Data Science and Artificial Intelligence

Probability and Statistics

Testing of Hypothesis

Lecture No.- 01

Topics to be Covered

χ^2 - Test Question

full marks

Topic Question Based on chi Square Test

CHI-Square TEST

10 question set

Chi square Testing

CHI-Square Test :-

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi^2 = \sum_{i=1}^{n} \frac{O_i^2}{E_i} - N$$

degree of Freedom = $(n-1)$

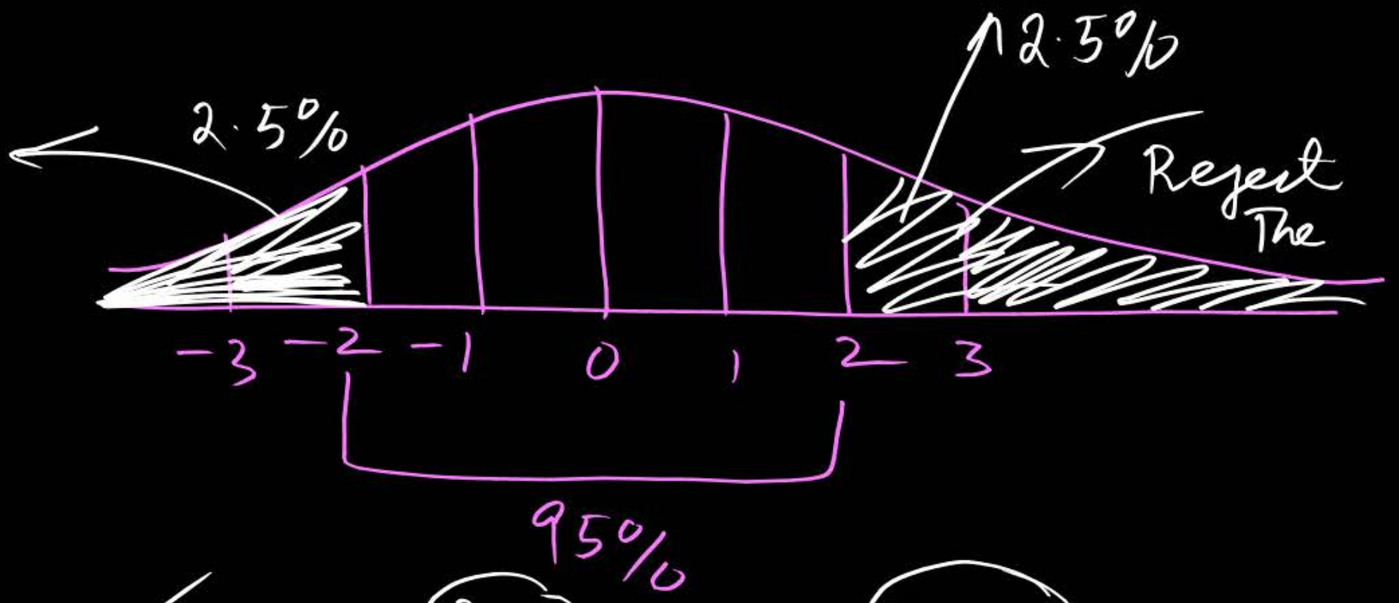
Confidence Interval = belief of hypothesis

Significance level = 1 − confidence interval

Accept / reject

$$\chi^2 = \text{Number} = \sum_{i=1}^{n} \frac{O_i^2}{E_i} = \text{Number}$$

→ degree of freedom



Null (Accepted / rejected) { ✓ Accept/reject — ✓ Accept / — Reject

You: Chi square ↓ degree of freedom ↑ / Chi square ↑ degree of freedom ↓

Table

Topic: Test of hypothesis

0·61 = mean

Q1. Fit a Poisson distribution to the following data and best the goodness of fit:

x:	0	1	2	3	4
f:	109	65	22	3	1

x	f
0	109
1	65
2	22
3	3
4	1

$$\text{mean} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum f_i} = \frac{0\times109+1\times65+2\times22+3\times3+4\times1}{109+65+22+3+1}$$

✓ Mean $\mu = 0.61$

Total frequency = 200

This is a Poisson Distributed

$$P(X=x) = \frac{e^{-\mu}(\mu)^x}{x!} \times 200$$

$\mu = 0.61$

$N = 200$

→ No. of expected $N(E) = \frac{e^{-\mu}(\mu)^x}{x!} \times 200$

$x = 0$

$x = 1$

$x = 2$

$x = 3$

$x = 4$

$= \frac{e^{-0.61}(0.61)^0}{0!} \times 200 = 109$

$= \frac{e^{-0.61}(0.61)^1}{1!} \times 200 = 66$

$= \frac{e^{-0.61}(0.61)^2}{2!} \times 200 = 20$

$= \frac{e^{-0.61}(0.61)^3}{3!} (200) = 4$

$= \frac{e^{-0.61}(0.61)^4}{4!} (200) = 1$

O_i	E_i	$\frac{O_i - E_i}{E_i}$
109	109	
65	66	
22	20	
3	4	
1	1	

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi^2 = 0.4651$$

→ def →

Accepted The Hypothesis

Degree of Freedom = 4

Topic: Test of hypothesis

Accident – Uniformly distributed ✓ Accept The hypothesis

Q2. The following table gives the number of accidents that took place in an industry during various days of the week. Test if accidents are uniformly distributed over the week.

degree of Freedom = 6 − 1 = (5)

$14+18+12+11+15+14 = \frac{84}{6} = 14$

$14\times\frac{1}{6}+18\times\frac{1}{6}+12\times\frac{1}{6}+11\times\frac{1}{6}+15\times\frac{1}{6}+14\times\frac{1}{6}$

Day	Mon	Tue	Wed	Thurs	Fri	Sat
No. of accidents	14 ✓	18 ✓	12 ✓	11 ✓	15 ✓	14 ✓

Day	Observed	Expected value	$\frac{(O-E)^2}{E}$
mon	14	14	
Tue	18	14	
wed	12	14	
Thurs	11	14	
Fri	15	14	
Sat	14	14	

$\chi^2 = \sum_{i=1}^{6} \frac{(O_i-E_i)^2}{E_i}$

$\chi^2 = 2.14$

Df n = 5 → 11.070

Df = 5 (0.5)

Topic: Test of hypothesis

Rejected The Hypothesis

Q3. A die is thrown 276 times and the results of these throws are given below:

No. appeared on the die	1	2	3	4	5	6
Frequency	40	32	29	59	57	59
	O_1	O_2	O_3	O_4	O_5	O_6

Test whether the die is biased or not.

Die is Biased or not

1	2	3	4	5	6
40	32	29	59	57	59
46	46	46	46	46	46

Reject The Hypothesis

$\frac{276}{6} = 46$

$\chi^2 = \frac{\Sigma(O-E)^2}{E} = 21.30$

$df = 6 - 1 = 5$

No : Die is Unbiased

No reject is rejected

Die is biased

$\chi^2_{0.5} = 11.(\quad)$ (5)

Topic: Test of hypothesis

$N(x) = {}^{n}C_x p^x q^{n-x}$

$N(0) = 800 \times {}^{4}C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^{4-0}$

E_0

Q4. Example 4. Records taken of the number of male and female births in 800 families having four children are as follows:

Rejected

No. of male births	0	1	2	3	4
No. of female birth	4	3	2	1	0
No. of families	32	178	290	236	64

O_1 O_2 O_3 O_4 O_5

Test whether the data are consistent with the hypothesis that the Binomial law holds and the chance of male birth is equal to that of female birth, namely p = q = 1/2.

E_3 $N(3) = 800 \times {}^{4}C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^{4-3}$ $N(1) = 800 \times {}^{4}C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^{4-1}$

E_4 $N(4) = 800 \times {}^{4}C_4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^{4-4}$ $N(2) = 800 \times {}^{4}C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{4-2}$

E_1 E_2

	O	E	$(O-E)^2/E$
✓	32	50	
✓	178	200	
✓	290	300	
✓	236	200	
✓	64	50	

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$\chi^2 = 19.4$

Degree of Freedom

$= 5 - 1$

$= (4)$

↓ 0.5 Significance level

= Rejected H_0

Topic: Test of hypothesis

Q5. The theory predicts the proportion of beans in the four groups, G_1, G_2, G_3, G_4, should be in the ratio 9:3:3:1. In an experiment with 1600 beans the numbers in the four groups were 882, 313, 287 and 118. Does the experimental result support the theory?

900/300/300/100

G_1, G_2, G_3, G_4

1600 Beans

9 3 3 1

Observed / Expected

O	E	$\frac{(O-E)^2}{E}$
882	900	
313	300	
287	300	
118	100	9:3:3:1

$$\chi^2 = \frac{\sum (O-E)^2}{E} = 4.72$$

$df = (4-1) = 3$

Accepted the Hypothesis

Topic: Test of hypothesis

$X^2 = 19.5$

$X^2 = 9.488$

Critical values of chi-square (right tail), Significance level (α)

Degrees of freedom (*df*)	.99	.975	.95	.9	.1	.05	.025	.01
1	-------	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217

significance level

degree of freedom

Topic: Test of hypothesis

13	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932
22	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642

Topic: Test of hypothesis

27	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892
40	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691
50	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154
60	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379
70	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425
80	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329
100	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116
1000	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807

THANK - YOU