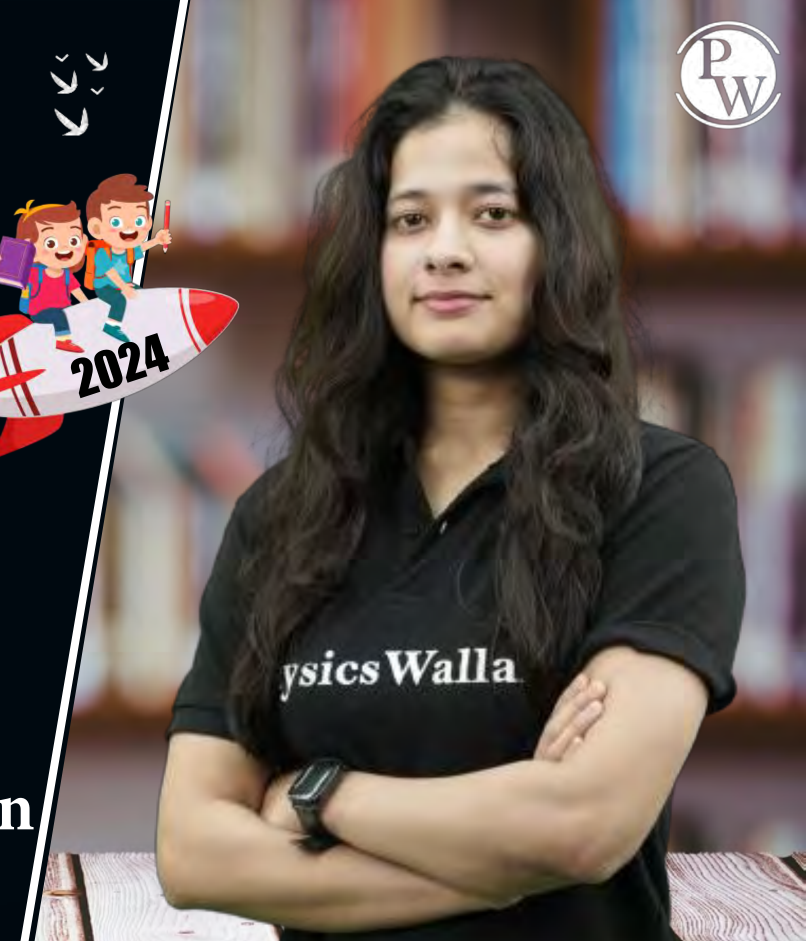
Data Science & AI

Probability & Statistics

Introduction to Sampling Distribution



DPP Discussion Notes

By- Rajshree mam

500 - 220 - 170 - 90

#Q. A sample analysis of examination results of500 students, it was found that 220 students have failed, 170 have secured a third class, 90have secured a second class and the rest, a first class. Do these figures support the general belief that above categories are in the ratio 4 :3 : 2 : 1 respectively ? (The tabular value of χ^2 for d .f . 3 at 5% level of significance is 7.81).

$\chi^2_{\text{calculate}} = \sum \frac{(O-E)^2}{E}$

Observed

Expected

$\chi^2_{cal} < \chi^2_{critical}$ ✗

χ^2 - test

$\chi^2_{cal} > \chi^2_{critical}$ ✗

$n \rightarrow$ sample $\longrightarrow$ d.o.f $(n-1)$

✓ $\chi^2 \rightarrow$ value $\rightarrow$ data's

✓ $\chi^2_{critical} \rightarrow$ table $(n-1)$ d.o.f

0.05% 0.01%

Table

face value	fail	3rd	2nd	1st
Observed	220	170	90	20
Expected	200	150	100	50

$\cancel{4}x + \cancel{3}n + \cancel{2}n + \cancel{x} = 500$

$10n = 500$

$n = 50$

H_0:- $\underline{O \to E}$ same

H_1:- $\underline{\text{diff}^{r}}$

$O-E$	20	20	-10	-30
$(O-E)^2$	400	400	100	900
$\frac{(O-E)^2}{E}$	2	2.667	1	18

$$\sum \frac{(O-E)^2}{E} = \underbrace{23.67}$$

$(V) = n-1 = 4-1 = 3$

H_0: Rejected at 5% Level of significy

$\chi^2 \rightarrow$ table

7.815

0.05

$n = 3$

#Q. What is χ^2 – test?

A die is thrown 90 times what the following results:

with

Face:	1	2	3	4	5	6	Total
Frequency	10	12	16	14	18	20	90

Use χ^2-test to test whether these data are consistent with the hypothesis that die is unbiased. → unbiased

Given χ^2 0.05 ≡ 11.07 for 5 degrees of freedom.

1, 2, 3, 4, 5, 6

$\frac{1}{6} = \cancel{\frac{90}{6}}\ 15$

Face	1	2	3	4	5	6
Observed	10	18	11	13	20	18
Expected	15	15	15	15	15	15
O − E	−5	3	−4	−2	5	3

$(O-E)^2$ 25   9   16   4   25   9

$\frac{(O-E)^2}{E} = 1.667$ 0.600 1.067 0.267 1.667 0.600

$$\sum \frac{(O-E)^2}{E} = \underline{\underline{5.868}}$$

$n = 6$

$$d.f = n-1 \Rightarrow \underline{\underline{\nu = 5}}$$

$\sim \chi^2 (d.o.f)$

0.05

11.07

n = 5

$\chi^2_{cal} = 5.878 < 11.07$

Accepted

H_0 :- die is unbiased.

H_1 :- die is biased.

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

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

Given that values of χ^2 for 5 degrees of freedom are 11.1 and 15.1at 0.05 and 0.01 significance level respectively , test the hypothesis that male and female birth are equally probable.

320 families

Bache ⟶ child

~~face~~ value

5 → G
0 → B

18

4 → G
1 → B

56

3 - G
2 - B

110

2 - G
3 - B

88

4 - B
1 → G

40

0 - G
5 - B

8

18

face value →

Expected →

Observed 18 52 110 48 90 8

320 → 5 child

9 1

5 0

320 → 5

family

320 × 5

1600

800 800

24

2B

320

24

2B

5

Expected observed

$$\chi^2 \rightarrow \frac{(O-E)^2}{E}$$

χ^2 critical $\rightarrow$ 11.1 15.1

0.05 0.01

Calculation for expected –

Topic : Probability & Statistics

#Q. A chemical extraction plant processes sea water to collect sodium chloride and magnesium. It is known that sea water contains sodium chloride, magnesium and other elements in the ratio 62:4: 34. A sample of 200 tonnes of sea water has resulted in 130 tonnes of sodium chloride and 6 tonnes of magnesium. Are these data consistent with the known composition of sea water at 5% level of significance? (Given that the tabular value of χ^2 is 5.991 for 2 degree of freedom).

200

Sea water

NaCl

Magnesium

other elements

Observed → 130 6 64

Expected →

$6\cancel{2}x + 4x + 3\cancel{4}x = 200$

$100x = 200$

$\boxed{x = 2}$

Expected	124	8	68
Observed	130	6	64
$O-E$	6	-2	-4
$(O-E)^2$	36	4	16
$\frac{(O-E)^2}{E}$	0.29	0.5	0.235

$$\chi^2 = 1.025$$

$\chi^2_{critical}$ (d.o.f) = n = 2

= $\underline{\underline{5.991}}$

$\underline{\underline{1.025 < 5.991}}$

Accepted

#Q. 4 coins were tossed at a time and this operation is repeated 160 times. It is found that 4 heads occur 6 times, 3 heads occur 43 times, 2 heads occur 69 times, one head occur 34 times. Discuss whether the coin may be regarded as unbiased?

4 coins 160 times

$X \Rightarrow$	4	3	2	1
$O \Rightarrow$	6	43	69	34
$E \Rightarrow$	10	40	60	40
$(O-E)$:	-4	3	9	-6
$(O-E)^2 =$	16	9	81	36

1 times

4 coin $\rightarrow$ 4 head $= \frac{1}{16}$

160 times $= \frac{1}{\cancel{16}} \times \cancel{160} = 10$

$$3 \text{ heads} = \frac{1}{\cancel{4}} \times \cancel{160} = 40$$

$$2 \text{ head} = \frac{3}{\cancel{8}} \times \cancel{160}^{20} = 60$$

$$1 \text{ head} = \frac{1}{\cancel{4}} \times \cancel{160}^{40} = 40$$

$$\frac{(O-E)^2}{E} = \quad 1.6 \quad 0.225 \quad 1.35 \quad 0.9$$

$$\sum \frac{(O-E)^2}{E} = \underline{\underline{4.075}}$$

$\chi^2_{critical} \sim \quad \nu = n-1 = 3$ (d.o.f)

Table

$$\chi^2_{calculated} < \chi^2_{critical}$$

Accepted

equal n^4

expected

#Q. 200 digits are chosen at random from a set of tables. The frequencies of the digits were:

20, 20

Digits	0	1	2	3	4	5	6	7	8	9
Frequency	18	19	23	21	16	25	22	20	21	15

Use χ^2-test to assess the correctness of the hypothesis that the digits were distributed in equal numbers in the table, given that the value of χ^2 are respectively 16.9, 18.3 and 19.7 for 9, 10 and 11 degrees of freedom at 5% level of significance.

$200 \to n = 10 \to \frac{200}{10} = 20$

digits	0	1	2	3	4	5	6	7	8	9
Expected	20	20	20	20	20	20	20	20	20	20
observed	18	19	23	21	16	25	22	20	21	15
$O - E$	-2	-1	3	1	-4	5	2	0	1	-5
$(O-E)^2$	4	1	9	1	16	25	4	0	1	25

$\frac{(O-E)^2}{E}$ 0.2 0.05 0.45 0.05 0.8 1.25 0.2 0

0.05

1.25

$$\sum \frac{(O-E)^2}{E} = \underline{\underline{4.3}}$$

$n-1 = d.o.f = 9$

$$\chi^2_{critical} = \underline{\underline{16.919}}$$

$\underline{table}$

$4.3 < 16.919$

Accepted

M N

#Q. A genetical law says that children having one parent of blood group M and the other parent of blood group N will always be one of the three blood groups M, MN, N and that the average no. of children in these groups will be in the ratio 1:2:1. The report on an experiment states as follows: "Of 162 children having one M parent and one N parent, 28.4% were found to be of group M. 42% of group MN and the rest of the group N." Do the data in the report conform to the expected genetic ratio 1:2:1?

1:2:1

M N MN

(100)%

(28.4 − 42)%

	MM	MN	N
face value	M	MN	N
Expected	n	$2n$	n
	$= 40.5$	81	40.5

$4x = 162$

$x = 40.5$

Observed	46.008	68.04	47.952
	$\frac{28.4}{100} \times 162$	$\frac{42}{100} \times 162$	$\frac{29.6}{100} \times 162$
$O - E$	5.508	−12.96	7.452
$(O-E)^2$	30.33	167.961	55.53
$\frac{(O-E)^2}{E}$	0.748	2.0735	1.3071

H_0: Accepted

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

$$= 4.192$$

1 : 2 : 1

$4 - 1 = 2$

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

$$4.192 < 5.991$$

#Q. Every clinical thermometer is classified into one of the four categories A, B. C and D on the basis of inspection and test. From past experience, it is known that thermometers produced by a certain manufacturer are distributed among the four categories in the following proportions:

Category	A	B	C	D
Proportion	0.87	0.09	0.03	0.01

A new lot of 1336 thermometers is submitted by the manufacturer for inspection and test and the following distribution into four categories results :

face value	A	B	C	D
observed	1188	91	47	10
Expected	$0.87 \times 1336 = 1162.32$	120.24	40.08	13.36
	0.87			
O - E	25.68	-29.24	6.92	-3.36

$(O-E)^2$ 659.9624 854.9002 920.8867

11.2896

$\frac{(O-E)^2}{E}$ 0.567 7.11 1.19482 0.8852

$$\sum \frac{(O-E)^2}{E} = 9.71 = \chi^2_{\text{calculated}}$$

$\chi^2_{critical}$ for table & d.o.f =

$n-1 = 3$

$n = 4$

↓

$\underline{\underline{7.815}}$

$\underline{\underline{9.71 > 7.815}} \rightarrow H_0$:- $\underline{\underline{\text{Rejected}}}$

Category	A	B	C	D
No of the thermometers	1188	91	47	10

Does this new lot of thermometers differ from the previous experience with regards to proportion of thermometers in each category?

H_0 :- differ nahi karega

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

1

χ^2

$n-1$

$= d.o.f$

$O \times E$

same

ones

(2) (3)

THANK - YOU

Topics to be Covered