $H_0 : \sigma_1^2 = \sigma_2^2$

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

Decision rule: at $\alpha = 5\%$ accept H_0 when $F < F_{n_1-1, n_2-1, \alpha} = F_{7,9,0.05} = 3.29$



Given, $\sum (x_{1i} - \bar{x}_1)^2 = 84.4$, $n_1 = 8$ and $\sum (x_{2i} - \bar{x}_2)^2 = 102.6$, $n_2 = 10$

$$s_1^2 = \frac{\sum (x_{1i} - \bar{x}_1)^2}{n_1 - 1} = \frac{84.4}{7} = 12.057$$

$$s_2^2 = \frac{102.6}{9} = 11.4$$

$$F = \frac{s_1^2}{s_2^2} = 1.1185$$

Conclusion: at $\alpha = 5\%$, $F = 1.1185 < F_{7,9,0.05} = 3.29$

$\Rightarrow$ accept H_0

12/1/22

2② Can we conclude that the two populations variance are equal for the following data of post graduates passed out of state and private universities. Let $\alpha = 1\%$.

State	8350	8260	8130	8340	8070	
Private	7890	8740	7900	7950	7840	7920

→ x_1 → no. of post graduates passed out of state university

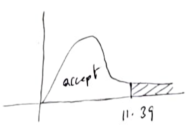
x_2 → no. of post graduates passed out of private uni.

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

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

Decision rule: at $\alpha = 1\%$ accept H_0 when $F < F_{n_1-1, n_2-1, \alpha}$

ie $F < F_{4,5,0.01} = 11.39$



<u>Numerical Computat^n</u>: $n_1 = 5$, $n_2 = 6$, $\alpha = 0.01$

$$S_1^2 = \frac{\sum (x_{1i} - \bar{x}_1)^2}{n_1 - 1}$$

$$\bar{x}_1 = \frac{8350 + 8260 + 8130 + 8340 + 8070}{5} = 8230$$

$$\bar{x}_2 = 7940$$

$$S_1^2 = \frac{(120)^2 + (30)^2 + (100)^2 + 8(110)^2 + (160)^2}{4} = 15750$$

$$S_2^2 = 10920$$

$S_1^2 > S_2^2$; $F = \frac{S_1^2}{S_2^2} = 1.4423$

<u>Conclusion</u>: at $\alpha = 1\%$ $F = 1.44 < 11.39$ $\therefore$ accept H_0

---

③ The household net expenditure on healthcare in south and north India in 2 samples of household, expressed as percentage of total income is shown in following table, LOS → 5%.

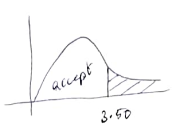
South	15	8	3.8	5.4	27.4	19	35.3	13.6	
North	18.8	23.1	10.3	8	18	10.2	15.2	19	20.2

→ x_1 → household net expenditure on healthcare in south.
x_2 → " " " " north

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

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

<u>Decision rule</u>: at $\alpha = 5\%$, accept H_0 when $F < F_{7,8,0.05} = 3.50$



<u>N.C</u>: $n_1 = 8$, $n_2 = 9$, $\alpha = 0.05$

$$S_1^2 = \frac{\sum (x_{1i} - \bar{x}_1)^2}{n_1 - 1}$$

$\bar{x}_2 = 15.867$, $\bar{x}_1 = 16.06$

$S_1^2 = 117.6$

$S_2^2 = 27.4375$

$S_1^2 > S_2^2$; $F = \frac{S_1^2}{S_2^2} = 4.286$

<u>Conclusion</u>: at $\alpha = 5\%$, $F = 4.2869 \nless 3.50$ ⇒ reject H_0

<u>Test - 9 : χ^2 Distribution</u> (not for checking of parameters)
(chi square dist^n) (checking for goodness of a fit)

This is non-parametric test

There are 3 applicat^ns of chi-square dist

- goodness of fit
- test of homogeneity
- independence of attributes

H_0 : Distribution is a good fit

H_1 : Distribution is not a good fit

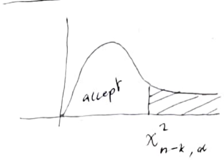
$$\chi^2 = \sum \left(\frac{(O_i - E_i)^2}{E_i} \right) \sim \chi^2_{n-k,\alpha}$$

where n is no. of samples and k = no. of constraints

<u>Decision rule</u>: at $\alpha\%$ accept H_0 when $\chi^2 < \chi^2_{n-k,\alpha}$

O → observed value
E → expected value

K - no. of constraints
n → no. of sample



Q) Fit a poisson distributⁿ to the following data and test for goodness of fit at $\alpha = 5\%$.

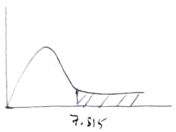
[For poisson distribⁿ take constraint $K = 2$]

x	0	1	2	3	4
f	419	352	154	56	19

→ H_0 : Poisson distributⁿ is a good fit

H_1 : " " is not " "

Decision rule : At $\alpha = 5\%$ accept H_0 when $\chi^2 < \chi^2_{n-k,\alpha} = \chi^2_{3,0.05}$ $(n=5, k=2)$ $= 7.815$



N.C

x	f = O	P(x)	Expected fre E	Expected fre E	$(O-E)^2/E$
0	419	0.4049	~~4045~~ ≃ ~~40~~	405	0.4839
1	352	0.366	~~3654 ≃ 37~~	367	0.6130
2	154	0.165	~~1656 ≃ 17~~	166	0.8674
3	56	0.0498	~~50 ≃ 5~~	50	0.72
4	19	0.011	~~12 ≃ 1~~	12	4.0833

$P(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$; $\lambda = \dfrac{\Sigma x_i f_i}{\Sigma f_i} = \dfrac{352 + 308 + 168 + 76 \cdots}{\Sigma f_i}$

$= 0.904$

$P(0) = e^{-0.904} = 0.404$

$P(1) = \dfrac{e^{-0.904} \times 0.904^1}{1} =$

$\chi^2 = \Sigma\left(\dfrac{(O_i - E_i)^2}{E_i}\right) = 6.7676$

Conclusion : at $\alpha = 5\%$ $\chi^2 = 6.7676 < 7.815 = \chi^2_{n-k,\alpha}$

⇒ accept H_0 ⇒ poisson is a good fit

Q) Fit a poisson distributⁿ for given frequency distⁿ and test for its goodness of fit at $\alpha = 1\%$.

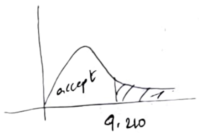
→

x	0	1	2	3
f	47	33	16	3

→ H_0 : Poisson is a good fit

H_1 : " " not " " "

Decision rule : At $\alpha = 0.01$ accept H_0 when $\chi^2 < \chi^2_{n-k,\alpha} = \chi^2_{2,0.01}$



N.C

x	f	P(x)	E = P(x)·99		$\frac{(O-E)^2}{E}$
0	47	0.4735	46.8765	≃ 47	0
1	33	0.3539	35.0361	≃ 36	0.25
2	16	0.1322	13.0878	≃ 13	0.6923
3	3	0.03295	3.262	≃ 3	0

$P(x) = \dfrac{e^{-\lambda}\lambda^x}{x!} = 0.7474$

$P(0) = 0.4735$

$\chi^2 = 0.25 + 0.6923 = 0.9423$

Conclusion : At $\alpha = 1\%$ $\chi^2 = 0.9423 < 9.210 \; (\chi^2_{n-k,\alpha})$

⇒ accept H_0

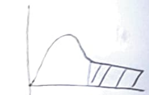
Q) A sample analysis of examination results of 500 students was made. It was found that 220 students failed, 170 secured third class, 90 secured second class and 20 secured first class. Do these figures support the general examination results which are in the ratio 4:3:2:1 for respective categories $\alpha = 5\%$.

→ H_0: The exam results are in stated ratio

H_1: " " " not " " "

Decision rule: at $\alpha = 5\%$ accept H_0 when

$$\chi^2 < \chi^2_{n-k,\alpha} = \chi^2_{3,0.05} = 7.815$$


7.815

N.C

O	E	$(O-E)^2/E$
220	$\frac{4}{10} \times 500 = 200$	$(20)^2/200 = 2$
170	$\frac{3}{10} \times 500 = 150$	$(20)^2/150 = 2.667$
90	$\frac{2}{10} \times 500 = 100$	$(10)^2/100 = 1$
20	$\frac{1}{10} \times 500 = 50$	$(30)^2/50 = 18$
500		

[In poisson it is assumed $\Sigma O = \Sigma E$]

Here also we need to assume

$$\chi^2 = \frac{\Sigma (O_i - E_i)^2}{\Sigma E_i} = 23.667$$

Conclusion: at $\alpha = 5\%$ $\chi^2 = 23.667 \nless 7.815$

$\Rightarrow$ reject H_0

Q) In mendelian Experiment 4 types of plants, are expected to occur in the proportion 9:3:3:1. The observed frequencies are 891 round & yellow, 316 wrinkled yellow, 290 round & green, 119 wrinkled & green. Find χ^2 & examine correspondance b/w theory & experiment. [$\alpha = 1\%$]

→ H_0: ~~Theory in correspondance with exp.~~ Plants are in stated ratio

H_1: " " not in " " .

Decision rule: At $\alpha = 1\%$, accept H_0 when

$$\chi^2 < \chi^2_{n-k,\alpha} = \chi^2_{4-1,0.01} = 11.345$$


11.345

N.C

O	E	$(O-E)^2/E$
891	$\frac{9}{16} \times 1616 = 909$	0.3564
316	$\frac{3}{16} \times 1616 = 303$	0.5577
290	$\frac{3}{16} \times 1616 = 303$	0.5577
119	$\frac{1}{16} \times 1616 = 101$	3.207
1616		

$$\chi^2 = \frac{\Sigma (O_i - E_i)^2}{E_i} = 4.6788$$

Conclusion: At $\alpha = 1\%$, $\chi^2 = 4.6788 < 11.345$

$\Rightarrow$ Accept H_0

28-1-22

UNIT - 5

{ Markov chain
{ Queuing theory

Markov chain

memoryless

$x(t) \rightarrow$ Discrete

chain

$f : S \rightarrow$ Real no.

sample points

$\Rightarrow$ Random var 'x'

when x depends on parameter, usually time $x(t)$

Range of $x = \{ \dots \}$

Stochastic Process

It is a family of random variables $\{ x(t) / t \in T \}$ defined on a common sample space S and indexed by parameter t which varies on an index set T.

State : The values assumed by random variables $x(t)$,

Set of all possible values from the state space of the process is denoted by I.

If state space is discrete, the stochastic process is called a chain.

A stochastic process consists of a sequence of experiments in which each experiment has finite no. of outcomes with given probability.

Markov chain

It is a markov process in which the state space is discrete. So Markov chain is a finite stochastic process consisting of a sequence of trials say $x_1, x_2 \dots$ satisfying two conditions.

① Each outcome belongs to state space $I = \{ a_1, a_2 \dots a_m \}$ which is finite.

② The outcome of any trial depends atmost on the outcome of immediately preceding trial and not upon any other previous outcome.

The system is said to be in state a_i at time n or at n^{th} step if a_i is outcome on n^{th} trial.

The no. P_{ij} gives probability that system changes from i^{th} state to j^{th} state.

Transition matrix :

P is a square matrix of transit^n probabilities ~~p~~ which has probability changes from one state to another.

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1m} \\ P_{21} & P_{22} & \dots & P_{2m} \\ \vdots & & & \\ P_{m1} & P_{m2} & \dots & P_{mm} \end{bmatrix}$$

The i^{th} row of P namely $(P_{i1}, P_{i2} \dots P_{im})$ represent probabilities of that system which changes state a_i to $a_1, a_2 \dots, a_m$.

Eg: A, B, C are playing passing the parcel. A passes the parcel only to B & B passes the parcel only to C and C passes the parcel equally likely to A as to B.

→ $S = \{A, B, C\}$

$A \to B$
$B \to C$
$C \to A/B$

T.P. matrix =

$$\begin{array}{c|ccc} & A & B & C \to \text{next} \\ \hline A & 0 & 1 & 0 \\ B & 0 & 0 & 1 \\ C & 1/2 & 1/2 & 0 \end{array}$$

↓ now

Probability Vector:

A vector $V = (v_1, v_2, \dots v_n)$ is called probability vector if $v_i \geq 0$ for every i & $\sum_{i=1}^{n} v_i = 1$.

$P_i \geq 0$
$\sum P_i = 1$

eg: $V = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$

$V = (0 \; 1 \; 0)$

$V = (\frac{1}{2} \; 0 \; \frac{1}{2})$

Note: $V = (1 \; 2 \; 3)$
$V = (\frac{1}{6} \; \frac{2}{6} \; \frac{3}{6})$

A vector whose components are non-negative but their sum is not 1, can be converted into a probability vector by dividing each vector component by sum of components.

Stochastic Matrix

P is a square matrix in which each row is a probability vector then P is a stochastic matrix.

A vector V is said to be a fixed vector if $VA = V$ & $V \neq 0$

Note: ① If P is a stochastic matrix then P^n is stochastic matrix for $n \in N$.

② A T.P.M is a stochastic matrix in a markov chain

③ Stochastic matrix is said to be regular if every element of P^n is non-negative & non-zero.

1⑤ A habitual gambler is member of 2 clubs A & B. He visits either of clubs every day for playing cards. He never visits club A on 2 consecutive days. But if he visits club B on a particular day, then next day he is as likely to visit A as to B. Find T.P.M & show that T.P.M is regular. Also find unique p fixed probability vector.

→ State space = $\{A, B\}$

T.P.M =
$$\begin{array}{c|cc} & A & B \to \text{next (tomorrow)} \\ \hline A & 0 & 1 \\ B & 1/2 & 1/2 \end{array}$$

↓ now (today)

$$P^2 = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$$

As all the values in this matrix are non-zero & non-negative, P is a regular stochastic matrix.

Let $V = (x \; y)$ be unique fixed prob vector.

$VP = V$ & $x + y = 1$

$$(x \quad y)\begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = (x \quad y)$$

$$\left(\frac{y}{2} \quad x + \frac{y}{2}\right) = (x \quad y)$$

$\Rightarrow$ $\frac{y}{2} = x \Rightarrow y = 2x$

& $x + y = 1 \Rightarrow 3x = 1$

$x = \frac{1}{3}$

$y = \frac{2}{3}$

F.P.V

$V = \left(\frac{1}{3} \quad \frac{2}{3}\right)$

---

(9) Find unique fixed probability vector for regular stochastic matrix,

$$A = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

$\rightarrow$ Let $V = (x \quad y \quad z)$ be a unique fixed probability vector (UFPV), taken $VA = V$ & $x + y + z = 1$

$$(x \quad y \quad z)\begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{bmatrix} = (x \quad y \quad z)$$

$$\left(\frac{y}{6} \quad x + \frac{y}{2} + \frac{z}{3} \quad \frac{y}{3} + \frac{z}{3}\right) = (x \quad y \quad z)$$

$\frac{y}{6} = x$ $\qquad$ $\frac{y}{3} + \frac{z}{3} = z$

$y = 6x$ $\qquad$ $y + z = 3z$ $\qquad$ $\frac{y}{2} = z$

$\frac{6x}{2} = z$

$x = 2z$

$x + y + z = 1$

$x + 6x + 3x = 1 \Rightarrow x = \frac{1}{10}, \; y = \frac{6}{10}, \; z = \frac{3}{10}$

(10) Find UFPV for regular stochastic matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

$\rightarrow$ Let $V = (x \quad y \quad z)$ be UFPV

$VA = V$ & $x + y + z = 1$

$$(x \quad y \quad z)\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} = (x \quad y \quad z)$$

$$\left(\frac{z}{2} \quad x + \frac{z}{4} \quad y + \frac{z}{4}\right) = (x \quad y \quad z)$$

$\frac{z}{2} = x$ $\qquad$ $y + \frac{z}{4} = z$

$z = 2x$ $\qquad$ $4y + z = 4z \Rightarrow 4y = 3z$

$4y = 6x$

$y = \frac{3}{2}x$

$x + y + z = 1$

$x + \frac{3x}{2} + 2x = 1$

$x = \frac{2}{9}, \quad y = \frac{3}{9}, \quad z = \frac{4}{9}$

UFPV $= \left(\frac{2}{9}, \frac{3}{9}, \frac{4}{9}\right)$

---

(11) Verify that the matrix $A = \begin{bmatrix} 0 & 1 \\ 0.3 & 0.7 \end{bmatrix}$ is a regular stochastic matrix.

$\rightarrow$ $A = \begin{bmatrix} 0 & 1 \\ 0.3 & 0.7 \end{bmatrix}$ $\quad A^2 = \begin{bmatrix} 0 & 1 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0.3 & 0.7 \end{bmatrix}$

$$= \begin{bmatrix} 0.3 & 0.7 \\ 0.21 & 0.3+0.49 \end{bmatrix}$$

As all entries of A^2 are non-zero, non-negative, it is a regular stochastic matrix.

Q) A student's study habits are as follows - If he studies one day, he is 70% sure not to study next day. On the other hand if he doesn't study one day, he is 60% sure not to study the next day. In the long run how often does he study.

$\rightarrow$ $S = \{A, B\}$

A $\rightarrow$ he's studying

B $\rightarrow$ he's not studying

Transit^n prob. matrix

$$\begin{array}{ccc} & A & B \\ P = \begin{array}{c} A \\ B \end{array} & \left[\begin{array}{c} 0.3 \\ 0.4 \end{array}\right. & \left.\begin{array}{c} 0.7 \\ 0.6 \end{array}\right] \end{array} \text{(tomorrow)}$$

(today)

$V = (x \ y)$ be UFPV & $x+y=1$

$$(x \ y) \begin{bmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{bmatrix} = (x \ y)$$

$$(0.3x + 0.4y \quad 0.7x + 0.6y) = (x \quad y)$$

$$0.3x + 0.4y = x$$

$$0.4y = 0.7x$$

$$y = \frac{7}{4}x$$

$x+y=1 \Rightarrow x + \frac{7}{4}x = 1 \quad \Rightarrow \quad 4x + 7x = 4$

$\cancel{x = \frac{4}{11}, y = 7}$

$$\Rightarrow x = \frac{4}{11}, \quad y = \frac{7}{11}$$

$$\text{UFPV} = \left(\frac{4}{11}, \frac{7}{11}\right)$$

In the long run, probability of him studying is $\frac{4}{11}$ & of not studying is $\frac{7}{11}$.

Q) A salesman's territory consists of 3 cities A, B, C. He never sells in the same city on successive days. If he sells in city A, then next day he sells in city B. However, if he sells in either B or C, the next day he is twice as likely to sell in city A as in other city. In long run how often does he sell in each of city?

$\rightarrow$ $S = \{A, B, C\}$

$$\begin{array}{cccc} & A & B & C \\ P = \begin{array}{c} A \\ B \\ C \end{array} & \left[\begin{array}{c} 0 \\ 2k_1 \\ 2k_2 \end{array}\right. & \begin{array}{c} 1 \\ 0 \\ k_2 \end{array} & \left.\begin{array}{c} 0 \\ k_1 \\ 0 \end{array}\right] \end{array} \text{tomorrow}$$

today

$2k_1 + k_1 = 1 \qquad 2k_2 + k_2 = 1$

$k_1 = \frac{1}{3} \qquad k_2 = \frac{1}{3}$

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 2/3 & 0 & 1/3 \\ 2/3 & 1/3 & 0 \end{bmatrix}$$

Let UFPV be $v = (x\ y\ z)$

$vP = v$ & $x + y + z = 1$

$$(x\ y\ z)\begin{bmatrix} 0 & 1 & 0 \\ 2/3 & 0 & 1/3 \\ 2/3 & 1/3 & 0 \end{bmatrix} = (x\ y\ z)$$

$$\left(\frac{2y+2z}{3} \quad x + \frac{z}{3} \quad \frac{y}{3}\right) = (x\ y\ z)$$

$x + \frac{z}{3} = y$, $y/3 = z$

$x + \frac{z}{3} = 3z$ $\qquad y = 3z$

$x = \frac{8z}{3}$ $\qquad y = 3z$

$x = \frac{8}{3}\left(y/3\right)$

$x = \frac{8y}{9}$

$x + y + z = 1$

$\frac{8z}{3} + 3z + z = 1$

$8z + 9z + 3z = 3$

$20z = 3$

$z = 3/20$, $y = 9/20$ & $x = \frac{8}{20}$

UFPV $v = \left(\frac{8}{20}, \frac{9}{20}, \frac{3}{20}\right)$

In long run P of selling in $A = 8/20$, $B = 9/20$, $C = 3/20$

③ In certain city, weather on a day is reported sunny (S), cloudy (C), or rainy (R). If a day is sunny, prob. that next day is sunny is 70%, cloudy 20% & rainy 10%. If a day is cloudy, prob that the next day is sunny is 30%, cloudy 20%, rainy 50%. If a day is rainy, prob. that next day is sunny is 30%, cloudy 30% & rainy 40%. If Sunday is sunny find probability that wednesday is rainy

→ $S = \{S, C, R\}$

$$\begin{array}{c|c} & \text{tomorrow} \\ P = & \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.2 & 0.5 \\ 0.3 & 0.3 & 0.4 \end{bmatrix} \\ \text{today} & \end{array}$$

Initial probability vector $p = (1, 0, 0)$ (of Sunday)

wt $p^{(3)} = p^{(1)} P \cdot P$

$p^{(3)} = p^{(0)} P \cdot P^2$

On wednesday : $p^{(3)} = p^{(0)} P^3$

$$= (1\ 0\ 0)\begin{bmatrix} 0.532 & 0.221 & 0.247 \\ 0.468 & 0.233 & 0.299 \\ 0.468 & 0.234 & 0.298 \end{bmatrix}$$

$$p^{(3)} = \begin{pmatrix} \overset{S}{0.532} & \overset{C}{0.221} & \overset{R}{0.247} \end{pmatrix}$$

On wednesday, probability of sunny $= 0.532$

cloudy $= 0.221$

rainy $= 0.247$

Q) 14/2/22 3 boys A, B, C are throwing ball to eachother. A always throws the ball to B and B always throws the ball to C. C is just as likely to throw the ball to B as to A. If C was the first person to throw the ball, find probability that after 3 throws ⓐ A has the ball, ⓑ B has the ball, ⓒ C has the ball.

→ State space = {A, B, C}

T.P.M

$$P = \begin{array}{c|c} & \begin{array}{ccc} A & B & C \end{array} \; \to \text{next} \\ \hline \begin{array}{c} A \\ B \\ C \end{array} & \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \end{array}$$

↓ now

Initial probability $p^{(0)} = \begin{array}{ccc} A & B & C \\ (0, & 0, & 1) \end{array}$

[C was first to throw ball]

$\therefore p^{(0)}$ ↓ after first throw $p^{(1)}$ ↓ $p^{(2)}$ ↓ $\boxed{p^{(3)}}$ → to calculate

$$p^{(3)} = p^{(2)}P = p^{(1)}P^2 = p^{(0)}P^3$$

$$p^{(3)} = (0 \;\; 0 \;\; 1)\begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

$$= (0.25 \quad 0.25 \quad 0.5)$$

Q) A gambler's luck follows a pattern. If he wins a game, the probability of winning the next game is 0.6. However if he loses a game, the prob. of losing the next game is 0.7. There is an even chance of the gambler winning the first game. If so

ⓐ what is the probability of he winning the third game?

ⓑ In long run how often will he win?

→ State space : {W, L} W-winning

T.P.M

$$P = \begin{array}{c|c} & \begin{array}{cc} W & L \end{array} \; \text{next} \\ \hline \begin{array}{c} W \\ L \end{array} & \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix} \end{array}$$

↓ now

Initial probability $p^{(0)} = \begin{array}{cc} W & L \\ (\frac{1}{2} & \frac{1}{2}) \end{array}$ [for first game]

$p^{(0)}$ → first
P ↓
$p^{(1)}$ → second
↓
$p^{(2)}$ → third

ⓐ $p^{(2)} = p^{(0)} \cdot P^2$

$$= (0.5 \quad 0.5)\begin{bmatrix} 0.48 & 0.52 \\ 0.39 & 0.61 \end{bmatrix}$$

$$= (0.435 \quad 0.565)$$

ⓑ Let $V = (x \quad y)$ be U.F.P.V

$VP = V$ & $x + y = 1$

$$(x \quad y)\begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix} = (x \quad y)$$

$$(0.6x + 0.3y \quad 0.4x + 0.7y) = (x \quad y)$$

$0.6x + 0.3y = x$

$0.3y = 0.4x$

$x = \frac{3y}{4}$

$x + y = 1$

$\frac{3}{4}y + y = 1$

$y = \frac{4}{7} \Rightarrow x = \frac{3}{7}$

$\therefore$ UFPV $= \begin{array}{cc} W & L \\ (\frac{3}{7} & \frac{4}{7}) \end{array}$

In long run prob of winning is $3/7$.

(8) A housewife buys 3 kinds of cereals A, B & C. She never buys the same cereal in successive weeks. If she buys cereal A the next week, she buys cereal B. However if she buys cereal B or C, the next week she is three times as likely to buy cereal A as to other cereal. In long run, how often she buys each of the 3 cereals?

→ State space : { A B C }

$$\text{T.P.M } P = \begin{array}{c|c} & \begin{array}{ccc} A & B & C \end{array} \\ \begin{array}{c} A \\ B \\ C \end{array} & \begin{bmatrix} 0 & 1 & 0 \\ 3k_1 & 0 & k_1 \\ 3k_2 & k_2 & 0 \end{bmatrix} \end{array}$$

→ next week

↓ now

$3k_1 + k_1 = 1$

$k_1 = 1/4$

$k_2 = 1/4$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 3/4 & 0 & 1/4 \\ 3/4 & 1/4 & 0 \end{bmatrix}$$

Let $V = (x \ y \ z)$ be U.F.P.V

$VP = V$ and $x + y + z = 1$

$$(x \ y \ z) \begin{bmatrix} 0 & 1 & 0 \\ 3/4 & 0 & 1/4 \\ 3/4 & 1/4 & 0 \end{bmatrix} = (x \ y \ z)$$

$$\left(\tfrac{3}{4}y + \tfrac{3}{4}z \quad x + \tfrac{z}{4} \quad \tfrac{y}{4} \right) = (x \ y \ z)$$

$y = 4z$

$x + z/4 = y \Rightarrow 4x + z = 4y$

$4x + z = 16z$

$x = 15/4 \, z$

$x + y + z = 1$

$\frac{15}{4}z + 4z + z = 1$

$(15 + 20)z = 4$

$z = 4/35$

$y = 16/35$, $x = 15/35$.

$\Rightarrow$ UFPV, $V = \left(\frac{15}{35} \quad \frac{16}{35} \quad \frac{4}{35} \right)$

In long run prob. of buying cereal A = $15/35$

B = $16/35$

C = $4/35$

---

(9) The pattern of sunny & rainy days on the planet rainbow is a homogenous Markov chain with two states. Every sunny day is followed by another sunny day with prob 0.8. Every rainy day is followed by another rainy day with prob 0.6. Today is sunny on Rainbow.

(a) What is prob of rain the day after tomorrow

(b) Compute the prob. that april 1st next year is rainy on rainbow?

→ State space = { S R }

S → Sunny R → Rainy

$$\text{T.P.M} \quad P = \begin{array}{c|c} & \begin{array}{cc} S & R \end{array} \\ \begin{array}{c} S \\ R \end{array} & \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} \end{array}$$

today $\{p^{(0)}$

↓

tomorrow $p^{(1)}$

↓

$p^{(2)}$

Initial prob. $p^{(0)} = \begin{pmatrix} S & R \\ 1 & 0 \end{pmatrix}$

$$p^{(2)} = p^{0} \cdot p^{2} = (1 \ 0) \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix}^2 = (1 \ 0) \begin{bmatrix} 0.72 & 0.2 \\ 0.56 & 0.4 \end{bmatrix}$$

$= (0.72 \quad 0.28)$

a) Prob of rainy the day after tomorrow = 0.28

b) Let $V = (x \quad y)$ be U.F.P.V

$VP = V$ & $x + y = 1$

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} = (x \quad y)$$

$$(0.8x + 0.4y \quad 0.2x + 0.6y) = (x \quad y)$$

$0.8x + 0.4y = x$

$0.4y = 0.2x$

$2y = x$

$x + y = 1$

$3y = 1$

$y = 1/3$, $x = 2/3$

UFPV $V = (2/3 \quad 1/3)$

Prob that april 1st next year is rainy is $1/3$.

16/2/22

QUEUEING THEORY

Arrival follows Poisson distribution. Service follows Poisson distribution.

$\lambda \to$ arrival rate ; $\frac{1}{\lambda} \to$ arrival time

$\mu \to$ service rate ; $\frac{1}{\mu} \to$ service time

M|M|1 Model

arrival pattern & service pattern both follow Poisson dist^n.

① At any instant of time, the probability that there are 'n' arrivals (customers) in the queue (waiting line) including those being served is given by

$$P_n = (1-\rho)\rho^n, \quad n \geq 0$$

Here $\rho = \frac{\lambda}{\mu}$. ρ is called as the traffic intensity or utilisat^n factor. Its unit is Erlang.

ρ - utilisat^n factor generally < 1

$\rho = \frac{\lambda}{\mu} < 1$

$\Rightarrow \lambda < \mu$

$\boxed{\text{arrival rate} < \text{service rate}}$

(Model / system fails if $\rho \geq 1$)

② Probability that there is no one in queue, (ie service unit is idle)

$$P_0 = (1-\rho) = \left(1 - \frac{\lambda}{\mu}\right)$$

$q \to$ length

$S \to$ system q + service

$S + 2$

③ The avg. (expected) queue length is given by L_q.

$L_q = L_s - \rho$ ie [Expected no. of units in system − Expected no. in service]

$$L_q = \frac{\rho^2}{1-\rho} = \frac{\lambda^2}{\mu(\mu-\lambda)}$$

$W_q \to$ waiting time in q

$W_s \to$ in system

$W_q = W_s - \frac{1}{\mu}$

④ The mean or expected waiting time of customers in the system (including service time) is

$$W_s = \frac{1}{\mu(1-\rho)}$$

$$W_s = \frac{1}{\mu-\lambda}$$

i.e expected time, a unit spends in the system.

⑤ $L_s = \frac{\lambda}{\mu - \lambda}$

⑤ The expected waiting time per unit in the queue
i.e. expected waiting time of customers in the queue.
(excluding service time) is

$W_q = W_s - \frac{1}{\mu}$ [∵ Expected time in system − time in service]

$W_q = \frac{\rho}{\mu(1-\rho)}$ $W_q = \frac{\lambda}{\mu(\mu-\lambda)}$

⑥ Avg. length of non-empty queue

$L_n = \frac{1}{1-\rho} = \frac{\mu}{\mu-\lambda}$

⑦ Avg. waiting time in a non-empty queue

$W_n = \frac{1}{\mu-\lambda}$

⑧ Probability of queue length being greater than or equal to n.

$P(\geq n) = \rho^n = \left(\frac{\lambda}{\mu}\right)^n$

⑨ Probability of more than k units in system, where n is the no. of units in system,

$P_{n>k} = \left(\frac{\lambda}{\mu}\right)^{k+1}$

Interrelations b/w L_s, L_q, W_s & W_q

① $L_s = \lambda W_s$

① The arrival rate of customer at a banking counter has a mean of 45 per hour. The service rate of the counter clerk has a mean of 60 per hour.

a) What is the prob. of having no customers in system

b) " " 5 " " "

c) Find L_s, L_q, W_s, W_q.

→ The arrival rate, $\lambda = 45/h$; $\mu = 60/h$

$\rho = \frac{\lambda}{\mu} = \frac{45}{60} = \frac{3}{4}$

a) Prob. of no customers $= P_0 = 1-\rho = 1 - 3/4$
$= 1/4$

b) Prob. of having 5 customers $= P_5 = (1-\rho)\rho^5 = (1-3/4)\left(\frac{3}{4}\right)^5$
$= 1/4 \cdot (3/4)^5$

c) $L_s = \frac{\rho}{1-\rho} = \frac{3/4}{1/4} = 3$ customers

$L_q = \frac{\rho^2}{1-\rho} = \frac{(3/4)^2}{1-3/4} = 9/4 = 2.25$ customers

$W_s = \frac{1}{\mu-\lambda} = \frac{1}{60-45} = \frac{1}{15}$ hours $= 0.0667$ hours

$W_q = \frac{\rho}{\mu(1-\rho)} = \frac{3/4}{60(1-3/4)} = \frac{3}{60} = \frac{1}{20} = 0.05$ hours

② The arrival rate of customers at single window booking counter of a two wheeler agency follows Poisson dist^n & service rate follows Poisson dist^n. The arrival rate & service rate are 25 customers/hour & 35 " " resp.

Find,

a) Utilizatⁿ of booking clerk

b) Avg. no. of waiting customers in q.

c) Avg. no. of waiting " " system.

d) Avg. waiting time per customer in q.

e) Avg. " " " " in system.

→ (M/M/1 Model)

$\lambda = 25/hr$ $\mu = 35/hr.$

a) $\rho = \frac{\lambda}{\mu} = \frac{25}{35} = \frac{5}{7}$ [out of 7 hours, counter is busy for 5 hours]

b) $L_q = \frac{\rho^2}{1-\rho} = \frac{25/49}{2/7} = \frac{25}{14} = 1.785$ customers

c) $L_s = \frac{\rho}{1-\rho} = \frac{5/7}{2/7} = 5/2 = 2.5$ customers

d) $W_q = \frac{\rho}{\mu(1-\rho)} = \frac{5/7}{35(2/7)} = \frac{1}{14}$ hrs $= 0.0714$ hrs

e) $W_s = \frac{1}{\mu-\lambda} = \frac{1}{35-25} = \frac{1}{10} = 0.1$ hrs

---

Q) A TV repairman finds that the average time spent on a job is 30 minutes. If he repairs sets in the order in which they come in & if the avg. arrival of sets is 10 per 8 hours a day,

a) What is the repairman's expected idle time each day?

b) How many jobs are ahead of the average set just brought in?

→ $\frac{1}{\mu} = 30$ minutes $\Rightarrow \mu = \frac{1}{30}$ sets per min $= 2$ sets per hour.

$\lambda = 10$ per 8 hours a day

$\lambda = \frac{10}{8}$ sets per hour

$\rho = \frac{\lambda}{\mu} = \frac{10/8}{2} = 5/8$

a) $P_0 = 1-\rho = 1-5/8 = 3/8$

$\Rightarrow$ His idle time is 3 hrs.

b) $L_s = \frac{\rho}{1-\rho} = \frac{5/8}{3/8} = 5/3$ sets. $= 1.67$ sets.

Q) In 24 hours service station, vehicles arrive @ of 30 per day on the avg.. The avg. servicing time for a vehicle is 36 minutes. Find

a) The mean no. of vehicles waiting in the system

b) The probability that queue size exceeds 9.

→ $\lambda = 30$ per day $= \frac{30}{24 \times 60} = \frac{1}{48}$ per min

$\frac{1}{\mu} = 36$ min $\Rightarrow \mu = \frac{1}{36}$ per min

$\rho = \frac{\lambda}{\mu} = \frac{36}{48} = 3/4 = 0.75$

a) Mean no. of vehicles waiting in the system:

$L_s = \frac{\rho}{1-\rho} = \frac{3/4}{1/4} = 3$ vehicles

b) $P(\geq 10) = \rho^n = \rho^{10} = \left(\frac{\lambda}{\mu}\right)^{10}$

$= (0.75)^{10} = 0.0563135$

5) In a railway marshalling yard, goods train arrive @ of 30 trains / day. Assuming that the inter-arrival time follows an exponential dist^n & service time is also exponential with an avg. 36 min, Calculate

a) Avg. no. of trains in the yard.

b) Prob. that queue size exceeds 6

c) Expected waiting time in the queue

d) Avg. no. of trains in the queue.

→ $\lambda = 30$ trains / day $= \frac{30}{24 \times 60} = \frac{1}{48}$ per min

$\frac{1}{\mu} = 36$ min $\Rightarrow \mu = \frac{1}{36}$ per min

$\rho = \frac{\lambda}{\mu} = \frac{36}{48} = 3/4 = 0.75$

a) $L_s = \frac{\rho}{1-\rho} = \frac{3/4}{1/4} = 3$

b) Probability that queue size exceeds 6 :

$P(\geq 7) = \rho^7 = (3/4)^7 =$

c) $W_q = \frac{\lambda}{\mu(\mu-\lambda)} = \frac{\rho}{\mu(1-\rho)} = \frac{3/4}{1/36\,(1/4)} = 3 \times 36 = 108$ min

d) $L_q = \frac{\rho^2}{1-\rho} = \frac{9/16}{1/4} = \frac{9}{4} = 2.25$ trains

---

6) Customers arrive in a telephone booth at intervals of 10 mins on the avg. The length of a phone call is 3 min on the avg.

a) What is the probability that a person arriving at the booth will have to wait

b) What is the average length of queue that forms time to time.

c) Estimate the fraction of a day that phone is in use

d) Find average no. of units in the system.

5) The telephone dept will install a second booth when convinced that an arrival would expect to wait atleast 3 min for the phone. By how much the flow of arrivals be increased in order to justify second booth ?

→ $\frac{1}{\lambda} = 10$ min, $\lambda = \frac{1}{10}$ per min ; $\frac{1}{\mu} = 3$ min, $\mu = \frac{1}{3}$ per min

$\rho = \frac{\lambda}{\mu} = \frac{1/10}{1/3} = \frac{3}{10} = 0.3$

a) $1 - P_0 = 1 - (1-\rho)$

$= \rho = 0.3$

b) $L_n = \frac{1}{1-\rho} = \frac{1}{0.7} = 1.428$

c) ρ (Utilisat^n factor) $= 0.3$

d) $L_s = \frac{\rho}{1-\rho} = \frac{0.3}{0.7} = \frac{3}{7} = 0.428$

e) (M/M/2 Model)

Let λ' be the arrival rate (new)

$W_q \geq 3$

$\frac{\lambda'}{\mu(\mu-\lambda')} \geq 3 \Rightarrow$

$\lambda' \geq 3\mu(\mu - \lambda')$

$\lambda' \geq 3\mu^2 - 3\mu\lambda'$

$\lambda' \geq \frac{3\mu^2}{1+3\mu}$

$$\lambda' \geq \frac{3\left(\frac{1}{9}\right)}{1+3\left(\frac{1}{3}\right)}$$

$$\lambda' \geq \frac{1}{6}$$

Thus the arrival rate should be atleast 0.16 person per min.

$\therefore$ Increase in arrival rate $= 0.16 - 0.1 = 0.06$ person/min

$(\because \lambda' - \lambda)$

POINTS TO REMEMBER

UNIT - 2

(Principle of Inclusion- Exclusion)

1. The number of elements in S that satisfy exactly m of the n conditions $(0 \leq m \leq n)$:

$$E_m = S_m - \binom{m+1}{1} S_{m+1} + \binom{m+2}{2} S_{m+2} - \cdots + (-1)^{n-m} \binom{n}{n-m} S_n$$

2. The number of elements in S that satisfy at least m of the n conditions $(1 \leq m \leq n)$

$$L_m = S_m - \binom{m}{m-1} S_{m+1} + \binom{m+1}{m-1} S_{m+2} - \cdots + (-1)^{n-m} \binom{n-1}{m-1} S_n$$

---

* Covariance can be -ve