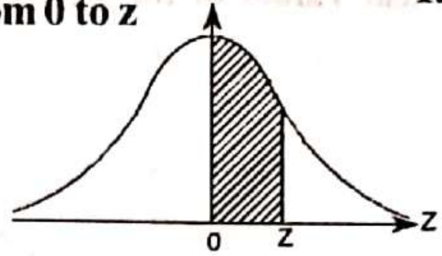
Table – 3

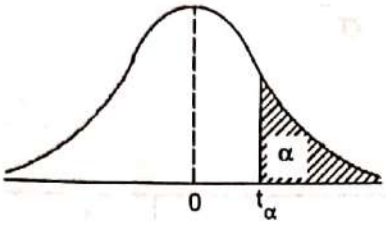
Areas under the Standard Normal Curve from 0 to z

$$Z = \frac{x - \mu}{\sigma}$$



z	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0938	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1256	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1916	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2649
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4654	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4979	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

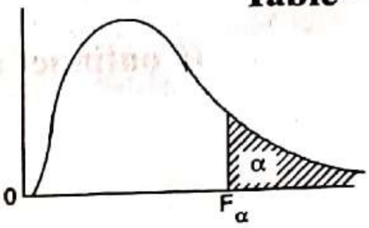
t_α - Critical Values of the t-Distribution with ν Degrees of Freedom Table – 4



ν	α = 0.40	0.30	0.20	0.15	0.10	0.05	0.025	0.01	0.005
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706	31.821	63.657
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306	2.896	3.355
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262	2.821	3.250
10	0.260	0.542	0.879	1.093	1.372	1.812	2.228	2.764	3.169
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201	2.718	3.106
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179	2.681	3.055
13	0.259	0.537	0.870	1.079	1.350	1.771	2.160	2.650	3.012
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145	2.624	2.977
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131	2.602	2.947
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120	2.583	2.921
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110	2.567	2.898
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101	2.552	2.878
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093	2.539	2.861
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086	2.528	2.845
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080	2.518	2.831
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074	2.508	2.819
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069	2.500	2.807
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064	2.492	2.797
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060	2.485	2.787
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056	2.479	2.779
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052	2.473	2.771
28	0.256	0.530	0.855	1.056	1.313	1.701	2.048	2.467	2.763
29	0.256	0.530	0.854	1.055	1.311	1.699	2.045	2.462	2.756
30	0.256	0.530	0.854	1.055	1.310	1.697	2.042	2.457	2.750
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021	2.423	2.704
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000	2.390	2.660
120	0.254	0.526	0.845	1.041	1.289	1.658	1.980	2.358	2.617
∞	0.253	0.524	0.842	1.036	1.282	1.645	1.960	2.326	2.576

Table – 5

Critical Values of the F-Distribution



Values of $F_{0.05}(\nu_1, \nu_2)$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

(Continued) Critical Values of the F-Distribution

	Values of $F_{0.05}(\nu_1, \nu_2)$									
	ν_1									
ν_2	10	12	15	20	24	30	40	60	120	∞
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.52
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.62
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.73	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.17	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.94	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.75	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

(Continued) Critical Values of the F-Distribution

Values of $F_{0.01}(\nu_1, \nu_2)$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	4052	4999.5	5403	5625	5764	5859	5928	5981	6022
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	3.79	2.66	2.56
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

(Continued) Critical Values of the F-Distribution

Values of $F_{0.01}(\nu_1, \nu_2)$

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
1	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Table – 6

χ_α^2 - Critical Values of the Chi-squared Distribution with ν Degrees of Freedom

ν	0.30	0.25	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.001
1	1.074	1.323	1.642	2.706	3.841	5.024	5.412	6.635	7.879	10.827
2	2.408	2.773	3.219	4.605	5.991	7.378	7.824	9.210	10.597	13.815
3	3.665	4.108	4.642	6.251	7.815	9.348	9.837	11.345	12.838	16.268
4	4.878	5.385	5.989	7.779	9.488	11.143	11.668	13.277	14.860	18.465
5	6.064	6.626	7.289	9.236	11.070	12.832	13.388	15.086	16.750	20.517
6	7.231	7.841	8.558	10.645	12.592	14.449	15.033	16.812	18.548	22.457
7	8.383	9.037	9.803	12.017	14.067	16.013	16.622	18.475	20.278	24.322
8	9.524	10.219	11.030	13.362	15.507	17.535	18.168	20.090	21.955	26.125
9	10.656	11.389	12.242	14.684	16.919	19.023	19.679	21.666	23.589	27.877
10	11.781	12.549	13.442	15.987	18.307	20.483	21.161	23.209	25.188	29.588
11	12.899	13.701	14.631	17.275	19.675	21.920	22.618	24.725	26.757	31.264
12	14.011	14.845	15.812	18.549	21.026	23.337	24.054	26.217	28.300	32.909
13	15.119	15.984	16.985	19.812	22.362	24.736	25.472	27.688	29.819	34.528
14	16.222	17.117	18.151	21.064	23.685	26.119	26.873	29.141	31.319	36.123
15	17.322	18.245	19.311	22.307	24.996	27.488	28.259	30.578	32.801	37.697
16	18.418	19.369	20.465	23.542	26.296	28.845	29.633	32.000	34.267	39.252
17	19.511	20.489	21.615	24.769	27.587	30.191	30.995	33.409	35.718	40.790
18	20.601	21.605	22.760	25.989	28.869	31.526	32.346	34.805	37.156	42.312
19	21.689	22.718	23.900	27.204	30.144	32.852	33.687	36.191	38.582	43.820
20	22.775	23.828	25.038	28.412	31.410	34.170	35.020	37.566	39.997	45.315
21	23.858	24.935	26.171	29.615	32.671	35.479	36.343	38.932	41.401	46.797
22	24.939	26.039	27.301	30.813	33.924	36.781	37.659	40.289	42.796	48.268
23	26.018	27.141	28.429	32.007	35.172	38.076	38.968	41.638	44.181	49.728
24	27.096	28.241	29.553	33.196	36.415	39.364	40.270	42.980	45.558	51.179
25	28.172	29.339	30.675	34.382	37.652	40.646	41.566	44.314	46.928	52.620
26	29.246	30.434	31.795	35.563	38.885	41.923	42.856	45.642	48.290	54.052
27	30.319	31.528	32.912	36.741	40.113	43.194	44.140	46.963	49.645	55.476
28	31.391	32.620	34.027	37.916	41.337	44.461	45.419	48.278	50.993	56.893
29	32.461	33.711	35.139	39.087	42.557	45.772	46.693	49.588	52.336	58.302
30	33.530	34.800	36.250	40.256	43.773	46.979	47.962	50.892	53.672	59.703

Unit-IV
Sampling theory

Model No 4.1: Introduction to population and sample

It is not easy to collect all the information about population and also it is not possible to study the characteristics of the entire population (finite or infinite) due to time factor, cost factor and other constraints. Thus we need sample. Sample is a finite subset of statistical individuals in a population and the number of individuals in a sample is called the sample size. Sampling is quite often used in our day-to-day practical life.

For example in a shop we assess the quality of rice, wheat or any other commodity by taking a handful of it from the bag and then to decide to purchase it or not.

Population

The population is a complete set of all possible observations of the type which is to be investigated. Total number of students studying in a school or college, total number of books in a library, total number of houses in a village or town is some examples of population.

Sometimes it is possible and practical to examine every person or item in the population we wish to describe. We call this a complete enumeration, or census. We use sampling when it is not possible to measure every item in the population. Statisticians use the word population to refer not only to people but to all items that have been chosen for study.

Finite population and infinite population:

A population is said to be finite if it consists of finite number of units. Number of workers in a factory, production of articles in a particular day for a company is examples of finite population. The total number of units in a population is called population size (N). A population is said to be infinite if it has infinite number of units. For example the number of stars in the sky, the number of people seeing the Television programmes etc.,

Sample

Statisticians use the word sample to describe a portion chosen from the population. A finite subset of statistical individuals defined in a population is called a sample. The number of units in a sample is called the sample size (n).

Types of sampling:

i) **Purposive sampling:** Purposive sampling is one in which sample units are selected with a definite purpose in view.
Ex: Suppose you want to collect feedback from students on the pedagogical methods in their school.

ii) **Random sampling:** Random sample is the one in which each unit of population has an equal chance of being included in it. And the sample obtained by this sampling is termed as random sample.
Ex: 25 students were selected in WIPRO from VVIT out of a hat from 3000 students who are studying in VVIT.

iii) **Simple sampling:** Simple sampling in which each unit of the population has an equal chance of being included in the sample and this probability is independent of the previous drawings.

Note:

1. Simple sampling may be regarded as random sampling but a random sampling is not necessarily a simple sampling.
2. For a finite population, random sampling with replacement is a simple sampling while random sampling without replacement is not a simple sampling.
3. For an infinite population, any random sampling is simple.

Example for Simple sampling: 25 students were selected in WIPRO from VVIT out of a hat from 3000 students who are studying in VVIT. In this case, the population is all 3000 students, and the sample is random because each STUDENT has an equal chance of being chosen.

iv) Stratified random sampling:

A method of sampling that involves dividing a population into smaller groups–called **strata**. The groups or strata are organized based on the shared characteristics or attributes of the members in the group. The process of classifying the population into groups is called **stratification**.

Examples for Stratified random sampling One might divide a sample of adults into subgroups by age, like 18–29, 30–39, 40–49, 50–59, and 60 and above.

Large Sample & Small Sample:

If the sample size
$n \geq 30$ i.e, referred as Large Sample
$n < 30$ i.e, referred as Small Sample

Sampling is done in 2 ways:

i) With replacement (infinite)
ii) Without replacement (finite)

Parameters and statistics:

We can describe samples and populations by using measures such as the mean, median, mode and standard deviation. When these terms describe the characteristics of a population, they are called parameters. When they describe the characteristics of a sample, they are called statistics. A parameter is a characteristic of a population and a statistic is a characteristic of a sample. Since samples are subsets of population statistics provide estimates of the parameters. That is, when the parameters are unknown, they are estimated from the values of the statistics.

	Parameters (Population)	Statistics (Sample)
Mean	μ	$\bar{x}$
Proportion	**P (Capital)**	**p (small)**
Variance	σ^2	s^2
Standard deviation	σ	**s**

Model No 4.2: Standard Error

The standard deviation of the sampling distribution of a statistic is known as its standard error. It is abbreviated as S.E.

For example, the standard deviation of the sampling distribution of the mean x known as the standard error of the mean.

S. No.	Standard error (S. E.)	With replacement Infinite Population	Without replacement Finite Population
1	**Standard error of sample mean (** $\bar{x}$ **)**	$\frac{\sigma}{\sqrt{n}}$	$\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$
2	**Standard error of sample Proportion (p)**	$\sqrt{\frac{PQ}{n}}$ Where $Q=1-P$	$\sqrt{\frac{PQ}{n}}\sqrt{\frac{N-n}{N-1}}$ Where $Q=1-P$
3	**Standard error of sample Standard deviation (s)**	$\frac{\sigma}{\sqrt{2n}}$	
4	**Standard error of the difference of two sample means** $\bar{x}_1$ **and** $\bar{x}_2$	$\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$	
5	**Standard error of the difference of two sample proportions** p_1 **and** p_2	$\sqrt{\frac{P_1Q_1}{n_1}+\frac{P_2Q_2}{n_2}}$	
6	**Standard error of the difference of two standard deviations** s_1 **and** s_2	$\sqrt{\frac{\sigma_1^2}{2n_1}+\frac{\sigma_2^2}{2n_2}}$	

Model No 4.3: Sampling distribution of means and variances

Parameters of the Population: Population size-N Finite population-Without replacement

Infinite population-With replacement

1. Mean of the population $\mu = \frac{\sum x}{N}$
2. Variance of the population $\sigma^2 = \frac{\sum (x-\mu)^2}{N}$
3. **3.** Standard deviation of the population $\sigma = \sqrt{\frac{\sum (x-\mu)^2}{N}}$

Statistics of the Sample: Sample size-n

Correction factor= $\frac{N-n}{N-1}$

1. The total number of samples with replacement (Infinite Population) is N^n
2. The total number of samples without replacement (Finite Population) is N_{C_n}.

Problem 1: What is the value of correction factor if n=5 and N=200.

Solution:

Problem 2: The size of the population is 2000 and the size of the sample is 200. Find the correction factor in the population.

Solution:

Problem 3: How many different samples of size two can be chosen from a finite population of size 25.

Solution:

Problem 4: In a random sample of 1000 packages shipped by air freight 13 had some damage. Find the standard error proportions.

Solution:

******* Problem 5:** A population consists of five numbers 2, 3, 6, 8 and 11. Consider all possible samples of size two which can be drawn from this population.

i) With replacement ii) Without replacement Find

(a) The mean of the population .

(b) The standard deviation of the population.

(c) The mean of the sampling distribution of means.

(d) The standard deviation of the sampling distribution of means.

(i.e. the standard error of means)

(e) Mean of the sampling distribution of variances

(f) Varience of the sampling distribution of variances.

Solution: Population: 2, 3, 6, 8, 11
Population size N=5

(a) Mean population $\mu = \frac{2+3+6+8+11}{5} = 6$

(b)Variance of the population $\sigma = \frac{\sum (x-\mu)^2}{N}$

$$\sigma^2 = \frac{\sum (2-6)^2 + (3-6)^2 + (6-6)^2 + (8-6)^2 + (11-6)^2}{5} = 10.8$$

Standard deviation of the population $\sigma = \sqrt{10.8} = 3.29.$

Sample: With replacement or Infinite Population

The total number of samples with replacement is $N^n = 5^2 = 25$

The samples with their corresponding means and variances sample size 2 is shown in the following table:

Samples	Sample Means	Sample Variances x_{s^2}
(2, 2)	2	$\frac{(2-2)^2+(2-2)^2}{2}=0$
(2, 3)	2.5	$\frac{(2-2.5)^2+(3-2.5)^2}{2}=0.25$
(2, 6)	4	$\frac{(2-4)^2+(6-4)^2}{2}=4$
(2, 8)	5	$\frac{(2-5)^2+(8-5)^2}{2}=9$
(2, 11)	6.5	$\frac{(2-6.5)^2+(11-6.5)^2}{2}=20.25$
(3, 2)	2.5	$\frac{(3-2.5)^2+(2-2.5)^2}{2}=0.25$
(3, 3)	3	$\frac{(3-3)^2+(3-3)^2}{2}=0$
(3, 6)	4.5	$\frac{(3-4.5)^2+(6-4.5)^2}{2}=2.25$

(3, 8)	5.5	$\frac{(3-5.5)^2+(8-5.5)^2}{2}=6.25$
(3, 11)	7	$\frac{(3-7)^2+(11-7)^2}{2}=16$
(6, 2)	4	$\frac{(6-4)^2+(2-4)^2}{2}=4$
(6, 3)	4.5	$\frac{(6-4.5)^2+(3-4.5)^2}{2}=2.25$
(6, 6)	6	$\frac{(6-6)^2+(6-6)^2}{2}=0$
(6, 8)	7	$\frac{(6-7)^2+(8-7)^2}{2}=1$
(6, 11)	8.5	$\frac{(6-8.5)^2+(11-8.5)^2}{2}=6.25$
(8, 2)	5	$\frac{(8-5)^2+(2-5)^2}{2}=9$
(8, 3)	5.5	$\frac{(8-5.5)^2+(3-5.5)^2}{2}=6.25$
(8, 6)	7	$\frac{(8-7)^2+(6-7)^2}{2}=1$
(8, 8)	8	$\frac{(8-8)^2+(8-8)^2}{2}=0$
(8, 11)	9.5	$\frac{(8-9.5)^2+(11-9.5)^2}{2}=2.25$
(11, 2)	6.5	$\frac{(11-6.5)^2+(2-6.5)^2}{2}=20.25$
(11, 3)	7	$\frac{(11-7)^2+(3-7)^2}{2}=16$
(11, 6)	8.5	$\frac{(11-8.5)^2+(6-8.5)^2}{2}=6.25$
(11, 8)	9.5	$\frac{(11-9.5)^2+(8-9.5)^2}{2}=2.25$
(11, 11)	11	$\frac{(11-11)^2+(11-11)^2}{2}=0$

(c)The mean of the sampling distribution of means $\mu_{\bar{x}} = \frac{\text{Sum of all sample means}}{25} = \frac{150}{25} = 6$.

**** The mean of the sampling distribution of means=Mean of the population

(d) The variance of the sampling distribution of means

$$\sigma_{\bar{x}}^2 = \frac{\begin{bmatrix} (2-6)^2+(2.5-6)^2+(4-6)^2+(5-6)^2+(6.5-6)^2+(2.5-6)^2+ \\ (3-6)^2+(4.5-6)^2+(5.5-6)^2+(7-6)^2+(4-6)^2+(4.5-6)^2+(6-6)^2+ \\ (7-6)^2+(8.5-6)^2+(5-6)^2+(5.5-6)^2+(7-6)^2+(8-6)^2+(9.5-6)^2 \\ (6.5-6)^2+(7-6)^2+(8.5-6)^2+(9.5-6)^2+(11-6)^2 \end{bmatrix}}{25} =$$

$$\sigma_{\bar{x}} = \sqrt{\quad} =$$

Finite population involving sampling with replacement or Infinite population

Standard error (S. E.) about mean $\bar{x} = \frac{\sigma}{\sqrt{n}} =$

(e) The mean of the sampling distribution of variances $\bar{s}^2 = \frac{\text{Sum of all } sample \text{ variances}}{25} =$

(f) The variance of the sampling distribution of variances

$$\sigma_{s^2} = \frac{\sum\left(x_{s^2} - \mu_{s^2}\right)}{25}$$

$$\frac{\begin{bmatrix} (0-\quad)^2+(0.25-\quad)^2+(4-\quad)^2+(9-\quad)^2+(20.25-\quad)^2+(0.25-\quad)^2+ \\ (0-\quad)^2+(2.25-\quad)^2+(6.25-\quad)^2+(16-\quad)^2+(4-\quad)^2+(2.25-\quad)^2+(1-\quad)^2+ \\ (0-\quad)^2+(6.25-\quad)^2+(9-\quad)^2+(6.25-\quad)^2+(1-\quad)^2+(0-\quad)^2+(2.25-\quad)^2 \\ (20.25-\quad)^2+(16-\quad)^2+(6.25-\quad)^2+(2.25-\quad)^2+(0-\quad)^2 \end{bmatrix}}{25}$$

$=$

Sample: Without replacement or Finite Population

The total number of samples with replacement is $N_{C_n} = 5_{C_2} = 10$

The samples with their corresponding means and variances sample size 2 is shown in the following table:

Samples	Sample Means	Sample Variances x_{s^2}
(2, 3)	2.5	$\frac{(2-2.5)^2+(3-2.5)^2}{2} = 0.25$

(2, 6)	4	$\frac{(2-4)^2+(6-4)^2}{2}=4$
(2, 8)	5	$\frac{(2-5)^2+(8-5)^2}{2}=9$
(2, 11)	6.5	$\frac{(2-6.5)^2+(11-6.5)^2}{2}=20.25$
(3, 6)	4.5	$\frac{(3-4.5)^2+(6-4.5)^2}{2}=2.25$
(3, 8)	5.5	$\frac{(3-5.5)^2+(8-5.5)^2}{2}=6.25$
(3, 11)	7	$\frac{(3-7)^2+(11-7)^2}{2}=16$
(6, 8)	7	$\frac{(6-7)^2+(8-7)^2}{2}=1$
(6, 11)	8.5	$\frac{(6-8.5)^2+(11-8.5)^2}{2}=6.25$
(8, 11)	9.5	$\frac{(8-9.5)^2+(11-9.5)^2}{2}=2.25$

(c)The mean of the sampling distribution of means $\mu_{\bar{x}}=\frac{\text{Sum of all sample means}}{10}=\frac{150}{25}=6$

**** The mean of the sampling distribution of means=Mean of the population

(d) The variance of the sampling distribution of means

$$\sigma_{\bar{x}}^2=\frac{\begin{bmatrix}(2.5-6)^2+(4-6)^2+(5-6)^2+(6.5-6)^2+(4.5-6)^2+(5.5-6)^2+(7-6)^2\\(7-6)^2+(8.5-6)^2+(9.5-6)^2\end{bmatrix}}{10}=$$

$\sigma_{\bar{x}}=\sqrt{\quad}=$

Finite population involving sampling with replacement or Infinite population

Standard error (S. E.) about mean $\bar{x}=\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}=.$

(e) The mean of the sampling distribution of variances

$$\mu_{s^2}=\frac{\text{Sum of all } sample \text{ variances}}{10}=$$

(f) The variance of the sampling distribution of variances

$$\sigma_{s^2} = \frac{\sum\left(x_{s^2} - \mu_{s^2}\right)}{10}$$

$$\frac{\begin{bmatrix}(0.25- \quad)^2 + (4- \quad)^2 + (9- \quad)^2 + (20.25- \quad)^2 + (2.25- \quad)^2 \\ +(6.25- \quad)^2 + (16- \quad)^2 + (1- \quad)^2 + (6.25- \quad)^2 + (2.25- \quad)^2\end{bmatrix}}{10}$$

$=$

****Problem 6: A population consists of six numbers 4, 8, 12, 16, 20, 24. Consider all possible samples of size two that can be drawn without replacement and with replacement from this population. Find

(a) The mean of the population .

(b) The standard deviation of the population.

(c) The mean of the sampling distribution of means.

(d) The standard deviation of the sampling distribution of means.

Solution: Do Practice at note book

****Problem 7: Find the mean and Standard deviation of sampling distribution of means for the population 2, 3, 4, 5 by drawing samples of size two with replacement and without replacement.

Solution: Do Practice at note book

Problem 8: Let $u_1 = (3, 7, 8)$, $u_2 = (2, 4)$. Find

(a) $\mu_{u_1}, \mu_{u_2}, \mu_{u_1-u_2}$ (*Mean of the sampling distribution of means*)

(b) $\sigma_{u_1}, \sigma_{u_2}, \sigma_{u_1-u_2}$ (*S tan dard deviations of the sampling distribution of means*)

Problem 9: The variance of a population is 2. The size of the sample collected from the population is 169. What is the standard error of mean.

Solution:

Problem 10: When a sample is taken from an infinite population, what happens the standard error of the mean if the sample size is decreased from 800 to 200.

Solution:

Model No 4.4: Central limit theorem

Central Limit Theorem: If $\bar{x}$ be the mean of a sample size n drawn from a population mean μ and standard deviation σ then the standardized normal variate $z = \dfrac{\bar{x}-\mu}{\left(\dfrac{\sigma}{\sqrt{n}}\right)}$ is asymptotically normal.

The central limit theorem (CLT) states that the distribution of sample means approximates a normal distribution as the sample size gets larger, regardless of the population's distribution.

Note: Sample sizes equal to or greater than 30 are often considered sufficient for the CLT to hold.

Problem 11: Determine the mean and standard deviation of the sampling distribution of means of 300 random samples each of size n=36 are drawn from the population N=1500 which is normally distributed with $\mu = 22.4$, $\sigma = 0.048$. If the sampling distribution is done

i) With replacement
ii) Without replacement

a) Between 22.39 and 22.41
b) Greater than 22.42
c) Less than 22.37
d) Less than 22.38 and greater than 22.41

Solution:

Problem 12: A random sample of size 144 is taken from an infinite population having mean $\mu = 75,\ \sigma^2 = 225.$ What is the probability that $\bar{x}$ will lies between 72 and 77.

Problem 13: The mean of certain normal population is equal to the standard error of the mean of the samples of 64 from that distribution. Find the probability that the mean of the sample size 36 will be negative.

Solution:

Problem 14: A random sample of size 100 is taken from an infinite population having the mean μ=76 and the variance σ^2 =256.What is the probability that $\bar{x}$ will be between 75 and 78.

Solution:

Problem 15: A random sample of size 64 is taken from an infinite population having the mean 45 and the Standard deviation 8. What is the probability that x will be between 46 and 47.5.

Solution:

Problem 16: A normal population has a mean of 0.1 and standard deviation 2.1. Find the probability that mean of a sample of size 900 will be negative.

Solution:

Problem 17: A random sample of size 64 is taken from a normal population with μ=51.4 and σ=68. What is the probability that the mean of the sample will (a) exceed 52.9
(b) fall between 50.5 and 52.3 (c) be less than 50.6.

Solution:

Problem 18: If the mean of breaking strength of copper wire is 575 lbs, with a standard deviation of 8.3 lbs. How large a sample must be used in order that there will be one chance in 100 that the mean breaking strength of the sample is less than 572 lbs?

Solution:

Model No 4.5: Problems on series

Formulae: If A, B, C, D are connected in series then

1. Mean $\mu_{A+B+C+D} = \mu_A + \mu_B + \mu_C + \mu_D$
2. Standard deviation $\sigma_{A+B+C+D} = \sqrt{\sigma_A^2 + \sigma_B^2 + \sigma_C^2 + \sigma_D^2}$
3. Mean $\mu_{A-B} = \mu_A - \mu_B$
4. Standard deviation $\sigma_{A-B} = \sqrt{\sigma_A^2 + \sigma_B^2}$.

Problem 19: The mean voltage of a battery is 15 and S.D is 0.2. Find the probability that four such batteries connected in series will have a combined voltage of 60.8 or more volts.
Solution:

Problem 20: Three masses of means are measured as 62.34 kgs, 20.48 kgs, 35.97 kgs with S.D 0.54kgs, 0.21kgs, 0.46kgs. Find the mean and S.D of the sum of the masses.

Solution:

Problem 21: The diameter of motor shafts in a lot has a mean of 0.249 inch and a S.D of 0.003 inch. The inner diameter of bearings in another lot have a mean of 0.225 inch and a S.D of 0.002 inch.
i) What are the mean and the S.D of the clearances between shafts and bearings selected from those lots?
ii) If a shafts and a bearing are selected at random, what is the probability that the shaft will not fit inside, the bearing ? Assume that both dimensions are normally distributed.
Solution:

Model No 4.6: Estimations & Point and interval estimations

Estimation: The judgment made by unknown parameter is called estimation.

(or)

Estimation is a procedure by which numerical values are assigned to parameter based on information collected by samples.

Estimate: A statement made to find an unknown population parameter is called estimate.

Estimator: The procedure or rule, to determine an unknown population parameter is called an estimator.

Example: Sample mean an estimator of population mean be sample mean is a method of determining the population mean.

Note:

1. An estimate must be a static and it must only depends on the sample.
2. A parameter can have one or two or more estimators.

Types of Estimation:

They are two kinds of estimations to determine the static of the population parameters they are

1. Point estimation
2. Interval estimation

1. Point estimation: If an estimate of the population parameter is given by a single values then the estimate is called Point estimation of the parameter.

Example :- The sample mean is a point estimate of the population mean μ .

2. Interval estimation: If an estimate of the population is given by two different values between which the parameter may consider to lie. Then the estimate is called interval estimate of the parameter.

Example: We say that a distance is 5.8 kms

In this case we are going to take point estimate.

The distance is in between 5.28 + or – 0.3

In this case we are going to take interval estimate.

Important properties of the estimator:

Reliability: A statement of error is known as reliability

Unbiased Estimator: If the mean of sampling distribution of the statistic is equal to the population parameter then the statistic is said to be unbiased estimation of population parameter.

The corresponding value of statistic is there called unbiased estimate of the parameter.

Efficient Estimator:

If the sampling distribution of the statistic with less mean then the statistic with less mean then the statistic with the smaller variance is called a more efficient estimation of the mean. The corresponding value of the statistic is called an efficient estimate.

Confidence interval: An interval estimate i.e., constructed based on the confidence level is called confidence interval.

Confidence level is denoted by $(1-\alpha)\times 100\%$.

$(1-\alpha)$ is called confidence coefficient or degree of coefficient and α is called significance level. Then confidence coefficient 1-α =1-0.01=0.99. If α =0.01.

Model No 4.7: Unbiased Estimations
Model No 4.8: Maximum error of estimate
Formulae:

	Large sample $n \geq 30$	Small sample $n < 30$
Confidence interval for population mean $\bar{x}$ = Sample mean $z_{\frac{\alpha}{2}}$ = The confident coefficient α = Confidence level σ = Standard deviation n = Sample size s= standard deviation of the sample.	$\mu = \bar{x} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ (or) $\bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ (or) $\left(\bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right)$.	$\mu = \bar{x} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$ or $\bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$ or $\left(\bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}, \bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \right)$. Degrees of freedom $\vartheta = n - 1$
Confidence interval for Proportions Limits for population parameter Proportion P for 99% or 1% only are given by		$p \pm z_{\frac{\alpha}{2}} \sqrt{\frac{PQ}{n}}$ $p \pm 2.58\sqrt{\frac{pq}{n}} \sim p \pm 3\sqrt{\frac{pq}{n}}$ depends on given data, replace p by 'P' in the formula.
Maximum error (**Atmost**) of the estimate E with $(1-\alpha)$ probability	$E = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$	$E = t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$
Maximum error **(Atmost)** of the estimate for proportions No mean-NoS.D.		$E = z_{\frac{\alpha}{2}} \sqrt{\frac{PQ}{n}}$
Sample size	$n = \left(\frac{z_{\frac{\alpha}{2}} \sigma}{E} \right)^2$	$n = \left(\frac{t_{\frac{\alpha}{2}} s}{E} \right)^2$

Note: If the proportions are not given then take $p = \frac{1}{2}$.

Problem 22: If $x_1, x_2 \ldots\ldots x_n$ is random sample from a given population with mean μ and variance σ^2. Show that the sample mean is an unbiased estimator of population mean μ.

Solution:

Problem 23: Show that the sample variance s^2 is an unbiased estimator of population variance σ^2.

Solution:

Problem 24: In a study of an automobile insurance a random sample of 80 body repair costs had a mean of ₹472.36 and the S.D of ₹62.35. If $\bar{x}$ is used as a point estimate to the true average repair costs, with what confidence we can assert that the maximum error doesn't exceed ₹10.

Solution:

Problem 25: It is desired to estimate the mean number of hours of continuous use until a certain computer will first require repairs. If it can be assumed that σ = 48 hours, how large a sample be needed so that one will be able to assert with 90% confidence that the sample mean is off by at most 10 hours.

Solution:

Problem 26: A random sample of size 100 has a standard deviation of 5. What can you say about the maximum error with 95% confidence.

(or)

If n = 100, σ = 5, find the maximum error with 95% confidence limits.

Solution:

Problem 27: The efficiency expert of a computer company tested 40 engineers to estimate the average time it takes to assemble a certain computer component, getting a mean of 12.73 minutes and S.D of 2.06 minutes.

(a) If $\bar{x}$ = 12.73 is used as a point estimate of the actual average time required to perform the task, determine the maximum error with 99% confidence.

(b) Construct 98% confidence intervals for the true average time it takes to do the job.

(c) With what confidence can we assert that the sample mean does not differ from the true mean by more than 30 seconds.

(or)

To estimate the average time it takes to assemble a certain computer component, the industrial engineer at an electronic firm timed 40 technicians in the performance of the task, getting a mean of 12.73 min and a S.D of 2.06 min.

(a) What can we say with 99% confidence about the maximum error if $\bar{x}$ =12.73 is used a point estimate of the actual average time required to do the job ?

(b) Use the given data to construct 98% confidence interval.

(c) With what confidence we can assert that sample mean does not differ from the true mean by more than 30 sec.

Solution:

Problem 28: The mean and standard deviation of a population are 11.795 and 14.054 respectively. What can one assert with 95% confidence about the maximum error if $\bar{x}$ = 11.795 and n = 50. And also construct 95% confidence interval for the true mean.

(or)

The mean and the standard deviation of a population are 11.795 and 14.054 respectively. If n = 50, find 95% confidence interval for the mean.

Solution:

Problem 29: The mean of random sample is an unbiased estimate of the mean of the population 3,6,9,15,27.
i) List of all possible samples of size 3 that can be taken without replacement from the finite population.
ii) Calculate the mean of each of the samples listed in (a) and assigning each sample a probability of 1/10. Verify that the mean of these $\bar{x}$ is equal to 12. Which is equal to the mean of the population θ i.e E($\bar{x}$) = θ i.e., prove that $\bar{x}$ is an unbiased estimate of θ.
Solution:

Problem 30: A professor's feelings about the mean mark in the final examination in "Probability" of a large group of students is expressed subjectively by normal distribution with $\mu_0 = 67.2$ and $\sigma_0 = 1.5$.
(a) If the mean mark lies in the interval (65.0, 70.0) determine the prior probability the professor should assign to the mean mark.
(b) Find the professor mean μ_1 and the posterior S.D σ_1 if the examinations are conducted on a random sample of 40 students yielding mean 74.9 and S.D 7.4. Use S =7.4 as an estimate σ .
(c) Determine the posterior probability which he will thus assign to the mean mark being in the interval (65.0,70.0) using results obtained in (b).
(d) Construct a 95% Bayesian interval for μ.

Solution: Hints: $\mu_1 = \dfrac{n\bar{x}\sigma_0^2 + \mu_0\sigma^2}{n\sigma_0^2 + \sigma^2}$, $\sigma_1 = \sqrt{\dfrac{\sigma^2\sigma_0^2}{n\sigma_0^2 + \sigma^2}}$

Problem31: A random sample of 100 teachers in a large metropolitan area revealed a mean weekly salary of Rs.487 with a standard deviation Rs.48 .With what degree of confidence can we assert that the average weekly salary of all teachers in the metropolitan area is between 472 to 502?

Solution:

Problem 32: Among 100 fish caught in a large lake, 18 were inedible due to the pollution of the environment. With what confidence can we assert the error of this estimate is at most 0.065?

Solution:

Problem 33: The mean mark in mathematics in common entrance that will vary from year to year. If this variation of the mean mark is expressed subjectively by a normal distribution with mean μ_0 = 72 and variance $\sigma_0{}^2$=5.76.

i) What probability can we assign to the actual mean mark being somewhere between 71.8 and 73.4 for the next years test ?

ii) Construct a 95% Bayesian interval for μ if the test is conducted for a random sample of 100 students from the next incoming class yielding a mean mark of 70 with S.D of 8.

iii) What posterior probability should we assign to the event of part (i).

UNIT-V
TEST OF SIGNIFICANCE

Part-A Large samples

Model No 5.1: Test of Significance of a single mean
Model No 5.2: Test of Significance for difference of means
Model No 5.3: Test of Significance for single Proportion
Model No 5.4: Test for equality of two Proportions (or)
Test of significance of difference between two sample Proportions

Test of hypothesis: In many circumstances, to arrive at decisions about the population on the basis of sample information, we make assumptions (or guesses) about the population parameters involved. Such an assumption (or statement) is called a statistical hypothesis which may or may not be true. The procedure which enables us to decide on the basis of sample results whether a hypothesis is true or not, is called Test of Hypothesis or Test of significance.

Null Hypothesis (H_0)**:** It is denoted by H_0, is a statement about the population parameter which is to be actually tested for acceptance or rejection.

Alternative Hypothesis (H_1)**:** It is denoted by H_1, is the opposite statement of null hypothesis.

Types of errors in test of hypothesis:

Type I error: The rejection of null hypothesis when it is true and should be accepted.

Type II error: The acceptance of null hypothesis when it is false and should be rejected.

	Accept H_0	Reject H_0
H_0 is true	Correct Decision	Type I error
H_0 is false	Type II error	Correct Decision

Examples:

- **Type I error (false positive):** the test result says you have corona virus, but you actually don't.
- **Type II error (false negative):** the test result says you don't have corona virus, but you actually do.

Level Of Significance (L.O.S.): It is denoted by α, is the probability of committing type I error. Thus L.O.S. measures the amount of risk or error associated in taking decisions. L.O.S. is expressed in percentage. Thus L.O.S. $\alpha = 5\%$ means that there are 5 chances in 100 that null hypothesis is rejected when it is true.

$$\alpha = probability \;\; of \;\; committing \;\; type\, I \;\; error = P(\quad)$$

$$\beta = probability \;\; of \;\; committing \;\; type\, II \;\; error = P(\quad)$$

Critical Region (C.R.): In any test of hypothesis, a test statistic S^*, calculated from the sample data is used to accept or reject the null hypothesis. Consider the area under the probability curve of the sampling distribution of the test statistic S^*. This area under the probability curve is divided into two regions, namely the region of rejection where N.H. is rejected and the region of acceptance where N.H. is accepted. Thus, critical region is the region of rejection of N.H. The area of the critical region equals to the level of significance α. Note that C.R. always lies on the tail of the distribution.

One tailed test and two tailed tests:

Right tailed test: When the alternative hypothesis H_1 is of the greater than type $i.e., H_1 : \mu > \mu_0 \text{ or } H_1 : \sigma_1^2 > \sigma_2^2 \text{ etc}.$ Then the entire critical region of area α lies on the right side of the curve as shown shaded in the fig. In such case the test of hypothesis is known as right tailed test.

Left tailed test: When the alternative hypothesis H_1 is of the less than type $i.e., H_1 : \mu < \mu_0 \text{ or } H_1 : \sigma_1^2 < \sigma_2^2. \text{ etc}$ Then the entire critical region of area α lies on the left side of the curve as shown shaded in the fig. In such case the test of hypothesis is known as left tailed test.

Two tailed test: When the alternative hypothesis H_1 is of the Not equals type $i.e., H_1 : \mu \neq \mu_0 \text{ or } H_1 : \sigma_1^2 \neq \sigma_2^2 \text{ etc}.$ Then the entire critical region of area α lies on the both sides of the curve as shown shaded in the fig. In such case the test of hypothesis is known as two tailed test.

PROCEDURE FOR TESTING OF HYPOTHESIS:

(i) Null Hypothesis (H_0): Define a Null Hypothesis H_0 taking into consideration the nature of the problem and data involved.

(ii)Alternative Hypothesis (H_1): Define an Alternative Hypothesis H_1 so that we could decide whether we should use one-tailed or two-tailed test.

(iii) Level of Significance (α): Select the appropriate level of significance α depending on the reliability of the estimates and permissible risk.

(iv) Test Statistic: Compute the test statistic $z_{cal} = \dfrac{t - E(t)}{S.E. \text{ of } t}$

(v) Conclusion: We compare the computed value of the test statistic Z with the critical value Z_α at a given level of significance α

(i) If $|z_{cal}| < z_{tab}$ we accept the Null Hypothesis H_0

(ii) If $|z_{cal}| > z_{tab}$ we reject the Null Hypothesis H_0 i.e., we accept the Alternative Hypothesis H_1.

Test statistic for Large samples

Large Samples: $n \geq 30$		Test Statistic	Identification
1	**Test of Significance of a single mean**		
a.	Direct	$z = \dfrac{\bar{x} - \mu}{\left(\sigma / \sqrt{n}\right)}$	One sample mean One population mean One population S. D. One sample size
b.	Population S. D. (σ) is not known. *Sample S. D. (S) is known	$z = \dfrac{\bar{x} - \mu}{\left(s / \sqrt{n}\right)}$	One sample mean One population mean One sample S. D. One sample size
2	**Test of Significance for difference of means**		
a.	Direct	$z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ or $z = \dfrac{\left(\bar{x}_1 - \bar{x}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ $\mu_1 - \mu_2 = 0$ $\quad$ $\mu_1 - \mu_2 \neq 0$	Two sample means Two population S. D.s Two sample sizes
b.	When the samples are taken from the same population. $\sigma_1 = \sigma_2 = \sigma$	$z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma^2}{n_1} + \dfrac{\sigma^2}{n_2}}}$	Two sample means One Population S. D.s Two sample sizes
c.	The sample variances s_1 & s_2 are given	$z = \dfrac{\left(\bar{x}_1 - \bar{x}_2\right)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$	$s_1^2 = \dfrac{\sum\left(x_1 - \bar{x}_1\right)^2}{n_1 - 1}$ $s_2^2 = \dfrac{\sum\left(x_2 - \bar{x}_2\right)^2}{n_2 - 1}$
3	**Test of Significance for Single Proportion**		
a.	Direct	$z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}}$	No sample mean No S. D.s One sample size Observations are given (Probability)
4	**Test for equality of two Proportions (or)** **Test of significance of difference b/w two sample Proportions**		
a	Direct	$z = \dfrac{P_1 - P_2}{\sqrt{\dfrac{P_1 Q_1}{n_1} + \dfrac{P_2 Q_2}{n_2}}}$	No sample mean No S. D.s Two sample sizes Two Observations are given (Probability)
b	**Method of pooling:** Two Sample proportions p_1 and p_2 into a single proportion p	$z = \dfrac{p_1 - p_2}{\sqrt{pq\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$ $\quad$ $z = \dfrac{\left(p_1 - p_2\right) - \left(P_1 - P_2\right)}{\sqrt{pq\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$ $P_1 - P_2 = 0$ $\quad$ $P_1 - P_2 \neq 0$	$p = \dfrac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \dfrac{x_1 + x_2}{n_1 + n_2}$ $q = 1 - p$

Critical Values of z			
Level of Significance α	**1%**	**5%**	**10%**
Critical values for two-tailed test	$\lvert Z_\alpha \rvert = 2.58$	$\lvert Z_\alpha \rvert = 1.96$	$\lvert Z_\alpha \rvert = 1.645$
Critical values for right-tailes test	$Z_\alpha = 2.33$	$Z_\alpha = 1.645$	$Z_\alpha = 1.28$
Critical values for left-tailed test	$Z_\alpha = -2.33$	$Z_\alpha = -1.645$	$Z_\alpha = -1.28$

TEST OF SIGNIFICANCE FOR LARGE SAMPLES

Model No. 5.1: Test of significance for single mean:

(i) Null Hypothesis (H_0)***:*** $\bar{x} = \mu$ i.e., "there is no significance difference between the sample mean and population mean" or "the sample has been drawn from the population"

(ii)Alternative Hypothesis (H_1)***:*** (i) $\bar{x} \neq \mu$ or (ii) $\bar{x} < \mu$ or(iii) $\bar{x} > \mu$

(iii) Level of Significance (α)***:*** Set a level of significance

(iv) Test Statistic:

Case(i): When the S.D. of the population σ is known, The test statistic $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

Case(ii): When the S.D. of the population is not known. In this case, we take S.D. of the sample The test statistic $z = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$

(v) Conclusion: (i) If $\lvert z_{cal} \rvert < z_{tab}$ we accept the Null Hypothesis H_0

(ii) If $\lvert z_{cal} \rvert > z_{tab}$ we reject the Null Hypothesis H_0 i.e., we accept the Alternative Hypothesis H_1.

Problem 1: A sample of 64 students have a mean weight of 70kgs. Can this be regarded as a sample from a population with mean weight 56kgs and standard deviation 25kgs.

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 2: A sample of 900 members has a mean of 3.4 cms and S.D. 2.61 cms. Is this sample has been from a large population of mean 3.25 cm and S.D. 2.61 cms. If the population is normal and its mean is unknown find the 95% confidence limits of true mean.
Solution:
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

The confidence limits are

Problem 3: A sample of 400 items is taken from a population whose S.D. is 10. The mean of the sample is 40. Test whether the sample has come from a population with mean 38. Also calculate 95% confidence interval for the population.
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

The confidence limits are

Problem 4: An ambulance service claims that it takes on the average less than 10 minutes to reach its destination in emergency calls. A sample of 36 calls has a mean of 11 minutes and the variance of 16 minutes. Test the claim at 0.05 level of significance.
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 5: In a random sample of 60 workers, the average time taken to set to work is 33.8 minutes with a standard deviation of 6.1 minutes. Can we reject the null hypothesis μ = 32.6 minutes in favor of alternative null hypothesis μ > 32.6 at α =0.025 level of significance .

Solution:

(i) Null Hypothesis $\left(H_0\right)$:

(ii)Alternative Hypothesis $\left(H_1\right)$:

(iii) Level of Significance $\left(\alpha\right)$:

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 6: An insurance agent has claimed that the average age of policy holders who issue through him is less than the average for all agents which is 30.5 years. A random sample of 100 policy holders who had issued through him gave the following age distribution.

Age	16-20	21-25	26-30	31-35	36-40
No. of persons	12	22	20	30	16

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Model No 5.2: <u>Test of significance for difference of means</u>:

(i) Null Hypothesis (H_0)***:*** $\overline{x}_1 = \overline{x}_2$ *or* $\mu_1 = \mu_2$ i.e., "there is no significance difference between means of the populations" or "the two samples have been drawn from the same population"

(ii)Alternative Hypothesis (H_1)***:*** $\overline{x}_1 \neq \overline{x}_2$ *or* $\mu_1 \neq \mu_2$

(iii) Level of Significance (α)***:*** Set a level of significance

(iv) Test Statistic:

<u>Case(i)</u>:(a)When the S.D. of the populations σ_1 , σ_2 are given then the test statistic

$$z = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}}}$$

b) When the samples are taken from the same population then the test statistic $z = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{\sigma^{\ 2}}{n_1} + \frac{\sigma^{\ 2}}{n_2}}}$

<u>Case(ii)</u>: When the S.D. of the population is not known then the test statistic $z = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^{\ 2}}{n_1} + \frac{s_2^{\ 2}}{n_2}}}$

(v) Conclusion: (i) If $|z_{cal}| < z_{tab}$ we accept the Null Hypothesis H_0

(ii) If $|z_{cal}| > z_{tab}$ we reject the Null Hypothesis H_0 i.e., we accept the Alternative Hypothesis H_1.

Problem 7: The average marks scored by 32 boys is 72 with a S.D. of 8. While that for 36 girls is 70 with a S.D. of 6. Does this indicate that the boys perform better than girls at level of significance 0.05?

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 8: Two types of new cars produced in U.S.A. are tested for petrol mileage, one sample is consisting of 42 cars averaged 15 kmpl while the other sample consisting of 80 cars averaged 11.5 kmpl with population variance as $\sigma_1^2 = 2.0 \ \ and \ \ \sigma_2^2 = 1.5$ **respectively. Test whether there is any significance difference in the petrol consumption of these two types of cars. (use level of significance 0.01)**

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 9: The means of two large samples of sizes 1000 and 2000 members are 67.5 inches and 68.0 inches respectively. Can the samples be regarded as drawn from the same population of S.D. 2.5 inches. Use 5% L.O.S

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 10: Samples of students were drawn from two universities and from their weights in kilograms, mean and S.D. are calculated and shown below. Make a large sample test to test the significance of the difference between the means.

	Mean	**S.D.**	**Size of the sample**
University A	**55**	**10**	**400**
University B	**57**	**15**	**100**

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 11: The mean yield of wheat from a district A was 210 pounds with S.D 10 pounds per acre from a sample of 100 pounds. In another district the mean yield was 200 pounds with S.D 12 pounds from a sample of 150 plots. Assuming that the S.D of yield in the entire was 11 pounds, test whether there is any significant difference between the mean yield of crops in the two districts.

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of   Tabulated value of

Problem 12: In a survey of buying habits, 400 women shoppers are chosen at random in super market 'A' located in a certain section of the city. Their average weekly food expenditure is Rs250 with a S.D of Rs40. For 400 women shoppers chosen at random in super market 'B' in another section of the city, the average weekly food expenditure is Rs220 with a S.D of Rs55. Test at 1% level of significance whether the average weekly food expenditure of the two populations of shoppers are equal.

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of   Tabulated value of

Problem 13: The nicotent in milligrams of two samples of tobacco were found to be as follows. Find the standard error and confidential limits for the difference between the means at 0.05 level.

Sample A	24	27	26	23	25	
Sample B	29	30	30	31	24	26

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Model No 5.3: **Test of significance for single proportion**:

(i) Null Hypothesis (H_0)***:*** $P = p$ i.e., "there is no significance difference between the sample proportion and population proportion" or "the sample has been drawn from the population"

(ii)Alternative Hypothesis (H_1)***:*** (i) $P \neq p$ or (ii) $P < p$ or (iii) $P > p$

(iii) Level of Significance (α)***:***Set a level of significance

(iv) Test Statistic: The test statistic $z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}}$

(v) Conclusion: (i) If $|z_{cal}| < z_{tab}$ we accept the Null Hypothesis H_0

(ii) If $|z_{cal}| > z_{tab}$ we reject the Null Hypothesis H_0 i.e., we accept the Alternative Hypothesis H_1

Problem 14: A manufacturer claimed that atleast 95% of the equipment which he supplied to a factory conformed to specifications. An examination of a sample of 200 pieces of equipment revealed that 18 were faulty. Test his claim at 5% level of significance. Also find the confidence interval.

Solution:

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

The confidence interval is

Problem 15: In a random sample of 125 cool drinkers, 68 said they prefer thumsup to pepsi. Test the null hypothesis P=0.5 against the alternative hypothesis P>0.5.
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 16: In a sample of 500 from a village in Rajasthan, 280 are found to be wheat eaters and the rest rice eaters. Can we assume that the both articles are equally popular.
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 17: A die was thrown 9000 times and of these 3220 yielded a 3 or 4. Is this consistent with the hypothesis that the die was unbiased?
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 18: In a big city 325 men out of 600 men found to be smokers Does this information support the conclusion that the majority of men in this city are smokers ?
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 19: A social worker believes that fewer than 25% of the couples in a certain area have used any form of birth control. A random sample of 120 couples was contacted. Twenty of them said that they have used. Test the belief of the social worker at 0.05 level.
Solution:

(i) Null Hypothesis $\left(H_0\right)$:

(ii)Alternative Hypothesis $\left(H_1\right)$:

(iii) Level of Significance $\left(\alpha\right)$:

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 20: Among 900 people in a state 90 are found to be chapati eaters. Construct 99% confidence interval for the true proportion.

Solution:

(i) Null Hypothesis $\left(H_0\right)$:

(ii)Alternative Hypothesis $\left(H_1\right)$:

(iii) Level of Significance $\left(\alpha\right)$:

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Model No 5.4: <u>Test of significance for difference of proportions</u>:

(i) Null Hypothesis (H_0)***:*** $p_1 = p_2$ *or* $P_1 = P_2$ i.e., "there is no significance difference between the proportions of the samples or proportions of populations" or "the two samples have been drawn from the same population"

(ii)Alternative Hypothesis (H_1)***:*** $p_1 \neq p_2$ *or* $P_1 \neq P_2$

(iii) Level of Significance (α)***:*** Set a level of significance

(iv) Test Statistic:

<u>Case(i)</u>: When the population proportions P_1 *and* P_2 are known

The test statistic $z = \dfrac{p_1 - p_2}{\sqrt{\dfrac{P_1Q_1}{n_1} + \dfrac{P_2Q_2}{n_2}}}$

<u>Case(ii)</u>: When the population proportions P_1 *and* P_2 are unknown, and the sample proportions p_1 *and* p_2 are known

The test statistic $z = \dfrac{p_1 - p_2}{\sqrt{\dfrac{p_1q_1}{n_1} + \dfrac{p_2q_2}{n_2}}}$ *or* $z = \dfrac{p_1 - p_2}{\sqrt{pq\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$ *where* $p = \dfrac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$, $q = 1 - p$

(v) Conclusion: (i) If $|z_{cal}| < z_{tab}$ we accept the Null Hypothesis H_0

(ii) If $|z_{cal}| > z_{tab}$ we reject the Null Hypothesis H_0

i.e., we accept the Alternative Hypothesis H_1

Problem 21: In two large populations, there are 30% and 25% are fair haired people. Is this difference likely to be hidden in samples of 1200 and 900 respectively from the two populations?

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 22: A cigarette manufacturing firm claims that its brand A line of cigarettes outsells its brand B by 8%. If it is found that 42 out of a sample of 200 smokers prefer brand A and 18 out of another samle of 100 smokers prefer brand B, test whether the 8% difference is a valid claim.

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 23: A machine puts out 9 imperfect articles ina sample of 200 articles. After the machine is over haulted it puts out 5 imperfect articles ina sample of 700 articles. Test at 5% level whether the machine is improved?

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 24: In a city A, 20% of a random sample of 900 school boys has a certain slight physical defect. In another city B, 18.5% of a random sample of 1600 school boys has the same defect. Is the difference between the proportions significant at 005 level of significance.

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 25: Random samples of 400 men and 600 women asked whether they would like to have a flyover near their residence. 200 men and 325 women were in favour of the proposal. Test the hypothesis that proportions of men and women in favour of the proposal are same, at 5% level.

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 26: A manufacturer of electronic equipment subjects sample of two completing brands of transistor to an accelerated performance test. If 45 of 180 transistors pf the first kind and 34 of 120 transistor of the second kind fail the test, what can he conclude at the level of significance a = 0.05 about the difference the corresponding sample proportions ?

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Problem 27: On the basis of their total scores, 200 candidates of a civil service examinations are divided into two groups, the upper 30% and the remaining 70%. Consider the first question of the examination. Among the first group, 40 had the correct answer, where as the second group, 80 had the correct answer. On the basis of these results, can one conclude that the first question is not good at discriminating ability of the type being examined here ?

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Tabulated value of

Calculated value of

Calculated value of Tabulated value of

Part-B **Formulas for small samples**

Small Samples: $n \le 30$		Test Statistic	Degrees of freedom
1	**Student's 't' test for single mean (Single sample)**		
a.	S.D. is not given	$t = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$	$\nu = n - 1$ $s = \sqrt{\dfrac{\sum (x_i - \bar{x})^2}{n - 1}}$
b.	S.D. is given directly	$t = \dfrac{\bar{x} - \mu}{s / \sqrt{n - 1}}$	$\nu = n - 1$
2	**Student's "t" test for difference of means (Two Sample Means)**		
a.	Direct	$t = \dfrac{\bar{x} - \bar{y}}{S\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ or $t = \dfrac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{S\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ $(\mu_1 - \mu_2) = 0$ $(\mu_1 - \mu_2) \ne 0$	$\nu = n_1 + n_2 - 2$ $s^2 = \dfrac{\sum (x_i - \bar{x})^2 + (y_i - \bar{y})^2}{n_1 + n_2 - 2}$
3	F-Test- Variances **(Two Sample Variances)**		
a	Direct	$F = \dfrac{Greater\ \ Variance}{Smaller\ \ Variance} = \dfrac{s_1^2}{s_2^2}$ $\nu = (n_1 - 1,\ n_2 - 1)$ Or $F = \dfrac{Greater\ \ Variance}{Smaller\ \ Variance} = \dfrac{s_2^2}{s_1^2}$ $\nu = (n_2 - 1,\ n_1 - 1)$	$S_1 = \sqrt{\dfrac{n_1 s_1^2}{n_1 - 1}} = \sqrt{\dfrac{\sum (x_i - \bar{x})^2}{n_1 - 1}}$ $S_2 = \sqrt{\dfrac{n_2 s_2^2}{n_2 - 1}} = \sqrt{\dfrac{\sum (y_i - \bar{y})^2}{n_2 - 1}}$
3	**CHI-SQUARE (χ^2) Test FOR GOODNESS OF FIT (Attributes)**		
a	Direct	$\chi^2 = \sum \dfrac{(O_i - E_i)^2}{E_i}$	$\nu = n - 1$ $E_i = \dfrac{Sum\ of\ all\ observations}{number\ of\ observations}$
b	Expected frequencies by Binomial distribution	$\chi^2 = \sum \dfrac{(O_i - E_i)^2}{E_i}$	$\nu = n - 1$
c	Expected frequencies by Poisson distribution	$\chi^2 = \sum \dfrac{(O_i - E_i)^2}{E_i}$	$\nu = n - 2$

4. CHI-SQUARE $\left(\chi^2\right)$ Test for Independence Attributes

(Matrix Type or Habitual activities)

Test Statistic: $\chi^2 = \sum \frac{\left(O_i - E_i\right)^2}{E_i}$

Let us consider two attributes A and B, and they are divided into two classes. The various frequencies can be expressed as follows:

A	a	b
B	c	d

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	$a+b+c+d=N$

The expected frequencies are given given by:

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

Degrees of freedom $=(n-1)(m-1)$

Let us consider three attributes A, B and C they are divided into three classes. The various frequencies can be expressed as follows: Degrees of freedom $=(n-1)(m-1)$

A	a	b	c
B	d	e	f
C	g	h	i

$E(a)=\frac{(a+d+g)(a+b+c)}{N}=$	$E(b)=\frac{(b+e+h)(a+b+c)}{N}=$	$E(b)=\frac{(c+f+i)(a+b+c)}{N}=$
$E(c)=\frac{(a+d+g)(d+e+f)}{N}=$	$E(b)=\frac{(b+e+h)(d+e+f)}{N}=$	$E(b)=\frac{(c+f+i)(d+e+f)}{N}=$
$E(c)=\frac{(a+d+g)(g+h+i)}{N}=$	$E(b)=\frac{(b+e+h)(g+h+i)}{N}=$	$E(b)=\frac{(c+f+i)(g+h+i)}{N}=$

Model No 5.5: Test of significance for single mean (Students's t- test)
Model No 5.6: Student's "t" test for difference of means (Two Sample Means)
Model No 5.7: F-Test- Variances **(Two Sample Variances)**
Model No 5.8: CHI-SQUARE $\left(\chi^2\right)$ Test FOR GOODNESS OF FIT (Attributes)
Model No 5.9: CHI-SQUARE $\left(\chi^2\right)$ Test for Independence Attributes

TEST OF SIGNIFICANCE FOR SMALL SAMPLES:

Model No 5.5: <u>Test of significance for single mean (Students's t- test)</u>:

(i) Null Hypothesis (H_0)***:*** $\bar{x} = \mu$ i.e., "there is no significance difference between the sample mean and population mean" or "the sample has been drawn from the population"

(ii)Alternative Hypothesis (H_1)***:*** (i) $\bar{x} \neq \mu$ or (ii) $\bar{x} < \mu$ or (iii) $\bar{x} > \mu$

(iii) Level of Significance (α)***:*** Set a level of significance

(iv) Test Statistic: The test statistic $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

(v) Conclusion: (i) If $|t_{cal}| < t_{tab}$ we accept the Null Hypothesis H_0

(ii) If $|t_{cal}| > t_{tab}$ we reject the Null Hypothesis H_0 i.e., we accept the Alternative Hypothesis H_1.

Problem 1: A machinist is making engine parts with axle diameters of 0.700 inch. A random sample of 10 parts shows a mean diameter of 0.742 inch with a standard deviation of 0.040 inch. Compute the test statistic you would use to test whether the work is meeting the specifications. Also state how you would proceed further.

Solution: Here we are given,

μ =0.700 inch, $\bar{x}$=0.742 inches, s=0.040 inch and n=10

Null Hypothesis: H_0: μ=0.700 inch, i.e., the product is confirming to specifications

Alternative hypothesis: H_1: $\mu \neq 0.700$inches

Level of significance: $\alpha = 0.05$

Test statistic: Under H_0, the test statistic is

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

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = \frac{0.742 - 0.700}{0.040/\sqrt{9}} = 3.15$$

How to proceed further: Here the test statistic't' follows student's t-distribution with 10-1=9 degrees of freedom. We will now compare this calculated value with the tabulated value for t for 9 degrees of freedom and at a certain level of significance, say 5%.

i) If calculated 't' =3.15> t-table value, we say that the value of t is significant. This implies that $\bar{x}$ differs significantly from μ and H_0 is rejected at this level of significance and we conclude that the product is not meeting the specifications.

ii) If calculated t<t-table value, we say that the value of t is not significant. There is no significant difference between $\bar{x}$and μ. We may take the product conforming to specifications.

$t_{0.05}$=2.26 , $t_{cal} > t_{tab}$

Therefore H_0 is rejected. Hence the product is not meeting the specification.

Problem 2: A random sample of 10 boys had the following I. Q's: 70, 120, 110, 101, 88, 83, 95, 98, 107, 100.

a) Do these data support the assumption of a population mean I. Q. of 100?

b) Also, Find a reasonable range in which most of the mean I. Q. values of samples of 10 boys lie.

Solution: Null Hypothesis H_0: The data are consistent with the assumption of a mean I. Q. of 100 in the population, i.e., H_0: $\mu=100$.

Alternative Hypothesis: H_1: $\mu\neq100$.

Level of Significance (α): 5%

Test statistic: Under H_0, the test statistic is: $t = \frac{\bar{x}-\mu}{S/\sqrt{n}}$

Where $\bar{x}$ and s^2 are to be computed from the sample values of I. Q.'s.

Calculations for Sample Mean and Standard deviation:

Here n=10, $\bar{x}$=972/10=97.2 and $s=\sqrt{\frac{\sum\left(x_i-\bar{x}\right)^2}{n-1}}=1833.60/9=203.73$

$$|t| = \frac{2.8}{14.27/\sqrt{10}} = 0.62$$

t-table value at 5% LOS for 9 degrees of freedom for two-tailed test is 2.262.

Conclusion:

Since calculated t is less than tabulated t ($t_{cal} < t_{tab}$). Null Hypothes is H_0, may be accepted at 5% level of significance. Hence we conclude that the data are consistent with the assumption of mean I.Q. of 100 in the population.

The 95% confidence limits within which the mean I. Q. values of samples of 10 boys will be are given by:

$\bar{x} \pm t_{0.05} S/\sqrt{n}$=97.2±2.262*4.514=107.41 and 86.99

Hence the required 95% confidence interval is [86.99, 107.41].

Problem 3: **Producer of gutkha, claims that the nicotine content in his "gutkha" on the average is 1.83 mg. Can this claim accepted if a random sample of 8 gutkha of this type have the nicotine contents of 2.0, 1.7, 2.1, 1.9, 2.2, 2.0, 1.6 mg? Use a 0.05 L.O.S.**

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of t_{tab} =

Calculated value of $|t_{cal}|$ =

Calculated value of $|t_{cal}|$ Tabulated value of t_{tab}

Problem 4: The life time of electric bulbs for a random sample of 10 from a large consignment gave the following data.

Item	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Life in 1000hrs	**1.2**	**4.6**	**3.9**	**4.1**	**5.2**	**3.8**	**3.9**	**4.3**	**4.4**	**5.6**

Can we accept the hypothesis that the average life time of bulbs is 4000hrs?
Use a 0.05 L.O.S.
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of t_{tab} =

Calculated value of $|t_{cal}|$ =

Calculated value of $|t_{cal}|$ Tabulated value of t_{tab}

Problem 5: A sample of 26 bulbs gives a mean life of 990 hours with a S.D of 20 hours. The manufacturer claims that the mean life of bulbs is 1000 hours. Is the sample not up to the standard?

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of t_{tab} =

Calculated value of $|t_{cal}|$ =

Calculated value of $|t_{cal}|$ Tabulated value of t_{tab}

Problem 6: The average breaking strength of the steel rods is specified to be 18.5 thousands pounds. To test this sample of 14 rods were tested. The mean and standard deviations obtained were 17.85 and 1.955 respectively. Is the result of experiment significant.

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of t_{tab} =

Calculated value of $|t_{cal}|$ =

Calculated value of $|t_{cal}|$ Tabulated value of t_{tab}

Problem 7: The heights of 10 males of a given locality are found to be 70, 67, 62, 68, 61, 68, 70, 64, 64, 66 inches. Is it reasonable to believe that the average height is greater than 64 inches ? Test at 5% significance level assuming that for 9 degree of freedom (t=1.833 at α =0.05) .

Item	1	2	3	4	5	6	7	8	9	10
Life in 1000hrs	1.2	4.6	3.9	4.1	5.2	3.8	3.9	4.3	4.4	5.6

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of t_{tab} =

Calculated value of $|t_{cal}|$ =

Calculated value of $|t_{cal}|$ Tabulated value of t_{tab}

Problem 8: A new process of producing synthetic diamonds can be operated at a profitable level only if the average weight of the diamonds is greater than 0.5 carat. To test the probability of the process, 6 diamonds are produced with weights 0.46, 0.60, 0.52, 0.49, 0.58 and 0.54 carat respectively. Do the 6 measurements present sufficient evidence to indicate that the average weight of the diamonds produced by the process is in excess of 0.5 carat?

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of $t_{tab} =$

Calculated value of $|t_{cal}| =$

Calculated value of $|t_{cal}|$ Tabulated value of t_{tab}

Model No 5.6: <u>Test of significance for difference of means (Students's t- test)</u>:

(i) Null Hypothesis** (H_0)**: $\bar{x} = \mu$ i.e., "there is no significance difference between the sample mean and population mean" or "the sample has been drawn from the population"

(ii)Alternative Hypothesis** (H_1)**: (i) $\bar{x} \neq \mu$ or (ii) $\bar{x} < \mu$ or (iii) $\bar{x} > \mu$

(iii) Level of Significance** (α)**: Set a level of significance

(iv) Test Statistic: The test statistic $t = \dfrac{\bar{x} - \bar{y}}{s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ *where* $s^2 = \dfrac{\sum (x_i - \bar{x})^2 + (y_i - \bar{y})^2}{n_1 + n_2 - 2}$

(v) Conclusion: (i) If $|t_{cal}| < t_{tab}$ we accept the Null Hypothesis H_0

(ii) If $|t_{cal}| > t_{tab}$ we reject the Null Hypothesis H_0 i.e., we accept the Alternative Hypothesis H_1

Problem 9: Two horses A and B were tested according to the time (in seconds) to run a particular track with the following results.

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:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of $t_{tab} =$

Calculated value of $|t_{cal}| =$

Calculated value of $|t_{cal}|$ Tabulated value of t_{tab}

Problem 10: To examine the hypothesis that the husbands are more intelligent than the wives, an investigator took a sample of 10 couples and admininistrated them a test which measures the I.Q. The results as follows:

Husbands	**117**	**105**	**97**	**105**	**123**	**109**	**86**	**78**	**103**	**107**
Wives	**106**	**98**	**87**	**104**	**116**	**95**	**90**	**69**	**108**	**85**

Test the hypothesis with a reasonable test at the level of significance of 0.05.
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of t_{tab} =

Calculated value of $|t_{cal}|$ =

Calculated value of $|t_{cal}|$ Tabulated value of t_{tab}

Problem 11: Ten soldiers participated in a shooting competition in the first week. After intensive training they participated in the competition in the second week. Their scores before and after training are given as follows:

Scores before	**67**	**24**	**57**	**55**	**63**	**54**	**56**	**68**	**33**	**43**
Scores after	**70**	**38**	**58**	**58**	**56**	**67**	**68**	**75**	**42**	**38**

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of t_{tab} =

Calculated value of $|t_{cal}|$ =

Calculated value of $|t_{cal}|$ Tabulated value of t_{tab}

Problem 12: Samples of two types of electric light bulbs were tested for length of life and following data were obtained

Type I	Type II
Sample number, n1 = 8 Sample mean, $\bar{x}$ = 1234 hrs Sample S.D., s_1=36 hrs	$n_2 = 7$ $\bar{y}$ =1036 hrs s_2 = 40 hrs

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

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of $t_{tab} =$

Calculated value of $\left|t_{cal}\right| =$

Calculated value of $\left|t_{cal}\right|$ Tabulated value of t_{tab}

Problem 13: The means of two random samples of sizes 9 and 7 are 196.42 and 198.82 respectively. The sum of the squares of the deviations from the mean are 26.94 and 18.73 respectively. Can the sample be considered to have been drawn from the same normal population.

Solution:

(i) Null Hypothesis $\left(H_0\right)$:

(ii)Alternative Hypothesis $\left(H_1\right)$:

(iii) Level of Significance $\left(\alpha\right)$:

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of $t_{tab} =$

Calculated value of $\left|t_{cal}\right| =$

Calculated value of $\left|t_{cal}\right|$ Tabulated value of t_{tab}

Model No 5.7: F-Test: Variances

This test is also called as variance ratio test. The objective of this test is to determine whether two independent estimates of the population variance differ significantly or whether the two samples may be regarded as drawn from the normal populations having the same variance, i.e., $\sigma_X^2 = \sigma_Y^2 = \sigma^2$. To carry out this test, we find the ratio F given by

$F = \frac{S_X^2}{S_Y^2}$ where $S_X^2 = \frac{\sum_{i=1}^{n_1}(x_i - \overline{x})^2}{n_1 - 1}$ and $S_Y^2 = \frac{\sum_{i=1}^{n_2}(y_j - \overline{y})^2}{n_2 - 1}$ and the test follows F-distribution with $\gamma_1 = n_1 - 1$ and $\gamma_2 = n_2 - 1$ degrees of freedom. It is to be noted that the numerator is greater than variance.

SNEDECOR'S F-TEST OF SIGNIFICANCE

(i) Null Hypothesis (H_0)**:** $\sigma_1^2 = \sigma_2^2 \ or \ s_1^2 = s_2^2$ i.e., the variances of the two populations are same.

(ii)Alternative Hypothesis (H_1)**:** $\sigma_1^2 \neq \sigma_2^2$

(iii) Level of Significance (α)**:** set a lelvel of significance

(iv)Test Statistic: The test statistic

$$F = \frac{larger\ variance}{smaller\ variance}, \quad where\ s_1^2 = \frac{\sum\left(x - \overline{x}\right)^2}{n_1 - 1},\ s_2^2 = \frac{\sum\left(y - \overline{y}\right)^2}{n_2 - 1}$$

(v) Conclusion: Degrees of freedom = $(n, m) = (n_1 - 1, n_2 - 1)$

(i) If Calculated value of F < Tabulated value of F, we accept H_0

(ii) If Calculated value of F > Tabulated value of F, we reject H_0

Problem 14: The time taken by the 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 from population from which these samples are drawn do not differ significantly?

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom =

Tabulated value of $F =$

Calculated value of $F =$

Calculated value of F Tabulated value of F

Problem 15: The measurements of the output of two units have given the following results. Assuming that both samples have been obtained from the normal population at 10% significant level, test whether the two populations have the same variance.

Unit-A	**14.1**	**10.1**	**14.7**	**13.7**	**14.0**
Unit-B	**14.0**	**14.5**	**13.7**	**12.7**	**14.1**

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom =

Tabulated value of $F =$

Calculated value of $F =$

Calculated value of F Tabulated value of F

Problem 16: In two independent samples of sizes 8 and 10 the sum of squares of deviations of the sample values from the respective means were 84.4 and 102.6. Test whether the difference of variances of the population is significant or not. Use a 0.05 level of significance.
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):
(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom =

Tabulated value of $F =$

Calculated value of $F =$
Calculated value of F Tabulated value of F

Problem 17: In one sample of 8 observations from a normal population, the sum of the squares of deviations of the sample values from the sample mean is 84.4 and in another sample of 10 observations it was 102.6 . Test at 5% level whether the populations have the same variance.

(or)

Problem 17: In two independent samples of sizes 8 and 10 the sum of squares of deviations of the samples values from the respective sample means were 84.4 and 102,6 . Test whether the difference of variances of the population is significant or not. Use a 0.05 level of significance.
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):
(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom =

Tabulated value of $F =$

Calculated value of $F =$
Calculated value of F Tabulated value of F

Problem 18: In one sample of 10 observations from a normal population, the sum of the squares of the deviations of the sample values from the sample mean is 102.4 and in another sample of 12 observations from another normal population, the sum of the squares of the deviations of the sample values from the sample mean is 120.5. Examine whether the two normal populations have the same variance.
Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):
(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom =

Tabulated value of $F =$

Calculated value of $F =$
Calculated value of F Tabulated value of F

Problem 19: Pumpkins were grown under two experimental conditions. Two random samples of 11 and 9 pumpkins, show the sample standard deviations of their weights as 0.8 and 0.5 respectively. Assuming that the weight distributions are normal, test hypothesis that the true variances are equal.

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom =

Tabulated value of $F =$

Calculated value of $F =$

Calculated value of F Tabulated value of F

Problem 20: The nicotine contents in milligrams of two samples of tobacco were found to be as follows. Test whether there is a significant difference between the two samples.

Sample A	**24**	**27**	**26**	**23**	**25**	**--**
Sample B	**29**	**30**	**30**	**31**	**24**	**36**

Solution: t-Test:

(i) Null Hypothesis (H_0):

(ii) Alternative Hypothesis (H_1):

(iii)Level of Significance (α)**:**

(iv)Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom=

Tabulated value of t_α =

Calculated value of $|t_\alpha|$ =

Calculated value of $|t_\alpha|$ Tabulated value of t_α

F-Test:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(ii) Level of Significance (α):

(iii) Test Statistic: The test statistic

(v) Conclusion: Degrees of freedom =

Tabulated value of F =

Calculated value of F =

Calculated value of F Tabulated value of F

Model No 5.8: CHI- SQUARE TEST (χ^2) FOR GOODNESS OF FIT

(i) Null Hypothesis (H_0): There is no significant difference between expected frequency and observed frequency

(ii)Alternative Hypothesis (H_1): There is a significant difference between expected frequency and observed frequency

(iii) Level of Significance (α): set a lelvel of significance

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

(v) Conclusion: Degrees of freedom = $n-1$

(i) If Calculated value of $\chi^2 <$ Tabulated value of χ^2, we accept H_0

(ii) If Calculated value of $\chi^2 >$ Tabulated value of χ^2, we reject H_0

Problem 21: A die is thrown 264 times with the following results. Show that the die is biased.

No. appeared on the die	1	2	3	4	5	6
Frequency	40	32	28	58	54	52

Solution: Given n=

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
40 32 28 58 54 52			
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} =$

(v) Conclusion: Degrees of freedom =

Calculated value of $\chi^2 =$

Tabulated value of $\chi^2 =$

Calculated value of χ^2 Tabulated value of χ^2

Problem 22: The following figures show the distribution of digits in numbers chosen at random from a telephone directory.

Digits	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Frequency	**1026**	**1107**	**997**	**966**	**1075**	**933**	**1107**	**972**	**964**	**853**

Test whether the digits may be taken to occur equally frequently in the directory.

Solution: Given n=

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(ii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
1026			
1107			
997			
966			
1075			
933			
1107			
972			
964			
853			
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} =$

(v) Conclusion: Degrees of freedom $= n - 1 =$

Calculated value of $\chi^2 =$

Tabulated value of $\chi^2 =$

Calculated value of χ^2 Tabulated value of χ^2

Problem 23: A sample analysis of examination results of 500 students was made. It was found that 220 students had failed, 170 had scored a third class, 90 were placed in second class and 20 got a first class. Do these figures commensurate with the general examination result which is in the ratio 4:3:2:1 for the various categories respectively.

Solution: Given n=

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} =$

(v) Conclusion: Degrees of freedom $= n-1 =$

Calculated value of $\chi^2 =$

Tabulated value of $\chi^2 =$

Calculated value of χ^2 Tabulated value of χ^2

Problem 24: A pair of dies are thrown 360 times and the frequency of each sum is indicated below:

$X = x_i$	2	3	4	5	6	7	8	9	10	11	12
Frequency	**8**	**24**	**35**	**37**	**44**	**65**	**51**	**42**	**26**	**14**	**14**

Would you say that the dice are fair on the basis of the chi-square test at 0.05 level of significance?

Solution: Given n=

$X = x_i$	2	3	4	5	6	7	8	9	10	11	12
$p(x_i)$											
Expected Frequencies $= 360\,p(x_i)$											

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
8 24 35 37 44 65 51 42 26 14 14			
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} =$

(v) Conclusion: Degrees of freedom $= n - 1 =$

Calculated value of $\chi^2 =$

Tabulated value of $\chi^2 =$

Calculated value of χ^2 Tabulated value of χ^2

Problem 25: 4 coins were tossed 160 times and the following results were obtained.

No. of Heads	0	1	2	3	4
Observed Frequency	17	52	54	31	6

Under the assumption that coins are balanced, find the expected frequencies of 0,1,2,3 or 4 heads, and test the goodness of fit at $\alpha = 0.05$

Solution: No. of coins =

Probability to get a head p = , q = 1 – p =

$X = x_i$	**0**	**1**	**2**	**3**	**4**
$p(x_i)$					
Expected Frequencies $= 160\,p(x_i)$					

Given n=

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
17 52 54 31 6			
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} =$

(v) Conclusion: Degrees of freedom $= n - 1 =$

Calculated value of $\chi^2 =$

Tabulated value of $\chi^2 =$

Calculated value of χ^2 Tabulated value of χ^2

Problem 26: A survey of 240 families with 4 children each revealed the following distribution.

Male Births	**4**	**3**	**2**	**1**	**0**
Observed Frequencies	**10**	**55**	**105**	**58**	**12**

Can we accept that the male and female births are equally distributed?

Solution: No. of families = , No. of children =

Probability to have a male birth p = , q = 1 – p =

$X = x_i$	**4**	**3**	**2**	**1**	**0**
$p(x_i)$					
Expected Frequencies $= 240\,p(x_i)$					

Given n=

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
10 55 105 58 12			
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} =$

(v) Conclusion: Degrees of freedom $= n - 1 =$

Calculated value of $\chi^2 =$

Tabulated value of $\chi^2 =$

Calculated value of χ^2 Tabulated value of χ^2

Problem 27: Fit a poisson distribution to the following data and for its goodness of fit at level of significance 0.05?

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

$X = x_i$					
$p(x_i)$					
Expected Frequencies =					

Given n=

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
10 55 105 58 12			
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} =$

(v) Conclusion: Degrees of freedom $= n - 2 =$

Calculated value of $\chi^2 =$

Tabulated value of $\chi^2 =$

Calculated value of χ^2 Tabulated value of χ^2

Model No 5.9: **CHI-SQUARE TEST FOR INDEPENDENT OF ATTRIBUTES**

Problem 28: **The following table gives the classification of 100 workers according to sex and nature of work. Test whether the nature of work is independent of the sex of the worker.**

	Stable	**Unstable**	**Totla**
Males	**40**	**20**	**60**
Females	**10**	**30**	**40**
Total	**50**	**50**	**100**

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

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

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
40 20 10 30			
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} =$

(v) Conclusion: Degrees of freedom $=(n-1)(m-1)=$

Calculated value of $\chi^2 =$

Tabulated value of $\chi^2 =$

Calculated value of χ^2 Tabulated value of χ^2

Problem 29: **Given the following contingency table for hair colour and eye colour. Find the value of χ^2**

Is there good association between the two?

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

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

$E(a)=\frac{(a+d+g)(a+b+c)}{N}=$	$E(b)=\frac{(b+e+h)(a+b+c)}{N}=$	$E(b)=\frac{(c+f+i)(a+b+c)}{N}=$
$E(c)=\frac{(a+d+g)(d+e+f)}{N}=$	$E(b)=\frac{(b+e+h)(d+e+f)}{N}=$	$E(b)=\frac{(c+f+i)(d+e+f)}{N}=$
$E(c)=\frac{(a+d+g)(g+h+i)}{N}=$	$E(b)=\frac{(b+e+h)(g+h+i)}{N}=$	$E(b)=\frac{(c+f+i)(g+h+i)}{N}=$

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
15 5 20 20 10 20 25 15 20			
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} =$

Conclusion: Degrees of freedom $=(n-1)(m-1)=$

Calculated value of $\chi^2 =$

Tabulated value of $\chi^2 =$

Calculated value of χ^2 Tabulated value of χ^2

Problem 30: **From the following data, find whether there is any significant liking in the habit of taking soft drinks among the categories of employees.**

Employees / Soft drinks	**Clerks**	**Teachers**	**Officers**
Pepsi	**10**	**25**	**65**
Thums Up	**15**	**30**	**65**
Fanta	**50**	**60**	**30**

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

$E(a) = \frac{(a+d+g)(a+b+c)}{N} =$	$E(b) = \frac{(b+e+h)(a+b+c)}{N} =$	$E(b) = \frac{(c+f+i)(a+b+c)}{N} =$
$E(c) = \frac{(a+d+g)(d+e+f)}{N} =$	$E(b) = \frac{(b+e+h)(d+e+f)}{N} =$	$E(b) = \frac{(c+f+i)(d+e+f)}{N} =$
$E(c) = \frac{(a+d+g)(g+h+i)}{N} =$	$E(b) = \frac{(b+e+h)(g+h+i)}{N} =$	$E(b) = \frac{(c+f+i)(g+h+i)}{N} =$

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
10 25 65 15 30 65 50 60 30			
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} =$

(v)

Conclusion: Degrees of freedom $= (n-1)(m-1) =$

Calculated value of $\chi^2 =$

Tabulated value of $\chi^2 =$

Calculated value of χ^2 Tabulated value of χ^2

Problem 31: **1000 students at college level were graded according to their I.Q. and the economic conditions of their home. Use χ^2 test to find out whether there is any association between condition at home and I.Q. Use 0.05 L.O.S.**

Economic Condition \ I.Q.	**High**	**Low**	**Total**
Rich	**460**	**140**	**600**
Poor	**240**	**160**	**400**
Total	**700**	**300**	**1000**

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

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

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
460			
140			
240			
160			
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} =$

(v) conclusion: Degrees of freedom $= (n-1)(m-1) =$

Calculated value of $\chi^2 =$

Tabulated value of $\chi^2 =$

Calculated value of χ^2 Tabulated value of χ^2

Part C:

Analysis of Variance (Anova)

Model No 5.10: One-way Anova

Model No 5.11: Two-way Anova

Analysis of Variance (ANOVA)

A test for Homogeneity of Mean.

①

The technique of "Analysis of variance is referred to as ANOVA. The technique of ANOVA is to split the variation into its various components. They are (i) Variance between samples.
(ii) Variance within samples.

The observations (or data) may be classified according to one factor or two factors. which are called one-way classification and two-way classification.

One-way ANOVA In This if we consider the influence of any one factor, then it is called. one-way classification.

Eg: The yields of several plots of land may be classified according to one or more types of fertilizati

The techniques for ANOVA one-way classification model are:

(i) Direct Method
(ii) Short-Cut Method
(iii) Coding method.

(i) Direct method

$H_0: \mu_1 = \mu_2 = \cdots = \mu_k$, where $\mu_1, \mu_2, \cdots, \mu_k$ are the arithmetic means of the k populations from which k samples are drawn at random.

$H_1: \mu_1 \neq \mu_2 \neq \cdots \neq \mu_k$.

a) Calculation of variance between the Samples

It is the sum of the squares of the deviations of the means of the various samples from the grand mean.

(i) Calculate the sample means $\bar{X}_1, \bar{X}_2, \bar{X}_3, \cdots \bar{X}_k$ of all ~~the~~ k samples.

(ii) Calculate the mean of the sample means:

$$\bar{X} = \frac{\bar{X}_1 + \bar{X}_2 + \cdots + \bar{X}_k}{k} \quad \text{or} \quad \boxed{\bar{X} = \frac{T}{N}}$$

(iii) Evaluate the deviations of the sample means from the grand mean i,e find

$\bar{X}_1 - \bar{X}, \ \bar{X}_2 - \bar{X}, \cdots, \ \bar{X}_k - \bar{X}$.

(IV) SSB (or SSC) = Sum of the squares of the variations between the samples or between the columns $= \sum_{i=1}^{k} n_i (\bar{X}_i - \bar{X})^2$.

MSB or MSC = Variance or the mean square between the samples (or between the columns)

$$= \frac{SSB}{\nu_1}, \quad \nu_1 = \text{degrees of freedom}$$

$= \text{no: of samples} - 1$

$= k - 1$

b) Calculation of variance within the samples.

SSW (or SSE) = Sum of the squares of the variations within the samples

(or)

Sum of the squares due to errors.

$$= \Sigma (x_1 - \bar{x})^2 + \Sigma (x_2 - \bar{x})^2 + \cdots + \Sigma (x_k - \bar{x}_k)^2$$

MSW or MSE = Variance or mean square within the samples.

$$= \frac{SSW}{\nu_2}$$

$\nu_2 = d.f = \text{total no: of observations} - \text{No: of Samples}$

$= N - K$.

c) $F = \frac{MSB}{MSW}$ or $\frac{MSC}{MSE} = \frac{\text{Variance between the Samples}}{\text{Variance Within the Samples}}$

D.f $= \nu_1 = K - 1, \quad \nu_2 = N - K$.

Annova table (One-way classification)

Source of Variation	Sum of Squares SS	Degrees of freedom	Mean Squares MS	Test Statistic (F-test)
Between Samples or columns.	SSB	K-1	$MSB = \frac{SSB}{K-1}$	
Within Samples (Errors)	SSW	N-K	$MSW = \frac{SSW}{N-K}$	$F = \frac{MSB}{MSW}$
Total	SST	N-1	—	—

Short-Cut Method

1. Caluclate $T = \Sigma X_1 + \Sigma X_2 + \cdots + \Sigma X_k$.
2. Calculate $\frac{T^2}{N}$.
3. Compute SST = Total sum of the squares of deviatn
$$= \Sigma X_1^2 + \Sigma X_2^2 + \cdots + \Sigma X_k^2 - \frac{T^2}{N}$$
4. Calculate $SSB = \left[\frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \cdots + \frac{(\Sigma X_k)^2}{n_k}\right] - \frac{T^2}{N}$
5. Calculate $SSW = SST - SSB$.
6. Now proceed as in Direct Method to obtain MSB, MSW and F and arrive at the final decision.

Annova for two-way classification
(Manifold Calssification)

In two-way classification, observations are classified according to two different factors or criteria.

Eg: Fertilizers may be tried on different soil textures.

Working Rule

1. Calculate SSC i,e the sum of squares (or variance) between the columns.
2. SSR = Sum of squares (or Variance) between the rows
3. SSE = the sum of squares for the residuals
4) SST = SSC + SSR + SSE

C – no: of Columns, R – no: of rows.

$\therefore$ total no: of degrees of freedom = $cr - 1$

D.f between columns = $c - 1$

" " rows = $r - 1$

D.f between residuals $= (cr-1)-(c-1)-(r-1)$

$= (c-1)(r-1)$

5) Calculate $MSC = \frac{SSC}{c-1}$

$$MSR = \frac{SSR}{r-1}$$

$$MSE = \frac{SSE}{(c-1)(r-1)}$$

6) Calc. $F_C = \frac{MSC}{MSE}$, $MSC > MSE$

$$F_R = \frac{MSR}{MSE} \quad MSR > MSE$$

if $MSE > MSC$, $F_C = \frac{MSE}{MSC}$

Illy if $MSE > MSR$, $F_R = \frac{MSE}{MSR}$.

7. Write conclusions for F_C and F_R :

Analysis Of Variance (ANOVA) — UNIT-5 Part-C, 26.05.2022

1 way Classification:
1. Short-cut Method
2. Direct Method
3. Coding Method

only 1 Parameter is Calculated.

Hint: ANOVA is More than 2 Samples

1) Short cut Method:

* K = No of Samples
* N = Total No of Observations
* $T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3$
* Correction Factor $[CF] = \frac{T^2}{N}$
* Sum of Squares Between Samples [SSB]:

$$SSB = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} - \frac{T^2}{N}$$

* Total Sum of Squares of Samples [SST]:

$$SST = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 - \frac{T^2}{N}$$

* $SSW = SST - SSB$
* $MSB = \frac{SSB}{\nu_1} = \frac{SSB}{K-1}$ * $MSW = \frac{SSW}{\nu_2} = \frac{SSW}{N-K}$
* $F_{cal} = \frac{MSB}{MSW}$ or $\frac{MSW}{MSB}$ * $F_{tab} = F_{0.05}\underbrace{(K-1, N-K)}_{\text{degrees of Freedom}}$

$\mu_1 = \mu_2 = \mu_3$

2) Direct Method

$$\begin{array}{ccc} x_1 & x_2 & x_3 \\ | & | & | \\ \bar{x}_1 & \bar{x}_2 & \bar{x}_3 \\ \searrow & \downarrow & \swarrow \\ & \bar{x} & \end{array}$$

$$SSB = n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + n_3(\bar{x}_3 - \bar{x})^2$$

$$SSW = \Sigma(x_1 - \bar{x}_1)^2 + \Sigma(x_2 - \bar{x}_2)^2 + \Sigma(x_3 - \bar{x}_3)^2$$

17. Three different Machines are used for Production. On Basis Of their Output the Machines are Equally effective?

MACHINE-I	MACHINE-II	MACHINE-III
10	9	20
5	5	16
11	7	10
10	6	4

Sol: It belongs to the Category Of one way ANOVA; Since It Contains More than 2 Samples & we Studied only 1 Parameter

* Null Hypothesis (H_0): $\mu_1 = \mu_2 = \mu_3$ (or)
  The three Machines are Equally effective (or) Homogenity Of Means.
* Alternative Hypothesis (H_1): $\mu_1 \neq \mu_2 \neq \mu_3$ (or)
  The Three Machines are Not Equally Working (or) Non-Homogenity Of Means.
* Level of Significance (α): 0.05
* Test Statistic:

1, SHORT-CUT METHOD: K = No. of samples = 3
N = Total No. Of observations = 12

MACHINE-I x_1	MACHINE-II x_2	MACHINE-III x_3	x_1^2	x_2^2	x_3^2
10	9	20	100	81	400
5	5	16	25	25	256
11	7	10	121	49	100
10	6	4	100	36	16
$\Sigma x_1 = 36$	$\Sigma x_2 = 27$	$\Sigma x_3 = 50$	$\Sigma x_1^2 = 346$	$\Sigma x_2^2 = 191$	$\Sigma x_3^2 = 7$

$T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3 = 36 + 27 + 50 = 113$

$CF = \frac{T^2}{N} = \frac{113}{12} = 1064.0833$

$SSB = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} - \frac{T^2}{N} = \frac{1296}{4} + \frac{729}{4} + \frac{2500}{4} - 1064.0833$

$= \frac{\cancel{5729}\ 4525}{4} - 1064.0833$

$= 1131.25 - 1064.0833$

$= 67.1667$

$$SST = \Sigma (x_1)^2 + \Sigma x_2^2 + \Sigma x_3^2 - \frac{T^2}{N}$$
$$= 346 + 191 + 772 - 1064.0833 =$$
$$= 1309 - 1064.0833 = 244.9167$$

$$SSW = SST - SSB = 244.9167 - 67.1667 = 177.75$$

$$MSB = \frac{SSB}{\vartheta_1} = \frac{SSB}{K-1} = \frac{67.1667}{2} = 33.5833$$

$$MSW = \frac{SSW}{\vartheta_2} = \frac{SSW}{N-K} = \frac{177.75}{12-3} = 19.75$$

$$F_{cal} = \frac{MSB}{MSW} = \frac{33.5833}{19.75} = \boxed{1.7004}$$

$$F_{tab} = F_{0.05}(K-1, N-K) = F_{0.05}(2,9) = \boxed{4.26}$$

$\therefore$ $F_{cal} < F_{tab}$ **Null Hypothesis is Accepted**

Hence, the 3 Machines are Equally Effective.

(or) II METHOD

DIRECT METHOD:

x_1	x_2	x_3	$(x_1-\bar{x}_1)^2$	$(x_2-\bar{x}_2)^2$	$(x_3-\bar{x}_3)^2$
10	9	20	1	5.0625	56.25
5	5	16	16	3.0625	12.25
11	7	10	4	0.0625	6.25
10	6	4	1	0.5625	72.25
36	27	50	22	8.75	147

$$\bar{x}_1 = \frac{\Sigma x_1}{n_1} = \frac{36}{4} = 9$$

$$\bar{x}_2 = \frac{\Sigma x_2}{n_2} = \frac{27}{4} = 6.75$$

$$\bar{x}_3 = \frac{\Sigma x_3}{n_3} = \frac{50}{4} = 12.5$$

Grand Mean $\bar{x} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3}{3}$

$$= \frac{9 + 6.75 + 12.5}{3} = 9.4167$$

$$SSB = n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + n_3(\bar{x}_3 - \bar{x})^2$$
$$= 4(9 - 9.4167)^2 + 4(6.75 - 9.4167)^2 + 4(12.5 - 9.4167)^2$$
$$= 0.6946 + 28.4452 + 38.0270 = 67.1668$$

$$SSW = \Sigma(x_1 - \bar{x}_1)^2 + \Sigma(x_2 - \bar{x}_2)^2 + \Sigma(x_3 - \bar{x}_3)^2$$
$$= [22] + [8.75] + [147] = 177.75$$

$$MSB = \frac{SSB}{K-1} = \frac{67.1667}{2} = 33.583$$

$$MSW = \frac{SSW}{N-K} = \frac{177.75}{9} = 19.75$$

ii) $F_{cal} = \frac{MSB}{MSW} = \frac{33.583}{19.75} = 1.7004$ $\boxed{F_{cal} < F_{tab}}$

$F_{tab} = F_{0.05}(2,9) = 4.26$

(Or)

III CODING METHOD

In this Method: We Have to ADD (or) SUBRACT (or) DIVIDE (or) MULTIPLY with a Constant Value with each of the Observat[ions]

In this Problem, we Subract 10 from each Observed Value, As '10' is repeated many times in Given Table.

M-I	M-II	M-III
0	-1	10
-5	-5	6
1	-3	0
0	-4	-6

Remaining Procedure can be done using Short-Cut Method Only

Q) Three Samples of five, five, four Motor Car tyres are drawn respectively from 3 Branches 'A', 'B', 'C', Manufact[ured] by 3 Machines. The Lifetime of 3 tyres in 1000 Miles is Given below.

Test whether Average lifetime of 3 Brands tyres are Equal or Not.

A	B	C
35	30	28
40	35	24
33	34	30
36	28	26
31	33	---

Sol $n_1 = 5$ $n_2 = 5$ $n_3 = 4$ $k = 3$ $N = 14$

SC Method

A x_1	B x_2	C x_3	x_1^2	x_2^2	x_3^2
35	30	28	1225	900	784
40	35	24	1600	1225	576
33	34	30	1089	1156	900
36	28	26	1296	784	676
31	33		961	1089	
175	160	108	6171	5154	2936

Null Hypothesis (H_0): $\mu_1 = \mu_2 = \mu_3$

Alternative Hypothesis (H_1): $\mu_1 \neq \mu_2 \neq \mu_3$

Level of Significance $\alpha = 0.05$

$T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3 = 175+160+108 = 443$, $CF = \frac{T^2}{N} = \frac{(443)^2}{14} = 14017.78$

$$SSB = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} - \frac{T^2}{N}$$

$$= \frac{30625}{5} + \frac{25600}{5} + \frac{11664}{4} - 14017.78$$

$$= 6125 + 5120 + 2916 - 14017.78 = 143.22$$

$$SST = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 - \left(\frac{T^2}{N}\right) = 6171 + 5154 + 2936 - 14017.78$$

$$= 243.22$$

$$SSW = SST - SSB = 243.22 - 143.22 = 100$$

$$MSB = \frac{SSB}{\gamma_1} = \frac{SSB}{K-1} = \frac{SSB}{2} = \frac{143.22}{2} = 71.61$$

$$MSW = \frac{SSW}{\gamma_2} = \frac{SSW}{N-K} = \frac{100}{14-3} = 9.0909$$

$$F_{cal} = \frac{MSB}{MSW} = \frac{71.61}{9.0909} = \boxed{7.8771}$$

$$F_{tab} = F_{0.05}(K-1, N-K) = F_{0.05}(\overset{\gamma_1}{2}, \overset{\gamma_2}{11}) = \boxed{3.98}$$

$\boxed{\therefore F_{cal} > F_{tab}}$

Null Hypothesis is Rejected

DIRECT-METHOD

x_1	x_2	x_3	$(x_1-\bar{x}_1)^2$	$(x_2-\bar{x}_2)^2$	$(x_3-\bar{x}_3)^2$
35	30	28	0	4	1
40	35	24	25	9	9
33	34	30	4	4	9
36	28	26	1	16	1
31	33	---	16	1	--
175	160	108	46	34	20

$$\bar{x}_1 = \frac{\Sigma x_1}{n_1} = \frac{175}{5} = 35$$

$$\bar{x}_2 = \frac{\Sigma x_2}{n_2} = \frac{160}{5} = 32$$

$$\bar{x}_3 = \frac{\Sigma x_3}{n_3} = \frac{108}{4} = 27$$

$$\bar{x} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3}{3} = \frac{35+32+27}{3} = \frac{94}{3} = 31.3333$$

$$SSB = n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + n_3(\bar{x}_3 - \bar{x})^2$$

$$= 5(35-31.3333)^2 + 5(32-31.3333)^2 + 4(27-31.3333)^2$$

$$= 144.5558$$

$$SSW = \Sigma(x_1-\bar{x}_1)^2 + \Sigma(x_2-\bar{x}_2)^2 + \Sigma(x_3-\bar{x}_3)^2$$

$$= 46 + 34 + 20 = 100$$

$$MSB = \frac{SSB}{\gamma_1} = \frac{SSB}{K-1} = \frac{144.5558}{2} = 72.2779$$

$$MSW = \frac{SSW}{\gamma_2} = \frac{SSW}{N-K} = \frac{100}{14-3} = 9.0909$$

$$F_{cal} = \frac{MSB}{MSW} = \boxed{7.9506}$$

$$F_{tab} = F_{0.05}(2, 11) = \boxed{3.98}$$

$\boxed{F_{cal} > F_{tab}}$

Null Hypothesis is Rejected.

ANOVA Two Way Classification:

N = No. of rows x No. of columns = $(r \times c)$

$T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3$

$CF = \frac{T^2}{N}$

$SSC = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} - \frac{T^2}{N}$

$SSR = \frac{(\Sigma T_1)^2}{n_1} + \frac{(\Sigma T_2)^2}{n_2} + \frac{(\Sigma T_3)^2}{n_3} - \frac{T^2}{N}$

$SST = \sum_i \sum_j x_{ij}^2 - \frac{T^2}{N}$

Error

$\boxed{SSE = SST - (SSC + SSR)}$

$MSC = \frac{SSC}{C-1}$

$MSR = \frac{SSR}{r-1}$

$MSE = \frac{SSE}{(C-1)(r-1)}$

$F_{cal} = \frac{MSC}{MSE}$ or $\frac{MSE}{MSC}$ $\underbrace{(C-1, (C-1)(r-1))}_{\text{degrees of Freedom}}$

$F_{cal} = \frac{MSR}{MSE}$ or $\frac{MSE}{MSR}$ $\underbrace{((r-1), (C-1)(r-1))}_{\text{degrees of Freedom}}$

1, A Farmer Applies 3 types of Fertilizers on four Separate plots. The Figure on yield per acre are tabulated below:

Find out if the Plots are Materially different in fertility also if 3 fertilizers make any Material difference in yields.

Yields:

Plots & Fertilizers	A	B	C	D
Nitrogen F_1	6	4	8	6
Potassium F_2	7	6	6	9
Phosphorus F_3	8	5	10	9

Sol: Here, we Study 2 Parameters: Plots A, B, C, D and Fertilizers F_1, F_2, F_3. So Here we have to apply 2 way classification of ANOVA.

Null Hypothesis (H_0): $A = B = C = D$, $F_1 = F_2 = F_3$

The effect of 3 Fertilizers are Same.

Alternative Hypothesis (H_1): $A \neq B \neq C \neq D$ & $F_1 \neq F_2 \neq F_3$

Level of Significance $\alpha = 0.05$

Test Statistics:

r = No. of rows = 3

C = No. of Columns = 4

Plots / Fertilizers	A (x_1)	B (x_2)	C (x_3)	D (x_4)	
Nitrogen (F_1) T_1	6	4	8	6	$\Sigma T_1 = 24$
Potassium (F_2) T_2	7	6	6	9	$\Sigma T_2 = 28$
Phosphorus (F_3) T_3	8	5	10	9	$\Sigma T_3 = 32$
	$\Sigma x_1 = 21$	$\Sigma x_2 = 15$	$\Sigma x_3 = 24$	$\Sigma x_4 = 24$	

(Yield)

$N = rC = 3 \times 4 = 12$

$T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3 = 21 + 15 + 24 + 24 = 84$

$CF = \frac{T^2}{N} = \frac{(84)^2}{12} = 588$

$SSC = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} + \frac{(\Sigma x_4)^2}{n_4} - \frac{T^2}{N} = \frac{(21)^2}{3} + \frac{(15)^2}{3} + \frac{(24)^2}{3} + \frac{(24)^2}{3} - \frac{588}{\cancel{12}} = 18$

$SSR = \frac{(\Sigma T_1)^2}{n_1} + \frac{(\Sigma T_2)^2}{n_2} + \frac{(\Sigma T_3)^2}{n_3} + \cancel{\frac{(\Sigma T_4)}{n_4}} - \frac{T^2}{N}$

$= \frac{(24)^2}{4} + \frac{(28)^2}{4} + \frac{(32)^2}{4} - 588 = 8$

$SST = \sum_i \sum_j x_{ij}^2 - \frac{T^2}{N} = 6^2 + 4^2 + 8^2 + 6^2 + 7^2 + 6^2 + 6^2 + 9^2 + 8^2 + 5^2 + 10^2 + 9^2 - 588$

$= 36$

Error $SSE = SST - (SSC + SSR) = 36 - (18 + 8) = 10$

$MSC = \frac{SSC}{C-1} = \frac{18}{3} = 6$

$MSR = \frac{SSR}{r-1} = \frac{8}{2} = 4$

$MSE = \frac{SSE}{(r-1)(C-1)} = \frac{10}{6} = 1.6666$

$F_C\,cal = \frac{MSC}{MSE} = \frac{6}{1.6666} = 3.6001$

$F_R\,cal = \frac{MSR}{MSE} = \frac{4}{1.6666} = 2.40009$

$F_C\ tab = F_C((c-1),(r-1)(c-1)) = F_{0.05}(3,6) = 4.76$

$F_R\ tab = F_R((r-1),(r-1)(c-1)) = F_{0.05}(2,6) = 5.14$

I $\boxed{\therefore F_C cal < F_C tab}$ $A = B = C = D$

All plots are equally effective.

II $\boxed{\therefore F_R cal < F_R tab}$ $F_1 = F_2 = F_3$

All Fertilizers are Equally Effective

2) To Study the Performance of 3 Detergents of 3 different temperatures, the following whiteners were observed. Perform a 2 way Analysis of Variance using 5% Level of Significance.

Water temperature	Detergent A	Detergent B	Detergent C
Cold water	57	55	67
Warm water	49	52	68
Hot water	54	46	58

Sol: Null Hypothesis: $D_A = D_B = D_C$

Cold water = Warm water = hot water

Alternative Hypothesis: $D_A \neq D_B \neq D_C$

Level of Significance :- $\alpha = 0.05$

Test Statistic: No of rows $(r) = 3$ $N = rc = 9$

No of columns $(c) = 3$

Water/ Temperature		Detergent-A (x_1)	Detergent B (x_2)	Det C (x_3)	Total
Cold water	T_1	57	55	67	$\Sigma T_1 = 179$
Warm water	T_2	49	52	68	$\Sigma T_2 = 169$
Hot water	T_3	54	46	58	$\Sigma T_3 = 158$
		$\Sigma x_1 = 160$	$\Sigma x_2 = 153$	$\Sigma x_3 = 193$	

By ANOVA 2 way Method:

$T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3 = 160 + 153 + 193 = 506$

$CF = \frac{T^2}{N} = \frac{(506)^2}{9} = 28448.4$

$SSC = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} - \frac{T^2}{N}$

$= \frac{(160)^2}{3} + \frac{(153)^2}{3} + \frac{(193)^2}{3} - 28448.4 = 304.2222$

$SSR = \frac{(\Sigma T_1)^2}{n_1} + \frac{(\Sigma T_2)^2}{n_2} + \frac{(\Sigma T_3)^2}{n_3} - \frac{T^2}{N}$

$= \frac{(179)^2}{3} + \frac{(169)^2}{3} + \frac{(158)^2}{3} - 28448.4 = 73.6$

$SST = \sum_i \sum_j x_{ij}^2 - \frac{T^2}{N} = 57^2 + 49^2 + 54^2 + 55^2 + 52^2 + 46^2 + 67^2 + 68^2 + 58^2 - 28448.4$

$= 439.6$

Error SSE:

$SSE = [SST - (SSC + SSR)] = \overset{439.6}{\cancel{28888}} - (304.2666 + 73.6) = 61.7782$

$MSC = \frac{SSC}{C-1} = \frac{304.2222}{3-1} = 152.1111$

$MSR = \frac{SSR}{R-1} = \frac{73.6}{3-1} = 36.7778$

$MSE = \frac{SSE}{(C-1)(r-1)} = \frac{\overset{61.7782}{\cancel{439.6}}}{(2)(2)} = 15.4445$

$F_C Cal = \frac{MSC}{MSE} = \frac{152.3333}{109.9} = 9.8438$

$F_R Cal = \frac{MSE}{MSR} = \frac{109.9}{36.8} = 2.98641$

$F_C tab = ((C-1), (r-1)(C-1))$

$= F_{0.05}(2, 4)$

$= 6.94$

$F_R tab = ((r-1), (C-1)(r-1))$

$= F_{0.05}(2, 4)$

$= 6.94$

$\boxed{F_C Cal > F_C tab}$ → Null Hypothesis is Rejected

The detergents are Not Equally effective. $D_A \neq D_B \neq D_C$

$\boxed{F_R Cal < F_R tab}$ → Null Hypothesis is Accepted

The Temperature of water is Equally Effective.

Coding Method:

In this Coding Method, we multiply or divide or Subtraction or Addition Of any value for Smaller Values.

Water/ Temperature	Det A	Det B	Det C
Cold water	5	3	15
Warm water	-3	0	16
Hot water	2	-6	6

SHORT-CUT Method

	A x_1	B x_2	C x_3	x_1^2	x_2^2	x_3^2	
T_1	5	3	15	25	9	225	$\Sigma T_1 = 23$
T_2	-3	0	16	9	0	256	$\Sigma T_2 = 13$
T_3	2	-6	6	4	36	36	$\Sigma T_3 = 2$
	$\Sigma x_1 = 4$	$\Sigma x_2 = -3$	$\Sigma x_3 = 37$	38	45	517	

Null Hypothesis: $D_A = D_B = D_C$

Cold water = Warm water = Hot water

All types of water has Equal effect on all Detergents.

Alternative Hypothesis: $D_A \neq D_B \neq D_C$

Level of Significance: $\alpha = 0.05$

Test Statistics: No. of rows $(r) = 3$ $\quad N = r \times c = 9$

No. of columns $(C) = 3$

$$T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3 = 4 - 3 + 37 = 38$$

$$CF = \frac{T^2}{N} = \frac{38^2}{9} = 160.4444$$

$$SSC = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} - \frac{T^2}{N} = \frac{16}{3} + \frac{9}{3} + \frac{1369}{3} - 160.4444 = 304.2222$$

$$SSR = \frac{(\Sigma T_1)^2}{n_1} + \frac{(\Sigma (T_2))^2}{n_2} + \frac{(\Sigma T_3)^2}{n_3} - \frac{T^2}{N} = \frac{(23)^2}{3} + \frac{13^2}{3} + \frac{4}{3} - 160.4444 = 73.5556$$

$$SST = \sum_i \sum_j x_{ij}^2 - \frac{T^2}{N} = 25 + 9 + 225 + 9 + 0 + 256 + 4 + 36 + 36 - 160.4444 = 439.5556$$

$$SSE = SST - (SSC + SSR) = 439.5556 - (304.2222 + 73.5556)$$
$$= 61.7778$$

$$MSC = \frac{SSC}{c-1} = \frac{304.2222}{2} = 152.1111$$

$$MSR = \frac{SSR}{r-1} = \frac{73.5556}{2} = 36.7778$$

$$MSE = \frac{SSE}{(c-1)(r-1)} = \frac{61.7778}{4} = 15.4444$$

$$F_C Cal = \frac{MSC}{MSE} = \frac{152.1111}{15.4444} = 9.8489 \qquad F_C tab = F_{0.05}(2,4) = 6.94$$

$$F_R Cal = \frac{MSR}{MSE} = \frac{36.7778}{15.4444} = 2.3813 \qquad F_R tab = F_{0.05}(2,4) = 6.94$$

$\boxed{\therefore F_C Cal > F_C tab} \rightarrow$ Null Hypothesis is Rejected

The Detergents are Not Equally Effective. $\boxed{D_A \neq D_B \neq D_C}$

$\boxed{\therefore F_R Cal > F_R tab} \rightarrow$ Null Hypothesis is Accepted.

The Temperature Of water is Equally Effective