Chapter 13: Chi-Square Tests

STAT 2601 – Business Statistics

Esam Mahdi

School of Mathematics and Statistics
Carleton University

January 21, 2026

Table of Contents

From Binomial to Multinomial

- **Binomial Experiment:** A trial has only 2 possible outcomes (Success/Failure).
- **Multinomial Experiment:** A trial has $k > 2$ possible outcomes.

Example: Tossing a Die

Suppose we toss a k-sided unfair die n times.

- Let f_i be the **observed frequency** of the i-th side.
- Naturally, $\sum_{i=1}^{k} f_i = n$.
- Let p_i be the **probability** of getting side i, where $\sum p_i = 1$.

The Trial

A single performance of an act (e.g., one toss of the die).

The Outcome

One of the k possible results of a trial.

The Multinomial Experiment

Consists of n independent, identical trials where the probability of each outcome $(p_1, p_2, \ldots, p_k)$ remains constant from trial to trial.

Example: Human Blood Types

- **Scenario:** Suppose we sample n individuals and record their blood type.
- **Possible Outcomes ($k = 4$):** $\{A, B, AB, O\}$
- **Observed Counts (f_i):**
  - f_1: Number of people with Type A
  - f_2: Number of people with Type B
  - f_3: Number of people with Type AB
  - f_4: Number of people with Type O
- **Constraint:** $f_1 + f_2 + f_3 + f_4 = n$
- **Probabilities (p_i):** The population proportions of each blood type (p_1, p_2, p_3, p_4).

Why is this not Binomial?

If we only asked "Is the person Type O or Not?", it would be Binomial ($k = 2$). Because we are tracking four distinct categories simultaneously, it is **Multinomial**.

Chi-Square Goodness-of-Fit Test

Goodness-of-Fit test compares the observed data f_i with the expected frequencies predicted by null hypothesis:

Case I: No preference (equal proportions) among categories.

$$H_0 : p_1 = p_2 = \cdots = p_k = p_0$$
$$H_a : \text{ At least one } p_i \text{ is different from the hypothesized value.}$$

where p_0 is the probability value that we want to test and k is the number of categories.

Case II: Specified proportions or preference from another known population.

$$H_0 : p_1 = p_{10}, p_2 = p_{20}, \ldots, p_k = p_{k0}$$
$$H_a : \text{ At least one } p_i \text{ is different from the hypothesized value.}$$

where $p_{10}, p_{20}, \cdots, p_{k0}$ are the probability values or relative preference of each category that we want to test.

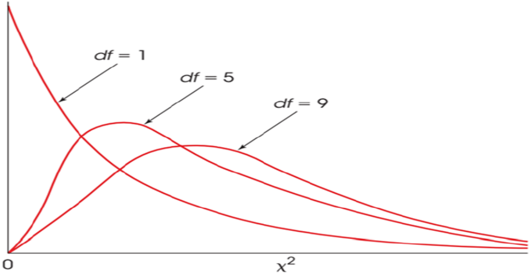
Chi-Square Goodness-of-Fit Test

Pearson's Chi-Square Test Statistic:

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(f_i - E_i)^2}{E_i}$$

Where $E_i = n \times p_{i0}, i = 1, 2, \cdots, k$ is the **expected frequency** under H_0.

Note $\sum_{i=1}^{k} E_i = \sum_{i=1}^{k} f_i = n$ and if we are testing for no preference then $p_{i0} = p_0$.



Rejection Region

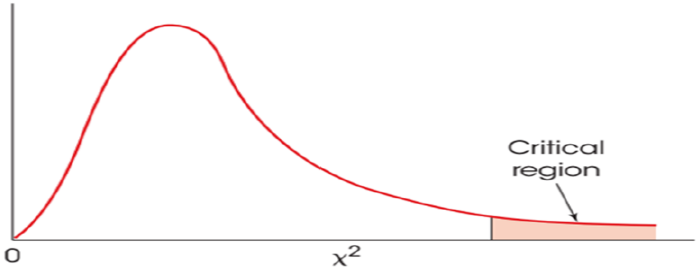
- The test statistic follows a Chi-square distribution with $df = k - 1$.
- **Decision Rule:** Reject H_0 if:

$$\chi^2 > \chi^2_{\alpha,k-1} \quad \text{(Right-tailed Test)}$$

or if

$$\text{p-value} < \alpha$$

- We only reject in the upper tail because large differences between f_i and e_i result in a large χ^2 value, indicating a poor fit.



Assumptions for the Chi-Square Test

To ensure the χ^2 approximation is valid, the following must hold:

1. **Counted Data Condition:** The data must be counts for the categories of a categorical variable.
2. **Independence Assumption:** The counts should be independent of each other.
3. **Randomization Condition:** The counted individuals should be a random sample of the population. Guard against auto-correlated samples.
4. **Sample Size Assumption:** We must have enough data for the methods to work.
5. **Expected Frequencies:** A common rule of thumb is that all expected frequencies e_i should be **at least 5** ($E_i \geq 5$).
6. If $E_i < 5$ for some cells, you may need to combine adjacent categories.

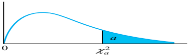


TABLE 5
Critical Values of Chi-Square

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$
1	.0000393	.0001571	.0009821	.0039321	.0157908
2	.0100251	.0201007	.0506356	.102587	.210720
3	.0717212	.114832	.215795	.351846	.584375
4	.206990	.297110	.484419	.710721	1.063623
5	.411740	.554300	.831211	1.145476	1.61031
6	.675727	.872085	1.237347	1.63539	2.20413
7	.989265	1.239043	1.68987	2.16735	2.83311
8	1.344419	1.646482	2.17973	2.73264	3.48954
9	1.734926	2.087912	2.70039	3.32511	4.16816
10	2.15585	2.55821	3.24697	3.94030	4.86518
11	2.60321	3.05347	3.81575	4.57481	5.57779
12	3.07382	3.57056	4.40379	5.22603	6.30380
13	3.56503	4.10691	5.00874	5.89186	7.04150
14	4.07468	4.66043	5.62872	6.57063	7.78953
15	4.60094	5.22935	6.26214	7.26094	8.54675
16	5.14224	5.81221	6.90766	7.96164	9.31223
17	5.69724	6.40776	7.56418	8.67176	10.0852
18	6.26481	7.01491	8.23075	9.39046	10.8649
19	6.84398	7.63273	8.90655	10.1170	11.6509
20	7.43386	8.26040	9.59083	10.8508	12.4426
21	8.03366	8.89720	10.28293	11.5913	13.2396
22	8.64272	9.54249	10.9823	12.3380	14.0415
23	9.26042	10.19567	11.6885	13.0905	14.8479
24	9.88623	10.8564	12.4011	13.8484	15.6587
25	10.5197	11.5240	13.1197	14.6114	16.4734
26	11.1603	12.1981	13.8439	15.3791	17.2919
27	11.8076	12.8786	14.5733	16.1513	18.1138
28	12.4613	13.5648	15.3079	16.9279	18.9392
29	13.1211	14.2565	16.0471	17.7083	19.7677
30	13.7867	14.9535	16.7908	18.4926	20.5992
40	20.7065	22.1643	24.4331	26.5093	29.0505
50	27.9907	29.7067	32.3574	34.7642	37.6886
60	35.5346	37.4848	40.4817	43.1879	46.4589
70	43.2752	45.4418	48.7576	51.7393	55.3290
80	51.1720	53.5400	57.1532	60.3915	64.2778
90	59.1963	61.7541	65.6466	69.1260	73.2912
100	67.3276	70.0648	74.2219	77.9295	82.3581

SOURCE: From "Tables of the Percentage Points of the χ^2-Distribution," *Biometrika Tables for Statisticians,* Vol. 1, 3rd ed. (1966). Reproduced by permission of the *Biometrika* Trustees.

TABLE 5
(continued)

$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$	*df*
2.70554	3.84146	5.02389	6.63490	7.87944	1
4.60517	5.99147	7.37776	9.21034	10.5966	2
6.25139	7.81473	9.34840	11.3449	12.8381	3
7.77944	9.48773	11.1433	13.2767	14.8602	4
9.23635	11.0705	12.8325	15.0863	16.7496	5
10.6446	12.5916	14.4494	16.8119	18.5476	6
12.0170	14.0671	16.0128	18.4753	20.2777	7
13.3616	15.5073	17.5346	20.0902	21.9550	8
14.6837	16.9190	19.0228	21.6660	23.5893	9
15.9871	18.3070	20.4831	23.2093	25.1882	10
17.2750	19.6751	21.9200	24.7250	26.7569	11
18.5494	21.0261	23.3367	26.2170	28.2995	12
19.8119	22.3621	24.7356	27.6883	29.8194	13
21.0642	23.6848	26.1190	29.1413	31.3193	14
22.3072	24.9958	27.4884	30.5779	32.8013	15
23.5418	26.2962	28.8485	31.9999	34.2672	16
24.7690	27.8571	30.1910	33.4087	35.7185	17
25.9894	28.8693	31.5264	34.8053	37.1564	18
27.2036	30.1435	32.8523	36.1908	38.5822	19
28.4120	31.4104	34.1696	37.5662	39.9968	20
29.6151	32.6705	35.4789	38.9321	41.4010	21
30.8133	33.9244	36.7807	40.2894	42.7956	22
32.0069	35.1725	38.0757	41.6384	44.1813	23
33.1963	36.4151	39.3641	42.9798	45.5585	24
34.3816	37.6525	40.6465	44.3141	46.9278	25
35.5631	38.8852	41.9232	45.6417	48.2899	26
36.7412	40.1133	43.1944	46.9630	49.6449	27
37.9159	41.3372	44.4607	48.2782	50.9933	28
39.0875	42.5569	45.7222	49.5879	52.3356	29
40.2560	43.7729	46.9792	50.8922	53.6720	30
51.8050	55.7585	59.3417	63.6907	66.7659	40
63.1671	67.5048	71.4202	76.1539	79.4900	50
74.3970	79.0819	83.2976	88.3794	91.9517	60
85.5271	90.5312	95.0231	100.425	104.215	70
96.5782	101.879	106.629	112.329	116.321	80
107.565	113.145	118.136	124.116	128.299	90
118.498	124.342	129.561	135.807	140.169	100

Example 1: Equal Preferences

Problem Statement

A researcher wants to see if three brands of coffee are equally preferred. Out of $n = 60$ people, the observed counts are: **Brand A: 25**, **Brand B: 15**, **Brand C: 20**. Test at $\alpha = 0.05$.

Step 1: State the Hypotheses

- $H_0 : p_1 = p_2 = p_3 = 1/3$ (no preference = equal proportions)
- H_a : At least one $p_i \neq 1/3$ (preference is not uniform = different proportions)

Step 2: Calculate Expected Frequencies ($E_i = n \times p_i$). Under H_0, we expect an equal distribution:

- $E_1 = np_{10} = 60 \times (1/3) = 20$
- $E_2 = np_{20} = 60 \times (1/3) = 20$
- $E_3 = np_{30} = 60 \times (1/3) = 20$

Note: All $E_i \geq 5$, $n > 30$ (large and independent random sample), so the Chi-square assumptions are met.

Example 1: Calculation, Decision and Conclusion

Step 3: Calculate the χ^2 Test Statistic

Category	**Observed** (f_i)	**Expected** (E_i)	$(f_i - E_i)^2$	$\frac{(f_i - E_i)^2}{E_i}$
Brand A	25	20	25	$25/20 = 1.25$
Brand B	15	20	25	$25/20 = 1.25$
Brand C	20	20	0	$0/20 = 0.00$
Total	$n = 60$	**60**	-	$\chi^2 = 2.5$

Step 4: Determine the Critical Value

- Degrees of Freedom: $df = k - 1 = 3 - 1 = 2$
- Level of significance: $\alpha = 0.05$. From Table: $\chi^2_{0.05,2} = 5.99147$

Step 5: Decision: Since the calculated $\chi^2 = 2.5$ is less than the critical value $\chi^2_{0.05,2} = 5.99147$, we **fail to reject** H_0.

Conclusion: At the 5% significance level, there is insufficient evidence to suggest that consumers have a preference for one brand over the others. The observed differences are likely due to sampling variation.

Examining the Residuals

- When we reject a null hypothesis, we can examine the residuals in each cell to discover which values are extraordinary (i.e., sample and model data are relatively discrepant) or contribute most to the value of aggregate chi-square.
- Because we might compare residuals for cells with very different counts, we should examine standardized residuals:

$$\frac{(f_i - E_i)}{\sqrt{E_i}}$$

- Note that standardized residuals from goodness-of-fit tests are actually z-scores (which we already know how to interpret and analyze).

Examining the Residuals in Example 1 (Equal Preferences)

Standardized residuals for the technical support data are as follows:

Category	**Standardized Residual** ($\frac{(f_i - E_i)}{\sqrt{E_i}}$)
Brand A	$1.12 = \frac{(25-20)}{\sqrt{20}}$
Brand B	$-1.12 = \frac{(15-20)}{\sqrt{20}}$
Brand C	$0 = \frac{(20-20)}{\sqrt{20}}$

There is no outlier value as all standardized residual values within the interval $(-3, 3)$.

Example 2: Genetics and Specific Proportions

Problem Statement

A genetics experiment expects a phenotypic ratio of **9:3:3:1** (Mendelian Inheritance). In a sample of $n = 160$ offspring, the observed counts are: **100, 20, 35, and 5**. Test at $\alpha = 0.05$ to see if the data fits the expected ratio.

Step 1: Hypotheses

- $H_0 : p_1 = \frac{9}{16}, p_2 = \frac{3}{16}, p_3 = \frac{3}{16}, p_4 = \frac{1}{16}$
- H_a : At least one p_i is different from the specified ratio.

Note that $n = \sum_{i=1}^{4} f_i = 100 + 20 + 35 + 5 = 160 = n$.

Step 2: Calculate Expected Frequencies ($E_i = n \times p_i$)

- $E_1 = np_{10} = 160 \times (9/16) = 90$
- $E_2 = np_{20} = 160 \times (3/16) = 30$
- $E_3 = np_{30} = 160 \times (3/16) = 30$
- $E_4 = np_{40} = 160 \times (1/16) = 10$

Check the assumptions: $n = 160 > 30$ *(large and independent random sample) and all* $E_i \geq 5$*, so the test is valid.*

Example 2: Calculation Table

Step 3: Calculate the χ^2 Test Statistic

Phenotype	**Observed** (f_i)	**Expected** (E_i)	$(f_i - E_i)$	$\frac{(f_i-E_i)^2}{E_i}$
Type 1	100	90	10	$100/90 \approx 1.111$
Type 2	20	30	-10	$100/30 \approx 3.333$
Type 3	35	30	5	$25/30 \approx 0.833$
Type 4	5	10	-5	$25/10 = 2.500$
Total	**160**	**160**	**0**	$\chi^2 = 7.777$

A good way to check your work: the sum of $(f_i - E_i)$ must always be 0.

Example 2: Conclusion

Step 4: Determine the Critical Value

- $df = k - 1 = 4 - 1 = 3$
- For $\alpha = 0.05$ and $df = 3$: $\chi^2_{0.05,3} = 7.81473$

Step 5: Decision Compare the test statistic $\chi^2 = 7.777$ to tabular value $\chi^2_{0.05,3} = 7.81473$.

- Since $7.777 < 7.815$, our test statistic falls just short of the rejection region.
- **Decision:** Reject H_0.

Conclusion: There is sufficient evidence at the $\alpha = 0.05$ level to reject the claim that the offspring follow the 9:3:3:1 phenotypic ratio. The genetic model is not supported by the data.

Two-Way Contingency Tables

Definition

An $r \times c$ contingency table displays the joint frequency distribution of two categorical variables:

- Row variable with r levels
- Column variable with c levels

Example

A 2×3 table:

	Column 1	**Column 2**	**Column 3**
Row 1	f_{11}	f_{12}	f_{13}
Row 2	f_{21}	f_{22}	f_{23}

Notation and Definitions

- f_{ij} = observed frequency in row i, column j
- $n = \sum_{i=1}^{r} \sum_{j=1}^{c} f_{ij}$ = total sample size
- Row totals: $r_i = \sum_{j=1}^{c} f_{ij}$ for $i = 1, \ldots, r$
- Column totals: $c_j = \sum_{i=1}^{r} f_{ij}$ for $j = 1, \ldots, c$
- p_{ij} = true probability of falling in cell $(i, j) \Rightarrow \sum_{i=1}^{r} \sum_{j=1} p_{ij} = 1$
- Expected frequency under independence: $E_{ij} = n \cdot p_{ij}$

Factor 1	**Factor 2**						**Row Totals**
	1	**2**	$\cdots$	**j**	$\cdots$	**c**	
1	f_{11}	f_{12}	$\cdots$	f_{1j}	$\cdots$	f_{1c}	$\mathbf{r_1}$
2	f_{21}	f_{22}	$\cdots$	f_{2j}	$\cdots$	f_{2c}	$\mathbf{r_2}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
i	f_{i1}	f_{i2}	$\cdots$	f_{ij}	$\cdots$	f_{ic}	$\mathbf{r_i}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
r	f_{r1}	f_{r2}	$\cdots$	f_{rj}	$\cdots$	f_{rc}	$\mathbf{r_r}$
Column Totals	$\mathbf{c_1}$	$\mathbf{c_2}$	$\cdots$	$\mathbf{c_j}$	$\cdots$	$\mathbf{c_c}$	$\mathbf{n}$

Important Relationship

Under independence: $p_{ij} = p_{i+} \cdot p_{+j}$, where $p_{i+} = p_{i1} + \cdots + p_{ic}$ and $p_{+j} = p_{1j} + \cdots + p_{rj}$

Thus: $\hat{E}_{ij} = n \cdot \hat{p}_{i+} \cdot \hat{p}_{+j} = \frac{r_i \cdot c_j}{n} = \frac{ith \text{ row total} \times jth \text{ column total}}{\text{grand total or sample size}}$

Hypotheses for Contingency Tables

Test of Independence

$$H_0: \text{ The two variables are independent}$$
$$H_a: \text{ The two variables are NOT independent}$$

Test of Association

$$H_0: \text{ Variable 1 is NOT associated with Variable 2}$$
$$H_a: \text{ Variable 1 is associated with Variable 2}$$

Often used when one factor represents different populations/groups

Chi-Square Test Statistic

Chi-Square Statistic

$$\chi_0^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(f_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$$

where $\hat{E}_{ij} = \dfrac{r_i \cdot c_j}{n} = \dfrac{ith \text{ row total} \times jth \text{ column total}}{\text{grand total or sample size}}$

- Measures discrepancy between observed and expected frequencies
- Large values indicate evidence against H_0
- Under H_0, the test statistics $\chi^2 \sim \chi^2_{(r-1)(c-1)}$ asymptotically, where df $= (r-1)(c-1)$.
- Reject H_0 if $\chi^2 > \chi^2_{\alpha,(r-1)(c-1)}$ or if p-value $< \alpha$.

Assumptions for Chi-Square Test

1. **Random Sampling**: Observations are independent
2. **Sample Size**:
   - All expected frequencies $\hat{E}_{ij} \geq 1$
   - At least 80% of cells have $\hat{E}_{ij} \geq 5$
3. **Categorical Data**: Variables are nominal or ordinal

Example 1: Test of Association

Research Question

Is there an association between gender (Male, Female) and preferred news source (TV, Online, Newspaper)? Data from $n = 200$:

	TV	Online	Newspaper	Total
Male	30	40	10	80
Female	20	70	30	120
Total	50	110	40	200

Example 1 Solution: Step 1

Step 1: State Hypotheses

- H_0: Gender and news source preference are independent
- H_a: Gender and news source preference are associated

Step 2: Calculate Expected Frequencies

$$\hat{E}_{ij} = \frac{r_i \cdot c_j}{n} = \frac{ith \text{ row total} \times jth \text{ column total}}{\text{grand total or sample size}}$$

	TV	Online	Newspaper
Male	$\frac{80\times50}{200} = 20$	$\frac{80\times110}{200} = 44$	$\frac{80\times40}{200} = 16$
Female	$\frac{120\times50}{200} = 30$	$\frac{120\times110}{200} = 66$	$\frac{120\times40}{200} = 24$

Assumptions check: All $\hat{E}_{ij} \geq 5$

Example 3 Solution: Step 2

Step 3: Calculate Test Statistic

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

Cell	f	$\hat{E}$	$(f-\hat{E})^2$	$(f-\hat{E})^2/\hat{E}$
Male, TV	30	20	100	5.00
Male, Online	40	44	16	0.36
Male, News	10	16	36	2.25
Female, TV	20	30	100	3.33
Female, Online	70	66	16	0.24
Female, News	30	24	36	1.50
Total				$\chi^2 = 12.68$

Example 3 Solution: Step 3

Step 4: Decision and Conclusion

- Degrees of freedom: $(r-1)(c-1) = (2-1)(3-1) = 2$
- Critical value: $\chi^2_{0.05,2} \approx 5.991$
- Since $12.68 > 5.991$, reject H_0

Conclusion

There is significant evidence ($\chi^2 = 12.68$, df = 2, p < 0.05) that gender and news source preference are associated.

Summary: Chi-Square Test Statistics

Goodness-of-Fit Test

- **Purpose:** Test if observed distribution matches hypothesized distribution
- **Test Statistic:**

$$\boxed{\chi_0^2 = \sum_{i=1}^{k} \frac{(f_i - E_i)^2}{E_i}}$$

  where $E_i = n \times p_{i0}$ (expected frequency)
- **Degrees of Freedom:** $df = k - 1$

Contingency Table Test (Independence/Association)

- **Purpose:** Test if two categorical variables are independent
- **Test Statistic:**

$$\boxed{\chi_0^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(f_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}}$$

  where $\hat{E}_{ij} = \dfrac{r_i \times c_j}{n}$ (expected frequency under independence)
- **Degrees of Freedom:** $df = (r - 1)(c - 1)$