Chapter 12
Tests of Goodness of Fit, Independence, and Multiple Proportions

- Goodness of Fit Test
- Test of Independence
- Testing for Equality of Three or More Population Proportions

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Tests of Goodness of Fit, Independence, and Multiple Proportions

- In this chapter we introduce three additional hypothesis-testing procedures.
- The test statistic and the distribution used are based on the chi-square (χ^2) distribution.
- In all cases, the data are categorical.

CENGAGE
© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Goodness of Fit Test: Multinomial Probability Distribution (1 of 4)

1. State the null and alternative hypotheses.

H_0: The population follows a multinomial distribution with specified probabilities for each of the *k* categories

H_a: The population does not follow a multinomial distribution with specified probabilities for each of the *k* categories

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Goodness of Fit Test: Multinomial Probability Distribution (2 of 4)

2. Select a random sample and record the observed frequency, f_i , for each of the k categories.

3. Assuming H_0 is true, compute the expected frequency, e_i , in each category by multiplying the category probability by the sample size

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Goodness of Fit Test: Multinomial Probability Distribution (3 of 4)

4. Compute the value of the test statistic.

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

where:

f_i = observed frequency for category i

e_i = expected frequency for category i

k = number of categories

Note: The test statistic has a chi-square distribution with $k - 1$ df provided that the expected frequencies are 5 or more for all categories.

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Goodness of Fit Test: Multinomial Probability Distribution (4 of 4)

5. Rejection rule:

   p-value approach: Reject H_0 if *p*-value $\leq \alpha$

   Critical-value approach: Reject H_0 if $\chi^2 \geq \chi^2_\alpha$

   where α is the significance level and there are $k - 1$ degrees of freedom.

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multinomial Distribution Goodness of Fit Test (1 of 10)

Example: Scott Marketing Research

A market share study conducted by Scott Marketing research has identified that the market for product X is shared by three companies: A, B and C. Company C plans to introduce a new and improved product to replace its current entry in the market. It wants Scott Marketing research to determine whether the new product will alter the market share for the three companies.

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multinomial Distribution Goodness of Fit Test (2 of 10)

Example: Scott Marketing Research

Using the historical market shares, we have a multinomial probability distribution with $p_A = .30$, $p_B = .50$, $p_C = .20$. Scott Marketing research conducts a sample study using a consumer panel of 200 customers.

A hypothesis test can be used to determine whether the new product of company C is likely to change the historical market shares for the three companies. We will use an $\alpha = .05$ level of significance.

CENGAGE
© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multinomial Distribution Goodness of Fit Test (3 of 10)

Example: Scott Marketing Research

- **Hypotheses**

H_0: $p_A = .30,\ p_B = .50,\ p_C = .20$

H_a: The probabilities are not $p_A = .30,\ p_B = .50,\ p_C = .20$

where:

$p_A =$ probability a customer purchases the company A product

$p_B =$ probability a customer purchases the company B product

$p_C =$ probability a customer purchases the company C product

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multinomial Distribution Goodness of Fit Test (4 of 10)

Example: Scott Marketing Research

- **Expected Frequencies**

Category	**Expected Frequency**
Company A	200(.30) = 60
Company B	200 (.50) = 100
Company C	200 (.20) = 40
Total	200

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multinomial Distribution Goodness of Fit Test (5 of 10)

Example: Scott Marketing Research

- **Observed Frequencies** (from the sample study)

Category	Observed Frequency
Company A	48
Company B	98
Company C	51
Total	200

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multinomial Distribution Goodness of Fit Test (6 of 10)

Example: Scott Marketing Research

- **Test Statistic**

$$\chi^2 = \frac{(60-48)^2}{60} + \frac{(100-98)^2}{100} + \frac{(40-51)^2}{40}$$

$$= 2.4 + .04 + 4.90$$

$$\chi^2 = 7.34$$

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multinomial Distribution Goodness of Fit Test (7 of 10)

Example: Scott Marketing Research

Conclusion Using the *p*-Value Approach

Area in Upper Tail	.10	.05	.025	.01	.005
χ^2 Value (2 *df*)	4.605	5.991	7.378	9.210	10.597

$\chi^2 = 7.34$

Because $\chi^2 = 7.34$ is between 5.991 and 7.378, the area in the upper tail of the distribution is between .05 and .025.

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

We reject the null hypothesis.

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multinomial Distribution Goodness of Fit Test (8 of 10)

Example: Scott Marketing Research

Conclusion Using the Critical-Value Approach

With $\alpha = .05$ and 2 degrees of freedom, the critical value for the test statistic is $\chi^2 = 5.991$.

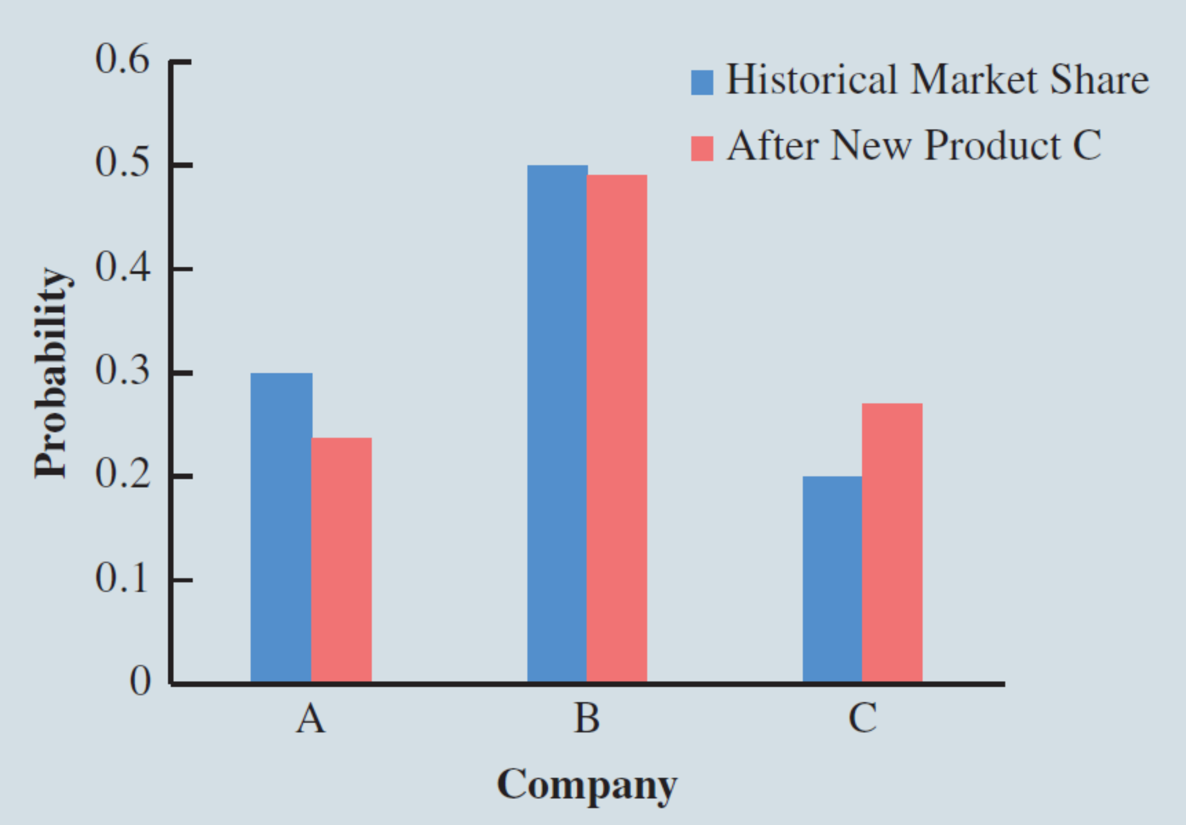
Rejection rule

Reject H_0 if $\chi^2 \geq 5.991$. With 7.34 > 5.991, we reject H_0.

CENGAGE
© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multinomial Distribution Goodness of Fit Test (9 of 10)

Example: Scott Marketing Research bar chart of market shares by company before and after the new product for company C.



CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multinomial Distribution Goodness of Fit Test (10 of 10)

<u>Example:</u> Scott Marketing Research: Excel Worksheet

	A	B	C	D	E	F
1	Customer	Preference				
2	1	B				
3	2	A		Categories	Obs. Frequency	
4	3	C		A	48	
5	4	C		B	98	
6	5	C		C	54	
7	6	A		Total	200	
8	7	A				
9	8	A		Hyp. Probability	Exp. Frequency	
10	9	C		0.3	=D10*E7	
11	10	A		0.5	=D11*E7	
12	11	C		0.2	=D12*E7	
13	12	B				
14	13	C		*p* -value	=CHISQ.TEST(E4:E6,E10:E12)	
15	14	A				
200	199	C				
201	200	C				
202						

	A	B	C	D	E	F
1	Customer	Preference				
2	1	B				
3	2	A		Categories	Obs. Frequency	
4	3	C		A	48	
5	4	C		B	98	
6	5	C		C	54	
7	6	A		Total	200	
8	7	A				
9	8	A		Hyp. Probability	Exp. Frequency	
10	9	C		0.3	60	
11	10	A		0.5	100	
12	11	C		0.2	40	
13	12	B				
14	13	C		*p* -value	0.0255	
15	14	A				
200	199	C				
201	200	C				
202						

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Test of Independence (1 of 11)

1. Set up the null and alternative hypotheses.

   H_0: The two categorical variables are independent

   H_a: The two categorical variables are not independent

2. Select a random sample and record the observed frequency, f_{ij} , for each cell of the contingency table.

3. Compute the expected frequency, e_{ij} , for each cell.

$$e_{ij} = \frac{(\text{Row } i \text{ Total})(\text{Column } j \text{ Total})}{\text{Sample Size}}$$

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Test of Independence (2 of 11)

4. Compute the test statistic.

$$\chi^2 = \sum_i \sum_i \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$$

5. Determine the rejection rule.

p-value approach: Reject H_0 if p-value $\leq \alpha$

Critical-value approach: Reject H_0 if $\chi^2 \geq \chi^2_\alpha$

where α is the significance level and, with r rows and c columns, there are $(r - 1)(c - 1)$ degrees of freedom.

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Test of Independence (3 of 11)

<u>Example:</u> Beer Industry Association

A sample of 200 beer drinkers is taken. The sample result is summarised in the following table.

- Observed frequencies

		Gender		
		Male	**Female**	**Total**
Beer Preference	**Light**	51	39	**90**
	Regular	56	21	**77**
	Dark	25	8	**33**
	Total	**132**	**68**	**200**

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Test of Independence (4 of 11)

Example: Beer Industry Association

- Hypotheses

  H_0: Beer preference is independent of gender

  H_a: Beer preference is not independent of gender

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Test of Independence (5 of 11)

Example: Beer Industry Association

- Expected frequencies

Gender				
		Male	**Female**	**Total**
Beer Preference	**Light**	59.4	30.6	**90**
	Regular	50.82	26.18	**77**
	Dark	21.78	11.22	**33**
	Total	**132**	**68**	**200**

CENGAGE
© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Test of Independence (6 of 11)

Example: Beer Industry Association

- Rejection Rule

$$\text{with } \alpha = .05 \text{ and } (3-1)(2-1) = 2 \text{ d.f., } \chi^2_\alpha = 5.991$$

$$\text{Reject } H_0 \text{ if } p\text{-value} \le .05 \text{ or } \chi^2 \ge 5.991$$

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Test of Independence (7 of 11)

Example: Beer Industry Association

- Computation of Chi-square test statistic

Beer Preference	Gender	Observed frequency f_{ij}	Expected frequency e_{ij}	Difference $(f_{ij} - e_{ij})$	Squared difference $(f_{ij} - e_{ij})^2$	Squared difference / Expected frequency
Light	Male	51	59.4	-8.4	70.56	1.19
Light	Female	39	30.6	8.4	70.56	2.31
Regular	Male	56	50.82	5.18	26.83	.53
Regular	Female	21	26.18	-5.18	26.83	1.02
Dark	Male	25	21.78	3.22	10.37	.48
Dark	Female	8	11.22	-3.22	10.37	.92
	Total	200	200.00			χ^2 = 6.45

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Test of Independence (8 of 11)

Example: Beer Industry Association

- Conclusion Using the *p*-Value Approach

Area in Upper Tail	.10	.05	.025	.01	.005
χ^2 Value (2 *df*)	4.605	5.991	7.378	9.210	10.597

$\chi^2 = 6.45$

Because $\chi^2 = 6.45$ is between 5.991 and 7.378, the area in the upper tail of the distribution is between .05 and .025.

The *p*-value $\le \alpha$. We can reject the null hypothesis.

(Actual *p*-value is .0398.)

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Test of Independence (9 of 11)

<u>Example:</u> Beer Industry Association

- Conclusion Using the Critical-Value Approach

$$\chi^2 = 6.45 \geq 5.991$$

We reject, at the .05 level of significance, the assumption that the beer preference is independent of the gender of the beer drinker.

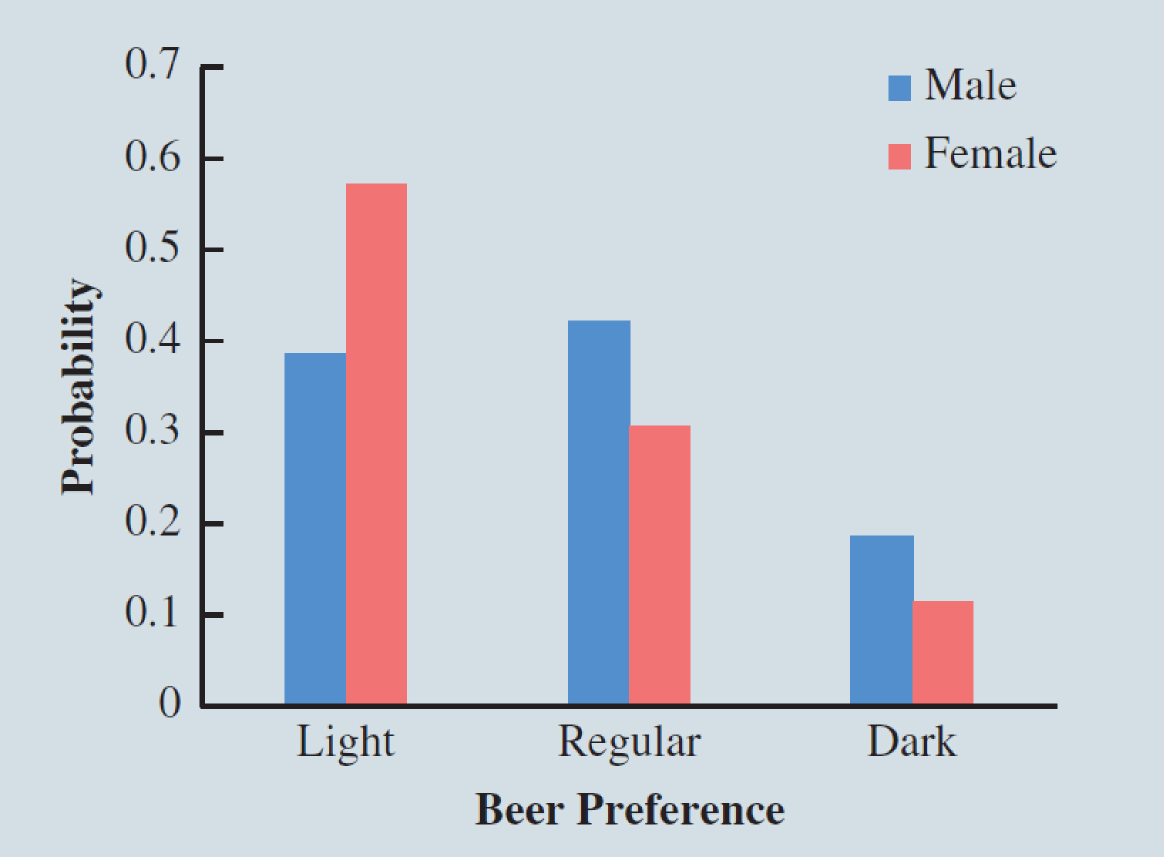
CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Test of Independence (10 of 11)

<u>Example:</u> Beer Industry Association

- Bar chart comparison of beer preference by gender



© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Test of Independence (11 of 11)

- **Example:** Beer Industry Association: Excel Worksheet

	A	B	C	D	E	F	G	H	I
1	Beer Drinker	Preference	Gender						
2	1	Regular	Male						
3	2	Light	Female		Count of Beer Drinker	Gender			
4	3	Regular	Male		Preference	Male	Female	Total	
5	4	Regular	Male		Light	51	39	90	
6	5	Regular	Female		Regular	56	21	77	
7	6	Regular	Male		Dark	25	8	33	
8	7	Dark	Male		Total	132	68	200	
9	8	Dark	Male						
10	9	Dark	Male			Gender			
11	10	Light	Female		Preference	Male	Female		
12	11	Light	Male		Light	=(H5*F$8)/H$8	=(H5*G$8)/H$8		
13	12	Dark	Female		Regular	=(H6*F$8)/H$8	=(H6*G$8)/H$8		
14	13	Regular	Male		Dark	=(H7*F$8)/H$8	=(H7*G$8)/H$8		
15	14	Regular	Male						
16	15	Light	Male				=CHISQ.TEST(F5:G7,F12:G14)		
17	16	Regular	Male						
200	199	Light	Male						
201	200	Light	Male						
202									

	A	B	C	D	E	F	G	H	I
1	Beer Drinker	Preference	Gender						
2	1	Regular	Male						
3	2	Light	Female		Count of Beer Drinker	Gender			
4	3	Regular	Male		Preference	Male	Female	Total	
5	4	Regular	Male		Light	51	39	90	
6	5	Regular	Female		Regular	56	21	77	
7	6	Regular	Male		Dark	25	8	33	
8	7	Dark	Male		Total	132	68	200	
9	8	Dark	Male						
10	9	Dark	Male			Gender			
11	10	Light	Female		Preference	Male	Female		
12	11	Light	Male		Light	59.40	30.60		
13	12	Dark	Female		Regular	50.82	26.18		
14	13	Regular	Male		Dark	21.78	11.22		
15	14	Regular	Male						
16	15	Light	Male				0.0398		
17	16	Regular	Male						
200	199	Light	Male						
201	200	Light	Male						
202									

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing for Equality of Population Proportions for Three or More Populations

Using the notation

$p_1 =$ population proportion for population 1

$p_2 =$ population proportion for population 2

$p_k =$ population proportion for population k

The hypotheses for the equality of population proportions for $k \geq 3$ populations are as follows:

$H_0: p_1 = p_2 = \ldots = p_k$

H_a: Not all population proportions are equal.

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (1 of 13)

- If H_0 cannot be rejected, we cannot detect a difference among the k population proportions.
- If H_0 can be rejected, we can conclude that not all k population proportions are equal.
- Further analyses can be done to conclude which population proportions are significantly different from others.

CENGAGE
© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (2 of 13)

Example: Customer loyalty for automobiles

Suppose in a particular study we want to compare the customer loyalty for three automobiles: Chevrolet Impala, Ford Fusion and Honda Accord.

$p_1 =$ proportion likely to repurchase for the population of Chevrolet Impala owners

$p_2 =$ proportion likely to repurchase for the population of Ford Fusion owners

$p_3 =$ proportion likely to repurchase for the population of Honda Accord owners

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (3 of 13)

Example: Customer loyalty for automobiles

- We begin by taking a sample of owners from each of the three populations.
- Each sample contains categorical data indicating whether the respondents are likely or not likely to repurchase the automobile.

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (4 of 13)

Example: Customer loyalty for automobiles

- Observed frequencies (Sample results)

Automobile Owners					
		Chevrolet Impala	**Ford Fusion**	**Honda Accord**	**Total**
Likely to Repurchase	**Yes**	69	120	123	**312**
	No	56	80	52	**188**
	Total	**125**	**200**	**175**	**500**

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (5 of 13)

- Next, we determine the expected frequencies under the assumption H_0 is correct

  Expected Frequencies under the assumption H_0 is true:

$$e_{ij} = \frac{(\text{Row } i \text{ Total})(\text{Column } j \text{ Total})}{\text{Sum of Sample Sizes}}$$

- If a significant difference exists between the observed and expected frequencies, H_0 can be rejected.

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (6 of 13)

Example: Customer loyalty for automobiles

- Expected frequencies (Computed)

		Automobile Owners			
		Chevrolet Impala	**Ford Fusion**	**Honda Accord**	**Total**
Likely to Repurchase	**Yes**	78	124.8	109.2	**312**
	No	47	75.2	65.8	**188**
	Total	**125**	**200**	**175**	**500**

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (7 of 13)

- Next, compute the value of the chi-square test statistic.

$$\chi^2 = \sum_i \sum_j \frac{\left(f_{ij} - e_{ij}\right)^2}{e_{ij}}$$

where

f_{ij} = observed frequency for the cell in row i and column j

e_{ij} = expected frequency for the cell in row i and column j under the assumption H_0 is true

Note: The test statistic has a chi-square distribution with $k - 1$ degrees of freedom, provided the expected frequency is 5 or more for each cell.

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (8 of 13)

- Computation of the Chi-Square Test Statistic.

Likely to repurchase	Automobile owner	Observed frequency f_{ij}	Expected frequency e_{ij}	Difference $(f_{ij} - e_{ij})$	Squared difference $(f_{ij} - e_{ij})^2$	Squared difference/ Expected frequency
Yes	Impala	69	78.0	–9.0	81.00	1.04
Yes	Fusion	120	124.8	–4.8	23.04	0.18
Yes	Accord	123	109.2	13.8	190.44	1.74
No	Impala	56	47.0	9.0	81.00	1.72
No	Fusion	80	75.2	4.8	23.04	0.31
No	Accord	52	65.8	-13.8	190.44	2.89
		500	500.0			$\chi^2 = 7.89$

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (9 of 13)

- Rejection Rule

p-value approach: Reject H_0 if p-value $\leq \alpha$.

Critical-value approach: Reject H_0 if $\chi^2 \geq \chi^2_\alpha$ where α is the significance level and there are $k-1$ degrees of freedom.

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (10 of 13)

Example: Customer loyalty for automobiles

- Conclusion Using the *p*-Value Approach

Area in Upper Tail	.10	.05	.025	.01	.005
χ^2 Value (2 *df*)	4.605	5.991	7.378	9.210	10.597

$\chi^2 = 7.89$ (between 7.378 and 9.210)

Because $\chi^2 = 7.89$ is between 7.378 and 9.210, the area in the upper tail of the distribution is between .025 and .01.

The *p*-value $\leq \alpha$. We can reject the null hypothesis.

(Actual *p*-value is .0193.)

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (11 of 13)

Example: Customer loyalty for automobiles

- Conclusion using the critical-value approach

With $\alpha - .05$ and 2 degrees of freedom, the critical value for the chi-square test statistic is $\chi^2 = 5.991$.

$$\chi^2 = 7.89 > 5.991\text{: We reject the hypothesis.}$$

Conclusion: There is a difference in brand loyalties among Chevrolet Impala, Ford Fusion, and Honda accord owners.

CENGAGE
© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (12 of 13)

- Excel Worksheets

	A	B	C	D	E	F	G	H	I	J
1	Owner	Automobile	Likely Repurchase							
2	1	Chevrolet Impala	Yes							
3	2	Chevrolet Impala	Yes		Count of Owner	Automobile ▾				
4	3	Chevrolet Impala	No		Likely Repurchase ▾	Chevrolet Impala	Ford Fusion	Honda Accord	Total	
5	4	Chevrolet Impala	Yes		Yes	69	120	123	312	
6	5	Chevrolet Impala	Yes		No	56	80	52	188	
7	6	Chevrolet Impala	Yes		Total	125	200	175	500	
8	7	Chevrolet Impala	Yes							
9	8	Chevrolet Impala	Yes							
10	9	Chevrolet Impala	No			Automobile				
11	10	Chevrolet Impala	Yes		Likely Repurchase	Chevrolet Impala	Ford Fusion	Honda Accord		
12	11	Chevrolet Impala	Yes		Yes	=(I5*F7)/I7	=(I5*G7)/I7	=(I5*H7)/I7		
13	12	Chevrolet Impala	No		No	=(I6*F7)/I7	=(I6*G7)/I7	=(I6*H7)/I7		
14	13	Chevrolet Impala	Yes							
15	14	Chevrolet Impala	No				*p*-value	=CHISQ.TEST(F5:H6,F12:H13)		
16	15	Chevrolet Impala	No							
500	499	Honda Accord	No							
501	500	Honda Accord	No							
502										

	A	B	C	D	E	F	G	H	I	J
1	Owner	Automobile	Likely Repurchase							
2	1	Chevrolet Impala	Yes							
3	2	Chevrolet Impala	Yes		Count of Owner	Automobile ▾				
4	3	Chevrolet Impala	No		Likely Repurchase ▾	Chevrolet Impala	Ford Fusion	Honda Accord	Total	
5	4	Chevrolet Impala	Yes		Yes	69	120	123	312	
6	5	Chevrolet Impala	Yes		No	56	80	52	188	
7	6	Chevrolet Impala	Yes		Total	125	200	175	500	
8	7	Chevrolet Impala	Yes							
9	8	Chevrolet Impala	Yes							
10	9	Chevrolet Impala	No			Automobile				
11	10	Chevrolet Impala	Yes		Likely Repurchase	Chevrolet Impala	Ford Fusion	Honda Accord		
12	11	Chevrolet Impala	Yes		Yes	78	124.8	109.2		
13	12	Chevrolet Impala	No		No	47	75.2	65.8		
14	13	Chevrolet Impala	Yes							
15	14	Chevrolet Impala	No				*p*-value	0.0193		
16	15	Chevrolet Impala	No							
500	499	Honda Accord	No							
501	500	Honda Accord	No							
502										

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Testing the Equality of Population Proportions for Three or More Populations (13 of 13)

- We have concluded that the population proportions for the three populations of automobile owners are not equal.
- To identify where the differences between population proportions exist, we will rely on a multiple comparisons procedure.

CENGAGE
© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multiple Comparisons Procedure (1 of 4)

- We begin by computing the three sample proportions.

$$\text{Chevrolet Impala: } \bar{p}_1 = \frac{69}{125} = .5520$$

$$\text{Ford Fusion: } \bar{p}_2 = \frac{120}{200} = .6000$$

$$\text{Honda Accord: } \bar{p}_3 = \frac{123}{175} = .7029$$

- We will use a multiple comparison procedure known as the Marascuillo procedure.

CENGAGE
© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multiple Comparisons Procedure (2 of 4)

- Marascuillo Procedure

  We compute the absolute value of the pairwise difference between sample proportions.

Chevrolet Impala and Ford Fusion	$\lvert \bar{p}_1 - \bar{p}_2 \rvert = \lvert .5520 - .6000 \rvert = .0480$
Chevrolet Impala and Honda Accord	$\lvert \bar{p}_1 - \bar{p}_3 \rvert = \lvert .5520 - .7029 \rvert = .1509$
Ford Fusion and Honda Accord	$\lvert \bar{p}_1 - \bar{p}_2 \rvert = \lvert .6000 - .7029 \rvert = .1029$

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multiple Comparisons Procedure (3 of 4)

- Critical Values for the Marascuillo Pairwise Comparison

  For each pairwise comparison compute a critical value as follows:

$$CV_{ij} = \sqrt{\chi^2_{\alpha,k-1}} \sqrt{\frac{\bar{p}_i(1-\bar{p}_i)}{n_i} + \frac{\bar{p}_j(1-\bar{p}_j)}{n_j}}$$

For $\alpha = .05$ and $k = 3$: $\chi^2 = 5.991$

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.

Multiple Comparisons Procedure (4 of 4)

- Pairwise Comparison Tests

Pairwise Comparison	$\lvert \bar{p}_i - \bar{p}_j \rvert$	CV_{ij}	Significant if $\lvert \bar{p}_i - \bar{p}_j \rvert > CV_{ij}$
Chevrolet Impala and Ford Fusion	.0480	.1380	Not significant
Chevrolet Impala and Honda Accord	.1509	.1379	Significant
Ford Fusion and Honda Accord	.1029	.1198	Not significant

CENGAGE

© 2021 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, or posted to a publicly accessible website, in whole or in part.