Department of Computer Engineering
Academic Year 2024-2025 (Even Sem)

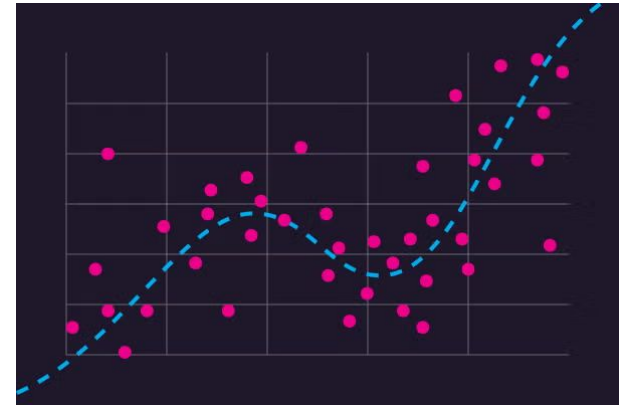# Chi-Square Test with Example

Ajit Gupta (A – 22)
Sidhenath Choudhari (A – 04)
Pruthviraj Ghatul (A – 20)

Dr. Jyoti Wadmare

Introduction

- A statistical method to test relationships between categorical variables.
- Helps determine if observed differences are due to chance or an actual association.
- Based on application, it includes tests like goodness-of-fit and test of independence.

Types of chi-square Test

- Two main types are used in categorical data analysis.
- Goodness-of-Fit checks how well data matches a known distribution.
- Test of Independence checks association between variables in a contingency table.

<u>Formula for chi-square Test</u>

- Used to measure difference between observed and expected frequencies.
- $\chi^2 = \Sigma - (O - E)^2 / E$
- Where O = Observed value, E = Expected value in each category.

Application for chi-square Test

- Commonly used in surveys, research, and experiments.
- Tests association between gender & preference, treatment & outcome, etc.
- Widely applied in social sciences, marketing, biology, and education studies.

Example

Chi-square (χ^2)

Table of Observed Values

Qualification / Marital Status	Middle School	High School	Bachelor's	Master's	Ph.D	Total
Never married	18	36	21	9	6	90
Married	12	36	45	36	21	150
Divorced	6	9	9	3	3	30
Widowed	3	9	9	6	3	30
Total	39	90	84	54	33	300

Null hypothesis: There is no relation between the marital status and educational qualification.

Alternate Hypothesis: There is significant relation between the marital status and educational qualification.

Significance level $(\alpha) = 0.05$

Example

Chi-square (χ^2)

Table of Observed Values

Qualification / Marital Status	Middle School	High School	Bachelor's	Master's	Ph.D	Total
Never married	18	36	21	9	6	90
Married	12	36	45	36	21	150
Divorced	6	9	9	3	3	30
Widowed	3	9	9	6	3	30
Total	39	90	84	54	33	300

Table of Expected Values

Qualification / Marital Status	Middle School	High School	Bachelor's	Master's	Ph.D
Never Married	$\frac{90 \times 39}{300} = 11.7$	$\frac{90 \times 90}{300} = 27$	25.2	16.2	9.9
Married	19.5	45	42	27	16.5
Divorced	3.9	9	8.4	5.4	3.3
Widowed	3.9	9	8.4	5.4	3.3

Example

Calculation of χ^2

Observed Values (O)	Expected Values (E)	$(O - E)$	$(O - E)^2$	$\dfrac{(O - E)^2}{E}$
18	11.7	6.3	39.69	3.39
36	27	9	81	3
21	25.2	-4.2	17.64	0.7
9	16.2	-7.2	51.84	3.2
6	9.9	-3.9	15.21	1.53
12	19.5	-7.5	56.25	2.88
36	45	-9	81	1.8
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
3	3.3	-0.3	0.09	0.02
				$\sum \dfrac{(O - E)^2}{E}$ $\chi^2 = 23.57$

Chi-square (χ^2)

Table of Observed Values

Qualification / Marital Status	Middle School	High School	Bachelor's	Master's	Ph.D	Total
Never married	18	36	21	9	6	90
Married	12	36	45	36	21	150
Divorced	6	9	9	3	3	30
Widowed	3	9	9	6	3	30
Total	39	90	84	54	33	300

Table of Expected Values

Qualification / Marital Status	Middle School	High School	Bachelor's	Master's	Ph.D
Never Married	$\dfrac{90 \times 39}{300} = 11.7$	$\dfrac{90 \times 90}{300} = 27$	25.2	16.2	9.9
Married	19.5	45	42	27	16.5
Divorced	3.9	9	8.4	5.4	3.3
Widowed	3.9	9	8.4	5.4	3.3

Inferences

Chi-square (χ^2)

Degrees of freedom= (columns − 1) (rows-1)

$$=(5\text{-}1)(4\text{-}1) \quad =4 \times 3 \quad =\mathbf{12}$$

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14

Inferences

Chi-square (χ^2)

Degrees of freedom= (columns $-$ 1) (rows-1)

$\qquad$ =(5-1) (4-1) $= 4 \times 3$ =**12**

Significance level $(\propto) = 0.05$

$$X^2_{tabular} = 21.03$$

$$\chi^2_{calculated} = 23.57$$

$$X^2_{calculated} > X^2_{tabular} \ (or \ called \ as \ X^2_{critical})$$

$\therefore$ we reject Null hypothesis, and accept alternate hypothesis

Alternate Hypothesis: There is significant relation between the marital status and educational qualification.

Thank You !