

# Chapter 11

## Multinomial Experiments and Contingency Tables

---

---

---

---

---

---

---

---

## Chapter 11 Multinomial Experiments and Contingency Tables

### 11-1 Overview

### 11-2 Multinomial Experiments: Goodness-of-fit

### 11-3 Contingency Tables: Independence and Homogeneity

---

---

---

---

---

---

---

---

## 11-1 Overview

- ❖ Focus on analysis of categorical (qualitative or attribute) data that can be separated into different categories (often called cells)
- ❖ Use a  $\chi^2$  (chi-square) test statistic and critical values from the Chi-Square Distribution (Table A-4)
- ❖ One-way frequency table (single row or column)
- ❖ Two-way frequency table or contingency table (two or more rows and columns)

---

---

---

---

---

---

---

---

# 11-2

## Multinomial Experiment

---

---

---

---

---

---

---

---

## Definition

### Multinomial Experiment

An experiment that meets the following conditions:

1. The number of trials is fixed.
2. The trials are independent.
3. All outcomes of each trial must be classified into exactly one of several different categories.
4. The probabilities for the different categories remain constant for each trial.

---

---

---

---

---

---

---

---

## Definition

### Goodness-of-fit test

used to test the hypothesis that an observed frequency distribution fits (or conforms to) some claimed distribution

---

---

---

---

---

---

---

---

## Goodness-of-Fit Test Notation

***O*** represents the observed frequency of an outcome

***E*** represents the expected frequency of an outcome

***k*** represents the number of different categories or outcomes

***n*** represents the total number of trials

---

---

---

---

---

---

---

---

## Expected Frequencies

If all expected frequencies are equal:

$$E = \frac{n}{k}$$

the sum of all observed frequencies divided  
by the number of categories

---

---

---

---

---

---

---

---

## Expected Frequencies

If all expected frequencies are not all equal:

$$E = n p$$

each expected frequency is found by multiplying  
the sum of all observed frequencies by the  
probability for the category

---

---

---

---

---

---

---

---

## Key Question

Are the differences between the observed values (O) and the theoretically expected values (E) statistically significant?

---

---

---

---

---

---

---

---

## Key Question

We need to measure the discrepancy between O and E; the test statistic will involve their difference:

$$O - E$$

---

---

---

---

---

---

---

---

## Multinomial Experiments: Goodness-of-Fit Assumptions

when testing hypothesis that the population proportion for each of the categories is as claimed:

1. The data have been randomly selected.
2. The sample data consist of frequency counts for each of the different categories.
3. The expected frequency is at least 5. (There is no requirement that the observed frequency for each category must be at least 5.)

---

---

---

---

---

---

---

---

## Test Statistic

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

## Critical Values

1. Found in Table A-4 using  $k-1$  degrees of freedom  
where  $k$  = number of categories
2. Goodness-of-fit hypothesis tests are always right-tailed.

---

---

---

---

---

---

---

---

## Multinomial Experiment: Goodness-of-Fit Test

$H_0$ : No difference between  
observed and expected  
probabilities

$H_1$ : at least one of the  
probabilities is different  
from the others

---

---

---

---

---

---

---

---

❖ A close agreement between observed and expected values will lead to a small value of  $\chi^2$  and a large  $P$ -value.

❖ A large disagreement between observed and expected values will lead to a large value of  $\chi^2$  and a small  $P$ -value.

❖ A significantly large value of  $\chi^2$  will cause a rejection of the null hypothesis of no difference between the observed and the expected.

---

---

---

---

---

---

---

---

**Relationships Among Components in Goodness-of-Fit Hypothesis Test**

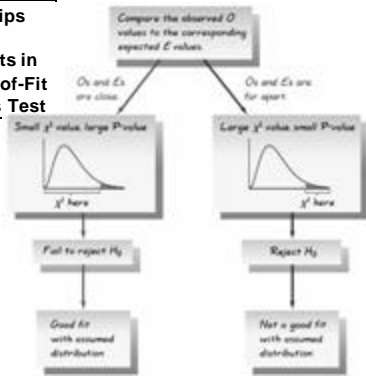


Figure 11-3

---

---

---

---

---

---

---

---

---

---

---

---

**Categories with Equal Frequencies**

(Probabilities)

$$H_0: p_1 = p_2 = p_3 = \dots = p_k$$

$H_1$ : at least one of the probabilities is different from the others

---

---

---

---

---

---

---

---

---

---

---

---

**Example:** A study was made of 147 industrial accidents that required medical attention. Test the claim that the accidents occur with equal proportions on the 5 workdays.

Frequency of Accidents					
Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31

---

---

---

---

---

---

---

---

---

---

---

---

**Example:** A study was made of 147 industrial accidents that required medical attention. Test the claim that the accidents occur with uniform distribution on the 5 workdays.

Frequency of Accidents					
Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31

---

---

---

---

---

---

---

---

---

---

**Example:** A study was made of 147 industrial accidents that required medical attention. Test the claim that the accidents occur with equal proportions on the 5 workdays.

Frequency of Accidents					
Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31

Claim: Accidents occur with the same proportion (frequency); that is,  $p_1 = p_2 = p_3 = p_4 = p_5$

---

---

---

---

---

---

---

---

---

---

**Example:** A study was made of 147 industrial accidents that required medical attention. Test the claim that the accidents occur with equal proportions on the 5 workdays.

Frequency of Accidents					
Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31

Claim: Accidents occur with the same proportion (frequency); that is,  $p_1 = p_2 = p_3 = p_4 = p_5$

$H_0: p_1 = p_2 = p_3 = p_4 = p_5$

---

---

---

---

---

---

---

---

---

---

**Example:** A study was made of 147 industrial accidents that required medical attention. Test the claim that the accidents occur with equal proportions on the 5 workdays.

Frequency of Accidents					
Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31

Claim: Accidents occur with the same proportion (frequency); that is,  $p_1 = p_2 = p_3 = p_4 = p_5$

$H_0$ :  $p_1 = p_2 = p_3 = p_4 = p_5$

$H_1$ : At least 1 of the 5 proportions is different from others

---

---

---

---

---

---

---

---

---

---

**Example:** A study was made of 147 industrial accidents that required medical attention. Test the claim that the accidents occur with equal proportions on the 5 workdays.

Frequency of Accidents					
Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31

$$E = n/k = 147/5 = 29.4$$

---

---

---

---

---

---

---

---

---

---

**Example:** A study was made of 147 industrial accidents that required medical attention. Test the claim that the accidents occur with equal proportions on the 5 workdays.

Frequency of Accidents					
Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31

$$E = n/k = 147/5 = 29.4$$

Observed and Expected Frequencies					
Day	Mon	Tues	Wed	Thurs	Fri
O: Observed accidents	31	42	18	25	31
E: Expected accidents	29.4	29.4	29.4	29.4	29.4

---

---

---

---

---

---

---

---

---

---

## Multinomial Experiment

### Goodness-of-fit Test Test Statistic

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

---

---

---

---

---

---

---

---

---

---

**Observed and Expected Frequencies of Industrial Accidents**

Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31
Expected accidents	29.4	29.4	29.4	29.4	29.4
(O - E) <sup>2</sup> /E					

$$\frac{(O - E)^2}{E} = \frac{(31 - 29.4)^2}{29.4} = 0.0871$$

---

---

---

---

---

---

---

---

---

---

**Observed and Expected Frequencies of Industrial Accidents**

Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31
Expected accidents	29.4	29.4	29.4	29.4	29.4
(O - E) <sup>2</sup> /E	0.0871				

$$\frac{(O - E)^2}{E} = \frac{(31 - 29.4)^2}{29.4} = 0.0871$$

---

---

---

---

---

---

---

---

---

---

**Observed and Expected Frequencies of Industrial Accidents**

Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31
Expected accidents	29.4	29.4	29.4	29.4	29.4
$(O - E)^2/E$	0.0871	5.4000	4.4204	0.6585	0.0871 (rounded)

---

---

---

---

---

---

---

---

**Observed and Expected Frequencies of Industrial Accidents**

Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31
Expected accidents	29.4	29.4	29.4	29.4	29.4
$(O - E)^2/E$	0.0871	5.4000	4.4204	0.6585	0.0871 (rounded)

**Test Statistic**

$$X^2 = \sum \frac{(O - E)^2}{E} = 0.0871 + 5.4000 + 4.4204 + 0.6585 + 0.0871 = 10.6531$$

---

---

---

---

---

---

---

---

## Multinomial Experiments Goodness-of-fit Test

### Critical Values

1. Found in Table A-4 using  $k - 1$  degrees of freedom  
where  $k$  = number of categories
2. Goodness-of-fit hypothesis tests are always right-tailed.

---

---

---

---

---

---

---

---

**Observed and Expected Frequencies of Industrial Accidents**

Day	Mon	Tues	Wed	Thurs	Fri
Observed accidents	31	42	18	25	31
Expected accidents	29.4	29.4	29.4	29.4	29.4
$(O - E)^2/E$	0.0871	5.4000	4.4204	0.6585	0.0871 (rounded)

**Test Statistic:**

$$\chi^2 = \sum \frac{(O - E)^2}{E} = 0.0871 + 5.4000 + 4.4204 + 0.6585 + 0.0871 = 10.6531$$

**Critical Value:  $\chi^2 = 9.488$**

**Table A-4 with  $k-1 = 5 - 1 = 4$   
and  $\alpha = 0.05$**

---

---

---

---

---

---

---

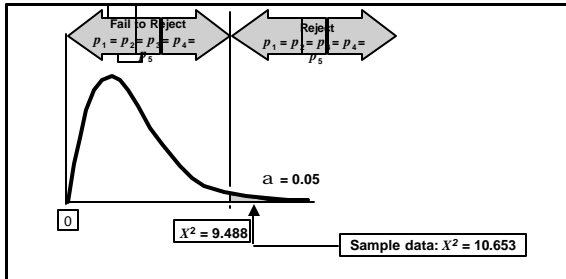
---

---

---

---

---




---

---

---

---

---

---

---

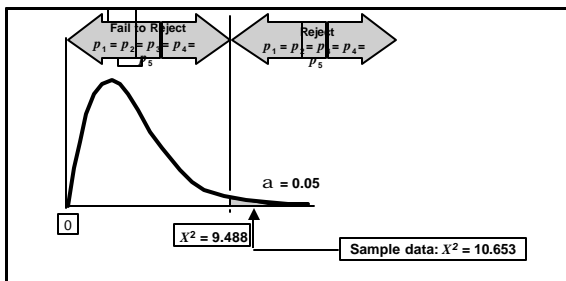
---

---

---

---

---



**Test Statistic falls within the critical region: REJECT the null hypothesis**

---

---

---

---

---

---

---

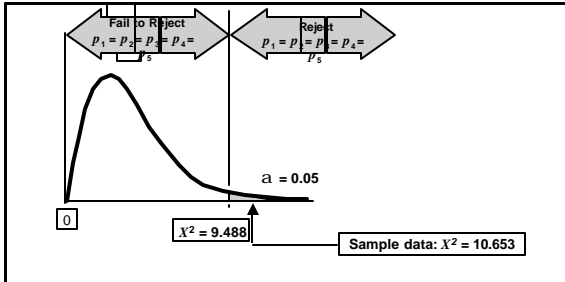
---

---

---

---

---



Test Statistic falls within the critical region: **REJECT** the null hypothesis

Claim: Accidents occur with the same proportion (frequency); that is,  $p_1 = p_2 = p_3 = p_4 = p_5$

$H_0$ :  $p_1 = p_2 = p_3 = p_4 = p_5$

$H_1$ : At least 1 of the 5 proportions is different from others

Triola, Essentials of Statistics, Third Edition. Copyright 2008. Pearson Education, Inc.

---

---

---

---

---

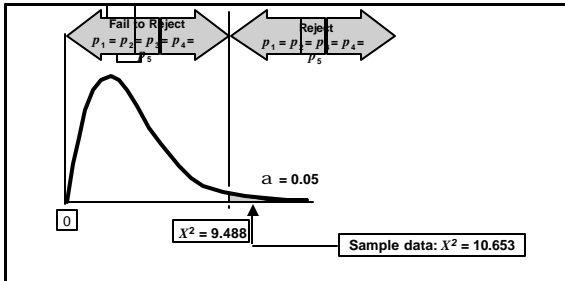
---

---

---

---

---



Test Statistic falls within the critical region: **REJECT** the null hypothesis

We reject claim that the accidents occur with equal proportions (frequency) on the 5 workdays. (Although it appears Wednesday has a lower accident rate, arriving at such a conclusion would require other methods of analysis.)

Triola, Essentials of Statistics, Third Edition. Copyright 2008. Pearson Education, Inc.

---

---

---

---

---

---

---

---

---

---

## Categories with Unequal Frequencies

(Probabilities)

$H_0$ :  $p_1, p_2, p_3, \dots, p_k$  are as claimed

$H_1$ : at least one of the above proportions is different from the claimed value

Triola, Essentials of Statistics, Third Edition. Copyright 2008. Pearson Education, Inc.

---

---

---

---

---

---

---

---

---

---

**Example:** Mars, Inc. claims its M&M candies are distributed with the color percentages of 30% brown, 20% yellow, 20% red, 10% orange, 10% green, and 10% blue. At the 0.05 significance level, test the claim that the color distribution is as claimed by Mars, Inc.

---

---

---

---

---

---

---

---

---

---

**Example:** Mars, Inc. claims its M&M candies are distributed with the color percentages of 30% brown, 20% yellow, 20% red, 10% orange, 10% green, and 10% blue. At the 0.05 significance level, test the claim that the color distribution is as claimed by Mars, Inc.

**Claim:**  $p_1 = 0.30, p_2 = 0.20, p_3 = 0.20, p_4 = 0.10,$   
 $p_5 = 0.10, p_6 = 0.10$

$H_0 :$   $p_1 = 0.30, p_2 = 0.20, p_3 = 0.20, p_4 = 0.10,$   
 $p_5 = 0.10, p_6 = 0.10$

$H_1 :$  At least one of the proportions is different from the claimed value.

---

---

---

---

---

---

---

---

---

---

**Example:** Mars, Inc. claims its M&M candies are distributed with the color percentages of 30% brown, 20% yellow, 20% red, 10% orange, 10% green, and 10% blue. At the 0.05 significance level, test the claim that the color distribution is as claimed by Mars, Inc.

Frequencies of M&Ms						
	Brown	Yellow	Red	Orange	Green	Blue
Observed frequency	33	26	21	8	7	5

$n = 100$

---

---

---

---

---

---

---

---

---

---

**Example:** Mars, Inc. claims its M&M candies are distributed with the color percentages of 30% brown, 20% yellow, 20% red, 10% orange, 10% green, and 10% blue. At the 0.05 significance level, test the claim that the color distribution is as claimed by Mars, Inc.

Frequencies of M&Ms						
	Brown	Yellow	Red	Orange	Green	Blue
Observed frequency	33	26	21	8	7	5

$n = 100$   
 Brown  $E = np = (100)(0.30) = 30$   
 Yellow  $E = np = (100)(0.20) = 20$   
 Red  $E = np = (100)(0.20) = 20$   
 Orange  $E = np = (100)(0.10) = 10$   
 Green  $E = np = (100)(0.10) = 10$   
 Blue  $E = np = (100)(0.10) = 10$

---

---

---

---

---

---

---

---

---

---

---

---

Frequencies of M&Ms						
	Brown	Yellow	Red	Orange	Green	Blue
Observed frequency	33	26	21	8	7	5
Expected frequency	30	20	20	10	10	10

---

---

---

---

---

---

---

---

---

---

---

---

Frequencies of M&Ms						
	Brown	Yellow	Red	Orange	Green	Blue
Observed frequency	33	26	21	8	7	5
Expected frequency	30	20	20	10	10	10
$(O - E)^2/E$	0.3	1.8	0.05	0.4	0.9	2.5

---

---

---

---

---

---

---

---

---

---

---

---

**Frequencies of M&Ms**

	Brown	Yellow	Red	Orange	Green	Blue
Observed frequency	33	26	21	8	7	5
Expected frequency	30	20	20	10	10	10
$(O - E)^2/E$	0.3	1.8	0.05	0.4	0.9	2.5

**Test Statistic**

$$X^2 = \sum \frac{(O - E)^2}{E} = 5.95$$

---

---

---

---

---

---

---

---

---

---

**Frequencies of M&Ms**

	Brown	Yellow	Red	Orange	Green	Blue
Observed frequency	33	26	21	8	7	5
Expected frequency	30	20	20	10	10	10
$(O - E)^2/E$	0.3	1.8	0.05	0.4	0.9	2.5

**Test Statistic**

$$X^2 = \sum \frac{(O - E)^2}{E} = 5.95$$

**Critical Value  $X^2 = 11.071$   
(with  $k-1 = 5$  and  $\alpha = 0.05$ )**

---

---

---

---

---

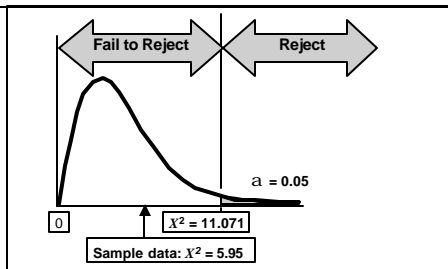
---

---

---

---

---



Test Statistic does not fall within critical region;  
Fail to reject  $H_0$ ; percentages are as claimed

There is not sufficient evidence to warrant rejection of the claim that the colors are distributed with the given percentages.

---

---

---

---

---

---

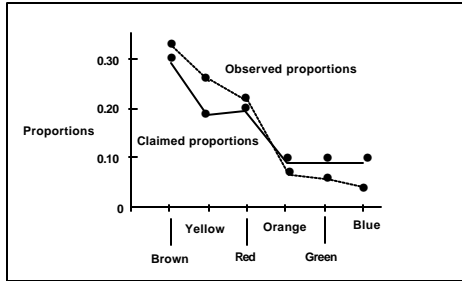
---

---

---

---

## Comparison of Claimed and Observed Proportions



Triola, Essentials of Statistics, Third Edition. Copyright 2008. Pearson Education, Inc.

---

---

---

---

---

---

---

---

---

---



Triola, Essentials of Statistics, Third Edition. Copyright 2008. Pearson Education, Inc.

---

---

---

---

---

---

---

---

---

---