ENGINEERING MATHEMATICS

UNIT-5 : *Statistical Techniques – III*

(Testing a Hypothesis and Statistical Quality Control)

Lec-06

Today's Target

- $\chi^2 - test\ of\ goodness\ of\ FIT\ (Part - 2)$
- PYQ
- DPP

Hpothesis Testing

Parametric test

(Sample drawn from normal population)

(i) $t-$ test

(ii) $z-$ test

(iii) $F-$ test

Non – Parametric test

(No information about population)

(i) χ^2- test

χ^2- test

χ^2- test of goodness of FIT

L-5, L-6

χ^2- test as a test of independence

L-7

χ^2 – test of goodness of Fit :

- Given Data can be fitted in

(i) Poisson Distribution

(ii) Binomial Distribution

~~(iii) Normal Distribution~~

- Fitting is good or not is test by χ^2 – test.
- In χ^2 – test $o_1, o_2 \ldots . o_n$ is a set of observed frequencies and $E_1, E_2, E_3 \ \ldots \ E_n$ are the corresponding set of expected frequencies such that $\sum o_i = \sum E_i$

Conditions for applying χ^2 – test

(i) Sum of frequencies must be greater than equal to 50.

(ii) The expected frequency E_i of any item must be greater than 5.

Note : If any item has frequency less than 5 then it should be combined with the next or preceding item until the total frequency exceed 5.

Steps to solve questions

Step – 1 : Null Hypothesis (H_o)

(H_o): There is no significant difference between observed and expected frequency.

Step – 2 : Calculate χ^2 under H_o

$$\chi^2 = \sum \left[\frac{(O_i - E_i)^2}{E_i}\right]$$

Where O_i = observed frequency

E_i = Expected frequency

Step – 3 : Degree of freedom = Number of variables − number of constraints

$$\vartheta = n - k$$

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

Step – 4 : From χ^2 – table

Calculate $\chi^2_{0.05}$ at the given level of significance for the given degree of freedom

Step – 5 :

(i) If $\chi^2 < \chi^2_{0.05}$, then H_o accepted

- There is no significants difference between observed frequency and expected frequency.
- Fit of the data is considered to be good.

(ii) $\chi^2 > \chi^2_{0.05}$, then H_o is rejected.

- There is significants difference between observed frequency and expected frequency.
- Fit of the data is not good.

Q.1 *A survey of 320 families with 5 children shows the following distribution:*

No of boys & girls :	5 boys & 0 girl	4 boys & 1 girl	3 boys & 2 girl	2 boys & 3 girl	1 boys & 4 girl	0 boys & 5 girl	Total
No. of families :	18	56	110	88	40	8	320

Test the hypothesis that data are binomially distributed and male and female births are equal probability. **(G.B.T.U. 2010)**

Null Hypothesis (H_0)

H_0: Data are binomially distributed and probability of Male and Female birth are equal

Let p = Prob. of Girl (success)

q = Prob. of boy (Failure)

$p = q = \frac{1}{2}$

$N = 320$

$n = 5$

$x = 0, 1, 2, 3, 4, 5$

Binomial Distribution

$P(X=x) = {}^{n}C_{x}\, p^{x} q^{n-x}$

Expected Frequencies

$$E_i = N\,{}^{n}C_{\gamma}\, p^{\gamma} q^{n-\gamma}$$

$$E_i = 320 \times {}^{5}C_{\gamma} \left(\frac{1}{2}\right)^{\gamma} \left(\frac{1}{2}\right)^{5-\gamma} \quad \text{— (1)}$$

$$E_1 = 320 \times {}^{5}C_{0} \times \left(\frac{1}{2}\right)^{0} \times \left(\frac{1}{2}\right)^{5}$$

$$= \overset{10}{\cancel{320}} \times 1 \times 1 \times \frac{1}{\cancel{32}}$$

$$\boxed{E_1 = 10}$$

$$E_2 = 320 \times {}^{5}C_{1} \times \left(\frac{1}{2}\right)^{1} \times \left(\frac{1}{4}\right)^{4} = \overset{10}{\cancel{320}} \times 5 \times \frac{1}{\cancel{32}} = 50$$

$$\boxed{E_2 = 50}$$

$$E_3 = 320 \times {}^{5}C_{2} \times \left(\frac{1}{2}\right)^{2} \times \left(\frac{1}{2}\right)^{3} = \overset{10}{\cancel{320}} \times \frac{5 \times \overset{2}{\cancel{4}}}{\cancel{2}} \times \frac{1}{\cancel{32}} = 100$$

$$\boxed{E_3 = 100}$$

$$E_4 = 320 \times {}^{5}C_{3} \times \left(\frac{1}{2}\right)^{3} \times \left(\frac{1}{2}\right)^{2} = \overset{10}{\cancel{320}} \times \frac{5 \times \overset{2}{\cancel{4}}}{\cancel{2}} \times \frac{1}{\cancel{32}}$$

$$\boxed{E_4 = 100}$$

$$E_5 = 320 \times {}^{5}C_{4} \times \left(\frac{1}{2}\right)^{4} \times \left(\frac{1}{2}\right)^{1} = \overset{10}{\cancel{320}} \times 5 \times \frac{1}{\cancel{32}} = 50$$

$E_5 = 50$

$$E_6 = 320 \times {}^{5}C_{5} \times \left(\frac{1}{2}\right)^{5} \times \left(\frac{1}{2}\right)^{0}$$

$$= \cancel{320}^{10} \times 1 \times \cancel{\frac{1}{32}}$$

$E_6 = 10$

calculation of χ^2, under H_0

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
18	10	8	64	6.4
56	50	6	36	0.72
110	100	10	100	1
88	100	−12	144	1.44
40	50	−10	100	2
8	10	−2	4	0.4
				$\sum\left[\frac{(O_i - E_i)^2}{E_i}\right] = 11.96$

$\chi^2_{cal} = 11.96$

$\chi_{tab} = 11.070$

Degree of freedom $= n - k$

$= 6 - 1$

$= 5$

Level of significance (α)

$\alpha = 5\% = 0.05$

From χ^2-table

$\chi_{tab} =$

conclusion

$\chi^2_{cal} = 11.96$ and $\chi_{tab} = 11.070$

$\because \chi^2_{cal} > \chi^2_{tab}$

$\therefore H_0$ is rejected

Hence,

Data are not Binomially Distributed and Probability of birth of boy and Girl are not equal

$\Rightarrow p \neq q$

GW
GATEWAY CLASSES

Table 4 : CHI-SQUARE

Significant Values $\chi^2(\alpha)$ of Chi-Square Distribution Right Tail Areas for Given Probability α,

$P = P_r(\chi^2 > \chi^2(\alpha)) = \alpha$

And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	*Probability (Level of Significance)*						
	0.99	*0.95*	*0.50*	*0.10*	*0.05*	*0.02*	*0.01*
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
2	.0201	.103	1.386	4.605	5.991	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

Q.1 *Fit a Poisson distribution to the following data and test the goodness of fit:*

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

Null Hypothesis (H_0)

H_0: Fitting of data in poisson Distribution is good

x	f	fx
0	109	0
1	65	65
2	22	44
3	3	9
4	1	4
	$\Sigma f = 200$	$\Sigma fx = 122$

$$\text{Mean} = \frac{\Sigma fx}{\Sigma f}$$

$$\lambda = \frac{122}{200} = 0.61$$

By Poisson Distribution

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

$$P(x) = \frac{e^{-0.61} \times (0.61)^x}{x!}$$

Expected Frequencies

$$E_i = \frac{N e^{-0.61} \times (0.61)^x}{x!}$$

$$E_i = 200 \times \frac{e^{-0.61} \times (0.61)^x}{x!} \quad \text{— (1)}$$

$$E_1 = \frac{200 \times e^{-0.61} \times (0.61)^0}{0!}$$

$$E_1 = \frac{200 \times e^{-0.61}}{1}$$

$$E_1 = 108.67 \approx 109$$

$$E_2 = 66.29 \approx 66$$

$$E_3 = 20.22 \approx 20$$

$$E_4 = 4.11 \approx 4$$

$$E_5 = 0.63 \approx 1$$

calculation of χ^2, under H_0

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
109	109	0	0	0
65	66	−1	1	0.01515
22	20	2	4	0.2
3, 1 (4)	4, 1 (5)	−1	1	0.2
				$\chi^2 = 0.41515$

$$\chi_{cal} = 0.41515$$

Degree of freedom $= n - K$

$$d.f = 5 - 1 - 1 - 1$$

$$d.f = 2$$

Level of significance (α)

$$\alpha = 5\% = \frac{5}{100} = 0.05$$

From χ^2-table

$$\chi^2_{tab} = 5.991$$

Conclusion

$\chi^2_{cal} = 0.41515$ and $\chi^2_{tab} = 5.991$

$\because \chi^2_{cal} < \chi^2_{tab}$

$\therefore$ H_0 is accepted

Hence, Fitting of data in poisson Distribution is good

GW
GATEWAY CLASSES

Table 4 : CHI-SQUARE

Significant Values $\chi^2(\alpha)$ of Chi-Square Distribution Right Tail Areas for Given Probability α,

$P = P_r(\chi^2 > \chi^2(\alpha)) = \alpha$

And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	*Probability (Level of Significance)*						
	0.99	*0.95*	*0.50*	*0.10*	*0.05*	*0.02*	*0.01*
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
2	.0201	.103	1.386	4.605	5.991	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

Q.2 *When the first proof of 392 pages of a book of 1200 pages were read, the distribution of printing mistakes were found to be as follows:*

No. of mistakes in a page (x)	:	0	1	2	3	4	5	6
No. of pages (f)	:	275	72	30	7	5	2	1

Fit a poisson distribution to the above data and test the goodness of fit.

Null Hypothesis (H_0)

H_0 : Poisson Distribution is a good fit to the data

$$\text{Mean} = \frac{\Sigma fx}{\Sigma f} = \frac{189}{392}$$

$$\lambda = 0.4821$$

x	f	fx
0	275	0
1	72	72
2	30	60
3	7	21
4	5	20
5	2	10
6	1	6
	$\Sigma f = 392$	$\Sigma fx = 189$

By Poisson Distribution

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

Expected Frequency

$$E_i = \frac{N \times e^{-\lambda} \lambda^{x}}{x!}$$

$$E_i = 392 \times \frac{e^{-0.4821} \times (0.4821)^{x}}{x!}$$

$\longrightarrow$ ①

$$E_1 = 392 \times \frac{e^{-0.4821} \times (0.4821)^{0}}{0!}$$

$E_1 = 242.05 \simeq 242.1$

$E_2 = 116.69 \simeq 116.7$

$E_3 = 28.13 \simeq 28.1$

$E_4 = 4.52 \simeq 4.5$

$E_5 = 0.5 \simeq 0.5$

$E_6 = 0.052 \simeq 0.1$

$E_7 = 0.0042 \simeq 0$

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
275	242.1	32.9	1082.41	4.471
72	116.7	-44.7	1998.09	17.121
30	28.1	1.9	3.61	0.128
7	4.5			
5 = 15	0.5 = 5.1	9.9	98.01	19.217
2	0.1			
1	0			
				$\chi^2 = 40.937$

$$\chi^2_{cal} = \sum \left[\frac{(O_i - E_i)^2}{E_i} \right]$$

$$\chi^2_{cal} = 40.937$$

Degree of Freedom $= n - K$

$$d.f = 7 - 1 - 1 - 3$$

$$d.f = 2$$

Level of significance (α)

$$\alpha = 5\% = 0.05$$

From χ^2- table

$$\chi^2_{tab} = 5.991$$

conclusion

$$\chi^2_{cal} = 40.937$$

$$\chi^2_{tab} = 5.991$$

$\therefore \chi^2_{cal} > \chi^2_{tab}$

$\therefore$ H_0 is rejected

Hence

Poisson distribution is not a good fit to the given data

GW
GATEWAY CLASSES

Table 4 : CHI-SQUARE

Significant Values $\chi^2(\alpha)$ of Chi-Square Distribution Right Tail Areas for Given Probability α,

$$P = P_r(\chi^2 > \chi^2(\alpha)) = \alpha$$

And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	*Probability (Level of Significance)*						
	0.99	*0.95*	*0.50*	*0.10*	*0.05*	*0.02*	*0.01*
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
2	.0201	.103	1.386	4.605	5.991	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

Topic : $\chi^2 - test\ of\ goodness\ of\ FIT\ (Part - 2)$

Q.1 *Test for goodness of fit of a poisson distribution at 5% level of significance to the following frequency distribution:*

x	:	0	1	2	3	4	5	6	7	8
f	:	52	151	130	102	45	12	5	1	2

Q.2 *Test for goodness of fit of a poisson distribution at 5% level of significance to the following frequency distribution:*

x	:	0	1	2	3	4	5
f	:	275	138	75	7	4	1

Q.3 *Records taken of the number of male and female births in 800 families having four children are as follows:*

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

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 $p = q = 1/2$. [U.P.T.U. 2009]

GATEWAY CLASSES

AKTU Full Courses (Paid)

Download Gateway Classes Application

From Google Play store

All Subjects

Link in Description

Thank You

GATEWAY CLASSES