

PROBABILITY & QUEUEING THEORY

(As per SRM UNIVERSITY Syllabus)

Dr. P. GODHANDARAMAN

M.Sc., M.A., M.B.A., Ph.D.

Assistant Professor (Senior Grade)

Department of Mathematics

SRM IST, Kattankulathur – 603 203

Email : godhanda@gmail.com

Mobile : 9941740168



UNIT – III : TESTING OF HYPOTHESIS SMALL SAMPLES ($n < 30$)

- **t – test**
- **F – test**
- **χ^2 – test (Chi Square – test)**

Students ‘t’ – Test - SINGLE MEAN

Standard deviation given directly: $t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n-1}}\right)}$, where $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$,

\bar{x} - sample mean, μ - population mean, n - sample size, s^2 - sample variance

Standard deviation not given directly: $t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n}}\right)}$, where $S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$

\bar{x} - sample mean, μ - population mean, n - sample size, S^2 - population variance σ^2 , **Degree of freedom : $n - 1$**

DIFFERENCE MEAN

$t = \frac{\bar{x} - \bar{y}}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$, where $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$, $\bar{y} = \frac{\sum_{j=1}^n y_j}{n}$, **Degree of freedom : $n_1 + n_2 - 2$**

Standard deviation given directly: $s^2 = \frac{[n_1 s_1^2 + n_2 s_2^2]}{n_1 + n_2 - 2}$

Standard deviation not given directly: $s^2 = \frac{\sum (x - \bar{x})^2 + \sum (y - \bar{y})^2}{n_1 + n_2 - 2}$,

Confidence or Fiducial Limits for μ :

95% Confidence limits for μ : $\bar{x} - t_{0.05} \left(\frac{s}{\sqrt{n}}\right) \leq \mu \leq \bar{x} + t_{0.05} \left(\frac{s}{\sqrt{n}}\right)$

99% Confidence limits for μ : $\bar{x} - t_{0.01} \left(\frac{s}{\sqrt{n}}\right) \leq \mu \leq \bar{x} + t_{0.01} \left(\frac{s}{\sqrt{n}}\right)$

Application of t – distribution:

The t – distribution has a wide number of applications in statistics, some of which are enumerated below

- (i) To test if the sample mean \bar{x} differs significantly from the hypothetical value μ of the population mean.
- (ii) To test the significance of the difference between two sample means.
- (iii) To test the significance of an observed sample correlation coefficient and sample regression coefficient
- (iv) To test the significance of observed partial correlation coefficient.

Assumption for students’ t – test:

- (i) The parent population from which the sample is drawn is normal.
- (ii) The sample observations are independent, that is, the sample is random.
- (iii) The population standard deviation σ is unknown.

PROBLEMS

Students ‘t’ – Test : Single Mean and Standard Deviation Given Directly

1. The mean weekly sales of soap bars in departmental stores was 146.3 bars per store. After an advertising campaign the mean weekly sales in 22 stores for a typical week increased to 153.7 and showed a S.D. of 17.2. Was the advertising campaign successful?

Solution: $n = 22$, $\bar{x} = 153.7$, $\mu = 146.3$, $s = 17.2$

Null Hypothesis H_0 : $\mu = 146.3$, i.e. The advertising campaign is not successful.

Alternate Hypothesis H_1 : $\mu > 146.3$ (Right tailed)

$t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n-1}}\right)} = \frac{153.7 - 146.3}{\left(\frac{17.2}{\sqrt{22-1}}\right)} = 1.96$, degree of freedom: $n - 1 = 21$ at 5% level of significance = 1.72.

Calculate value $t >$ Tabulated t . **H_0 is rejected.**

Conclusion: The advertising campaign was definitely successful in promoting sales.

2. A sample of 20 items has mean 42 units and S.D. 5 units. Test the hypothesis that it is a random sample from a normal population with mean 45 units.

Solution: $n = 20$, $\bar{x} = 42$, $\mu = 45$, $s = 5$

Null Hypothesis H_0 : $\mu = 45$, **Alternate Hypothesis** H_1 : $\mu \neq 45$ (Two tailed)

$t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n-1}}\right)} = \frac{42 - 45}{\left(\frac{5}{\sqrt{20-1}}\right)} = 2.615$, degree of freedom: $n - 1 = 19$ at 5% level of significance = 2.09.

Calculate value $t >$ Tabulated t . **H_0 is rejected.** The sample could not have come from this population.

3. A spare part manufacturer is making spare parts with axle diameter of 0.700 inch. A random sample of 10 parts shows a mean diameter of 0.742 inch with a S.D. of 0.040 inch. Verify whether the work satisfies the specifications.

Solution: $n = 10$, $\bar{x} = 0.742$, $\mu = 0.700$, $s = 0.040$

Null Hypothesis $H_0 : \mu = 0.700$, **Alternate Hypothesis** $H_1 : \mu \neq 0.700$ (Two tailed)

$$t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n-1}}\right)} = \frac{0.742 - 0.700}{\left(\frac{0.040}{\sqrt{10-1}}\right)} = 3.15, \text{ degree of freedom: } n - 1 = 9 \text{ at } 5\% \text{ level of significance} = 2.26.$$

Calculate value $t >$ Tabulated t . **H_0 is rejected. Conclusion:** The product is not meeting the specifications.

4. The mean life time of a sample of 25 bulbs is found as 1550 hours with a S.D. of 120 hours. The company manufacturing the bulbs claims that the average life of their bulbs is 1600 hours. Is the claim acceptable at 5% level of significance?

Solution: $n = 25$, $\bar{x} = 1550$, $\mu = 1600$, $s = 120$

Null Hypothesis $H_0 : \bar{x} = \mu$, **Alternate Hypothesis** $H_1 : \bar{x} < \mu$ (Left Tailed)

$$t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n-1}}\right)} = \frac{1550 - 1600}{\left(\frac{120}{\sqrt{25-1}}\right)} = -2.04, \text{ degree of freedom: } n - 1 = 24 \text{ at } 5\% \text{ level of significance} = 1.71.$$

Calculate value $t >$ Tabulated t . **H_0 is rejected.** The claim of the company cannot be accepted at 5% LOS.

Students 't' – Test : Single Mean and Standard Deviation Not Given Directly

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

Solution: $n = 10$, $\mu = 100$, $H_0 : \mu = 100$, $H_1 : \mu \neq 100$ (Two tail)

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{972}{10} = 97.2, \quad S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{1833.6}{10-1} = 203.73$$

x	$x - \bar{x}$	$(x - \bar{x})^2$
70	-27.2	739.84
120	22.8	519.84
110	12.8	163.84
101	3.8	14.44
88	-9.2	84.64
83	-14.2	201.64
95	-2.2	4.84
98	0.8	0.64
107	9.8	96.04
100	2.8	7.84
$\sum x = 972$		$\sum (x - \bar{x})^2 = 1833.60$

$$t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{97.2 - 100}{\left(\frac{\sqrt{203.73}}{\sqrt{10}}\right)} = \frac{-2.8}{(\sqrt{20.373})} = -0.62 \Rightarrow |t| = 0.62 ; \text{ d.f.} = n - 1 = 10 - 1 = 9 \text{ at } 5\% \text{ LOS} = 2.262.$$

Calculate value $t <$ Tabulated t . **H_0 is accepted.**

Conclusion: The data are consistent with the assumption for a mean I.Q. of 100 in population.

95% Confidence Limits for μ : $\bar{x} - t_{0.05} \left(\frac{s}{\sqrt{n}}\right) \leq \mu \leq \bar{x} + t_{0.05} \left(\frac{s}{\sqrt{n}}\right)$

$$97.2 - 2.262 \left(\sqrt{\frac{203.73}{10}}\right) \leq \mu \leq 97.2 + 2.262 \left(\sqrt{\frac{203.73}{10}}\right)$$

$86.99 \leq \mu \leq 107.41$, Hence the required 95% confidence interval is [86.99, 107.41].

6. The wages of 10 workers taken at random from a factory are given as Wages: 578, 572, 570, 568, 572, 578, 570, 572, 596, 584. Is it possible that the mean wage of all workers of this factory could be Rs. 580

Solution: $n = 10$, $\mu = 580$, $H_0 : \mu = 580$, $H_1 : \mu \neq 580$ (Two tailed)

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{5760}{10} = 576, \quad S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{656}{10-1} = 72.89$$

d.f. = $n - 1 = 10 - 1 = 9$ at 5% LoS = 2.262.

x	$x - \bar{x}$	$(x - \bar{x})^2$
578	2	4
572	-4	16
570	-6	36
568	-8	64
572	-4	16
578	2	4
570	-6	36
572	-4	16
596	20	400
584	8	64
$\sum x = 5760$		$\sum (x - \bar{x})^2 = 656$

$$t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{576 - 580}{\left(\frac{\sqrt{72.89}}{\sqrt{10}}\right)} = \frac{-4}{(\sqrt{7.289})} = -1.48 \Rightarrow |t| = 1.48 ; \text{ Calculate value } t < \text{Tabulated } t. \text{ } H_0 \text{ is accepted.}$$

Conclusion: Is it possible that the mean wage of all workers of this factory could be Rs. 580.

Students 't' – Test : Difference Mean and Standard Deviation Given Directly

7. Sample of two types of electric light bulbs were tested for length of life and following data were obtained

	Type I	Type II
Sample Number	$n_1 = 8$	$n_2 = 7$
Sample Means	$\bar{x}_1 = 1234 \text{ hrs}$	$\bar{x}_2 = 1036 \text{ hrs}$
Sample S.D.	$s_1 = 36 \text{ hrs}$	$s_2 = 40 \text{ hrs}$

Is the difference in the means sufficient to warrant that type I is superior to type II regarding length of life

Solution: Null Hypothesis : $H_0: \bar{x}_1 = \bar{x}_2$, i.e. The two types I and II of electric bulbs are identical.

Alternate Hypothesis : $H_1: \bar{x}_1 > \bar{x}_2$ (Right tail)

Standard deviation given directly: $s^2 = \frac{[n_1 s_1^2 + n_2 s_2^2]}{n_1 + n_2 - 2} = \frac{[8(36)^2 + 7(40)^2]}{8+7-2} = 1659.08$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{1234 - 1036}{\sqrt{1659.08 \left(\frac{1}{8} + \frac{1}{7}\right)}} = 9.39, \text{ degree of freedom } = n_1 + n_2 - 2 = 3 \text{ at } 5\% = 1.77.$$

Calculate value $t > \text{Tabulated } t. H_0 \text{ is rejected. The type I is definitely superior to type II regarding length of life.}$

Students 't' – Test : Difference Mean and Standard Deviation Not Given Directly

8. Below are given the gain in weights (in kgs) of pigs fed on two diets A and B.

Diet A	25	32	30	34	24	14	32	24	30	31	35	25	-	-	-
Diet B	44	34	22	10	47	31	40	30	32	35	18	21	35	29	22

Test if the two diets differ significantly as regards their effect on increase in weight.

Solution: Null Hypothesis $H_0: \bar{x}_1 = \bar{x}_2$, i.e. There is no significant difference between the mean increase in weight due to diets A & B, Alternate Hypothesis : $H_1: \bar{x}_1 \neq \bar{x}_2$ (Two tailed)

x	$x - \bar{x}$	$(x - \bar{x})^2$	y	$y - \bar{y}$	$(y - \bar{y})^2$
25	-3	9	44	14	196
32	4	16	34	4	16
30	2	4	22	-8	64
34	6	36	10	-20	400
24	-4	16	47	17	289
14	-14	196	31	1	1
32	4	16	40	10	100
24	-4	16	30	0	0
30	2	4	32	2	4
31	3	9	35	5	25
35	7	49	18	-12	144
25	-3	9	21	-9	81
$\sum x = 336$		$\sum (x - \bar{x})^2 = 380$	35	5	25
			29	-1	1
			22	-8	64
			$\sum y = 450$		$\sum (y - \bar{y})^2 = 1410$

$$n_1 = 12, n_2 = 15, \bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{336}{12} = 28, \bar{y} = \frac{\sum_{j=1}^n y_j}{n} = \frac{450}{15} = 30$$

$$\text{Standard deviation not given directly: } s^2 = \frac{\sum(x-\bar{x})^2 + \sum(y-\bar{y})^2}{n_1+n_2-2} = 71.6$$

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{28-30}{\sqrt{71.6\left(\frac{1}{12}+\frac{1}{15}\right)}} = -0.609 \Rightarrow |t| = \mathbf{0.609}, \text{ d.f.} = n_1 + n_2 - 2 = 25 \text{ at } 5\% \text{ LOS} = 2.06.$$

Calculate value $t < \text{Tabulated } t$. **H_0 is accepted.**

Conclusion: The two diets do not differ significantly as regards their effect on increase in weight.

9. A group of five patients treated with medicine A weigh 42, 39, 48, 60 and 41 kg.; a second group of 7 patients from the same hospital treated with medicine B weigh 38, 42, 56, 64, 68, 69 and 62 kg. Do you agree with the claim that medicine B increases the weight significantly?

Solution: Null Hypothesis : $H_0: \mu_1 = \mu_2$, Alternate Hypothesis : $H_1: \mu_1 < \mu_2$

x	$x - \bar{x}$	$(x - \bar{x})^2$	y	$y - \bar{y}$	$(y - \bar{y})^2$
42	-4	16	38	-19	361
39	-7	49	42	-15	225
48	2	4	56	-1	1
60	14	196	64	7	49
41	-5	25	68	11	121
$\Sigma x = 230$		$\Sigma(x - \bar{x})^2 = 290$	69	12	144
			62	5	25
			$\Sigma y = 399$		$\Sigma(y - \bar{y})^2 = 926$

$$n_1 = 5, n_2 = 7, \bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{230}{5} = 46, \bar{y} = \frac{\sum_{j=1}^n y_j}{n} = \frac{399}{7} = 57$$

$$\text{Standard deviation not given directly: } s^2 = \frac{\sum(x-\bar{x})^2 + \sum(y-\bar{y})^2}{n_1+n_2-2} = \frac{290+926}{5+7-2} = \frac{1216}{10} = 121.6$$

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{46-57}{\sqrt{121.6\left(\frac{1}{5}+\frac{1}{7}\right)}} = \frac{46-57}{\sqrt{41.69}} = \frac{-11}{6.46} = -1.703 \Rightarrow |t| = \mathbf{1.703}, \text{ d.f.} = n_1 + n_2 - 2 = 10 \text{ at } 5\% = 1.81.$$

Calculate value $t < \text{Tabulated } t$. **H_0 is accepted.** Medicines A and B do not differ significantly.

10. The table below represent the values of protein content from cow's milk and buffalo's milk at a certain level. Examine if these differences are significant.

Cow's milk	1.82	2.02	1.88	1.61	1.81	1.54
Buffalo's milk	2	1.83	1.86	2.03	2.19	1.88

Solution: Null Hypothesis $H_0: \bar{x}_1 = \bar{x}_2$, Alternate Hypothesis : $H_1: \bar{x}_1 \neq \bar{x}_2$ (Two tailed)

x_1	x_1^2	x_2	x_2^2
1.82	3.3124	2	4.0040
2.02	4.0804	1.83	3.3489
1.88	3.5344	1.86	3.4596
1.61	2.5921	2.03	4.1209
1.81	3.2761	2.19	4.761
1.54	2.3716	1.88	3.5344
$\Sigma x_1 = 10.68$	$\Sigma x_1^2 = 19.1670$	$\Sigma x_2 = 11.79$	$\Sigma x_2^2 = 23.2599$

$$n_1 = 6, n_2 = 6, \bar{x}_1 = \frac{\Sigma x_1}{n_1} = \frac{10.68}{6} = 1.78; \bar{x}_2 = \frac{\Sigma x_2}{n_2} = \frac{11.79}{6} = 1.965$$

$$\text{Sample S.D.: } s_1^2 = \frac{\Sigma x_1^2}{n_1} - (\bar{x}_1)^2 = \frac{19.167}{6} - (1.78)^2 = 0.0261; s_2^2 = \frac{\Sigma x_2^2}{n_2} - (\bar{x}_2)^2 = \frac{23.2599}{6} - (1.965)^2 = 0.0154$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{1.78 - 1.965}{\sqrt{\left(\frac{6 \times 0.0261 + 6 \times 0.0154}{6 + 6 - 2}\right)\left(\frac{1}{6} + \frac{1}{6}\right)}} = -2.03 \Rightarrow |t| = \mathbf{2.03}, \text{ d.f.} = n_1 + n_2 - 2 = 10 \text{ at } 5\% = 2.26.$$

Calculate value $t < \text{Tabulated } t$. **H_0 is accepted.**

Conclusion: The difference between the mean protein values of the two varieties of milk is not significant at 5%.

Paired Student 't' – Test

$t = \frac{\bar{d}}{\left(\frac{s}{\sqrt{n-1}}\right)}$, where $s^2 = \frac{\sum d^2}{n} - (\bar{d})^2$; $d = x_1 - x_2$; $\bar{d} = \frac{\sum d}{n}$; d.f. = $n - 1$; Degree of freedom : $n - 1$

11. The following data relate to the marks obtained by 11 students in two tests, one held at the beginning of a year and the other at the end of the year after intensive coaching. Do the data indicate that the students have benefited by coaching?

Test I	19	23	16	24	17	18	20	18	21	19	20
Test II	17	24	20	24	20	22	20	20	18	22	19

Solution: Null Hypothesis $H_0: \bar{x}_1 = \bar{x}_2$; Alternate Hypothesis : $H_1: \bar{x}_1 < \bar{x}_2$ (left tailed)

x_1	x_2	$d = x_1 - x_2$	d^2
19	17	2	4
23	24	-1	1
16	20	-4	16
24	24	0	0
17	20	-3	9
18	22	-4	16
20	20	0	0
18	20	-2	4
21	18	3	9
19	22	-3	9
20	19	1	1
		$\sum d = -11$	$\sum d^2 = 69$

$n_1 = 11$, $n_2 = 11$, $\bar{d} = \frac{\sum d}{n} = \frac{-11}{11} = -1$; $s^2 = \frac{\sum d^2}{n} - (\bar{d})^2 = \frac{69}{11} - (-1)^2 = 5.27$; $s = 2.296$

$t = \frac{\bar{d}}{\left(\frac{s}{\sqrt{n-1}}\right)} = -\frac{1}{\left(\frac{2.296}{\sqrt{11-1}}\right)} = -1.38$; $\Rightarrow |t| = 1.38$, d.f. = $n - 1 = 10$ at 5% = 1.81.

Calculate value $t < \text{Tabulated } t$. **H_0 is accepted. Conclusion:** The students have not benefitted by coaching.

F – Test

$F = \frac{S_X^2}{S_Y^2}$, if $S_X^2 > S_Y^2$ (or) $F = \frac{S_Y^2}{S_X^2}$, if $S_X^2 < S_Y^2$,

$S_X^2 = \frac{\sum (x - \bar{x})^2}{n_1 - 1}$, $S_Y^2 = \frac{\sum (y - \bar{y})^2}{n_2 - 1}$, $S_1^2 = \frac{n_1 s_1^2}{n_1 - 1}$, $S_2^2 = \frac{n_2 s_2^2}{n_2 - 1}$,

degree of freedom : ($\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$).

Applications: (i) To test whether two independent samples have been drawn from the normal populations with the same variance σ^2 (ii) To test whether two independent estimates of the population variance are homogeneous or not.

PROBLEMS

12. If one sample of 8 observations the sum of the squares of deviations of the sample values from the sample mean was 84.4 and in the other sample of 10 observations it was 102.6. Test whether this difference is significant at 5% level.

Solution: $n_1 = 8$, $n_2 = 10$, $\sum (x - \bar{x})^2 = 84.4$, $\sum (y - \bar{y})^2 = 102.6$

Null Hypothesis : $H_0: \sigma_X^2 = \sigma_Y^2$, Alternate Hypothesis : $H_1: \sigma_X^2 \neq \sigma_Y^2$

$S_X^2 = \frac{\sum (x - \bar{x})^2}{n_1 - 1} = \frac{84.4}{8 - 1} = 12.057$, $S_Y^2 = \frac{\sum (y - \bar{y})^2}{n_2 - 1} = \frac{102.6}{10 - 1} = 11.4$, $F = \frac{S_X^2}{S_Y^2} = \frac{12.057}{11.4} = 1.057$,

d.f. = ($n_1 - 1$, $n_2 - 1$) = (7, 9) at 5% LOS = 3.29. Calculate value $F < \text{Tabulated } F$. **H_0 is accepted.**

13. Two random samples of 11 and 9 items show the sample standard deviations of their weights as 0.8 and 0.5 respectively. Assuming that the weight distributions are normal, test the hypothesis that the true variances are equal, against the alternative hypothesis that they are not.

Solution: $n_1 = 11$, $n_2 = 9$, $s_1^2 = 0.8$, $s_2^2 = 0.5$, Null Hypothesis: $H_0: \sigma_X^2 = \sigma_Y^2$, Alternate Hypothesis : $H_1: \sigma_X^2 \neq \sigma_Y^2$

$S_1^2 = \frac{n_1 s_1^2}{n_1 - 1} = \frac{11(0.8)^2}{11 - 1} = 0.704$, $S_2^2 = \frac{n_2 s_2^2}{n_2 - 1} = \frac{9(0.5)^2}{9 - 1} = 0.28125$, $F = \frac{S_1^2}{S_2^2} = \frac{0.704}{0.28125} = 2.503$,

d.f. = ($n_1 - 1$, $n_2 - 1$) = (10, 8) at 5% LOS = 3.34. Calculate value $F < \text{Tabulated } F$. **H_0 is accepted.**

14. The time taken by workers in performing a job by Method I and Method II is given below

Method I	20	16	26	27	23	22	-
Method II	27	33	42	35	32	34	38

Do the data show that the variances of time distribution form population from which these samples are drawn do not differ significantly?

Solution: $n_1 = 6, n_2 = 7, \bar{x} = \frac{\sum_{i=1}^n x_i}{n_1} = \frac{134}{6} = 22.3, \bar{y} = \frac{\sum_{j=1}^n y_j}{n_2} = \frac{241}{7} = 34.4$

x	$x - \bar{x}$	$(x - \bar{x})^2$	y	$y - \bar{y}$	$(y - \bar{y})^2$
20	-2.3	5.29	27	-7.4	54.76
16	-6.3	39.69	33	-1.4	1.96
26	3.7	13.69	42	7.6	57.76
27	4.7	22.09	35	0.6	0.36
23	0.7	0.49	32	-2.4	5.76
22	-0.3	0.09	34	-0.4	0.16
$\Sigma x = 134$		$\Sigma(x - \bar{x})^2 = 81.34$	38	3.6	12.96
			$\Sigma y = 241$		$\Sigma(y - \bar{y})^2 = 132.72$

Null Hypothesis : $H_0: \sigma_x^2 = \sigma_y^2$, Alternate Hypothesis : $H_1: \sigma_x^2 \neq \sigma_y^2$

$$S_x^2 = \frac{\Sigma(x - \bar{x})^2}{n_1 - 1} = \frac{81.34}{6 - 1} = 16.268, S_y^2 = \frac{\Sigma(y - \bar{y})^2}{n_2 - 1} = \frac{132.72}{7 - 1} = 22.29, F = \frac{S_y^2}{S_x^2} = \frac{22.29}{16.268} = 1.37,$$

d.f. = $(n_2 - 1, n_1 - 1) = (6, 5)$ at 5% LOS = 4.95. Calculate value $F < \text{Tabulated } F$. **H_0 is accepted.**

15. Two horses A and B were tested according to the time (in seconds) to run a particular track with the following

Horse A	28	30	32	33	33	29	34
Horse B	29	30	30	24	27	29	-

Test whether the two horses have the same running capacity.

Solution: $n_1 = 7, n_2 = 6, \bar{x} = \frac{\sum_{i=1}^n x_i}{n_1} = \frac{219}{7} = 31.28, \bar{y} = \frac{\sum_{j=1}^n y_j}{n_2} = \frac{169}{6} = 28.2$

x	$x - \bar{x}$	$(x - \bar{x})^2$	y	$y - \bar{y}$	$(y - \bar{y})^2$
28	-3.28	10.75	29	0.8	0.64
30	-1.28	1.64	30	1.8	3.24
32	0.72	0.52	30	1.8	3.24
33	1.72	2.96	24	-4.2	17.64
33	1.72	2.96	27	-1.2	1.44
29	-2.28	5.20	29	0.8	0.64
34	2.72	7.40	-	-	-
$\Sigma x = 219$		$\Sigma(x - \bar{x})^2 = 21.2$	$\Sigma y = 169$		$\Sigma(y - \bar{y})^2 = 108$

F - Test

Null Hypothesis : $H_0: \sigma_x^2 = \sigma_y^2$, Alternate Hypothesis : $H_1: \sigma_x^2 \neq \sigma_y^2$

$$S_x^2 = \frac{\Sigma(x - \bar{x})^2}{n_1 - 1} = \frac{21.2}{7 - 1} = 3.533, S_y^2 = \frac{\Sigma(y - \bar{y})^2}{n_2 - 1} = \frac{108}{6 - 1} = 21.6, F = \frac{S_y^2}{S_x^2} = \frac{21.6}{3.533} = 6.11,$$

degree of freedom = $(n_2 - 1, n_1 - 1) = (5, 6)$ at 5% LOS = 4.39.

Calculate value $F < \text{Tabulated } F$. **H_0 is accepted. Conclusion:** The two horses have the same running capacity.

t - test and F - Test

16. The nicotine contents in milligrams in two samples of tobacco were found to be as follows

Sample A	24	27	26	21	25	-
Sample B	27	30	28	31	22	36

Can it be said that two samples come from normal populations.

Solution: $n_1 = 5, n_2 = 6, \bar{x} = \frac{\sum_{i=1}^n x_i}{n_1} = \frac{123}{5} = 24.6, \bar{y} = \frac{\sum_{j=1}^n y_j}{n_2} = \frac{174}{6} = 29$

x	$x - \bar{x}$	$(x - \bar{x})^2$	y	$y - \bar{y}$	$(y - \bar{y})^2$
24	-0.6	0.36	27	-2	4
27	2.4	5.76	30	1	1
26	1.4	1.96	28	-1	1
21	-3.6	12.96	31	2	4
25	0.4	0.16	22	-7	49
$\sum x = 123$		$\sum (x - \bar{x})^2 = 21.2$	36	7	49
			$\sum y = 174$		$\sum (y - \bar{y})^2 = 108$

Student's t test

Null Hypothesis : $H_0: \bar{x}_1 = \bar{x}_2$, **Alternate Hypothesis** : $H_1: \bar{x}_1 \neq \bar{x}_2$ (Two tailed)

$$s^2 = \frac{\sum (x - \bar{x})^2 + \sum (y - \bar{y})^2}{n_1 + n_2 - 2} = \frac{21.2 + 108}{5 + 6 - 2} = 14.35, t = \frac{\bar{x} - \bar{y}}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{24.6 - 29}{\sqrt{14.35 \left(\frac{1}{5} + \frac{1}{6} \right)}} = -1.92 \Rightarrow |t| = 1.92,$$

degree of freedom = $n_1 + n_2 - 2 = 9$ at 5% LOS for two tailed test = 2.26.

Calculate value $t < \text{Tabulated } t$. **H_0 is accepted.**

F - Test

Null Hypothesis : $H_0: \sigma_X^2 = \sigma_Y^2$, **Alternate Hypothesis** : $H_1: \sigma_X^2 \neq \sigma_Y^2$

$$S_X^2 = \frac{\sum (x - \bar{x})^2}{n_1 - 1} = \frac{21.2}{5 - 1} = 5.3, S_Y^2 = \frac{\sum (y - \bar{y})^2}{n_2 - 1} = \frac{108}{6 - 1} = 21.6, F = \frac{S_Y^2}{S_X^2} = \frac{21.6}{5.3} = 4.07,$$

degree of freedom = $(n_2 - 1, n_1 - 1) = (5, 4)$ at 5% LOS = 6.26.

Calculate value $F < \text{Tabulated } F$. **H_0 is accepted.**

Conclusion: We conclude that the two samples come from same normal population.

χ^2 - test

The χ^2 distribution function is one of the most extensively used distribution functions in statistics.

Application (or uses) of χ^2 distribution

1. To test the goodness of fit.
2. To test the independence of attributes.
3. To test if the hypothetical value of the population variance is σ^2 .
4. To test the homogeneity of independent estimates of the population variance.
5. To test the homogeneity of independent estimates of the population correlation coefficient.

Condition for validity of χ^2 - test

1. The sample observations should be independent.
2. Constraints on the cell frequencies, if any, should be linear, e.g., $\sum O_i = \sum E_i$
3. N, the total frequency should be reasonably large, say, greater than 50.
4. No theoretical cell frequency should be less than 5. (The chi square distribution is frequency is less than 5). If any theoretical cell frequency is less than 5, then for the application of χ^2 test, it is pooled with the preceding or succeeding frequency so that the pooled frequency is more than 5 and finally adjust for the degree of freedom lost in pooling.

χ^2 - test of Goodness of fit

This is a powerful test for testing the significance of the discrepancy between theory and experiment – discovered by Karl Pearson in 1900. It helps us to find if the deviation of the experiment from theory is just by chance or it is due to the inadequacy of the theory to fit the observed data.

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}, \text{ where } O_i - \text{set of observed frequencies, } E_i - \text{Set of expected frequencies. d.f.} = n - 1$$

1. The following table gives the number of aircraft accidents that occurred during the various days of the week. Test whether the accidents are uniformly distributed over the week.

Days	Mon	Tue	Wed	Thu	Fri	Sat
No. of Accidents	14	18	12	11	15	14

Solution: H_0 : Accidents occur uniformly over the week. The expected no. of accidents on any day = $\frac{83}{6} = 14$.

O	E	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
14	14	0	0	0
18	14	4	16	1.143
12	14	-2	4	0.286
11	14	-3	9	0.643
15	14	1	1	0.071
14	14	0	0	0
				2.143

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} = 2.143, \text{ d.f.} = n - 1 = 5 \text{ at } 5\% \text{ LOS} = 11.07, \text{ Calculated } \chi^2 < \text{tabulated } \chi^2,$$

H_0 is accepted. The accidents are uniformly distributed over the week.

2. The theory predicts the proportion of beans, in the four groups A, B, C, and D should be 9 : 3 : 3 : 1. In an experiment with 1600 beans the numbers in the four groups were 882, 313, 287, and 118. Does the experimental result support the theory.

Solution: H_0 : The experimental result support the theory.

O	E	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
882	$\frac{9}{16} \times 1600 = 900$	-18	324	0.36
313	$\frac{3}{16} \times 1600 = 300$	13	169	0.563
287	$\frac{3}{16} \times 1600 = 300$	-13	169	0.563
118	$\frac{1}{16} \times 1600 = 100$	18	324	3.24
				4.726

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} = 4.726, \text{ d.f.} = n - 1 = 4 - 1 = 3 \text{ at } 5\% \text{ LOS} = 7.81$$

Calculated $\chi^2 < \text{tabulated } \chi^2$, H_0 is accepted. The experimental results support the theory.

3. Fit a binomial distribution for the following data and also test the goodness of fit. Find the parameters of the distribution.

x	0	1	2	3	4	5	6	Total
f	5	18	28	12	7	6	4	80

Solution: To find the binomial frequency distribution $N(q + p)^n$, which fits the given data, we require N, n and p .

We assume $N = \text{total frequency} = 80$ and $n = \text{no. of trials} = 6$ from the given data.

x	0	1	2	3	4	5	6	Total
f	5	18	28	12	7	6	4	80
fx	0	18	56	36	28	30	24	192

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{192}{80} = 2.4, np = 2.4 \Rightarrow 6p = 2.4, p = 0.4, q = 0.6, N P(X = x) = N n C_x p^x q^{n-x}, x = 0, 1, \dots, n$$

$$80 P(X = 0) = 80 \times {}^6C_0 (0.4)^0 (0.6)^{6-0} = 3.73, \quad 80 P(X = 1) = 80 \times {}^6C_1 (0.4)^1 (0.6)^{6-1} = 14.93$$

$$80 P(X = 2) = 80 \times {}^6C_2 (0.4)^2 (0.6)^{6-2} = 24.88, \quad 80 P(X = 3) = 80 \times {}^6C_3 (0.4)^3 (0.6)^{6-3} = 22.12$$

$$80 P(X = 4) = 80 \times {}^6C_4 (0.4)^4 (0.6)^{6-4} = 11.06, \quad 80 P(X = 5) = 80 \times {}^6C_5 (0.4)^5 (0.6)^{6-5} = 2.95$$

$$80 P(X = 6) = 80 \times {}^6C_6 (0.4)^6 (0.6)^{6-6} = 0.33$$

X	0	1	2	3	4	5	6	Total
E_i	4	15	25	22	11	3	0	80

The 1st class is combined with the second and the last 2 classes are combined with the last but 2nd class in order to make the expected frequency in each class ≥ 5 . Thus, after regrouping, we have

O_i	23	28	12	17
E_i	19	25	22	14

H_0 : The given distribution is approximately binomial distribution.

O	E	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
23	19	4	16	0.8421
28	25	3	9	0.36
12	22	-10	100	4.5455
17	14	3	9	0.6429
				6.39

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} = \mathbf{6.39}, \text{ d.f.} = n - k = 4 - 2 = 2 \text{ at } 5\% \text{ LOS} = 5.99,$$

Calculated $\chi^2 >$ tabulated χ^2 , H_0 is **rejected**. The binomial fit for the given distribution is not satisfactory.

4. **Fit a Poisson distribution for the following distribution and also test the goodness of fit.**

x	0	1	2	3	4	5	Total
f	142	156	69	27	5	1	400

Solution : $P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, \dots, \infty; \lambda = \bar{x} = \frac{\sum fx}{\sum f} = \frac{400}{400} = 1, N = 400$

x	0	1	2	3	4	5	Total
f	142	156	69	27	5	1	400
fx	0	156	138	81	20	5	400

Theoretical frequencies are given by $N P(X = x) = \frac{N e^{-\lambda} \lambda^x}{x!} = \frac{400 e^{-1} 1^x}{x!}, x = 0, 1, \dots, \infty$

$$400 P(X = 0) = \frac{400 e^{-1} 1^0}{0!} = 147.15, \quad 400 P(X = 1) = \frac{400 e^{-1} 1^1}{1!} = 147.15$$

$$400 P(X = 2) = \frac{400 e^{-1} 1^2}{2!} = 73.58, \quad 400 P(X = 3) = \frac{400 e^{-1} 1^3}{3!} = 24.53$$

$$400 P(X = 4) = \frac{400 e^{-1} 1^4}{4!} = 6.13, \quad 400 P(X = 5) = \frac{400 e^{-1} 1^5}{5!} = 1.23$$

O_i	142	156	69	27	5	1
E_i	147	147	74	25	6	1

The last 3 classes are combined into one, so that the expected frequency in that class may be ≥ 10 .

O_i	142	156	69	33
E_i	147	147	74	32

O	E	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
142	147	-5	25	0.17
156	147	9	81	0.55
69	74	-5	25	0.34
33	32	1	1	0.03
				1.09

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} = \mathbf{1.09}, \text{ d.f.} = n - k = 4 - 2 = 2 \text{ at } 5\% \text{ LOS} = 5.99$$

Calculated $\chi^2 <$ tabulated χ^2 , H_0 is **accepted**. The Poisson fit for the given distribution is satisfactory.

χ^2 – test of Independence of Attributes

Literally, an attribute means a quality or characteristic. E.g.: drinking, smoking, blindness, honesty, etc.

a	b	$(a + b)$
c	d	$(c + d)$
$(a + c)$	$(b + d)$	N

$E(a) = \frac{(a + c)(a + b)}{N}$	$E(b) = \frac{(b + d)(a + b)}{N}$	$(a + b)$
$E(c) = \frac{(a + c)(c + d)}{N}$	$E(d) = \frac{(b + d)(c + d)}{N}$	$(c + d)$
$(a + c)$	$(b + d)$	N

Degree of freedom = $(r - 1)(c - 1)$, where r – number of rows, c – number of columns.

5. On the basis of information given below about the treatment of 200 patients suffering from a disease, state whether the new treatment is comparatively superior to the conventional treatment.

	Favourable	Not Favourable	Total
New	60	30	90
Conventional	40	70	110
	100	100	200

Solution: H_0 : New and conventional treatment are independent.

$E(60) = \frac{90 \times 100}{200} = 45$	$E(30) = \frac{90 \times 100}{200} = 45$	90
$E(40) = \frac{110 \times 100}{200} = 55$	$E(70) = \frac{110 \times 100}{200} = 55$	110
100	100	200

O	E	O - E	(O - E) ²	$\frac{(O-E)^2}{E}$
60	45	15	225	5
30	45	-15	225	5
40	55	-15	225	4.09
70	55	15	225	4.09
				18.18

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} = 18.18, \text{ d.f.} = (r - 1)(c - 1) = (2 - 1)(2 - 1) = 1 \text{ at } 5\% \text{ LOS} = 3.841.$$

Calculated $\chi^2 >$ tabulated χ^2 , H_0 is rejected. New and conventional treatment are not independent.

6. Given the following contingency table for hair colour and eye colour. Find the value of chi square. Is there good association between the two.

		Hair Colour			Total
		Fair	Brown	Black	
Eye Colour	Blue	15	5	20	40
	Grey	20	10	20	50
	Brown	25	15	20	60
	Total	60	30	60	150

Solution: H_0 : The two attributes Hair colour and eye colour are independent.

O	E	O - E	(O - E) ²	$\frac{(O-E)^2}{E}$
15	16	-1	1	0.0625
5	8	-3	9	1.125
20	16	4	16	1
20	20	0	0	0
10	10	0	0	0
20	20	0	0	0
25	24	1	1	0.042
15	12	3	9	0.75
20	24	-4	16	0.666
				3.6458

$E(15) = \frac{60 \times 40}{150} = 16$	$E(5) = \frac{30 \times 40}{150} = 8$	$E(20) = \frac{60 \times 40}{150} = 16$	40
$E(20) = \frac{60 \times 50}{150} = 20$	$E(10) = \frac{30 \times 50}{150} = 10$	$E(20) = \frac{60 \times 50}{150} = 20$	50
$E(25) = \frac{60 \times 60}{150} = 24$	$E(15) = \frac{30 \times 60}{150} = 12$	$E(20) = \frac{60 \times 60}{150} = 24$	60
60	30	60	150

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} = 3.6458, \text{ d.f.} = (r - 1)(c - 1) = (3 - 1)(3 - 1) = 4 \text{ at } 5\% \text{ LOS} = 9.488.$$

Calculated $\chi^2 <$ tabulated χ^2 , H_0 is accepted. The hair colour and eye colour are independent.

t - test table

Values of $|t|$ with probability P and degrees of freedom ν

$P \backslash \nu$	0.50	0.10	0.05	0.02	0.01
1	1.000	6.34	12.71	31.82	63.66
2	0.816	2.92	4.30	6.96	9.92
3	0.765	2.35	3.18	4.54	5.84
4	0.741	2.13	2.78	3.75	4.60
5	0.727	2.02	2.57	3.36	4.03
6	0.718	1.94	2.45	3.14	3.71
7	0.711	1.90	2.36	3.00	3.50
8	0.706	1.86	2.31	2.90	3.36
9	0.703	1.83	2.26	2.82	3.25
10	0.700	1.81	2.23	2.76	3.17
11	0.697	1.80	2.20	2.72	3.11
12	0.695	1.78	2.18	2.68	3.06
13	0.694	1.77	2.16	2.65	3.01
14	0.692	1.76	2.14	2.62	2.98
15	0.691	1.75	2.13	2.60	2.95
16	0.690	1.75	2.12	2.58	2.92
17	0.689	1.74	2.11	2.57	2.90
18	0.688	1.73	2.10	2.55	2.88
19	0.688	1.73	2.09	2.54	2.86
20	0.687	1.72	2.09	2.53	2.84
21	0.686	1.72	2.08	2.52	2.83
22	0.686	1.72	2.07	2.51	2.82
23	0.685	1.71	2.07	2.50	2.81
24	0.685	1.71	2.06	2.49	2.80
25	0.684	1.71	2.06	2.48	2.79
26	0.684	1.71	2.06	2.48	2.78
27	0.684	1.70	2.05	2.47	2.77
28	0.683	1.70	2.05	2.47	2.76
29	0.683	1.70	2.04	2.46	2.76
30	0.683	1.70	2.04	2.46	2.75

Regards!

Dr. P. Godhanda Raman, Assistant Professor(Senior Grade), Department of Mathematics, SRM IST, Kattankulathur
Email : godhanda@gmail.com, Mobile : 9941740168

F test Table

5% and 1% points of F.

$v_1 \backslash v_2$	1	2	3	4	5	6	8	12	24	∞
2	18.51	19.00	19.16	19.25	19.30	19.32	19.37	19.41	19.45	19.50
	98.49	99.00	99.17	99.25	99.30	99.33	99.36	99.42	99.46	99.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.84	8.74	8.64	8.53
	34.12	30.82	29.46	28.71	28.24	27.91	27.49	27.05	26.60	26.12
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.91	5.77	5.63
	21.20	18.00	16.69	15.98	15.52	15.21	14.80	14.37	13.93	13.46
5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.68	4.53	4.36
	16.26	13.27	12.06	11.39	10.97	10.67	10.27	9.89	9.47	9.02
6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.00	3.84	3.67
	13.74	10.92	9.78	9.15	8.75	8.47	8.10	7.72	7.31	6.88
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.57	3.41	3.23
	12.25	9.55	8.45	7.85	7.46	7.19	6.84	6.47	6.07	5.65
8	5.32	4.46	4.07	3.84	3.69	3.58	3.44	3.28	3.12	2.93
	11.26	8.65	7.59	7.01	6.63	6.37	6.03	5.67	5.28	4.86
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.07	2.90	2.71
	10.56	8.02	6.99	6.42	6.06	5.80	5.47	5.11	4.73	4.31
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.91	2.74	2.54
	10.04	7.56	6.55	5.99	5.64	5.39	5.06	4.71	4.33	3.91
12	4.75	3.88	3.49	3.26	3.11	3.00	2.85	2.69	2.50	2.30
	9.33	6.93	5.95	5.41	5.06	4.82	4.50	4.16	3.78	3.36
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.53	2.35	2.13
	8.86	6.51	5.56	5.03	4.69	4.46	4.14	3.80	3.43	3.00
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.42	2.24	2.01
	8.53	6.23	5.29	4.77	4.44	4.20	3.89	3.55	3.18	2.75
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.34	2.15	1.92
	8.28	6.01	5.09	4.58	4.25	4.01	3.71	3.37	3.01	2.57
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.28	2.08	1.84
	8.10	5.85	4.94	4.43	4.10	3.87	3.56	3.23	2.86	2.42
25	4.24	3.38	2.99	2.76	2.60	2.49	2.34	2.16	1.96	1.71
	7.77	5.57	4.68	4.18	3.86	3.63	3.32	2.99	2.62	2.17
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.09	1.89	1.62
	7.56	5.39	4.51	4.02	3.70	3.47	3.17	2.84	2.47	2.01
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.00	1.79	1.51
	7.31	5.18	4.31	3.83	3.51	3.29	2.99	2.66	2.29	1.81
60	4.00	3.15	2.76	2.52	2.37	2.25	2.10	1.92	1.70	1.39
	7.08	4.98	4.13	3.65	3.34	3.12	2.82	2.50	2.12	1.60

Regards!

Dr. P. Godhanda Raman, Assistant Professor(Senior Grade), Department of Mathematics, SRM IST, Kattankulathur
Email : godhanda@gmail.com, Mobile : 9941740168

Chi-Square Test Table

Values of χ^2 with probability P and df v

$P \backslash v$	0.99	0.95	0.50	0.30	0.20	0.10	0.05	0.01
1	0.0002	0.004	0.46	1.07	1.64	2.71	3.84	6.64
2	0.020	0.103	1.39	2.41	3.22	4.60	5.99	9.21
3	0.115	0.35	2.37	3.66	4.64	6.25	7.82	11.34
4	0.30	0.71	3.36	4.88	5.99	7.78	9.49	13.28
5	0.55	1.14	4.35	6.06	7.29	9.24	11.07	15.09
6	0.87	1.64	5.35	7.23	8.56	10.64	12.59	16.81
7	1.24	2.17	6.35	8.38	9.80	12.02	14.07	18.48
8	1.65	2.73	7.34	9.52	11.03	13.36	15.51	20.09
9	2.09	3.32	8.34	10.66	12.24	14.68	16.92	21.67
10	2.56	3.94	9.34	11.78	13.44	15.99	18.31	23.21
11	3.05	4.58	10.34	12.90	14.63	17.28	19.68	24.72
12	3.57	5.23	11.34	14.01	15.81	18.55	21.03	26.22
13	4.11	5.89	12.34	15.12	16.98	19.81	22.36	27.69
14	4.66	6.57	13.34	16.22	18.15	21.06	23.68	29.14
15	5.23	7.26	14.34	17.32	19.31	22.31	25.00	30.58
16	5.81	7.96	15.34	18.42	20.46	23.54	26.30	32.00
17	6.41	8.67	16.34	19.51	21.62	24.77	27.59	33.41
18	7.02	9.39	17.34	20.60	22.76	25.99	28.87	34.80
19	7.63	10.12	18.34	21.69	23.90	27.20	30.14	36.19
20	8.26	10.85	19.34	22.78	25.04	28.41	31.41	37.57
21	8.90	11.59	20.34	23.86	26.17	29.62	32.67	38.93
22	9.54	12.34	21.34	24.94	27.30	30.81	33.92	40.29
23	10.20	13.09	22.34	26.02	28.43	32.01	35.17	41.64
24	10.86	13.85	23.34	27.10	29.55	33.20	36.42	42.98
25	11.52	14.61	24.34	28.17	30.68	34.68	37.65	44.31
26	12.20	15.38	25.34	29.25	31.80	35.56	38.88	45.64
27	12.88	16.15	26.34	30.32	32.91	36.74	40.11	46.96
28	13.56	16.93	27.34	31.39	34.03	37.92	41.34	48.28
29	14.26	17.71	28.34	32.46	35.14	39.09	42.56	49.59
30	14.95	18.49	29.34	33.53	36.25	40.26	43.77	50.89

Regards!

Dr. P. Godhanda Raman, Assistant Professor(Senior Grade), Department of Mathematics, SRM IST, Kattankulathur
Email : godhanda@gmail.com, Mobile : 9941740168

UNIT – 4 : QUEUEING THEORY

Syllabus

- Introduction to Markovian queueing models
- Single Server Model with Infinite system capacity - $(M/M/1) : (\infty/FIFO)$
- Single Server Model with Finite System Capacity - $(M/M/1) : (K/FIFO)$

INTRODUCTION

History : A.K.Erlang (1909) – “The Theory of probabilities and telephone conversations”.

All of us have experienced the annoyance of having to wait in line.

Example: 1. We wait in line in our cars in traffic jams. 2. We wait in line at supermarket to check out.

Why then is there waiting? There is more demand for service than there is facility for service available.

Why is this so? 1. There may be a shortage of available servers. 2. There may be a space limit to the amount of service that can be provided.

Question : 1. How long must a customer wait? 2. How many people will form in the line?

Answer: Queuing theory attempts to answer these questions through detailed mathematical analysis.

Customer: The term ‘Customer’ is used in a general sense and does not imply necessarily a human customer.

E.g.: 1. An Air plane waiting in line to take off. 2. A Computer program waiting to be run on a time shared basis.

Characteristics of Queuing Process

1. Arrival pattern of Customers
2. Service pattern of Servers
3. Queue discipline
4. System capacity
5. Number of service channels
6. Number of service stages

1. Arrival Pattern of Customer : (i) Bulk or Batches (ii) balked (iii) Reneged (iv) Jockey

(i) **Bulk or Batches** : More than one arrival can be entering the system simultaneously, the input is said to occur in bulk or batches.

(ii) **Balked** : If customer decides not to enter the queue upon arrival, he is said to have balked.

(iii) **Reneged** : A Customer may enter the queue, but after a time loses patience and decide to leave. In this case he is said to renege.

(iv) **Jockey** : Two or more parallel waiting lines, customers may switch from one to another.

2. Service Pattern of Services : If the system is empty, the service facility is idle. Service may also be deterministic (or) probabilistic. Service may also be single (or) batch one generally thinks of one customer being served at a time by a given server, but there are many situations where customer may be served simultaneously by the same server. **E.g.:** 1. Computer with parallel processing. 2. People boarding a train.

The service rate may depend on the number of customer waiting for service. A server may work faster if sees that the queue is building up (or) conversely, he may get flustered and become less efficient. The situation in which service depends on the no. of customers waiting is referred to as state dependent service.

3. Queue Discipline :

(i) First Come First Served (FCFS) or First In First Out (FIFO)

(ii) Last Come First Served (LCFS) or Last In First Out (LCFO)

(iii) Random Selection for Services (RSS) (iv) Priority (a) Preemptive (b) Non- Preemptive

(a) **Preemptive** : The customer with the highest priority is allowed to enter service immediately even if a customer with lower priority is already in service when the higher priority customer enters system.

(b) **Non - preemptive** : The highest priority customer goes to the head of the queue but cannot get into service until the customer presently in service is completely, even though this customer has a lower priority.

4. System Capacity : (i) Finite (ii) Infinite

(i) **Finite**: A queue with limited waiting room, so that when the time reaches a certain length, no further customer are allowed to enter until space becomes available by a service completion.

(ii) **Infinite**: A queue with unlimited waiting room.

5. Number of Service Channels : (i) Single channel system (ii) Multiple channel systems.

(i) **Single Channel System**



(ii) Multiple Channel System



Eg: 1. Barber Shop 2. Supermarket 3. Ticket Counters.

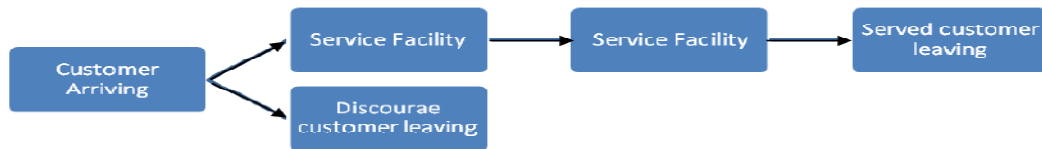
6. Number of Service Stages : (i) Single stage (ii) Multiple stage

(i) Single Stage



Eg: 1. Barber Shop 2. Supermarket 3. Theater.

(ii) Multiple Stage



Eg: 1. Medical History 2. Bank A/c opening 3.Canteen

Kendall Notation (A/B/X/Y/Z)

- A - Inter arrival time : M - Exponential, D - Deterministic, Ek - Erlang type k, Hk – Hyper exponential type k, Ph - phase type, G – General.
- B - Service time : M - Exponential, D - Deterministic, Ek - Erlang type k, Hk - Hyper exponential type k, Ph – phase type, G – General.
- X - No. of parallel servers : 1, 2 ...∞
- Y - System capacity : 1, 2 ...∞
- Z - Queue discipline : FCFS, LCFS, RSS, PR, GD

Queuing Models : 1. Probabilistic or stochastic models 2. Deterministic models 3. Mixed models

Probabilistic model: When there is uncertainty in both arrival rate and service rate (i.e. not treated a customer or not know) and are assumed to be random variables.

Deterministic model: Both arrival rate and service rate are constants (exactly known).

Mixed model: When either the arrival rate or the service rate is exactly known and the other is not known.

We look for probabilistic model only

1. (M/M/1) : (∞ / FIFO)

2. (M/M/1) : (K / FIFO)

SINGLE SERVER MODEL WITH INFINITE SYSTEM CAPACITY - (M/M/1) : (∞ / FIFO)

1. $P_0 = 1 - \frac{\lambda}{\mu}$ and $P_n = \left(\frac{\lambda}{\mu}\right)^n P_0$, λ = arrival rate, μ = service rate
2. Probability that the system is busy = $1 - P_0 = \frac{\lambda}{\mu}$
3. Expected number of customers in the system : $L_s = \frac{\lambda}{\mu - \lambda}$
4. Expected number of customers in the queue: $L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}$
5. Expected number of customers in non empty queues: $L_n = \frac{\mu}{\mu - \lambda}$

6. Expected waiting time a customer in the system: $W_s = \frac{1}{\mu - \lambda}$
7. Average waiting time that a customer in the queue: $W_q = \frac{\lambda}{\mu(\mu - \lambda)}$
8. Probability that the number of customers in the system exceeds k : $P(n > k) = \left(\frac{\lambda}{\mu}\right)^{k+1}$
9. Probability that the number of customers in the system greater than or equal to k : $P(n \geq k) = \left(\frac{\lambda}{\mu}\right)^k$
10. Probability that the waiting time of a customer in the system exceeds t : $P(w > t) = e^{-(\mu - \lambda)t}$
11. Probability that the waiting time of a customer in the queue exceeds t : $P(w > t) = \frac{\lambda}{\mu} e^{-(\mu - \lambda)t}$
12. Probability density function of the waiting time in the system: $f(w) = (\mu - \lambda)e^{-(\mu - \lambda)w}$
Which is the probability density function of an exponential distribution with parameter $\mu - \lambda$.
13. Probability density function of the waiting time in the queue: $g(w) = \begin{cases} \frac{\lambda}{\mu}(\mu - \lambda)e^{-(\mu - \lambda)w}, & w > 0 \\ 1 - \frac{\lambda}{\mu}, & w = 0 \end{cases}$

Little's formula: $L_s = \lambda W_s$, $L_q = \lambda W_q$, $W_s = W_q + \frac{1}{\mu}$, $L_s = L_q + \frac{\lambda}{\mu}$

SINGLE SERVER MODEL WITH INFINITE SYSTEM CAPACITY - (M/M/1) : (∞/FIFO)

1. Customers arrive at a one man barber shop according to a Poisson process with a mean inter arrival time of 20 minutes. Customers spend an average of 15 minutes in the barber chair. If an hour is used as the unit of time,

- (i) What is the probability that a customer need not wait for a hair cut?
- (ii) What is the expected number of customers in the barber shop and in the queue?
- (iii) How much time can a customer expect to spend in the barbershop?
- (iv) Find the average time that a customer spends in the queue?
- (v) What is the probability that there will be more than 6 customers?
- (vi) What is the probability that there will be 6 or more customers waiting for service?
- (vii) What is the probability that the waiting time in the (a) system (b) queue, is greater than 12 minutes?

Solution: $\frac{1}{\lambda} = \frac{20}{60} = \frac{1}{3} \Rightarrow \lambda = 3$ customers/hour, $\frac{1}{\mu} = \frac{15}{60} = \frac{1}{4} \Rightarrow \mu = 4$ customers/hour.

- (i) $P(\text{a customer need not wait}) = P(\text{no customer in the system})$: $P_0 = 1 - \frac{\lambda}{\mu} = 1 - \frac{3}{4} = 0.25$
- (ii) Expected number of customers in the barber shop : $L_s = \frac{\lambda}{\mu - \lambda} = \frac{3}{4 - 3} = 3$ customers
Expected number of customers in the queue: $L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{3^2}{4(4 - 3)} = \frac{9}{4} = 2.25$ customers
- (iii) Expected time a customer spend in the barbershop: $W_s = \frac{1}{\mu - \lambda} = \frac{1}{4 - 3} = 1$ hour
- (iv) Average time that a customer spends in the queue: $W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{3}{4(4 - 3)} = \frac{3}{4} = 0.75$ hour
- (v) The probability that there will be more than 6 customers:
 $P(n > k) = \left(\frac{\lambda}{\mu}\right)^{k+1} \Rightarrow P(n > 6) = \left(\frac{3}{4}\right)^{6+1} = 0.1335$
- (vi) The probability that there will be 6 or more customers waiting for service:
 $P(n \geq k) = \left(\frac{\lambda}{\mu}\right)^k \Rightarrow P(n \geq 6) = \left(\frac{3}{4}\right)^6 = 0.1779$
- (vii) The probability that the waiting time in the system is greater than 12 minutes?
 $P(w > t) = e^{-(\mu - \lambda)t} \Rightarrow P(w > 12) = e^{-(4 - 3) \times \frac{12}{60}} = e^{-0.2} = 0.8187$
- (viii) The probability that the waiting time in the queue, is greater than 12 minutes?
 $P(w > t) = \frac{\lambda}{\mu} e^{-(\mu - \lambda)t} \Rightarrow P(w > 12) = \left(\frac{3}{4}\right) e^{-(4 - 3) \times \frac{12}{60}} = \left(\frac{3}{4}\right) e^{-0.2} = 0.61405$

2. If People arrive to purchase cinema tickets at the average rate of 6 per minute at a one man counter, and it takes an average of 7.5 seconds to purchase a ticket. If a person arrives 2 minutes before the picture starts and it takes exactly 1.5 minutes to reach the correct seat after purchasing the ticket,

- (i) can he expect to be seated for the start of the picture?
- (ii) What is the probability that he will be seated for the start of the picture?
- (iii) How early must he arrive in order to be 99% sure of being seated for the start of the picture?

Solution: $\lambda = 6$ /minute, $\mu = 8$ /minute

$$(i) W_s = \frac{1}{\mu - \lambda} = \frac{1}{8 - 6} = \frac{1}{2} = 0.5 \text{ minute}$$

E(total time required to purchase the ticket and to reach the seat) = $0.5 + 1.5 = 2 \text{ min}$

$$(ii) P(\text{total time} < 2 \text{ minute}) = P(w < t) = 1 - P(w > t) = 1 - e^{-(\mu - \lambda)t}$$

$$P\left(w < \frac{1}{2}\right) = 1 - e^{-(8-6) \times \frac{1}{2}} = 1 - e^{-1} = 0.63$$

$$(iii) P(w < t) = 99\% = 0.99 \Rightarrow 1 - P(w > t) = 0.99 \Rightarrow P(w > t) = 0.01 \Rightarrow e^{-(\mu - \lambda)t} = 0.01 \Rightarrow e^{-(8-6)t} = 0.01 \Rightarrow e^{-2t} = 0.01 \Rightarrow -2t = \ln(0.01) \Rightarrow -2t = -4.6 \Rightarrow t = 2.3 \text{ minute}$$

$$P(\text{ticket purchasing time} < 2.3) = 0.99$$

$$P[\text{total time to get the ticket and to go to the seat} < (2.3 + 1.5)] = 0.99$$

\therefore The person must arrive at least 2.64 minutes early so as to be 99% sure of seeing the start of the picture.

- 3. The arrivals at the counter in a bank occur in accordance with a Poisson process at an average rate of 8 per hour. The duration of service of a customer has an exponential distribution with a mean of 6 minutes. Find the probability that an arriving customer (i) Has to wait (ii) Finds 4 customers in the system (iii) Has to spend less than 15 minutes in the bank.**

Solution: $\lambda = 8 / \text{hour}$, $\mu = \frac{1}{6} / \text{minute} = 10 / \text{hour}$

$$(i) \text{ Probability that a customer has to wait} = \text{Probability that the system is busy} = \frac{\lambda}{\mu} = \frac{8}{10} = 0.8$$

$$(ii) \text{ Probability that there are 4 customers in the system} = P_4 = \left(\frac{8}{10}\right)^4 \left(1 - \frac{8}{10}\right) = 0.08192$$

$$(iii) \text{ Probability that a customer has to spend less than 15 minutes in the bank: } P(w > t) = e^{-(\mu - \lambda)t}$$

$$P(W_s < 15 \text{ minutes}) = P\left(W_s < \frac{1}{4} \text{ hour}\right) = 1 - P\left(W_s > \frac{1}{4} \text{ hour}\right) = 1 - e^{-(10-8)\left(\frac{1}{4}\right)} = 0.3935$$

- 4. Customers arrive at a watch repair shop according to a Poisson process at a rate of one per every 10 minutes and the service time is an exponential random variable with mean 8 minutes.**

(i) Find the average number of customers in the shop.

(ii) Find the average time a customer spends in the shop.

(iii) Find the average number of customers in the queue.

(iv) What is the probability that the server is idle?

Solution: $\lambda = 6 / \text{hour}$, $\mu = \frac{60}{8} = \frac{15}{2} / \text{hour}$.

$$(i) \text{ The average number of customers in the shop: } L_s = \frac{\lambda}{\mu - \lambda} = \frac{6}{\frac{15}{2} - 6} = 4 \text{ customers}$$

$$(ii) \text{ The average time a customer spends in the shop: } W_s = \frac{1}{\mu - \lambda} = \frac{1}{\frac{15}{2} - 6} = \frac{2}{3} \text{ hour}$$

$$(iii) \text{ The average number of customers in the queue: } L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{6^2}{\frac{15}{2}(\frac{15}{2} - 6)} = \frac{16}{5}$$

$$(iv) P(\text{system is empty}): P_0 = 1 - \frac{\lambda}{\mu} = 1 - \frac{6}{\frac{15}{2}} = \frac{1}{5}$$

- 5. A departmental store has a single cashier. During the rush hours, customers arrive at the rate of 20 customers per hour. The average no. of customers that can be processed by the cashier is 24 per hour. Calculate the following**

(i) What is the probability that the cashier is idle?

(ii) What is the average number of customers in the queueing system.

(iii) What is the average time a customer spends in the system.

(iv) What is the average number of customers in the queue?

(v) What is the average time a customer spends in the queue, waiting for service?

Solution: $\lambda = 20 / \text{hour}$, $\mu = 24 / \text{hour}$.

$$(i) P(\text{cashier is idle}) = P_0 = 1 - \frac{\lambda}{\mu} = 1 - \frac{20}{24} = 0.1674$$

$$(ii) \text{ Average number of customers in the queueing system: } L_s = \frac{\lambda}{\mu - \lambda} = \frac{20}{24 - 20} = 5 \text{ customers}$$

$$(iii) \text{ Average time a customer spend in the system: } W_s = \frac{1}{\mu - \lambda} = \frac{1}{24 - 20} = \frac{1}{4} \text{ hour}$$

$$(iv) \text{ Average number of customers in the queueing system: } L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{20^2}{24(24 - 20)} = 4.167 \text{ customers}$$

$$(v) \text{ Average time that a customer spends in the queue: } W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{20}{24(24 - 20)} = \frac{20}{96} \text{ hour}$$

SINGLE SERVER MODEL WITH FINITE SYSTEM CAPACITY - (M/M/1) : (k/FIFO)

$$P_0 = \begin{cases} \frac{(1-\frac{\lambda}{\mu})}{1-(\frac{\lambda}{\mu})^{k+1}}, & \lambda \neq \mu \\ \frac{1}{k+1}, & \lambda = \mu \end{cases} \quad \text{and} \quad P_n = \begin{cases} \left(\frac{\lambda}{\mu}\right)^n P_0, & \lambda \neq \mu \\ \frac{1}{k+1}, & \lambda = \mu \end{cases}$$

$$\text{Average number of customers in the system : } L_s = \begin{cases} \left(\frac{\lambda}{\mu-\lambda}\right) - \frac{(k+1)(\frac{\lambda}{\mu})^{k+1}}{1-(\frac{\lambda}{\mu})^{k+1}}, & \lambda \neq \mu \\ \frac{k}{2}, & \lambda = \mu \end{cases}$$

$$\text{Average number of customers in the queue : } L_q = L_s - \frac{\lambda'}{\mu}$$

$$\text{Effective arrival rate : } \lambda' = \mu(1 - P_0)$$

$$\text{Average waiting time of a customers in the system : } W_s = \frac{L_s}{\lambda'}$$

$$\text{Average waiting time of a customers in the queue : } W_q = \frac{L_q}{\lambda'}$$

SINGLE SERVER MODEL WITH FINITE SYSTEM CAPACITY - (M/M/1) : (k/FIFO)

1. Patients arrive at a clinic according to Poisson distribution at a rate of 60 patients per hour. The waiting room does not accommodate more than 14 patients. Investigation time per patient is exponential with mean rate of 40 per hour.

(i) Determine the effective arrival rate at the clinic.

(ii) What is the probability that an arriving patient will not wait?

(iii) What is the expected waiting time until a patient is discharged from the clinic?

Solution: $\lambda = 60$ patients/hr, $\mu = 40$ patients/hr, $k = 14 + 1 = 15$ (14 waiting patients + 1 patient under investigation)

(i) Effective arrival rate $\lambda' = \mu(1 - P_0)$, Where $P_0 = \frac{(1-\frac{\lambda}{\mu})}{1-(\frac{\lambda}{\mu})^{k+1}}, \lambda \neq \mu$

$$P_0 = \frac{(1-\frac{60}{40})}{1-(\frac{60}{40})^{15+1}} = 0.0007624, \quad \lambda' = 40(1 - 0.0007624) = 39.9695 \text{ per hour.}$$

(ii) $P(\text{a patient will not wait}) = P_0 = 0.0007624$

(iii) Expected waiting time until a patient is discharged from the clinic: $W_s = \frac{L_s}{\lambda'}$

$$L_s = \left(\frac{\lambda}{\mu-\lambda}\right) - \frac{(k+1)(\frac{\lambda}{\mu})^{k+1}}{1-(\frac{\lambda}{\mu})^{k+1}} = \left(\frac{60}{40-60}\right) - \frac{(15+1)(\frac{60}{40})^{15+1}}{1-(\frac{60}{40})^{15+1}} = 13 \text{ patients,} \quad W_s = \frac{14}{39.9695} = 0.3203 \text{ hour}$$

2. At a railway station, only one train is handled at a time. The railway yard is sufficient only for two trains to wait while other is given signal to leave the station. Trains arrive at the station at an average rate of 6 per hour and the railway station can handle them on an average of 12 per hour. Assuming Poisson arrivals and exponential service distribution, find the steady state probabilities for the number of trains in the system. Also find the average waiting time of a new train coming into the yard. If the handling rate is reduced to half, what is the effect of the above results?

Solution: Case (i): $\lambda = 6$ trains/hr, $\mu = 12$ trains/hr, $k = 2 + 1 = 3$

Steady state probabilities for the number of trains in the system = P_1, P_2 and P_3

$$P_n = \left(\frac{\lambda}{\mu}\right)^n P_0, \text{ Where } P_0 = \frac{(1-\frac{\lambda}{\mu})}{1-(\frac{\lambda}{\mu})^{k+1}} = \frac{(1-\frac{6}{12})}{1-(\frac{6}{12})^4} = 0.5333, \lambda \neq \mu$$

$$P_1 = \left(\frac{6}{12}\right)^1 (0.5333) = 0.2667, \quad P_2 = \left(\frac{6}{12}\right)^2 (0.5333) = 0.1333, \quad P_3 = \left(\frac{6}{12}\right)^3 (0.5333) = 0.0667$$

Average waiting time of a new train coming in the yard : $W_s = \frac{L_s}{\lambda'}$

$$\lambda' = \mu(1 - P_0) = 12(1 - 0.5333) = 5.6004$$

$$L_s = \left(\frac{\lambda}{\mu-\lambda}\right) - \frac{(k+1)(\frac{\lambda}{\mu})^{k+1}}{1-(\frac{\lambda}{\mu})^{k+1}} = \left(\frac{6}{12-6}\right) - \frac{(3+1)(\frac{6}{12})^{3+1}}{1-(\frac{6}{12})^{3+1}} = 0.7333 \text{ train,} \quad W_s = \frac{0.7333}{5.6004} = 0.1309 \text{ hour}$$

Case (ii): If the handling rate is reduced to half, then $\lambda = 6$ trains/hr, $\mu = 6$ trains/hr, $k = 2 + 1 = 3$

Steady state probabilities for the number of trains in the system = P_1, P_2 and P_3

$$P_n = \frac{1}{k+1}, P_1 = \frac{1}{3+1} = \frac{1}{4}, P_2 = \frac{1}{4}, P_3 = \frac{1}{4}$$

Average waiting time of a new train coming in the yard : $W_s = \frac{L_s}{\lambda'}$

$$P_0 = \frac{1}{k+1} = \frac{1}{4}, \lambda' = \mu(1 - P_0) = 6 \left(1 - \frac{1}{4}\right) = 4.5, L_s = \frac{k}{2} = \frac{3}{2} = 1.5 \text{ train}, W_s = \frac{1.5}{4.5} = 0.3333 \text{ hours}$$

3. A petrol pump with only one pump can accommodate 5 cars. The arrival of cars is Poisson with a mean rate of 10 per hour. The service time is exponentially distributed with a mean 2 minutes. How many cars are in the petrol pump on an average? What is the probability of a newly arriving customer finding the system full and leaving without availing service?

Solution: $\lambda = 10/\text{hr}, \mu = 30/\text{hr}, k = 5, P_0 = \frac{\left(1 - \frac{\lambda}{\mu}\right)}{\left[1 - \left(\frac{\lambda}{\mu}\right)^{k+1}\right]} = \frac{\left(1 - \frac{10}{30}\right)}{\left[1 - \left(\frac{10}{30}\right)^{5+1}\right]} = 0.667, \lambda \neq \mu$

$$L_s = \left(\frac{\lambda}{\mu - \lambda}\right) - \frac{(k+1)\left(\frac{\lambda}{\mu}\right)^{k+1}}{1 - \left(\frac{\lambda}{\mu}\right)^{k+1}} = \left(\frac{10}{30-10}\right) - \frac{(5+1)\left(\frac{10}{30}\right)^{5+1}}{1 - \left(\frac{10}{30}\right)^{5+1}} = 0.492$$

$$P(\text{System full}) = P(5 \text{ cars in the system}) = \left(\frac{\lambda}{\mu}\right)^5 P_0 = \left(\frac{10}{30}\right)^5 (0.667) = 0.00274.$$

4. A one person barber shop has 6 chairs to accommodate people waiting for a hair cut. Assume that customers who arrive when all the 6 chairs are full leave without entering the barber shop. Customers arrive at the average rate of 3 per hour and spend an average of 15 minutes in the barber chair.

- What is the probability that a customer can get directly into the barber chair upon arrival?
- What is the expected number of customers waiting for a hair cut?
- How much time can a customer expect to spend in the barber shop?
- What fraction of potential customers are turned away?

Solution: $\lambda = 3/\text{hr}, \mu = 4/\text{hr}, k = 6 + 1 = 7$

$$P_0 = \frac{\left(1 - \frac{\lambda}{\mu}\right)}{\left[1 - \left(\frac{\lambda}{\mu}\right)^{k+1}\right]} = \frac{\left(1 - \frac{3}{4}\right)}{\left[1 - \left(\frac{3}{4}\right)^{7+1}\right]} = 0.2778, \lambda \neq \mu,$$

Effective arrival rate $\lambda' = \mu(1 - P_0) = 4(1 - 0.2778) = 2.89/\text{hour}$

- (i) Probability of empty system = $P_0 = 0.2778$

$$(ii) L_s = \left(\frac{\lambda}{\mu - \lambda}\right) - \frac{(k+1)\left(\frac{\lambda}{\mu}\right)^{k+1}}{1 - \left(\frac{\lambda}{\mu}\right)^{k+1}} = \left(\frac{3}{4-3}\right) - \frac{(7+1)\left(\frac{3}{4}\right)^{7+1}}{1 - \left(\frac{3}{4}\right)^{7+1}} = 2.11, L_q = L_s - \frac{\lambda'}{\mu} = 2.11 - \frac{2.89}{4} = 1.3875$$

$$(iv) W_s = \frac{L_s}{\lambda'} = \frac{2.11}{2.89} = 0.7301/\text{hour}$$

$$(v) P(\text{a customer is turned away}) = P(\text{system is full}) = P_7 = \left(\frac{\lambda}{\mu}\right)^7 P_0 = \left(\frac{3}{4}\right)^7 (0.2778) = 0.037$$

Hence 3.7% of potential customers are turned away.

All the Best

Regards!

Dr. P. Godhandaraman

Assistant Professor (Senior Grade)

Department of Mathematics

SRM IST, Kattankulathur – 603 203

Email : godhanda@gmail.com, Mobile : 9941740168

UNIT – 5 : MARKOV CHAINS

Syllabus

- Introduction to Stochastic process, Markov process, Markov chain one step & n-step Transition Probability.
- Transition Probability Matrix and Applications
- Chapman Kolmogorov theorem (Statement only) – Applications.
- Classification of states of a Markov chain – Applications

INTRODUCTION

Random Processes or Stochastic Processes : A random process is a collection of random variables $\{X(s, t)\}$ which are functions of a real variable t (time). Here $s \in S$ (sample space) and $t \in T$ (index set) and each $\{X(s, t)\}$ is a real valued function. The set of possible values of any individual member is called state space.

Classification : Random processes can be classified into 4 types depending on the continuous or discrete nature of the state space S and index set T .

1. Discrete random sequence : If both S and T are discrete
2. Discrete random process : If S is discrete and T is continuous
3. Continuous random sequence : If S is continuous and T is discrete
4. Continuous random process : If both S and T are continuous.

Markov Process : If, for $t_1 < t_2 < t_3 < \dots < t_n$, we have $P\{X(t) \leq x / X(t_1) = x_1, X(t_2) = x_2 \dots X(t_n) = x_n\} = P\{X(t) \leq x / X(t_n) = x_n\}$ then the process $\{X(t)\}$ is called a Markov process. That is, if the future behaviour of the process depends only on the present state and not on the past, then the random process is called a Markov process.

Markov Chain : If, for all n , $P\{X_n = a_n / X_{n-1} = a_{n-1}, X_{n-2} = a_{n-2} \dots X_0 = a_0\} = P\{X_n = a_n / X_{n-1} = a_{n-1}\}$ then the process $\{X_n; n = 0, 1, 2, \dots\}$ is called a Markov chain.

One Step Transition Probability : The conditional probability $P_{ij}(n-1, n) = P(X_n = a_j / X_0 = a_i)$ is called one step transition probability from state a_i to state a_j in the n^{th} step.

Homogeneous Markov Chain : If the one step transition probability does not depend on the step.

That is, $P_{ij}(n-1, n) = P_{ij}(m-1, m)$ the Markov chain is called a homogeneous markov chain or the chain is said to have stationary transition probabilities.

n - Step Transition Probability : $P_{ij}^{(n)} = P(X_n = a_j / X_0 = a_i)$

Chapman Kolmogorov Equations : If P is the tpm of a homogeneous Markov chain, the n^{th} step tpm $P^{(n)}$ is equal to P^n . That is, $[P_{ij}^{(n)}] = [P_{ij}]^n$.

Regular Matrix : A stochastic matrix P is said to be a regular matrix, if all the entries of P^m are positive. A homogeneous Markov chain is said to be regular if its tpm is regular.

Classification of States of a Markov Chain

Irreducible Chain and Non - Irreducible (or) Reducible: If for every i, j we can find some n such that $P_{ij}^{(n)} > 0$, then every state can be reached from every other state, and the Markov chain is said to be irreducible. Otherwise the chain is non - irreducible or reducible.

Return State : State i of a Markov chain is called a return state, if $P_{ii}^{(n)} > 0$ for some $n > 1$.

Periodic State and Aperiodic State : The period d_i of a return state i is the greatest common divisor of all m such that $P_{ii}^{(m)} > 0$. That is, $d_i = \text{GCD} \left(m: P_{ii}^{(m)} > 0 \right)$. State i is periodic with period d_i if $d_i > 1$ and aperiodic if $d_i = 1$.

Recurrent (Persistent) State and Transient : If $\sum_{n=1}^{\infty} f_{ii}^{(n)} = 1$, the return to state i is certain and the state i is said to be persistent or recurrent. Otherwise, it is said to be transient.

Null Persistent and Non - null Persistent State : $\mu_{ii} = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$ is called the mean recurrence time of the state i . If μ_{ii} is finite, the state i is non null persistent. If $\mu_{ii} = \infty$ the state i is null persistent.

Ergodic State: A non null persistent and aperiodic state are called ergodic.

Theorem used to classify states

1. If a Markov chain is irreducible, all its states are of the same type. They are all transient, all null persistent or all non null persistent. All its states are either aperiodic or periodic with the same period.
2. If a Markov chain is finite irreducible, all its states are non null persistent.

PROBLEMS

6. A man either drives a car or catches a train to go to office each day. He never goes 2 days in a row by train but if he drives one day, then the next day he is just as likely to drive again as he is to travel by train. Now suppose that on the first day of the week, the man tossed a fair die and drove to work if and only if a 6 appeared. (i) Find the probability that he takes a train on the 3rd day (ii) Find the probability that he drives to work in the long run.

Solution : Let T – Train and C - Car. If today he goes by train, next day he will not go by train.

$$P(\text{Train} \rightarrow \text{Train}) = 0, P(\text{Train} \rightarrow \text{Car}) = 1, P(\text{Car} \rightarrow \text{Train}) = \frac{1}{2}, P(\text{Car} \rightarrow \text{Car}) = \frac{1}{2}$$

$$P = \begin{matrix} & \begin{matrix} T & C \end{matrix} \\ \begin{matrix} T \\ C \end{matrix} & \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \end{matrix}. \text{ The first day of the week, the man tossed a fair die and drove to work if and only if a 6 appeared.}$$

Initial state probability distribution is obtained by throwing a die.

Probability of going by car = $\frac{1}{6}$ and Probability of going by train = $1 - \frac{1}{6} = \frac{5}{6}$

$$\text{The 1st day state distribution is } P^{(1)} = \begin{bmatrix} \frac{5}{6} & \frac{1}{6} \end{bmatrix}$$

$$\text{The 2nd day state distribution is } P^{(2)} = P^{(1)}P = \begin{bmatrix} \frac{5}{6} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{12} & \frac{11}{12} \end{bmatrix}$$

$$\text{The 3rd day state distribution is } P^{(3)} = P^{(2)}P = \begin{bmatrix} \frac{1}{12} & \frac{11}{12} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{11}{24} & \frac{13}{24} \end{bmatrix}$$

$$(i) \quad P(\text{he travels by train on 3rd day}) = \frac{11}{24}$$

(ii) The limiting form or long run probability distribution. $\pi p = \pi$

$$\begin{bmatrix} \pi_1 & \pi_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \pi_1 & \pi_2 \end{bmatrix}$$

$$\frac{\pi_2}{2} = \pi_1 \Rightarrow \pi_2 = 2\pi_1 \quad (1)$$

$$\pi_1 + \frac{\pi_2}{2} = \pi_2 \Rightarrow \pi_2 = 2\pi_1 \quad (2)$$

$$\pi_1 + \pi_2 = 1 \quad (3)$$

$$\text{Sub. (1) in (3), } \pi_1 + 2\pi_1 = 1 \Rightarrow \pi_1 = \frac{1}{3} \quad (4)$$

$$\text{Sub. (4) in (3), } \frac{1}{3} + \pi_2 = 1 \Rightarrow \pi_2 = \frac{2}{3}, \quad P(\text{driving in the long run}) = \frac{2}{3}.$$

7. A college student X has the following study habits. If he studies one night, he is 70% sure not to study the next night. If he does not study one night, he is only 60% sure not to study the next night also. Find (i) the transition probability matrix (ii) how often he studies in the long run.

Solution : Let S – Studying and N - Not Studying. If he studies one night, next night he is 70% not studying.

$$P(\text{Studying} \rightarrow \text{Not Studying}) = 0.7, \quad P(\text{Studying} \rightarrow \text{Studying}) = 0.3,$$

$$P(\text{Not Studying} \rightarrow \text{Not Studying}) = 0.6, \quad P(\text{Not Studying} \rightarrow \text{Studying}) = 0.4,$$

$$P = \begin{matrix} & \begin{matrix} S & N \end{matrix} \\ \begin{matrix} S \\ N \end{matrix} & \begin{bmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{bmatrix} \end{matrix}. \text{ The limiting form or long run probability distribution. } \pi p = \pi$$

$$\begin{bmatrix} \pi_1 & \pi_2 \end{bmatrix} \begin{bmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{bmatrix} = \begin{bmatrix} \pi_1 & \pi_2 \end{bmatrix}$$

$$0.3\pi_1 + 0.4\pi_2 = \pi_1 \Rightarrow 0.4\pi_2 = 0.7\pi_1 \Rightarrow 4\pi_2 = 7\pi_1 \quad (1)$$

$$0.7\pi_1 + 0.6\pi_2 = \pi_2 \Rightarrow 0.4\pi_2 = 0.7\pi_1 \Rightarrow 4\pi_2 = 7\pi_1 \quad (2)$$

$$\pi_1 + \pi_2 = 1 \quad (3)$$

$$\text{Sub. (1) in (3), } \frac{4}{7}\pi_2 + \pi_2 = 1 \Rightarrow \pi_2 = \frac{7}{11} \quad (4)$$

$$\text{Sub. (4) in (3), } \pi_1 + \frac{7}{11} = 1 \Rightarrow \pi_1 = \frac{4}{11}, \quad P(\text{he studies in the long run}) = \frac{4}{11}.$$

8. Suppose that the probability of a dry day following a rainy day is $\frac{1}{3}$ and that the probability of a rainy day following a dry day is $\frac{1}{2}$. Given that May 1 is a dry day, find the prob. that (i) May 3 is also a dry day (ii) May 5 is also a dry day.

Solution : Let D – Dry day and R - Rainy day.

$$P = \begin{matrix} & D & R \\ \begin{matrix} D \\ R \end{matrix} & \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} \end{matrix}$$
 Initial state probability distribution is the probability distribution on May 1. Since May 1 is a dry day.

$P(D) = 1$ and $P(R) = 0$.

$$P^{(1)} = [1 \ 0], \quad P^{(2)} = P^{(1)} P = [1 \ 0] \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad P^{(3)} = P^{(2)} P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{5}{12} & \frac{7}{12} \end{bmatrix},$$

$$P(\text{May 3 is a dry day}) = \frac{5}{12}$$

$$P^{(4)} = P^{(3)} P = \begin{bmatrix} \frac{5}{12} & \frac{7}{12} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{29}{72} & \frac{43}{72} \end{bmatrix}, \quad P^{(4)} P = \begin{bmatrix} \frac{29}{72} & \frac{43}{72} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{173}{432} & \frac{259}{432} \end{bmatrix},$$

$$P(\text{May 5 is a dry day}) = \frac{173}{432}$$

9. A salesman territory consists of 3 cities A, B and C. He never sells in the same city on successive days. If he sells in city A, then the 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 the other city. In the long run, how often does he sell in each of the cities?

$$\begin{matrix} & A & B & C \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{bmatrix} 0 & 1 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{bmatrix} \end{matrix}$$
 The limiting form or long run prob. distribution.

$\pi p = \pi$

$$[\pi_1 \ \pi_2 \ \pi_3] \begin{bmatrix} 0 & 1 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{bmatrix} = [\pi_1 \ \pi_2 \ \pi_3]$$

$$\frac{2}{3} \pi_2 + \frac{2}{3} \pi_3 = \pi_1 \Rightarrow 2 \pi_2 + 2 \pi_3 = 3 \pi_1 \Rightarrow 3 \pi_1 - 2 \pi_2 - 2 \pi_3 = 0 \quad (1)$$

$$\pi_1 + \frac{1}{3} \pi_3 = \pi_2 \Rightarrow 3 \pi_1 + \pi_3 = 3 \pi_2 \Rightarrow 3 \pi_1 - 3 \pi_2 + \pi_3 = 0 \quad (2)$$

$$\frac{1}{3} \pi_2 = \pi_3 \Rightarrow \pi_2 = 3 \pi_3 \quad (3)$$

$$\pi_1 + \pi_2 + \pi_3 = 1 \quad (4)$$

$$\text{Sub. (3) in (1), } 3 \pi_1 - 2 \pi_2 - 2 \pi_3 = 0 \Rightarrow 3 \pi_1 - 2 (3 \pi_3) - 2 \pi_3 = 0 \Rightarrow 3 \pi_1 = 8 \pi_3 \Rightarrow \pi_1 = \frac{8}{3} \pi_3 \quad (5)$$

$$\text{Sub. (3) and (5) in (4), } \pi_1 + \pi_2 + \pi_3 = 1 \Rightarrow \frac{8}{3} \pi_3 + 3 \pi_3 + \pi_3 = 1 \Rightarrow \pi_3 = \frac{3}{20} \quad (6)$$

$$\text{Sub. (6) in (3), } \pi_2 = \frac{9}{20}, \quad \text{Sub. (6) in (5), } \pi_1 = \frac{8}{20}, \quad \pi_1 = 0.40, \pi_2 = 0.45, \pi_3 = 0.15$$

Thus in the long run, he sells 40% of the time in city A, 45% of the time in the city B and 15% of the time in city C.

10. A fair die is tossed repeatedly. If X_n denotes the maximum of the numbers occurring in the first n tosses, find the transition probability matrix P of the Markov chain $\{X_n\}$. Find also P^2 and $P(X_2 = 6)$.

Solution: The state space is $\{1, 2, 3, 4, 5, 6\}$. X_n = maximum of the numbers occurring in the first n trials.

X_{n+1} = maximum of the numbers occurring in the first $(n + 1)$ trials = $\max[X_n, \text{number in the } (n + 1)^{\text{th}} \text{ trial}]$.

Let us see how the First Row of the tpm is filled.

$$X_n = 1$$

$X_{n+1} = 1$ if 1 appears in $(n + 1)^{\text{th}}$ trial

= 2 if 1 appears in $(n + 1)^{\text{th}}$ trial

= 3 if 1 appears in $(n + 1)^{\text{th}}$ trial

= 4 if 1 appears in $(n + 1)^{\text{th}}$ trial

= 5 if 1 appears in $(n + 1)^{\text{th}}$ trial

= 6 if 1 appears in $(n + 1)^{\text{th}}$ trial

Now, in the $(n + 1)^{\text{th}}$ trial, each of the numbers 1, 2, 3, 4, 5, 6 occurs with probability $\frac{1}{6}$.

Let us see how the Second Row of the tpm is filled.

Here $X_n = 2$

If $(n+1)^{th}$ trial results in 1 or 2, $X_{n+1} = 2$

If $(n+1)^{th}$ trial results in 3, $X_{n+1} = 3$

If $(n+1)^{th}$ trial results in 4, $X_{n+1} = 4$

If $(n+1)^{th}$ trial results in 5, $X_{n+1} = 5$

If $(n+1)^{th}$ trial results in 6, $X_{n+1} = 6$

If $X_n = 2$, $P(X_{n+1} = 2) = \frac{2}{6}$ and $P(X_{n+1} = k) = \frac{1}{6}$, $k = 3, 4, 5, 6$.

Proceeding similarly, the tpm is

$$P = n^{th} \begin{matrix} (n+1)^{th} \text{ state} \\ \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & \frac{2}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & \frac{3}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 0 & \frac{4}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & \frac{5}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{6} \end{bmatrix} \end{matrix}, P^2 = \begin{matrix} \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & \frac{2}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & \frac{3}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 0 & \frac{4}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & \frac{5}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{6} \end{bmatrix} \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & \frac{2}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & \frac{3}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 0 & \frac{4}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & \frac{5}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{6} \end{bmatrix} = \begin{bmatrix} \frac{1}{36} & \frac{3}{36} & \frac{5}{36} & \frac{7}{36} & \frac{9}{36} & \frac{11}{36} \\ 0 & \frac{4}{36} & \frac{5}{36} & \frac{7}{36} & \frac{9}{36} & \frac{11}{36} \\ 0 & 0 & \frac{9}{36} & \frac{7}{36} & \frac{9}{36} & \frac{11}{36} \\ 0 & 0 & 0 & \frac{16}{36} & \frac{9}{36} & \frac{11}{36} \\ 0 & 0 & 0 & 0 & \frac{25}{36} & \frac{11}{36} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{36} \end{bmatrix} \end{matrix}$$

The initial probability distribution is $P^{(0)} = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$

$$P(X_2 = 6) = \sum_{i=1}^6 P(X_2 = 6/X_0 = i)P(X_0 = i) = \frac{1}{6} \sum_{i=1}^6 P_{i6}^2 = \frac{1}{6} [P_{16}^2 + P_{26}^2 + P_{36}^2 + P_{46}^2 + P_{56}^2 + P_{66}^2]$$

$$P(X_2 = 6) = \frac{1}{6} \left[\frac{11}{36} + \frac{11}{36} + \frac{11}{36} + \frac{11}{36} + \frac{11}{36} + 1 \right] = \frac{91}{216}$$

11. The transition probability matrix of a Markov chain $\{X_n\}, n = 1, 2, \dots$ having 3 states 1, 2, 3 is

$$P = \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix} \text{ and the initial distribution } P^{(0)} = (0.7, 0.2, 0.1).$$

Find (i) $P(X_2 = 3, X_1 = 3, X_0 = 2)$ (ii) $P(X_2 = 3)$ (iii) $P(X_3 = 2, X_2 = 3, X_1 = 3, X_0 = 2)$.

Solution:

$$\begin{aligned} \text{(i)} \quad P(X_2 = 3, X_1 = 3, X_0 = 2) &= P(X_2 = 3/X_1 = 3, X_0 = 2)P(X_1 = 3, X_0 = 2) \\ &= P(X_2 = 3/X_1 = 3)P(X_1 = 3, X_0 = 2) = P_{33}^{(1)} P(X_1 = 3, X_0 = 2) \\ &= P_{33}^{(1)} P(X_1 = 3/X_0 = 2)P(X_0 = 2) = P_{33}^{(1)} P_{23}^{(1)} P(X_0 = 2) \\ &= (0.3)(0.2)(0.2) = 0.012 \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad P(X_2 = 3) &= P(X_2 = 3, X_0 = 1) + P(X_2 = 3, X_0 = 2) + P(X_2 = 3, X_0 = 3) \\ &= P(X_2 = 3/X_0 = 1)P(X_0 = 1) + P(X_2 = 3/X_0 = 2)P(X_0 = 2) + P(X_2 = 3/X_0 = 3)P(X_0 = 3) \\ &= P_{13}^{(2)} P(X_0 = 1) + P_{23}^{(2)} P(X_0 = 2) + P_{33}^{(2)} P(X_0 = 3) \\ &= (0.26)(0.7) + (0.34)(0.2) + (0.29)(0.1) = 0.279 \end{aligned}$$

$$P^2 = \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix} \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix} = \begin{bmatrix} 0.43 & 0.31 & 0.26 \\ 0.24 & 0.42 & 0.34 \\ 0.36 & 0.35 & 0.29 \end{bmatrix}$$

$$\begin{aligned} \text{(iii)} \quad P(X_3 = 2, X_2 = 3, X_1 = 3, X_0 = 2) &= P(X_3 = 2/X_2 = 3, X_1 = 3, X_0 = 2)P(X_2 = 3, X_1 = 3, X_0 = 2) \\ &= P(X_3 = 2/X_2 = 3)P(X_2 = 3, X_1 = 3, X_0 = 2) \\ &= P_{32}^{(1)} P(X_2 = 3/X_1 = 3, X_0 = 2)P(X_1 = 3, X_0 = 2) \\ &= P_{32}^{(1)} P(X_2 = 3/X_1 = 3)P(X_1 = 3, X_0 = 2) \\ &= P_{32}^{(1)} P_{33}^{(1)} P(X_1 = 3/X_0 = 2)P(X_0 = 2) = P_{32}^{(1)} P_{33}^{(1)} P_{23}^{(1)} P(X_0 = 2) \\ &= (0.4)(0.3)(0.2)(0.2) = 0.0048 \end{aligned}$$

12. The transition probability matrix of a Markov chain $\{X_n\}$, $n = 1, 2, \dots$ having 3 states 1, 2, 3 is $P = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix}$

& the initial distribution $P^{(0)} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Find (i) $P(X_3 = 2/X_2 = 1)$ (ii) $P(X_2 = 2)$

(iii) $P(X_2 = 2, X_1 = 1, X_0 = 2)$ (iv) $P(X_3 = 1, X_2 = 2, X_1 = 1, X_0 = 2)$

Solution:

$$(i) P(X_3 = 2/X_2 = 1) = P_{12}^{(1)} = \frac{1}{4}$$

$$(ii) P(X_2 = 2) = P(X_2 = 2, X_0 = 1) + P(X_2 = 2, X_0 = 2) + P(X_2 = 2, X_0 = 3) \\ = P(X_2 = 2/X_0 = 1)P(X_0 = 1) + P(X_2 = 2/X_0 = 2)P(X_0 = 2) + P(X_2 = 2/X_0 = 3)P(X_0 = 3) \\ = P_{12}^{(2)}P(X_0 = 1) + P_{22}^{(2)}P(X_0 = 2) + P_{32}^{(2)}P(X_0 = 3) = \left(\frac{5}{16}\right)\left(\frac{1}{3}\right) + \left(\frac{8}{16}\right)\left(\frac{1}{3}\right) + \left(\frac{9}{16}\right)\left(\frac{1}{3}\right) = \frac{1}{6}$$

$$P^2 = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{10}{16} & \frac{5}{16} & \frac{1}{16} \\ \frac{5}{16} & \frac{8}{16} & \frac{3}{16} \\ \frac{3}{16} & \frac{9}{16} & \frac{1}{16} \end{bmatrix}$$

$$(iii) P(X_2 = 2, X_1 = 1, X_0 = 2) = P(X_2 = 2/X_1 = 1, X_0 = 2)P(X_1 = 1, X_0 = 2) \\ = P(X_2 = 2/X_1 = 1)P(X_1 = 1, X_0 = 2) = P_{12}^{(1)}P(X_1 = 1, X_0 = 2) \\ = P_{12}^{(1)}P(X_1 = 1/X_0 = 2)P(X_0 = 2) = P_{12}^{(1)}P_{21}^{(1)}P(X_0 = 2) \\ = \left(\frac{1}{4}\right)\left(\frac{1}{4}\right)\left(\frac{1}{3}\right) = 0.0208$$

$$(iv) P(X_3 = 1, X_2 = 2, X_1 = 1, X_0 = 2) = P(X_3 = 1/X_2 = 2, X_1 = 1, X_0 = 2)P(X_2 = 2, X_1 = 1, X_0 = 2) \\ = P(X_3 = 1/X_2 = 2)P(X_2 = 2, X_1 = 1, X_0 = 2) \\ = P_{21}^{(1)}P(X_2 = 2/X_1 = 1, X_0 = 2)P(X_1 = 1, X_0 = 2) \\ = P_{21}^{(1)}P(X_2 = 2/X_1 = 1)P(X_1 = 1, X_0 = 2) \\ = P_{21}^{(1)}P_{12}^{(1)}P(X_1 = 1, X_0 = 2) = P_{21}^{(1)}P_{12}^{(1)}P(X_1 = 1/X_0 = 2)P(X_0 = 2) \\ = P_{21}^{(1)}P_{12}^{(1)}P_{21}^{(1)}P(X_0 = 2) = \left(\frac{1}{4}\right)\left(\frac{1}{4}\right)\left(\frac{1}{4}\right)\left(\frac{1}{3}\right) = 0.0052$$

13. Find the nature of the states of the Markov chain with the transition probability matrix $P = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}$.

$$\text{Solution : } P = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}, P^2 = P P = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix},$$

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

$$P_{11}^{(2)} > 0, P_{12}^{(1)} > 0, P_{13}^{(2)} > 0, P_{21}^{(1)} > 0, P_{22}^{(2)} > 0, P_{23}^{(3)} > 0, P_{31}^{(2)} > 0, P_{32}^{(1)} > 0, P_{33}^{(2)} > 0$$

The chain is irreducible. Also since there are only 3 states, the chain is finite. i.e., the chain is finite & irreducible.

All the states are non null persistent.

State 1: $P_{11}^{(2)} > 0, P_{11}^{(4)} > 0, P_{11}^{(6)} > 0, \dots$, Period of state 1 = $GCD(2, 4, 6, \dots) = 2$

State 2: $P_{22}^{(2)} > 0, P_{22}^{(4)} > 0, P_{22}^{(6)} > 0, \dots$, Period of state 2 = $GCD(2, 4, 6, \dots) = 2$

State 3: $P_{33}^{(2)} > 0, P_{33}^{(4)} > 0, P_{33}^{(6)} > 0, \dots$, Period of state 3 = $GCD(2, 4, 6, \dots) = 2$

All the states 1, 2, 3 have period 2. That is, they are periodic. All the states are non null persistent, but the states are periodic. **Hence all the states are not ergodic.**

14. Three boys A, B, C are throwing a ball to each other. A always throw the ball to B & B always throws to C but C is just as likely to throw the ball to B as to A. Show that the process is Markovian. Find the transition matrix & classify the states.

Solution:

$$P = \begin{matrix} & \begin{matrix} A & B & C \end{matrix} \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \end{matrix},$$

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

$$P^4 = P^3 P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}, \quad P^5 = P^4 P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{8} & \frac{1}{8} & \frac{3}{8} \\ \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \end{bmatrix}$$

$$P_{11}^{(3)} > 0, P_{12}^{(1)} > 0, P_{13}^{(2)} > 0, \quad P_{21}^{(2)} > 0, P_{22}^{(2)} > 0, P_{23}^{(1)} > 0, \quad P_{31}^{(1)} > 0, P_{32}^{(1)} > 0, P_{33}^{(2)} > 0$$

The chain is irreducible. Also since there are only 3 states, the chain is finite. That is, the chain is finite and irreducible. All the states are non null persistent.

$$1^{st} \text{ state A: } P_{11}^{(3)} > 0, P_{11}^{(5)} > 0, \dots \quad \text{Period of A} = GCD(3, 5, \dots) = 1$$

$$2^{nd} \text{ state B: } P_{22}^{(2)} > 0, P_{22}^{(3)} > 0, P_{22}^{(4)} > 0 \dots \quad \text{Period of B} = GCD(2, 3, 4, \dots) = 1$$

$$3^{rd} \text{ state C: } P_{33}^{(2)} > 0, P_{33}^{(3)} > 0, P_{33}^{(4)} > 0, \dots \quad \text{Period of C} = GCD(2, 3, 4, \dots) = 1$$

All the states A, B, C have period 1. That is, they are aperiodic.

All the states are aperiodic and non null persistent, they are ergodic.

All the Best

Dr. P. Godhanda Raman, M.Sc., M.A., M.B.A., Ph.D.

Assistant Professor (Senior Grade)

Department of Mathematics

SRM Institute of Science and Technology

Kattankulathur – 603 203, Chengalpattu District

Email : godhanda@gmail.com, Mobile : 9941740168

