統計學

Spring 2026

政治大學統計系余清祥
2026年02月24日
第十二章：卡方檢定
http://csyue.nccu.edu.tw

卡方檢定(Chi-Square Test)

■卡方檢定主要在處理類別資料，依序為：

→ Goodness of Fit Test (適合度檢定)

可用來檢查資料是否服從某分配，以及樣本是否與母體類似(樣本代表性)。

→Tests of Independence (獨立性檢定)

可用於問卷分析，像是檢查兩個問項間是否相關。

→Tests of Homogeneity(齊一性檢定)

可用於檢查幾個母體是否有相同比例分配。

卡方檢定的用途

https://haosquare.com/chi-squared-test/

	配適度檢定 （Goodness-of-fit Test）	同質性檢定 （Test of Homogeneity）	獨立性檢定 （Test of Independence）
用途	預期分佈 是否**適合**資料分佈	各組分佈 是否全都**相同**	兩類別變數 是否存在**關聯**
列聯表	單因子 （One-way）	二因子 （Two-way）	二因子 （Two-way）
預設比例數字	**有**	無	無
資料來源	同一群體	**兩組或多組群體**	同一群體

Chapter Contents

Introduction

In this chapter, we introduce three hypothesis testing procedures that extend our ability to make statistical inferences about populations described by categorical data.

In cases in which data are quantitative, we define categories and consider the observation count in each category.

The test statistic and the distribution used are based on the chi-square (X^2) distribution.

Chi-square tests are versatile and expand hypothesis testing with the following applications:

- Testing the equality of population proportions for three or more populations
- Testing the independence of two categorical variables
- Testing whether a probability distribution for a population follows a specific probability distribution

12.1 Chi-Square Test for the Equality of Three or More Population Proportions

1. State null and alternative hypotheses

   H_o: $p_1 = p_2 = \cdots = p_k$ with $k \geq 3$

   H_a: Not all population proportions are equal
2. Select one random sample from each population and collect data for every element in each sample. Record the *observed frequencies*, f_{ij}, with $i = 1{,}2$ and $j = 1 \ldots k$.
3. Assume the null hypothesis is true and compute the *expected frequencies*, e_{ij}.
4. If the expected frequency, e_{ij}, is 5 or more for each cell, compute the test statistic, X^2, as

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

5. Use a chi-square distribution with $\mathrm{df} = k - 1$ to compute the *p*-value or the critical value X_α^2.

CENGAGE

12.1 Testing for Automobile Brand Loyalty

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

The current owners of each of the three automobiles form the three populations for the study.

Thus, we define the three population proportions of interest as follows:

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

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

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

The hypotheses are stated as follows:

H_o: $p_1 = p_2 = p_3$

H_a: Not all population proportions are equal

An $\alpha = 0.05$ level of significance is specified for the study.

CENGAGE

12.1 The Chi-Square Test Statistic

As we saw in a previous slide, the Chi-square test statistic is

$$\boldsymbol{X^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}}} \quad \text{with } i = 1,2, j = 1 \ldots k, \text{and } k \geq 3$$

Where

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

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

Given a table of observed frequencies, f_{ij}, with *i* rows and *j* columns, the expected frequencies, e_{ij}, under the assumption H_0 is true, are calculated as follows:

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

Let us apply this methodology to the customer loyalty example.

12.1 Calculation of the Expected Frequencies

The data for samples of 125 Chevrolet Impala owners, 200 Ford Fusion owners, and 175 Honda Accord owners are summarized in the following crosstabulation.

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

e_{11}, the expected frequency of Impala owners likely to repurchase, is calculated as follows

$$e_{11} = \frac{\text{(Row 1 Total)(Column 1 Total)}}{\text{Total Sample Size}} = \frac{(312)(125)}{500} = (0.624)125 = 78$$

Calculations of the remaining e_{ij} provides the full table of expected frequencies for those customers likely to repurchase their car model if H_0 is true.

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

12.1 Calculation of the Chi-Square Statistic

A review of the expected frequencies table reveals that each $e_{ij} \geq 5$.

X_{11}^2, the chi-square component of Impala owners likely to repurchase, is calculated as follows

$$X_{11}^2 = \frac{(f_{11} - e_{11})^2}{e_{11}} = \frac{(69 - 78)^2}{78} = \frac{(-9)^2}{78} = \frac{81}{78} = 1.04$$

The chi-square test statistic, $X^2 = 7.89$, is the sum of the six chi-square components.

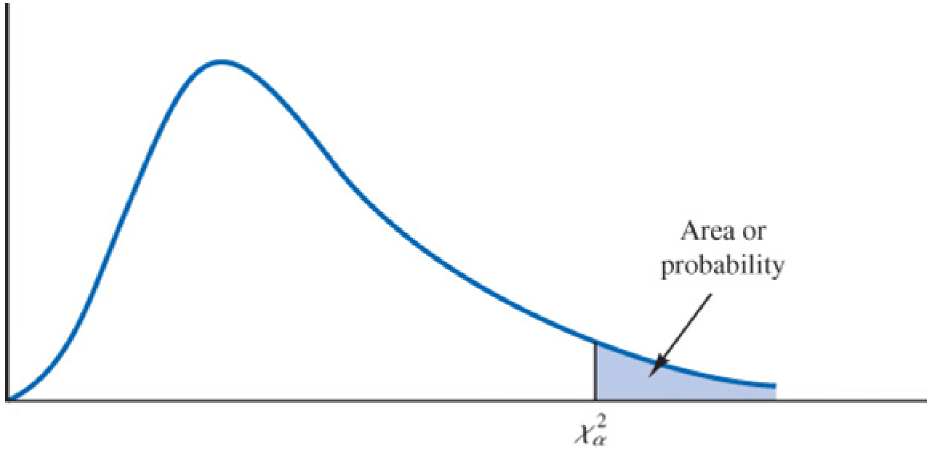
Likely to Repurchase?	Automobile Owner	Observed Frequency f_{ii}	Expected Frequency e_{ii}	Difference $f_{ii} - e_{ii}$	Squared Difference $(f_{ii} - e_{ii})^2$	Squared Difference Divided by Expected Frequency $(f_{ii} - e_{ii})^2 / e_{ii}$
Yes	Impala	69	78	-9.0	81.00	1.04
Yes	Fusion	120	124.8	-4.8	23.04	0.18
Yes	Accord	123	109.2	13.8	190.44	1.74
No	Impala	56	47	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
	Total	500	500			$X^2 = 7.89$

12.1 *p*-Value Approach to a Chi-Square Test for the Equality of Three Population Proportions

In the customer loyalty example, the *p*-value is the upper tail of a X^2 distribution with $k-1=3-1=2$ degrees of freedom: $p\text{−value} = P(X^2 \geq 7.89)$

Using Table 3 in Appendix B, the X^2 distribution with 2 degrees of freedom provides the following.

Area in upper tail	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
X^2 **Value** ($df=2$)	4.605	5.991	7.378	9.210	10.597



Because $X^2 = 7.89$ is between 7.378 and 9.210, the values in the “Area in Upper Tail” row show that the *p*-value range must be: $0.01 \leq p\text{−value} \leq 0.025$ (*see notes.)

Because p−value $\leq \alpha$, we reject the null hypothesis and conclude that the three population proportions are not all equal and that there is a difference in brand loyalties.

12.1 Critical Value Approach to a Chi-Square Test for the Equality of Three Population Proportions

With $\alpha = 0.05$, $X_{\alpha}^{2} = X_{0.05}^{2}$ provides the critical value for the upper tail chi-square test. Using Table 3 in Appendix B, with 2 degrees of freedom, we have

Area in upper tail	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
X^2 **Value** $(df = 2)$	4.605	5.991	7.378	9.210	10.597

We obtain the critical value, $\mathrm{X}_{0.05}^{2} = 5.991$ (*see notes.)

Thus, the rejection rule for the test of multiple proportions is:

Reject H_0 if $X^2 \geq \mathrm{X}_{0.05}^{2} = 5.991$

Because $X^2 = 7.89$, we reject the null hypothesis and conclude that the three population proportions are not all equal and that there is a difference in brand loyalties among the Chevrolet Impala, Ford Fusion, and Honda Accord owners.

CENGAGE

12.1 Comparison of Multiple Population Proportions

In the **Marascuilo pairwise comparison procedure**, we compute a critical value for each pair of sample proportions to identify which population proportions are different.

$$\boldsymbol{CV_{ij} = \sqrt{X_\alpha^2}\sqrt{\frac{\bar{p}_i(1-\bar{p}_i)}{n_i} + \frac{\bar{p}_j\left(1-\bar{p}_j\right)}{n_j}}}$$

Where

α is the significance level

X_α^2 is the critical value for a chi-square distribution with $k-1$ degrees of freedom, such that $P(X^2 \geq X_\alpha^2) = \alpha$

$\bar{p}_i$ and $\bar{p}_j$ are the sample proportions for populations *i* and *j*.

n_i and n_j are the sample sizes for populations *i* and *j*.

If $\left|\boldsymbol{\bar{p}_i - \bar{p}_j}\right| \geq \boldsymbol{CV_{ij}}$, the difference between population proportions p_i and p_j is significant.

CENGAGE

12.1 Calculation of the Pairwise Differences Between Sample Proportions

First, we compute the three sample proportions in the automobile brand loyalty study.

Chevrolet Impala: $\bar{p}_1 = 69/125 = 0.5520$
Ford Fusion: $\bar{p}_2 = 120/200 = 0.6000$
Honda Accord: $\bar{p}_3 = 124/175 = 0.7029$

Second, we compute the absolute value of the pairwise difference between sample proportions for each pair of populations in the study.

Chevrolet Impala and Ford Fusion: $|\bar{p}_1 - \bar{p}_2| = |0.5520 - 0.6000| = 0.0480$
Chevrolet Impala and Honda Accord: $|\bar{p}_1 - \bar{p}_3| = |0.5520 - 0.7029| = 0.1509$
Ford Fusion and Honda Accord: $|\bar{p}_2 - \bar{p}_3| = |0.6000 - 0.7029| = 0.1029$

Using $\alpha = 0.05$, and $k - 1 = 2$ degrees of freedom, we can now perform the Marascuilo procedure for the comparison of multiple population proportions.

CENGAGE

12.1 Pairwise Comparison Tests for the Automobile Brand Loyalty Study

We calculate the critical values, CV_{ij}, starting with CV_{12} for the pairwise comparison of Chevrolet Impala vs. Ford Fusion population proportions. Knowing that $X_{\alpha}^{2} = 5.991$, we have

$$CV_{12} = \sqrt{X_{\alpha}^{2}}\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_i} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_j}} = \sqrt{5.991}\sqrt{\frac{0.552(1-0.552)}{125} + \frac{0.6(1-0.6)}{200}} = 0.1380$$

Because $|\bar{p}_1 - \bar{p}_2| = 0.0480 \leq 0.1380$, p_1 and p_2 are not significantly different.

Similar calculations for the other pairwise comparisons show that the only significant comparison is between the populations of Chevrolet Impala and Honda Accord owners.

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 vs. Ford Fusion	0.0480	0.1380	Not significant
Chevrolet Impala vs. Honda Accord	0.1509	0.1379	Significant
Ford Fusion vs. Honda Accord	0.1029	0.1198	Not significant

CENGAGE

12.2 Chi-Square Test for Independence of Two Categorical Variables

1. State null and alternative hypotheses

   H_o: The two categorical variables are independent

   H_a: The two categorical variables are not independent
2. Select a random sample from the population and for every element, collect data for both variables. Record the observed frequencies, f_{ij}, in a table with *r* rows and *c* columns.
3. Assume the null hypothesis is true and compute the expected frequencies, e_{ij}.
4. If the expected frequency, e_{ij}, is 5 or more for each cell, compute the test statistic, X^2, as

$$\boldsymbol{X^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}}}$$

5. Use a chi-square distribution with $\mathrm{df} = (r-1)(c-1)$ to compute the *p*-value or X^2_α.

CENGAGE

12.2 Testing for the Independence of Type of Beer and Beer Drinker's Sex

A beer industry association conducts a survey to determine whether preference for three types of beer (light, regular, and dark) is independent of the sex of the beer drinker.

The hypotheses for this test of independence are

H_o: Beer preference is independent of drinker's sex

H_a: Beer preference is not independent of drinker's sex

An $\alpha = 0.05$ level of significance is specified for the study.

A sample of 200 beer drinkers is taken with each person in the sample asked to indicate their sex and preference for one of the three types of beers: light, regular, or dark.

The results of the survey are summarized in the crosstabulation to the right.

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

12.2 Calculation of the Expected Frequencies for the Test of Independence

Survey results indicate that $132/200 = 66\%$ of the sample beer drinkers are males, and $68/200 = 34\%$ females. Also, $90/200 = 45\%$ of the beer drinkers prefer light beer, $77/200 = 38.5\%$ regular beer, and $33/200 = 16.5\%$ prefer dark beer.

e_{11}, the expected frequency of males who drink light beer, is calculated as follows

$$e_{11} = \frac{(\text{Row 1 Total})(\text{Column 1 Total})}{\text{Total Sample Size}} = \frac{(90)(132)}{200} = (0.45)132 = 59.40$$

Calculations of the remaining e_{ij} shown to the right provides the full table of expected frequencies for preference of type of beer and sex of beer drinker under the assumption that H_0 is true.

	Sex		
Beer Preference	**Male**	**Female**	**Total**
Light	59.40	30.60	90
Regular	50.82	26.18	77
Dark	21.78	11.22	33
Total	132	68	200

12.1 Calculation of the Chi-Square Statistic for the Test of Independence

A review of the expected frequencies table reveals that each $e_{ij} \geq 5$. The chi-square component of males who drink light beer, X_{11}^2, is calculated as follows

$$X_{11}^2 = \frac{(f_{11} - e_{11})^2}{e_{11}} = \frac{(51 - 59.40)^2}{59.40} = \frac{70.56}{59.40} = 1.19$$

The chi-square test statistic, $X^2 = 6.45$, is the sum of the six chi-square components.

Beer Preference	**Sex**	**Observed Frequency** f_{ii}	**Expected Frequency** e_{ii}	**Difference** $f_{ii} - e_{ii}$	**Squared Difference** $(f_{ii} - e_{ii})^2$	**Squared Difference Divided by Expected Frequency** $(f_{ii} - e_{ii})^2 / e_{ii}$
Light	Male	51	59.40	-8.40	70.56	1.19
Light	Female	39	30.60	8.40	70.56	2.31
Regular	Male	56	50.82	5.18	26.83	0.53
Regular	Female	21	26.18	-5.18	26.83	1.02
Dark	Male	25	21.78	3.22	10.37	0.48
Dark	Female	8	11.22	-3.22	10.37	0.92
	Total	200	200			$X^2 = 6.45$

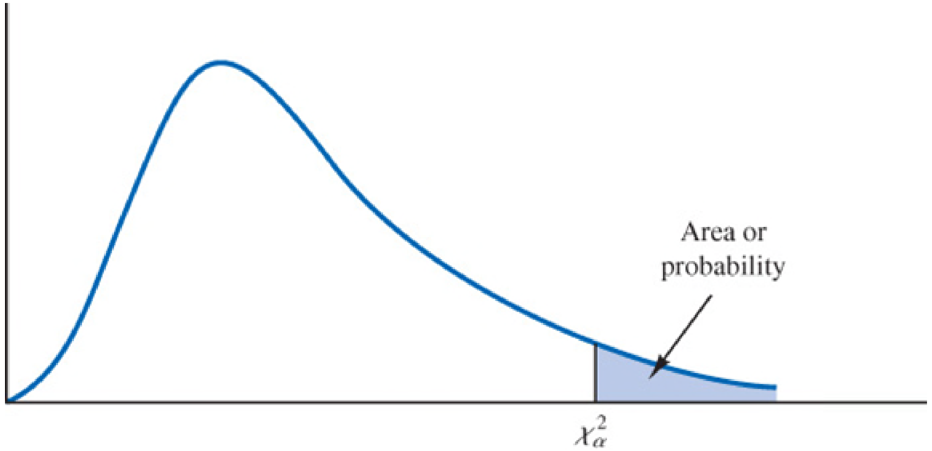
CENGAGE

12.2 *p*-Value Approach to a Chi-Square Test for Independence

In the beer preference example, the *p*-value is the upper tail of a X^2 distribution with $(r-1)(c-1) = 2$ degrees of freedom: $p\text{–value} = P(X^2 \geq 6.45)$

Using Table 3 in Appendix B, the X^2 distribution with 2 degrees of freedom provides the following.

Area in upper tail	0.10	0.05	0.025	0.01	0.005
X^2 Value $(df = 2)$	4.605	5.991	7.378	9.210	10.597



Because $X^2 = 6.45$ is between 5.991 and 7.378, the values in the “Area in Upper Tail” row show that the *p*-value range must be: $0.025 \leq p\text{–value} \leq 0.05$ (*see notes.)

Because p–value $\leq \alpha$, we reject the null hypothesis and conclude that the preference of type of beer is not independent of the sex of the beer drinker.

12.2 Critical Value Approach to a Chi-Square Test for Independence

With $\alpha = 0.05$, $X_{\alpha}^{2} = X_{0.05}^{2}$ provides the critical value for the upper tail chi-square test. Using Table 3 in Appendix B, with 2 degrees of freedom, we have

Area in upper tail	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
$\boldsymbol{X^2}$ **Value** $(\boldsymbol{df} = 2)$	4.605	5.991	7.378	9.210	10.597

We obtain the critical value, $\mathrm{X}_{0.05}^{2} = 5.991$.

We can also use Excel to compute the critical value: $=\mathrm{CHISQ.INV.RT}(0.05,2) = 5.991$.

Thus, the rejection rule for the test of independence is:

Reject H_0 if $X^2 \geq \mathrm{X}_{0.05}^{2} = 5.991$

Because $X^2 = 6.45$, we reject the null hypothesis and conclude that the preference of type of beer is not independent of the sex of the beer drinker.

CENGAGE

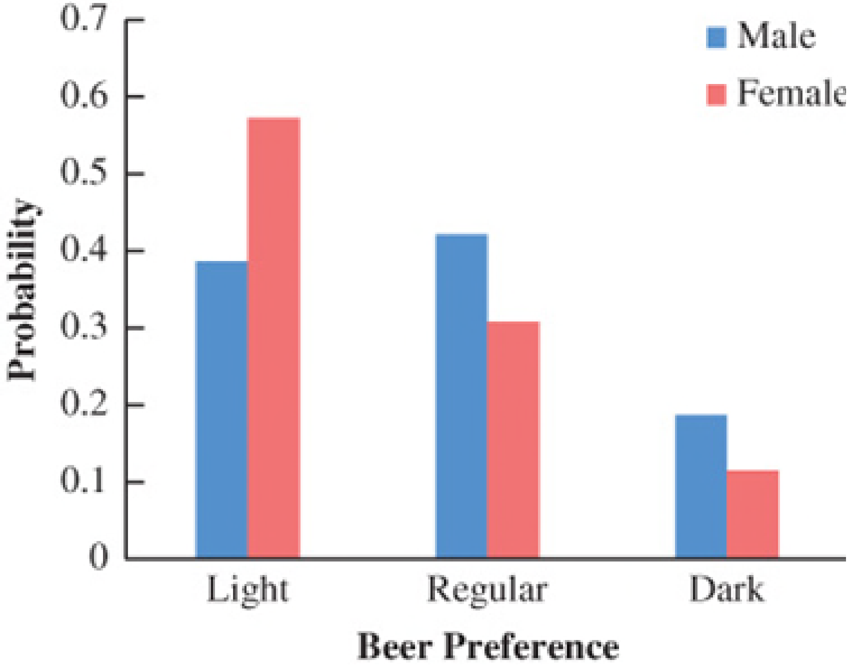
12.2 Bar Chart Comparison of Beer Preference by Sex

To assess the nature of the association between beer type and beer drinker’s sex, we compute the probability of the beer preference responses for males and females separately.

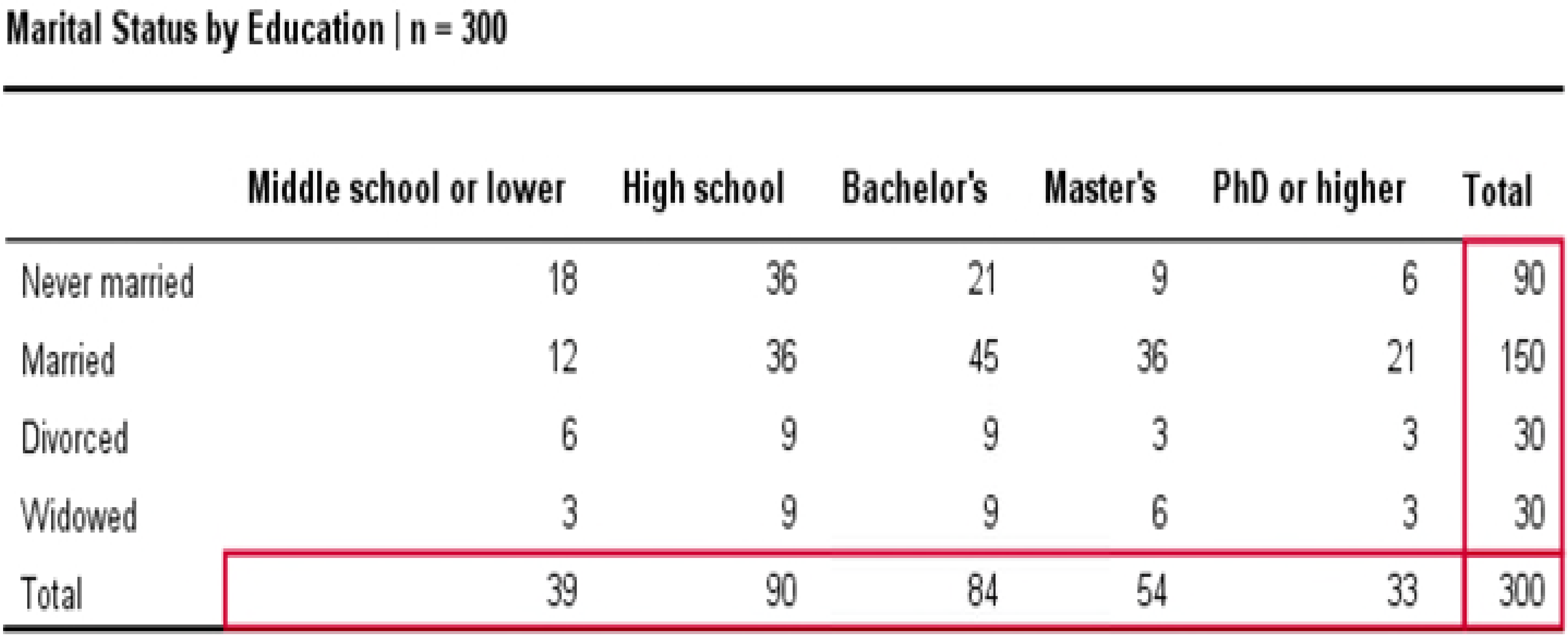
Beer Preference	**Male**	**Female**
Light	$51/132 = 38.64\%$	$39/68 = 57.35\%$
Regular	$56/132 = 42.42\%$	$21/68 = 30.88\%$
Dark	$25/132 = 18.94\%$	$8/68 = 11.76\%$

The bar chart for male and female beer drinkers of the three kinds of beer allows us to make observations about the association between beer preference and sex.

- Female beer drinkers have a higher preference for light beer than males.
- Male beer drinkers have a higher preference for both regular beer and dark beer.



- 卡方檢定通常以列聯表(Contingency Table)方式呈現，獨立性檢定可用於行列的類別不只兩種！



Marital Status by Education | n = 300

	Middle school or lower	High school	Bachelor's	Master's	PhD or higher	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

https://spss-tutorials.com/img/chi-square-test-contingency-table.png

An Example of Test of Independence

Hypotheses

H_0: Price of the home is independent of the style of the home that is purchased

H_a: Price of the home is not independent of the style of the home that is purchased

Price	Colonial	Log	Split-Level	A-frame	Total
< \$200 K	18	6	19	12	55
≥ \$200 K	12	14	16	3	45
Total	30	20	35	15	100

12.3 Multinomial Probability Distribution Goodness of Fit Test

1. State null and alternative hypotheses

   H_o: The population follows a multinomial probability distribution with probabilities, p_k, for each of the *k* categories.

   H_a: The population does not follow a multinomial probability distribution with probabilities, p_k, for each of the *k* categories.

2. Select a random sample and record the observed frequencies, f_i, for each category.
3. Assume the null hypothesis is true and compute the expected frequencies as: $e_i = np_i$.
4. If the expected frequency, e_i, is 5 or more for each cell, compute the test statistic, X^2, as

$$\boldsymbol{X^2 = \sum_i \sum_j \frac{(f_i - e_i)^2}{e_i}}$$

5. Use a chi-square distribution with $df = k - 1$ to compute *p*-value and X^2_α.

12.3 Testing a Multinomial Probability Distribution

With a **multinomial probability distribution**, each element of a population is assigned to one and only one of three or more categories.

As an example, suppose that the historical market shares for a certain product are 30% for company A, 50% for company B, and 20% for company C. Thus, the multinomial probability distribution of the product market share is $p_A = 0.30$, $p_B = 0.50$, and $p_C = 0.20$.

A marketing research firm plans to survey 200 existing product customers to determine whether the introduction of a “new and improved” product by company C will change market shares.

Sample results indicated the following product preferences: 48 customers for company A, 98 customers for company B, and 54 customers for company C.

The following hypotheses are used to test the sample results at a significance level, $\alpha = 0.05$.

H_o: $p_A = 0.30$, $p_B = 0.50$, and $p_C = 0.20$

H_a: The population proportions are not $p_A = 0.30$, $p_B = 0.50$, and $p_C = 0.20$

CENGAGE

12.3 Calculation of the Chi-Square Statistic for Goodness of Fit Test

Like other chi-square tests, the goodness of fit test is based on a comparison of observed frequencies with the expected frequencies under the assumption of a true null hypothesis.

Frequencies	**Company A**	**Company B**	**Company C**	**Total**
Observed	48	98	54	200
Expected	$200(0.30) = 60$	$200(0.50) = 100$	$200(0.20) = 40$	200

With the expected frequencies all 5 or more, the computation of the chi-square test statistic is shown in the following table, resulting in $X^2 = 7.34$.

Category	**Hypothesized Proportion**	**Observed Frequency f_i**	**Expected Frequency e_i**	**Difference $f_i - e_i$**	**Squared Difference $(f_i - e_i)^2$**	**Squared Difference Divided by Expected Frequency $(f_i - e_i)^2 / e_i$**
Company A	0.30	48	60	-12	144	2.40
Company B	0.50	98	100	-2	4	0.04
Company C	0.20	54	40	14	196	4.90
	Total	200	200			$X^2 = 7.34$

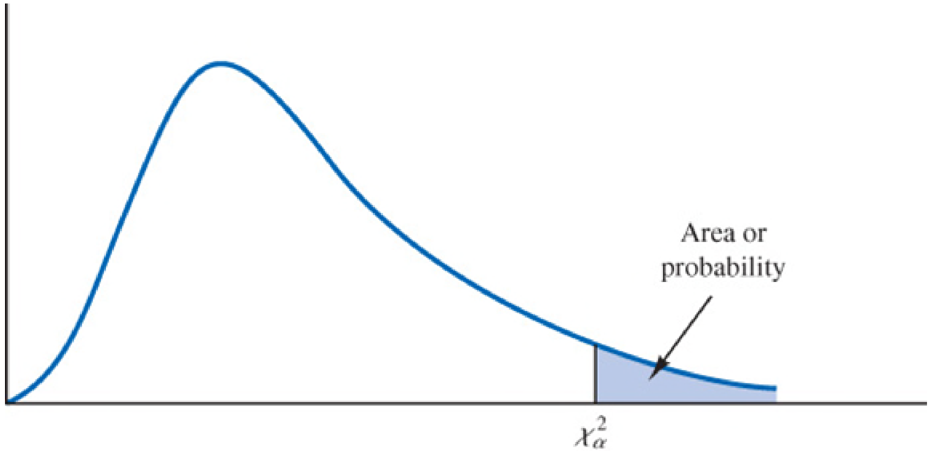
CENGAGE

12.3 *p*-Value Approach to a Goodness of Fitness Test of a Multinomial Probability Distribution

In the product market share example, the *p*-value is the upper tail of a X^2 distribution with $k = 2$ degrees of freedom: $p\text{–value} = P(X^2 \geq 7.34)$.

Using Table 3 in Appendix B, the X^2 distribution with 2 degrees of freedom provides the following

Area in upper tail	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
X^2 **Value** $(df = 2)$	4.605	5.991	7.378	9.210	10.597



Because $X^2 = 7.34$ is between 5.991 and 7.378, the values in the “Area in Upper Tail” row show that the *p*-value range must be: $0.025 \leq p\text{–value} \leq 0.05$ (*see notes.)

Because p−value $\leq \alpha$, we reject the null hypothesis and conclude that the introduction of the new product by company C will alter the historical market shares.

12.3 Critical Value Approach to a Goodness of Fitness Test of a Multinomial Probability Distribution

With $\alpha = 0.05$, $X^2_\alpha = X^2_{0.05}$ provides the critical value for the upper tail chi-square test. Using Table 3 in Appendix B, with 2 degrees of freedom, we have

Area in upper tail	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
X^2 **Value** $(df = 2)$	4.605	5.991	7.378	9.210	10.597

We obtain the critical value, $X^2_{0.05} = 5.991$.

We can also use Excel to compute the critical value: $=\text{CHISQ.INV.RT}(0.05,2) = 5.991$

Thus, the rejection rule for the goodness of fit test of a multinomial probability distribution is:

Reject H_0 if $X^2 \geq X^2_{0.05} = 5.991$

Because $X^2 = 7.34$, we reject the null hypothesis and conclude that the introduction of the new product by company C will alter the historical market shares.

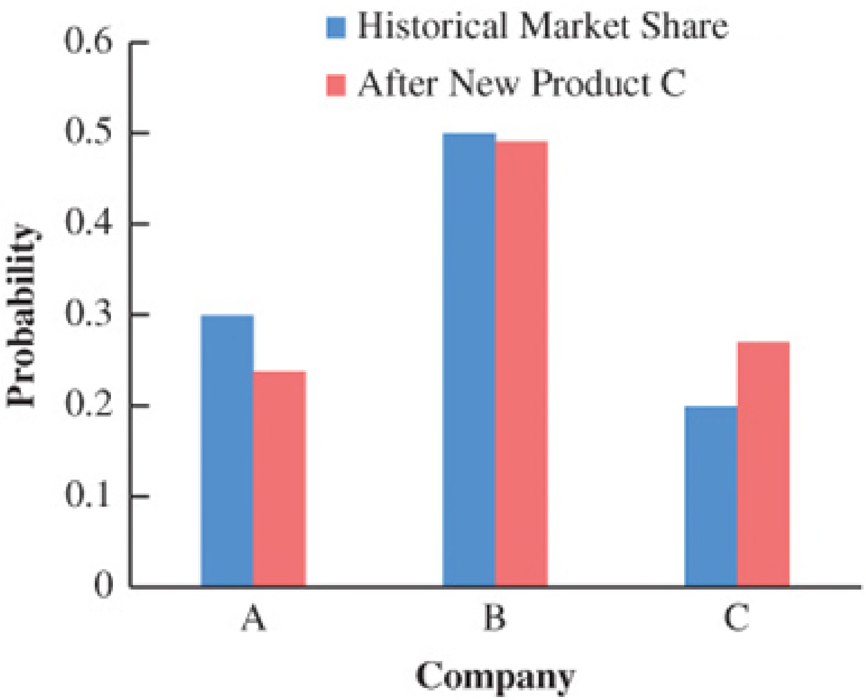
12.3 Bar Chart Comparison of Product Market Shares

Using the historical market shares and the sample data, we can summarize how the product market shares are likely to change due to the introduction of the new product by company C.

Company	Historical Market Share	Sample Data Market Share
A	30%	$48/200 = 24\%$
B	50%	$98/200 = 49\%$
C	20%	$54/200 = 27\%$

The historical market shares and the sample market shares are shown in the bar chart to the right.

- Company C will gain market share at the expense of company A.
- Market share of company B will remain relatively steady.



12.3 Normal Probability Distribution Goodness of Fit Test (1 of 2)

Because the normal probability distribution is continuous, to use a chi-square test we must modify the way the categories are defined and how the expected frequencies are computed.

To perform a goodness of fit test for a normal probability distribution, follow these steps, which for convenience have been broken down over this slide and the next one.

1. State null and alternative hypotheses

   H_o: The population has a normal probability distribution.

   H_a: The population does not have a normal probability distribution.

2. Select a random sample and
   a. Compute the sample mean and sample standard deviation
   b. Define *k* intervals of values so that the expected frequency e_i is at least five for each interval. Using equal probability intervals is a good approach.
   c. Record the observed frequency of data values f_i in each interval defined.

12.3 Normal Probability Distribution Goodness of Fit Test (2 of 2)

3. Compute the expected number of occurrences e_i for each interval of values defined in step 2b. Multiply the sample size by the probability of a normal random variable being in the interval.
4. Compute the value of the chi-square test statistic.

$$\boldsymbol{X^2 = \sum_i \sum_j \frac{(f_i - e_i)^2}{e_i}}$$

5. For the rejection rule, use a chi-square distribution with $k - 3$ (*see notes) degrees of freedom.

p-value approach:	Reject H_0 if $p\text{–value} \leq \alpha$
Critical value approach:	Reject H_0 if $X^2 \geq X_\alpha^2$

Where α is the significance level.

12.3 Testing a Normal Probability Distribution

The personnel director at Chemline wants to learn whether a normal distribution applies to the population of aptitude test scores for new job applicants (DATAfile: *Chemline.)*

The aptitude test scores for 50 randomly selected job applicants reveal the following statistics.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{3421}{50} = 68.42 \qquad s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{5310.0369}{49}} = 10.41$$

The personnel director uses the following hypotheses to test the distribution of the job applicant test scores, at a significance level, $\alpha = 0.10$.

H_o: The population of test scores has a normal distribution with mean 68.42 and standard deviation 10.41.

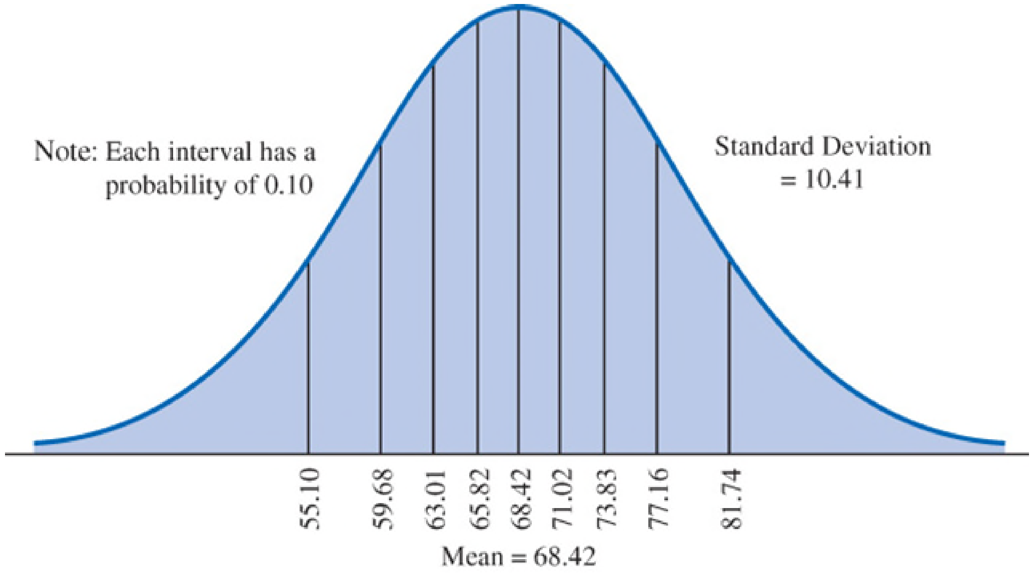
H_a: The population of test scores does not have a normal distribution with mean 68.42 and standard deviation 10.41.

12.3 Normal Distribution with Equal-Probability Intervals

We divide the normal probability distribution into 10 equal-probability intervals (*see notes) using the standard scores for each incremental 10th percentile: 10%, 20%, 30%, and so on.

The standard score for the 10th percentile is $z = -1.28$. Thus, the cutoff value for the lowest 10% is: $x = \bar{x} + zs = 68.42 - 1.28(10.41) = 55.10$, and so on for the other percentages.

Percentage	z	Test Score
10%	−1.28	68.42 − 1.28(10.41) = 55.10
20%	−0.84	68.42 − 0.84(10.41) = 59.68
30%	−0.52	68.42 − 0.52(10.41) = 63.01
40%	−0.25	68.42 − 0.25(10.41) = 65.82
50%	0.00	68.42 + 0(10.41) = 68.42
60%	+0.25	68.42 + 0.25(10.41) = 71.02
70%	+0.52	68.42 + 0.52(10.41) = 73.83
80%	+0.84	68.42 + 0.84(10.41) = 77.16
90%	+1.28	68.42 + 1.28(10.41) = 81.74



12.3 Computation of the Chi-Square Test Statistic for the Chemline Job Applicant Example

With the intervals of test scores now defined and the known expected frequency $e_i = 50/10 = 5$ per category, we can return to the sample data to determine the observed frequencies, f_i, for each of the test score intervals.

For example, there are 5 outcomes that are less than 55.10. Thus, we have $f_1 = 5$, and so on for f_2 to f_{10}.

The table shows the computations of the test statistic, $X^2 = 7.2$.

Test Score Interval	Observed Frequency f_i	Expected Frequency e_i	Difference $f_i - e_i$	Squared Difference $(f_i - e_i)^2$	Sq. Diff. Div. by Exp. Freq. $(f_i - e_i)^2/e_i$
Less than 55.10	5	5	0	0	0.0
55.10 < 59.68	5	5	0	0	0.0
59.68 < 63.01	9	5	4	16	3.2
63.01 < 65.82	6	5	1	1	0.2
65.82 < 68.42	2	5	-3	9	1.8
68.42 < 71.02	5	5	0	0	0.0
71.02 < 73.83	2	5	-3	9	1.8
73.83 < 77.16	5	5	0	0	0.0
77.16 < 81.74	5	5	0	0	0.0
81.74 and over	6	5	1	1	0.2
Total	50	50			$X^2 = 7.2$

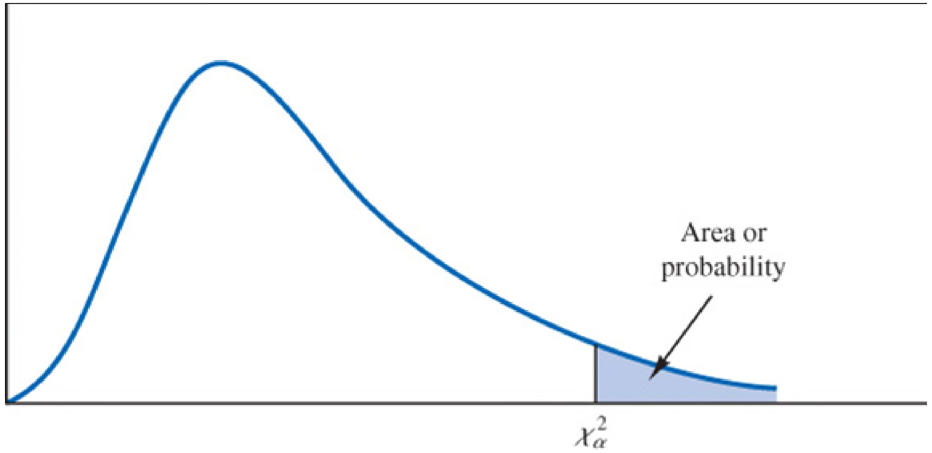
CENGAGE

12.3 *p*-Value Approach to a Goodness of Fitness Test of a Normal Probability Distribution

In the Chemline normal distribution example, the *p*-value is the upper tail of a X^2 distribution with $k = n - 3 = 7$ degrees of freedom: $p\text{−value} = P(X^2 \geq 7.2)$

Using Table 3 in Appendix B, the X^2 distribution with 7 degrees of freedom provides the following.

Area in upper tail	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
X^2 **Value** $(df = 7)$	12.017	14.067	16.013	18.475	20.278



Because $X^2 = 7.2$ is greater than 12.017, the values in the “Area in Upper Tail” row show that the *p*-value range must be: $p\text{−value} > 0.10$ (*see notes.)

Because p−value $> \alpha$, we cannot reject the null hypothesis that the probability distribution for the Chemline job applicant test scores is normal. The normal probability distribution may be applied to assist in the interpretation of test scores.

12.3 Critical Value Approach to a Goodness of Fitness Test of a Normal Probability Distribution

With $\alpha = 0.10$, $X_{\alpha}^{2} = X_{0.10}^{2}$ provides the critical value for the upper tail chi-square test. Using Table 3 in Appendix B, with 7 degrees of freedom, we have

Area in upper tail	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
$\boldsymbol{X^2}$ **Value** $(\boldsymbol{df} = 7)$	12.017	14.067	16.013	18.475	20.278

We obtain the critical value, $\mathrm{X}_{0.10}^{2} = 12.017$.

We can also use Excel to compute the critical value: $=\mathrm{CHISQ.INV.RT}(0.10{,}7) = 12.017$

Thus, the rejection rule for the goodness of fit test of a normal probability distribution is:

Reject H_0 if $X^2 \geq \mathrm{X}_{0.10}^{2} = 12.017$

Because $X^2 = 7.2$, we cannot reject the null hypothesis that the probability distribution for the Chemline job applicant test scores is normal.

Summary

- In this chapter, we have introduced hypothesis tests for the following applications.
  1. Testing the equality of population proportions for three or more populations.
  2. Testing the independence of two categorical variables.
  3. Testing whether a probability distribution for a population follows a specific historical or theoretical probability distribution (goodness of fit test.)
- All tests apply to categorical variables and use a chi-square (X^2) test statistic that is based on the differences between the observed frequencies from the sample data, and the expected frequencies, computed under the assumption that the null hypothesis is true.
- For the test about the equality of population proportions, we also showed how the Marascuilo procedure may help identify which pairs of population proportions are different.
- Finally, we showed how to categorize continuous data and use the goodness of fit test to check whether a continuous probability distribution follows the normal distribution.