Chapter 12

Comparing Multiple Proportions, Test of Independence and Goodness of Fit

Learning Objectives

1. Know how to conduct a test for the equality of three or more population proportions.
2. Be able to use the Marascuilo procedure to do multiple pairwise comparisons tests for three or more population proportions.
3. Understand the role of the chi-square distribution in conducting the tests in this chapter and be able to compute the chi-square test statistic for each application.
4. Understand the purpose of a test of independence.
5. Be able to set up tables, determine the observed and expected frequencies, and compute the chi-square test statistic for a test of independence.
6. Understand what a goodness-of-fit test is and be able to conduct the test for cases where the population is hypothesized to have either a multinomial probability distribution or a normal probability distribution.
7. Be able to use *p*-values based on the chi-square distribution to make the hypothesis testing conclusions in this chapter.

Solutions

1. H_0: $p_1 = p_2 = p_3$

   H_a: Not all population proportions are equal

   Observed Frequencies (f_{ij})

1	2	3	Total

Yes	150	150	96	396
No	100	150	104	354
Total	250	300	200	750

Expected Frequencies (e_{ij})

	1	2	3	Total
Yes	132.0	158.4	105.6	396
No	118.0	141.6	94.4	354
Total	250	300	200	750

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

	1	2	3	Total
Yes	2.45	.45	.87	3.77
No	2.75	.50	.98	4.22

$\chi^2 = 7.99$

Degrees of freedom = $k - 1 = (3 - 1) = 2$.

Using the χ^2 table with $df = 2$, $\chi^2 = 7.99$ shows the p-value is between .025 and .01.

Using software, the p-value corresponding to $\chi^2 = 7.99$ is .0184.

p-value $\leq$ .05; reject H_0. Conclude not all population proportions are equal.

2. a. $\bar{p}_1 = 150 / 250 = .60$

$\bar{p}_2 = 150/300 = .50$

$\bar{p}_3 = 96/200 = .48$

b. Multiple comparisons

For 1 versus 2

$$CV_{12} = \sqrt{\chi_\alpha^2}\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}} = \sqrt{5.991}\sqrt{\frac{.60(1-.60)}{250} + \frac{.50(1-.50)}{300}} = .1037$$

$df = k - 1 = 3 - 1 = z \quad \chi^2_{.05} = 5.991$

Comparison	p_i	p_j	Difference	n_i	n_j	Critical Value	Significant Diff > CV
1 vs. 2	.60	.50	.10	250	300	.1037	
1 vs. 3	.60	.48	.12	250	200	.1150	Yes
2 vs. 3	.50	.48	.02	300	200	.1117	

Only one comparison is significant, 1 versus 3. The others are not significant. We can conclude that the population proportions differ for populations 1 and 3.

3. a. H_0: $p_1 = p_2 = p_3$

H_a: Not all population proportions are equal

b. Observed Frequencies (f_{ij})

Flight	Delta	United	US Airways	Total
Delayed	39	51	56	146
On time	261	249	344	854
Total	300	300	400	1000

Expected Frequencies (e_{ij})

Flight	Delta	United	US Airways	Total
Delayed	43.8	43.8	58.4	146
On time	256.2	256.2	341.6	854
Total	300	300	400	1000

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Flight	Delta	United	US Airways	Total
Delayed	.53	1.18	.10	1.81
On time	.09	.20	.02	.31

$\chi^2 = 2.12$

Degrees of freedom = $k - 1 = (3 - 1) = 2$.

Using the χ^2 table with $df = 2$, $\chi^2 = 2.12$ shows the p-value is greater than .10.

Using software, the p-value corresponding to $\chi^2 = 2.12$ is .3465.

p-value > .05; do not reject H_0. We are unable to reject the null hypothesis that the population proportions are the same.

c. $\bar{p}_1 = 39/300 = .13$

$\bar{p}_2 = 51/300 = .17$

$\bar{p}_3 = 56/400 = .14$

Overall $\bar{p} = 146/1000 = .146$

4. a. H_0: $p_1 = p_2 = p_3$

H_a: Not all population proportions are equal

b. Observed Frequencies (f_{ij})

Component	A	B	C	Total
Defective	15	20	40	75
Good	485	480	460	1425
Total	500	500	500	1500

Expected Frequencies (e_{ij})

Component	A	B	C	Total
Defective	25	25	25	75
Good	475	475	475	1425
Total	500	500	500	1500

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Component	A	B	C	Total
Defective	4.00	1.00	9.00	14.00
Good	.21	.05	.47	0.74

$\chi^2 = 14.74$

Degrees of freedom = $k - 1 = (3 - 1) = 2$

Using the χ^2 table with $df = 2$, $\chi^2 = 14.74$ shows the p-value is less than .005.

Using software, the p-value corresponding to $\chi^2 = 14.74$ is .0006.

Because the p-value < .05; reject H_0. Conclude that the three suppliers do not provide equal proportions of defective components.

c. $\bar{p}_1 = 15/500 = .03$

$\bar{p}_2 = 20/500 = .04$

$\bar{p}_3 = 40/500 = .08$

Multiple comparisons

For Supplier A Versus Supplier B

$df = k - 1 = 3 - 1 = z \; \chi^2_{.05} = 5.991$

$$CV_{ij} = \sqrt{\chi^2_\alpha}\sqrt{\frac{\bar{p}_i(1-\bar{p}_i)}{n_i} + \frac{\bar{p}_j(1-\bar{p}_j)}{n_j}} = \sqrt{5.991}\sqrt{\frac{.03(1-.03)}{500} + \frac{.04(1-.04)}{500}} = .0284$$

Comparison	p_i	p_j	Difference	n_i	n_j	Critical Value	Significant Diff > CV
A vs. B	.03	.04	.01	500	500	.0284	
A vs. C	.03	.08	.05	500	500	.0351	Yes
B vs. C	.04	.08	.04	500	500	.0366	Yes

Supplier A and supplier B are both significantly different from supplier C. Supplier C can be eliminated on the basis of a significantly higher proportion of defective components. Since suppliers A and supplier B are not significantly different in terms of the proportion defective components, both suppliers should remain candidates for use by Benson.

5. H_0: $p_1 = p_2 = p_3$

H_a: Not all population proportions are equal

Observed Frequencies (f_{ij})

	Carnegie Classification			
Type of University	Moderate Research Activity	Higher Research Activity	Highest Research Activity	Total
Public	38	76	81	195
Not-for-profit private	58	31	34	123
Total	96	107	115	318

Expected Frequencies (e_{ij})

	Carnegie Classification			
Type of University	Moderate Research Activity	Higher Research Activity	Highest Research Activity	Total
Public	58.87	65.61	70.52	195
Not-for-profit private	37.13	41.39	44.48	123
Total	96	107	115	318

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Type of University	Carnegie Classification Moderate Research Activity	Higher Research Activity	Highest Research Activity	Total
Public	7.40	1.64	1.56	10.60
Not-for-profit private	11.73	2.61	2.47	16.80
Total	19.13	4.25	4.03	27.40

$\chi^2 = 27.40$

Degrees of freedom = $(r - 1)(c - 1) = (2 - 1)(3 - 1) = 2$

Using the χ^2 table with $df = 2$, $\chi^2 = 27.40$ shows the p-value is less than .005.

Using software, the p-value corresponding to $\chi^2 = 27.40$ is .000001.

p-value < .05; reject H_0. The proportion of public universities is not equal in each Carnegie category. The largest differences between actual and expected frequencies are in the moderate research activity classification, for which the number of public schools is much less than expected and the number of not-for-profit private schools is much greater than expected.

6. a. $\bar{p}_1 = 35/250 = .14$ 14% error rate

$\bar{p}_2 = 27/300 = .09$ 9% error rate

b. H_0: $p_1 - p_2 = 0$

H_a: $p_1 - p_2 \neq 0$

Observed Frequencies (f_{ij})

Return	Office 1	Office 2	Total
Error	35	27	62
Correct	215	273	488
	250	300	550

Expected Frequencies (e_{ij})

Return	Office 1	Office 2	Total
Error	28.18	33.82	62
Correct	221.82	266.18	488
	250	300	550

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Return	Office 1	Office 2	Total
Error	1.65	1.37	3.02
Correct	.21	.17	.38

$\chi^2 = 3.41$

$df = k - 1 = (2 - 1) = 1$.

Using the χ^2 table with $df = 1$, $\chi^2 = 3.41$ shows the p-value is between .10 and .05.

Using software, the p-value corresponding to $\chi^2 = 3.41$ is .0648.

p-value < .10; reject H_0. Conclude that the two offices do not have the same population proportion error rates.

c. With two populations, a chi-square test for equal population proportions has one degree of freedom. In this case, the test statistic χ^2 is always equal to z^2. This relationship between the two test statistics always provides the same *p*-value and the same conclusion when the null hypothesis involves equal population proportions. However, the use of the *z*-test statistic provides options for one–tailed hypothesis tests about two population proportions while the chi-square test is limited to two–tailed hypothesis tests about the equality of the two population proportions.

7. a. H_0: $p_1 = p_2 = p_3 = p_4$

H_a: Not all population proportions are equal

Observed Frequencies (f_{ij})

Social Media	United Kingdom	China	Russia	USA	Total
Yes	480	215	343	640	1,678
No	320	285	357	360	1,322
	800	500	700	1,000	3,000

Expected Frequencies (e_{ij})

Social Media	United Kingdom	China	Russia	USA	Total
Yes	447.47	279.67	391.53	559.33	1,678
No	352.53	220.33	308.47	440.67	1,322
	800	500	700	1,000	3,000

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Social Media	United Kingdom	China	Russia	USA	Total
Yes	2.36	14.95	6.02	11.63	34.96
No	3.00	18.98	7.63	14.77	44.38

$$\chi^2 = 79.34$$

Degrees of freedom = $df = k - 1 = (4 - 1) = 3$.

Using the χ^2 table with $df = 3$, $\chi^2 = 79.34$ shows the p-value is less than .005.

Using software, the p-value corresponding to $\chi^2 = 79.34$ is essentially 0.

p-value $\leq$ .05; reject H_0. Conclude the population proportions are not all equal.

b. United Kingdom 480/800 = .60

China 215/500 = .43

Russia 343/700 = .49

United States 640/1000 = .64 (Largest with 64% of adults)

c. Multiple pairwise comparisons

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

where $df = k - 1 = 4 - 1 = 3$ and $\chi^2_{.05} = 7.815$

Comparison	p_i	p_j	Difference	n_i	n_j	CV_{ij}	Diff > CV_{ij}
UK vs. C	0.60	0.43	0.17	800	500	0.0786	Yes
UK vs. R	0.60	0.49	0.11	800	700	0.0717	Yes
UK vs US	0.60	0.64	0.04	800	1000	0.0644	
C vs R	0.43	0.49	0.06	500	700	0.0814	
C vs US	0.43	0.64	0.21	500	1000	0.0750	Yes
R vs US	0.49	0.64	0.15	700	1000	0.0678	Yes

Only two comparisons are not significant: the difference in the proportion of adults that use social media in the United Kingdom and the United States is not significantly different, nor is the difference in the proportion of adults that use social media in China and Russia. All other comparisons show a significant difference.

8. H_0: The distribution of defects is the same for all suppliers

H_a: The distribution of defects is not the same all suppliers

Observed Frequencies (f_{ij})

Part Tested	A	B	C	Total
Minor defect	15	13	21	49
Major defect	5	11	5	21
Good	130	126	124	380
Total	150	150	150	450

Expected Frequencies (e_{ij})

Part tested	A	B	C	Total
Minor defect	16.33	16.33	16.33	49
Major defect	7.00	7.00	7.00	21
Good	126.67	126.67	126.67	380
Total	150	150	150	450

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Part tested	A	B	C	Total
Minor defect	.11	.68	1.33	2.12
Major defect	.57	2.29	.57	3.43
Good	.09	.00	.06	.15

$\chi^2 = 5.70$

Degrees of freedom = $(r - 1)(k - 1) = (3 - 1)(3 - 1) = 4$

Using the χ^2 table with $df = 4$, $\chi^2 = 5.70$ shows the p-value is greater than .10

Using software, the p-value corresponding to $\chi^2 = 5.70$ is .2227

p-value > .05; do not reject H_0. Conclude that we are unable to reject the hypothesis that the population distribution of defects is the same for all three suppliers. There is no evidence that quality of parts from one suppliers is better than either of the others two suppliers.

9. H_0: The column variable is independent of the row variable

H_a: The column variable is not independent of the row variable

Observed Frequencies (f_{ij})

	A	B	C	Total
P	20	44	50	114
Q	30	26	30	86
Total	50	70	80	200

Expected Frequencies (e_{ij})

	A	B	C	Total
P	28.5	39.9	45.6	114
Q	21.5	30.1	34.4	86
Total	50	70	80	200

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

	A	B	C	Total
P	2.54	.42	.42	3.38
Q	3.36	.56	.56	4.48

$\chi^2 = 7.86$

Degrees of freedom = (2–1)(3–1) = 2.

Using the χ^2 table with $df = 2$, $\chi^2 = 7.86$ shows the p-value is between .01 and

.025.

Using software, the p-value corresponding to $\chi^2 = 7.86$ is .0196.

p-value $\leq .05$; reject H_0. Conclude that there is an association between the column variable and the row variable. The variables are not independent.

10. H_0: The column variable is independent of the row variable

H_a: The column variable is dependent on the row variable

Observed Frequencies (f_{ij})

	A	B	C	Total
P	20	30	20	70
Q	30	60	25	115
R	10	15	30	55
Total	60	105	75	240

Expected Frequencies (e_{ij})

	A	B	C	Total
P	17.50	30.63	21.88	70
Q	28.75	50.31	35.94	115
R	13.75	24.06	17.19	55
Total	60	105	75	240

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

	A	B	C	Total
P	.36	.01	.16	.53
Q	.05	1.87	3.33	5.25
R	1.02	3.41	9.55	13.99

$\chi^2 = 19.77$

Degrees of freedom = $(r - 1)(c - 1) = (3 - 1)(3 - 1) = 4$.

Using the χ^2 table with $df = 4$, $\chi^2 = 19.77$ shows the p-value is less than .005.

Using software, the p-value corresponding to $\chi^2 = 19.77$ is .0006.

p-value $\leq .05$; reject H_0. Conclude that the column variable is not independent of the row variable.

11. a. H_0: Type of ticket purchased is independent of the type of flight

H_a: Type of ticket purchased is not independent of the type of flight

Expected Frequencies

$e_{11} = 35.59$ $e_{12} = 15.41$

$e_{21} = 150.73$ $e_{22} = 65.27$

$e_{31} = 455.68$ $e_{32} = 197.32$

Ticket	Flight	Observed Frequency (f_i)	Expected Frequency (e_i)	Chi-Square $(f_i - e_i)^2 / e_i$
First	Domestic	29	35.59	1.22

	International	22	15.41	2.82
Business	Domestic	95	150.73	20.61
	International	121	65.27	47.59
Full fare	Domestic	518	455.68	8.52
	International	135	197.32	19.68
	Totals:	920		$\chi^2 = 100.43$

Degrees of freedom = $(r - 1)(c - 1) = (3 - 1)(2 - 1) = 2$.

Using the χ^2 table with $df = 2$, $\chi^2 = 100.43$ shows the p-value is less than .005.

Using software, the p-value corresponding to $\chi^2 = 100.43$ is .0000.

p-value $\leq .05$; reject H_0. Conclude that the type of ticket purchased is not independent of the type of flight. We can expect the type of ticket purchased to depend upon whether the flight is domestic or international.

b. Column Percentages

	Type of Flight	
Type of Ticket	Domestic	International
First class	4.5%	7.9%
Business class	14.8%	43.5%
Economy class	80.7%	48.6%

A higher percentage of first class and business class tickets are purchased for international flights compared to domestic flights. Economy class tickets are purchased more for domestic flights. The first class or business class tickets are purchased for more than 50% of the international flights; 7.9% + 43.5% = 51.4%.

12. a. H_0: Employment plan is independent of the type of company

H_a: Employment plan is not independent of the type of company

Observed Frequency (f_{ij})

Employment Plan	Private	Public	Total
Add employees	37	32	69
No change	19	34	53
Lay off employees	16	42	58
Total	72	108	180

Expected Frequency (e_{ij})

Employment plan	Private	Public	Total
Add employees	27.6	41.4	69
No change	21.2	31.8	53
Lay off employees	23.2	34.8	58
Total	72.0	108.0	180

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Employment Plan	Private	Public	Total
Add employees	3.20	2.13	5.34

No change	0.23	0.15	0.38
Lay off employees	2.23	1.49	3.72
		$\chi^2 =$	9.44

Degrees of freedom = $(r - 1)(c - 1) = (3 - 1)(2 - 1) = 2$

Using the χ^2 table with $df = 2$, $\chi^2 = 9.44$ shows the p-value is between .01 and 0.005.

Using software, the p-value corresponding to $\chi^2 = 9.44$ is .0089.

Because the p-value $\leq .05$; reject H_0. Conclude the employment plan is not independent of the type of company. Thus, we expect employment plan to differ for private and public companies.

b. Column probabilities: For example, 37/72 = .5139

Employment plan	Private	Public
Add employees	.5139	.2963
No change	.2639	.3148
Lay off employees	.2222	.3889

Employment opportunities look to be much better for private companies with over 50% of private companies planning to add employees (51.39%). Public companies have the greater proportions of no change and lay off employees planned. 38.89% of public companies are planning to lay off employees over the next 12 months. 69/180 = .3833, or 38.33% of the companies in the survey are planning to hire and add employees during the next 12 months.

13. H_0: Interest in leaving job for more money is independent of the employee generation

H_a: Interest in leaving job for more money is not independent of the employee generation

Observed Frequencies (f_{ij})

Leave Job for More Money?	Baby Boomer	Generation X	Millennial	Total
Yes	129	152	164	445
No	207	183	171	561
Total	336	335	335	1,006

Expected Frequencies (e_{ij})

Health Insurance	Small	Medium	Large	Total
Yes	148.6	148.2	148.2	445
No	187.4	186.8	186.8	561
Total	336	335	335	225

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Health Insurance	Small	Medium	Large	Total
Yes	2.59	.10	1.69	4.38
No	2.06	.08	1.34	3.47

$\chi^2 = 7.85$

Degrees of freedom = $(r - 1)(c - 1) = (2 - 1)(3 - 1) = 2$.

Using the χ^2 table with $df = 2$, $\chi^2 = 7.85$ shows the p-value is between .01 and .025.

Using software, the p-value corresponding to $\chi^2 = 7.85$ with $df = 2$ is .0197.

p-value $\leq .05$; reject H_0. Conclude interest in leaving job for more money is not independent of the employee generation.

14. a. H_0: Quality rating is independent of the education of the owner

H_a: Quality rating is not independent of the education of the owner

Observed Frequencies (f_{ij}) Quality Rating	Some HS	HS Grad	Some College	College Grad	Total
Average	35	30	20	60	145
Outstanding	45	45	50	90	230
Exceptional	20	25	30	50	125
Total	100	100	100	200	500

Expected Frequencies (e_{ij})

Quality Rating	Some HS	HS Grad	Some College	College Grad	Total
Average	29	29	29	58	145
Outstanding	46	46	46	92	230
Exceptional	25	25	25	50	125
Total	100	100	100	200	500

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Quality Rating	Some HS	HS Grad	Some College	College Grad	Total

Average	1.24	.03	2.79	.07	4.14
Outstanding	.02	.02	.35	.04	.43
Exceptional	1.00	.00	1.00	.00	2.00

$\chi^2 = 6.57$

Degrees of freedom = $(r – 1)(c – 1) = (3 – 1)(4 – 1) = 6$.

Using the χ^2 table with $df = 6$, $\chi^2 = 6.57$ shows the p-value is greater than .10.

Using software, the p-value corresponding to $\chi^2 = 6.57$ is .3624.

p-value > .05; do not reject H_0. We are unable to conclude that the quality rating is not independent of the education of the owner. Thus, quality ratings are not expected to differ with the education of the owner.

b. Average: 145/500 = 29%

Outstanding	230/500 = 46%
Exceptional	125/500 = 25%

New owners look to be pretty satisfied with their new automobiles with almost 50% rating the quality outstanding and over 70% rating the quality outstanding or exceptional.

15. a. H_0: Quality of management is independent of the reputation of the company

H_a: Quality of management is not independent of the reputation of the company

Observed Frequencies (f_{ij})

Quality of Management	Excellent	Good	Fair	Total
Excellent	40	25	5	70
Good	35	35	10	80
Fair	25	10	15	50
Total	100	70	30	200

Expected Frequencies (e_{ij})

Quality of Management	Excellent	Good	Fair	Total
Excellent	35.0	24.5	10.5	70
Good	40.0	28.0	12.0	80
Fair	25.0	17.5	7.5	50
Total	100	70	30	200

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Quality of Management	Excellent	Good	Fair	Total
Excellent	.71	.01	2.88	3.61
Good	.63	1.75	.33	2.71
Fair	.00	3.21	7.50	10.71

$\chi^2 = 17.03$

Degrees of freedom = $(r - 1)(c - 1) = (3 - 1)(3 - 1) = 4$.

Using the χ^2 table with $df = 4$, $\chi^2 = 17.03$ shows the p-value is less than .005.

Using software, the p-value corresponding to $\chi^2 = 17.03$ is .0019.

p-value < .05; reject H_0. Conclude that the rating for the quality of management is not independent of the rating for the reputation of the company.

b. Using the highest column probabilities, if the reputation of the company is:

Excellent—There is a 40/100 = .40 chance the quality of management will also be excellent.

Good—There is a 35/70 = .50 chance the quality of management will also be good.

Fair—There is a 15/30 = .50 chance the quality of management will also be fair.

The highest probabilities are that the two variables will have the same ratings. Thus, the two ratings are associated.

16. a. Observed Frequency (fij)

	Age of Respondent				
Actress	18–30	31–44	45–58	Over 58	Totals
Jessica Chastain	51	50	41	42	184
Jennifer Lawrence	63	55	37	50	205
Emmanuelle Riva	15	44	56	74	189
Quvenzhané Wallis	48	25	22	31	126
Naomi Watts	36	65	62	33	196
Totals	213	239	218	230	900

The sample size is 900.

b. The sample proportion of movie fans who prefer each actress is:

$$\bar{p}_1 = \frac{184}{900} = .2044$$

$$\bar{p}_2 = \frac{205}{900} = .2278$$

$$\bar{p}_3 = \frac{189}{900} = .2100$$

$$\bar{p}_4 = \frac{126}{900} = .1400$$

$$\bar{p}_5 = \frac{196}{900} = .2178$$

The movie fans favored Jennifer Lawrence, but three other nominees (Jessica Chastain, Emmanuelle Riva, and Naomi Watts) each were favored by almost as many of the fans.

c. Expected Frequency (e_{ij})

	Age of Respondent				
Actress	18–30	31–44	45–58	Over 58	Totals
Jessica Chastain	43.5	48.9	44.6	47.0	184
Jennifer Lawrence	48.5	54.4	49.7	52.4	205
Emmanuelle Riva	44.7	50.2	45.8	48.3	189
Quvenzhané Wallis	29.8	33.5	30.5	32.2	126
Naomi Watts	46.4	52.0	47.5	50.1	196
Totals	213	239	218	230	900

Calculate $\frac{(f_{ij} - e_{ij})^2}{e_{ij}}$ for each cell in the table.

Age of Respondent

Actress	18–30	31–44	45–58	More Than 58	Totals
Jessica Chastain	1.28	0.03	0.29	0.54	2.12
Jennifer Lawrence	4.32	0.01	3.23	0.11	7.66
Emmanuelle Riva	19.76	0.76	2.28	13.67	36.48
Quvenzhané Wallis	11.08	2.14	2.38	0.04	15.65
Naomi Watts	2.33	3.22	4.44	5.83	15.82
Totals	38.77	6.16	12.61	20.20	77.74

$$\chi^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 77.74$$

With (5 – 1)(4 – 1) = 12 degrees of freedom, the *p*-value is approximately 0.

p-value $\leq$ .05; reject H_0. Attitude toward the actress who was most deserving of the 2013 Academy Award for actress in a leading role is not independent of age.

17. a. H_0: Hours of sleep per night is independent of age

H_a: Hours of sleep per night is not independent of age

Observed Frequencies (f_{ij})

Hours of Sleep	39 or Younger	40 or Older	Total
Fewer than 6	38	36	74
6 to 6.9	60	57	117
7 to 7.9	77	75	152

8 or more	65	92	157
Total	240	260	500

Expected Frequencies (e_{ij})

Hours of Sleep	39 or Younger	40 or Older	Total
Fewer than 6	35.52	38.48	74
6 to 6.9	56.16	60.84	117
7 to 7.9	72.96	79.04	152
8 or more	75.36	81.64	157
Total	240	260	500

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Hours of Sleep	39 or Younger	40 or Older	Total
Fewer than 6	.17	.16	.33
6 to 6.9	.26	.24	.50
7 to 7.9	.22	.21	.43
8 or more	1.42	1.31	2.74

$\chi^2 = 4.01$

Degrees of freedom = $(r - 1)(c - 1) = (4 - 1)(2 - 1) = 3$

Using the χ^2 table with $df = 3$, $\chi^2 = 4.01$ shows the p-value is greater than .10.

Using software, the p-value corresponding to $\chi^2 = 4.01$ is .2604.

p-value > .05; do not reject H_0. Cannot reject the assumption that age and hours of sleep are independent.

b. Because age does not appear to have an association on hours of sleep, use the overall row percentages.

Fewer than 6	74/500	= .148	14.8%
6 to 6.9	117/500	= .234	23.4%
7 to 7.9	152/500	= .304	30.4%
8 or more	157/500	= .314	31.4%

30.4% + 31.4% = 61.8% of individuals get seven or more hours of sleep a night.

18. Expected frequencies:

$e_{11} = 11.81$	$e_{12} = 8.44$	$e_{13} = 24.75$
$e_{21} = 8.40$	$e_{22} = 6.00$	$e_{23} = 17.60$
$e_{31} = 21.79$	$e_{32} = 15.56$	$e_{33} = 45.65$

		Observed Frequency	Expected Frequency	Chi-Square
Host A	Host B	(f_i)	(e_i)	$(f_i - e_i)^2 / e_i$
Con	Con	24	11.81	12.57
Con	Mixed	8	8.44	.02
Con	Pro	13	24.75	5.58
Mixed	Con	8	8.40	.02
Mixed	Mixed	13	6.00	8.17

Mixed	Pro	11	17.60	2.48
Pro	Con	10	21.79	6.38
Pro	Mixed	9	15.56	2.77
Pro	Pro	64	45.65	7.38

$\chi^2 = 45.36$

Degrees of freedom = $(r - 1)(c - 1) = (3 - 1)(3 - 1) = 4$

Using the χ^2 table with $df = 2$, $\chi^2 = 45.36$ shows the p-value is less than .005.

Using software, the p-value corresponding to $\chi^2 = 45.36$ is .0000.

p-value $\leq$ .01; reject H_0. Conclude that the ratings of the two hosts are not independent. The host responses are more similar than different and they tend to agree or be close in their ratings.

19. a. Expected frequencies:

$e_1 = 200\ (.40) = 80, \quad e_2 = 200\ (.40) = 80, \quad e_3 = 200\ (.20) = 40$

Observed frequencies:

$f_1 = 60, f_2 = 120, f_3 = 20$

$$\begin{aligned}\chi^2 &= \frac{(60-80)^2}{80} + \frac{(120-80)^2}{80} + \frac{(20-40)^2}{40} \\ &= \frac{400}{80} + \frac{1600}{80} + \frac{400}{40} \\ &= 5 + 20 + 10 \\ &= 35\end{aligned}$$

$k - 1 = 2$ degrees of freedom

Using the χ^2 table with $df = 2$, $\chi^2 = 35$ shows the p-value is less than .005.

Using software, the p-value corresponding to $\chi^2 = 35$ is approximately 0.

p-value $\leq .01$; reject H_0. Conclude the proportions differ from .40, .40, and .20.

b. $\chi_{.01} = 9.210$

Reject H_0 if $\chi^2 \geq 9.210$

$\chi^2 = 35$; reject H_0. Conclude the proportions differ from .40, .40, and .20.

20. With $n = 30$ we will use six classes, each with the probability of .1667.

$\bar{x} = 22.8$ $s = 6.27$

The z values that create six intervals, each with probability .1667, are –.97, –.43, 0, .43, .97.

z	Cutoff Value of x
–.97	22.8 – .97 (6.27) = 16.74
–.43	22.8 – .43 (6.27) = 20.10
0	22.8 + 0 (6.27) = 22.80
.43	22.8 + .43 (6.27) = 25.50
.97	22.8 + .97 (6.27) = 28.86

Interval	Observed Frequency	Expected Frequency	Difference
Less than 16.74	3	5	–2
16.74–20.10	7	5	2
20.10–22.80	5	5	0
22.80–25.50	7	5	2
25.50–28.86	3	5	–2

28.86 and up	5	5	0

$$\chi^2 = \frac{(-2)^2}{5} + \frac{(2)^2}{5} + \frac{(0)^2}{5} + \frac{(2)^2}{5} + \frac{(-2)^2}{5} + \frac{(0)^2}{5} = \frac{16}{5} = 3.20$$

Degrees of freedom = $k - p - 1 = 6 - 2 - 1 = 3$

Using the χ^2 table with $df = 3$, $\chi^2 = 3.20$ shows the p-value is greater than .10.

Using software, the p-value corresponding to $\chi^2 = 3.20$ is .3618.

p-value > .05; do not reject H_0. The claim that the data come from a normal distribution cannot be rejected.

21. H_0: $p_{ABC} = .29$, $p_{CBS} = .28$, $p_{NBC} = .25$, $p_{IND} = .18$

H_a: The proportions are not $p_{ABC} = .29$, $p_{CBS} = .28$, $p_{NBC} = .25$, $p_{IND} = .18$

Expected frequencies:

300(.29) = 87, 300(.28) = 84

300(.25) = 75, 300(.18) = 54

$e_1 = 87$, $e_2 = 84$, $e_3 = 75$, $e_4 = 54$

Observed frequencies:

$f_1 = 95$, $f_2 = 70$, $f_3 = 89$, $f_4 = 46$

$$\chi^2 = \frac{(95-87)^2}{87} + \frac{(70-84)^2}{84} + \frac{(89-75)^2}{75} + \frac{(46-54)^2}{54}$$
$$= 6.87$$

$k - 1 = 3$ degrees of freedom

Using the χ^2 table with $df = 3$, $\chi^2 = 6.87$ shows the p-value is between .05 and .10.

Using software, the p-value corresponding to $\chi^2 = 6.87$ is .0762.

p-value > .05; do not reject H_0. There has not been a significant change in the viewing audience proportions.

22.

Category	Hypothesized Proportion	Observed Frequency (f_i)	Expected Frequency (e_i)	Chi-Square $(f_i - e_i)^2 / e_i$
Blue	.24	105	120	1.88
Brown	.13	72	65	.75
Green	.20	89	100	1.21
Orange	.16	84	80	.20
Red	.13	70	65	.38
Yellow	.14	80	70	1.43
	Total:	500		$\chi^2 = 5.85$

$k - 1 = 6 - 1 = 5$ degrees of freedom

Using the χ^2 table with $df = 5$, $\chi^2 = 5.85$ shows the p-value is greater than .10

Using software, the p-value corresponding to $\chi^2 = 5.85$ is .3211

p-value > .05; do not reject H_0. We cannot reject the hypothesis that the overall percentages of colors in the population of M&M milk chocolate candies are .24 blue, .13 brown, .20 green, .16 orange, .13 red and .14 yellow.

23. Expected frequencies:

20% each $n = 60$

$e_1 = 12$, $e_2 = 12$, $e_3 = 12$, $e_4 = 12$, $e_5 = 12$

Observed frequencies:

$f_1 = 5, f_2 = 8, f_3 = 15, f_4 = 20, f_5 = 12$

$$\chi^2 = \frac{(5-12)^2}{12} + \frac{(8-12)^2}{12} + \frac{(15-12)^2}{12} + \frac{(20-12)^2}{12} + \frac{(12-12)^2}{12}$$
$$= 11.50$$

$k - 1 = 4$ degrees of freedom.

Using the χ^2 table with $df = 4$, $\chi^2 = 11.50$ shows the p-value is between .01 and .025.

Using software, the p-value corresponding to $\chi^2 = 11.50$ is .0215.

p-value < .05; reject H_0. Conclude the largest companies differ in performance from the 1000 companies. In general, the largest companies did not do as well as others. 15 of 60 companies (25%) are in the middle group and 20 of 60 companies (33%) are in the next lower group. These both are greater than the 20% expected. Relative few large companies are in the top A and B categories.

24. a. H_0: $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = 1/7$

H_a: Not all proportions are equal

Observed Frequency (f_i)

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
66	50	53	47	55	69	80

Expected Frequency (e_i) $e_i = 1/7(420) = 60$

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
60	60	60	60	60	60	60

Chi-Square Calculations $(f_i - e_i)^2 / e_i$

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
.60	1.67	.82	2.82	.42	1.35	6.67

χ^2 = 14.33

Degrees of freedom = (k – 1) = (7 – 1) = 6

Using the χ^2 table with df = 6, χ^2 = 14.33 shows the p-value is between .05 and .025.

Using software, the p-value corresponding to χ^2 = 14.33 is .0262

p-value $\leq$ .05; reject H_0. Conclude the proportion of traffic accidents is not the same for each day of the week.

b. Percentage of traffic accidents by day of the week

Sunday	66/420 = .1571	15.71%
Monday	50/420 = .1190	11.90%
Tuesday	53/420 = .1262	12.62%
Wednesday	47/420 = .1119	11.19%
Thursday	55/420 = .1310	13.10%
Friday	69/420 = .1643	16.43%
Saturday	80/420 = .1905	19.05%

Saturday has the highest percentage of traffic accident (19%). Saturday is typically the late night and more social day/evening of the week. Alcohol, speeding and distractions are more likely to affect driving on Saturdays. Friday is the second highest with 16.43%.

25. $\bar{x} = 21.7$ $s = 9.4$ $n = 25$

Use five classes.

Percentage	z	Data Value
20.00%	–.84	21.7–.84(9.4) = 13.80
40.00%	–.25	21.7–.25(9.4) = 19.35
60.00%	.25	21.7+.25(9.4) = 24.05
80.00%	.84	21.7+.84(9.4) = 29.60

Interval	Observed Frequency	Expected Frequency
Less than 13.80	7	5
13.80–19.35	7	5
19.35–24.05	1	5
24.05–29.60	1	5
29.60 and up	9	5

$\chi^2 = 11.20$

Degrees of freedom = $k - p - 1 = 5 - 2 - 1 = 2$.

Using the χ^2 table with $df = 2$, $\chi^2 = 11.20$ shows the p-value is less than .005.

Using software, the p-value corresponding to $\chi^2 = 11.20$ is .0037.

p-value $\leq .01$; reject H_0. Conclude the distribution does not have a normal probability distribution.

26. $\bar{x} = 24.5$ $s = 3$ $n = 30$

Use six classes,

Percentage	z	Data Value
16.67%	–.97	24.5–.97(3) = 21.59
33.33%	–.43	24.5–.43(3) = 23.21
50.00%	.00	24.5+.00(3) = 24.50
66.67%	.43	24.5+.43(3) = 25.79
83.33%	.97	24.5+.97(3) = 27.41

Interval	Observed Frequency	Expected Frequency
Less than 21.59	5	5
21.59–23.21	4	5
23.21–24.50	3	5
24.50–25.79	7	5
25.7927.41	7	5
27.41 up	4	5

$\chi^2 = 2.80$

Degrees of freedom = $(k - p - 1) = 6 - 2 - 1 = 3$.

Using the χ^2 table with $df = 3$, $\chi^2 = 2.80$ shows the p-value is greater than .10.

Using software, the p-value corresponding to $\chi^2 = 2.80$ is .4235.

p-value > .10; do not reject H_0. The assumption of a normal distribution cannot be rejected.

27. a. $\overline{p}_1 = 44/400 = .1100$

$\overline{p}_2 = 35/385 = .0909$

$\overline{p}_3 = 35/399 = .0877$

$\overline{p}_4 = 34/400 = .0850$

Washington, D.C. 8.8%; Bridgeport, CT 11.7%; San Jose, CA 9%, Lexington Park, MD 8.5%

b. H_0: $p_1 = p_2 = p_3 = p_4$

H_a: Not all population proportions are equal

Observed Frequencies (f_{ij})

Millionaire	Bridgeport, CT	San Jose, CA	Washington, D.C.	Lexington Park, MD	Total
Yes	44	35	35	34	148
No	356	350	364	366	1,436
Total	400	385	399	400	1,584

Expected Frequencies (e_{ij})

Millionaire	Bridgeport, CT	San Jose, CA	Washington, D.C.	Lexington Park, MD	Total
Yes	37.37	35.97	37.28	37.37	148
No	362.63	349.03	361.72	362.63	1,436

Total	400	385	399	400	1,584

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Millionaire	Bridgeport, CT	San Jose, CA	Washington, D.C.	Lexington Park, MD	Total
Yes	1.18	.03	.14	.30	1.65
No	.12	.00	.01	.03	.16

$\chi^2 = 1.81$

Degrees of freedom = $k - 1 = (4 - 1) = 3$

Using the χ^2 table with $df = 3$, $\chi^2 = 1.81$ shows the p-value is greater than .10

Using software, the p-value corresponding to $\chi^2 = 1.81$ is .6117

p-value > .05; do not reject H_0. Cannot conclude that there is a difference among the population proportion of millionaires for these four cities.

28. a. H_0: $p_1 = p_2 = p_3$

H_a: Not all population proportions are equal

Observed Frequencies (f_{ij})

Quality	First	Second	Third	Total
Good	285	368	176	829
Defective	15	32	24	71
Total	300	400	200	900

Expected Frequencies (e_{ij})

Quality	First	Second	Third	Total
Good	276.33	368.44	184.22	829
Defective	23.67	31.56	15.78	71
Total	300	400	200	900

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

Quality	First	Second	Third	Total
Good	.27	.00	.37	.64
Defective	3.17	.01	4.28	7.46

$\chi^2 = 8.10$

Degrees of freedom = $k - 1 = (3 - 1) = 2$.

Using the χ^2 table with $df = 2$, $\chi^2 = 8.10$ shows the p-value is between .025 and .01.

Using software, the p-value corresponding to $\chi^2 = 8.10$ is .0174.

p-value $\leq$ .05; reject H_0. Conclude the population proportion of good parts is not equal for all three shifts. The shifts differ in terms of production quality.

b. $\bar{p}_1 = 285 / 300 = .95$

$\bar{p}_2 = 368 / 400 = .92$

$\bar{p}_3 = 176 / 200 = .88$

$df = k - 1 = 3 - 1 = 2$ $\chi^2_{.05} = 5.991$

Comparison	p_i	p_j	Difference	n_i	n_j	Critical Value	Significant Diff > CV
1 vs. 2	.95	.92	.03	300	400	.0453	
1 vs. 3	.95	.88	.07	300	200	.0641	Yes
2 vs. 3	.92	.88	.04	400	200	.0653	

Shifts 1 and 3 differ significantly with shift 1 producing better quality (95%) than shift 3 (88%). The study cannot identify shift 2 (92%) as better or worse quality than the other two shifts. Shift 3, at 7% more defectives than shift 1 should be studied to determine how to improve its production quality.

29. Let p_1 = population proportion of visitors who rate the Louvre Museum as spectacular.

p_2 = population proportion of visitors who rate the National Museum in China as spectacular

p_3 = population proportion of visitors who rate the Metropolitan Museum of Art as spectacular

p_4 = population proportion of visitors who rate the Vatican Museums as spectacular

p_5 = population proportion of visitors who rate the British Museums as spectacular

a. Point estimates of the population proportion of visitors who rated each of these museums as spectacular are:

$\bar{p}_1$ = 113/150 = .7533 is the point estimate of the population proportion of visitors who rated the Louvre Museum as spectacular.

$\bar{p}_2$ = 88/132 = .6667 is the point estimate of the population proportion of visitors who rated the National Museum in China as spectacular.

$\bar{p}_3 = 94/140 = .6714$ is the point estimate of the population proportion of visitors who rated the Metropolitan Museum of Art as spectacular.

$\bar{p}_4 = 98/170 = .5765$ is the point estimate of the population proportion of visitors who rated the Vatican Museums as spectacular.

$\bar{p}_5 = 96/160 = .6000$ is the point estimate of the population proportion of visitors who rated the British Museum as spectacular.

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

H_a: Not all population proportions are equal

Observed Frequency (f_{ij})

	Louvre Museum	National Museum in China	Metropolitan Museum of Art	Vatican Museums	British Museum	Total
Rated spectacular	113	88	94	98	96	469
Did not rate spectacular	37	44	46	72	64	263
Totals	150	132	140	170	160	752

Expected Frequency (e_{ij})

	Louvre Museum	National Museum in China	Metropolitan Museum of Art	Vatican Museums	British Museum	Total
Rated spectacular	97.54	85.84	91.04	110.55	104.04	469
Did not rate spectacular	52.46	46.16	48.96	59.45	55.96	263
Totals	150	140	160	170	160	752

Chi-Square $(f_{ij} - e_{ij})^2 / e_{ij}$

	Louvre Museum	National Museum in China	Metropolitan Museum of Art	Vatican Museums	British Museum	Total
Rated spectacular	2.45	.05	.10	1.42	.62	4.65
Did not rate spectacular	4.56	.10	.18	2.65	1.16	8.64

$\chi^2 = 13.29$

Degrees of freedom = $k - 1 = 5 - 1 = 4$.

Using the χ^2 table with $df = 4$, $\chi^2 = 13.29$ shows the p-value is between .005 and .01.

Using software, the p-value corresponding to $\chi^2 = 13.29$ is .00996.

p-value $\leq .05$; reject H_0. We conclude that the population proportion of visitors who rated the museum as spectacular differs for these five museums.

30. a. H_0: The preferred pace of life is independent of gender

H_a: The preferred pace of life is not independent of gender

Observed Frequency (f_{ij})

	Gender		
Preferred Pace of Life	Male	Female	Total
Slower	230	218	448
No Preference	20	24	44
Faster	90	48	138
Total	340	290	630

Expected Frequency (e_{ij})

	Gender		
Preferred Pace of Life	Male	Female	Total
Slower	241.78	206.22	448
No Preference	23.75	20.25	44
Faster	74.48	63.52	138
Total	340	290	630

Chi-Square Calculations $(f_{ij} - e_{ij})^2 / e_{ij}$

	Gender		
Preferred Pace of Life	Male	Female	Total
Slower	.57	.67	1.25
No preference	.59	.69	1.28
Faster	3.24	3.79	7.03

$\chi^2 = 9.56$

Degrees of freedom = $(r - 1)(c - 1) = (3 - 1)(2 - 1) = 2$.

Using the χ^2 table with $df = 2$, $\chi^2 = 9.56$ shows the p-value is less than .01.

Using software, the p-value corresponding to $\chi^2 = 9.56$ is .0084.

p-value < .05; reject H_0. The preferred pace of life is not independent of gender.

Thus, we expect men and women differ with respect to the preferred pace of life.

b. Percentage responses for each gender

	Gender	
Preferred Pace of Life	Male	Female
Slower	67.65	75.17
No preference	5.88	8.28
Faster	26.47	16.55

The highest percentages are for a slower pace of life by both men and women.

However, 75.17% of women prefer a slower pace compared to 67.65% of men and

26.47% of men prefer a faster pace compared to 16.55% of women. More women prefer a slower pace while more men prefer a faster pace.

31. H_0: Church attendance is independent of age

H_a: Church attendance is not independent on age

Observed Frequencies (f_{ij})

	Age				
Church Attendance	20–29	30–39	40–49	50–59	Total
Yes	31	63	94	72	260
No	69	87	106	78	340
Total	100	150	200	150	600

Expected Frequencies (e_{ij})

	Age				
Church Attendance	20–29	30–39	40–49	50–59	Total
Yes	43	65	87	65	260
No	57	85	113	85	340
Total	100	150	200	150	600

Chi-Square $(f_{ij} - e_{ij})^2/ e_{ij}$

	Age				
Church Attendance	20–29	30–39	40–49	50–59	Total

Yes	3.51	.06	.62	.75	4.94
No	2.68	.05	.47	.58	3.78

$\chi^2 = 8.72$

Degrees of freedom = $(r - 1)(c - 1) = (2 - 1)(4 - 1) = 3$

Using the χ^2 table with $df = 3$, $\chi^2 = 8.72$ shows the p-value is between .025 and .05.

Using software, the p-value corresponding to $\chi^2 = 8.72$ is .0333.

p-value $\leq$ .05; reject H_0. Conclude church attendance is not independent of age.

Church Attendance by Age Group

20–29	$31/100 \rightarrow 31\%$
30–39	$63/150 \rightarrow 42\%$
40–49	$94/200 \rightarrow 47\%$
50–59	$72/150 \rightarrow 48\%$

Church attendance increases as individuals grow older.

32. H_0: The county with the emergency call is independent of the day of week

H_a: The county with the emergency call is not independent of the day of week

Observed Frequencies (f_{ij})

	Day of Week							
County	Sun	Mon	Tues	Wed	Thu	Fri	Sat	Total
Urban	61	48	50	55	63	73	43	393

Rural	7	9	16	13	9	14	10	78
Total	68	57	66	68	72	87	53	471

Expected Frequencies (e_{ij})

	Day of Week							
County	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Total
Urban	56.74	47.56	55.07	56.74	60.08	72.59	44.22	393
Rural	11.26	9.44	10.93	11.26	11.92	14.41	8.78	78
Total	68	57	66	68	72	87	53	471

Chi-Square $(f_{ij} - e_{ij})^2/ e_{ij}$

	Day of Week							
County	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Total
Urban	.32	.00	.47	.05	.14	.00	.03	1.02
Rural	1.61	.02	2.35	.27	.72	.01	.17	5.15

$\chi^2 = 6.17$

Degrees of freedom = $(r - 1)(c - 1) = (2 - 1)(7 - 1) = 6$.

Using the χ^2 table with $df = 6$, $\chi^2 = 6.17$ shows the p-value is greater than .10.

Using software, the p-value corresponding to $\chi^2 = 6.17$ is .4044.

p-value > .05; do not reject H_0. The assumption of independence cannot be rejected. The county with the emergency call does not vary or depend upon the day of the week.

33. a. The sample size is very large: 6448

b.

Observed Frequency (f_{ij})

Response	Country Great Britain	France	Italy	Spain	Germany	United States	Total
Strongly favor	141	161	298	133	128	204	1,065
Favor	348	366	309	222	272	326	1,843
Oppose	381	334	219	311	322	316	1,883
Strongly oppose	217	215	219	443	389	174	1,657
Total	1,087	1,076	1,045	1,109	1,111	1,020	6,448

Expected Frequency (e_{ij})

Response	Country Great Britain	France	Italy	Spain	Germany	United States	Total
Strongly favor	180	178	173	183	183	168	1,065
Favor	311	307	299	317	318	291	1,843

Oppose	317	315	305	324	324	298	1,883
Strongly oppose	279	276	268	285	286	263	1,657
Total	1,087	1,076	1,045	1,109	1,111	1,020	6,448

Chi-Square $(f_{ij} - e_{ij})^2 / e_{ij}$

	Country						
Response	Great Britain	France	Italy	Spain	Germany	United States	Total
Strongly favor	8.45	1.62	90.32	13.66	16.53	7.71	138.29
Favor	4.40	11.34	0.33	28.47	6.65	4.21	55.40
Oppose	12.92	1.15	24.25	0.52	0.01	1.09	39.94
Strongly oppose	13.78	13.48	8.96	87.59	37.09	30.12	191.02

$\chi^2 = 424.65$

$$\chi^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 424.65$$

Degrees of freedom = $(r - 1)(c - 1) = (4 - 1)(6 - 1) = 15$

The p-value is approximately 0.

p-value $\leq .05$; reject H_0. The attitude toward building new nuclear power plants is not independent of the country. Attitudes can be expected to vary with the country.

c. Use column percentages from the observed frequencies table to help answer this question.

	Country					
Response	Great Britain	France	Italy	Spain	Germany	United States
Strongly favor	13.0	15.0	28.5	12.0	11.5	20.0
Favor	32.0	34.0	29.5	20.0	24.5	32.0
Oppose	35.0	31.0	21.0	28.0	29.0	31.0
Strongly oppose	20.0	20.0	21.0	40.0	35.0	17.0
Total	100	100	100	100	100	100

Adding together the percentages of respondents who "Strongly favor" and those who "Favor", we find the following: Great Britain 45%, France 49%, Italy 58%, Spain 32%, Germany 36% and United States 52%. Italy shows the most support for nuclear power plants with 58% in favor. Spain shows the least support with only 32% in favor. Only Italy and the United States show more than 50% of the respondents in favor of building new nuclear power plants.

34.

Expected Frequencies for n = 344

Professional football	$e_1 = .33*344 = 113.52$
Baseball	$e_2 = .15*344 = 51.60$
Men's college football	$e_3 = .10*344 = 34.40$
Auto racing	$e_4 = .06*344 = 20.64$
Men's professional basketball	$e_5 = .05*344 = 17.20$
Ice hockey	$e_6 = .05*344 = 17.20$
Other sports	$e_7 = .26*344 = 89.44$

Actual Frequencies

Professional football	$f_1 = 111$
Baseball	$f_2 = 39$
Men's college football	$f_3 = 46$
Auto racing	$f_4 = 14$
Men's professional basketball	$f_5 = 6$
Ice hockey	$f_6 = 20$
Other sports	$f_7 = 108$

$$\chi^2 = \frac{(111-113.52)^2}{113.52} + \frac{(39-51.6)^2}{51.6} + \frac{(46-34.4)^2}{34.4} + \frac{(14-20.64)^2}{20.64} + \frac{(6-17.2)^2}{17.2} + \frac{(20-17.2)^2}{17.2} + \frac{(108-89.44)^2}{89.44}$$
$$= 20.78$$

$k - 1 = 6$ degrees of freedom

Using the χ^2 table with $df = 6$, $\chi^2 = 20.78$ shows the *p*-value is less than .005.

Using Excel, the *p*-value corresponding to $\chi^2 = 20.78$ with $df = 6$ is .0020.

p-value < .05; reject H_0. Yes, undergraduate students differ from the general public with regard to their favorite sports. Men's college football is more popular among undergraduate students, and baseball and auto racing are less popular among undergraduate students. The difference between expected and actual number of undergraduate students who responded *other sports* also suggests that undergraduate students are interested in a broader range of sports.

35. H_0: The market shares for the seven small-car categories in Chicago are .20, .17, .12, .10, .10, .08, .23

H_a: The market shares for the seven small-car categories in Chicago differ from the above shares

Compact Car	Hypothesized Market Share	Observed Frequency	Expected Frequency	Chi-Square $(f_i - e_i)^2 / e_i$
Honda Civic	.20	98	80	4.05
Toyota Corolla	.17	72	68	0.24
Nissan Sentra	.12	54	48	0.75
Hyundai Elantra	.10	44	40	0.40
Chevrolet Cruze	.10	42	40	0.10
Ford Focus	.08	25	32	1.53
Other	.23	65	92	7.92

$\chi^2 = 14.99$

Degrees of freedom = $k - 1 = 7 - 1 = 6$.

Using the χ^2 table with $df = 6$, $\chi^2 = 14.99$ shows the p-value is between .01 and .025.

Using software, the p-value corresponding to $\chi^2 = 14.99$ is .02.

p-value < .05; reject H_0. Conclude that the markets shares for the five compact cars in Chicago differ from the market shares reported. In particular, the Chicago market appears to have fewer purchases of the "Other" category and more purchases of the Honda Civic.

36. $\bar{x} = 76.83$ $s = 12.43$

Interval	Observed Frequency	Expected Frequency
Less than 62.54	5	5
62.54–68.50	3	5
68.50–72.85	6	5
72.85–76.83	5	5
76.83–80.81	5	5
80.81–85.16	7	5
85.16–91.12	4	5
91.12 up	5	5

$\chi^2 = 2$

Degrees of freedom = $k - p - 1 = 8 - 2 - 1 = 5$.

Using the χ^2 table with $df = 5$, $\chi^2 = 2.00$ shows the p-value is greater than .10.

Using software, the p-value corresponding to $\chi^2 = 2.00$ is .8491.

p-value > .05; do not reject H_0. The assumption of a normal distribution cannot be rejected.

37. a.

x	Observed Frequencies	Binomial Probability $n = 4, p = .30$	Expected Frequencies
0	30	.2401	24.01

1	32	.4116	41.16
2	25	.2646	26.46
3	10	.0756	7.56
4	3	.0081	.81
	100		100.00

The expected frequency of $x = 4$ is .81. Combine $x = 3$ and $x = 4$ into one category so that all expected frequencies are 5 or more.

x	Observed Frequencies	Expected Frequencies
0	30	24.01
1	32	41.16
2	25	26.46
3 or 4	13	8.37
	100	100.00

b. $\chi^2 = 6.17$

Degrees of freedom = $k - 1 = 4 - 1 = 3$.

Using the χ^2 table with $df = 3$, $\chi^2 = 6.17$ shows the p-value is greater than .10.

Using software, the p-value corresponding to $\chi^2 = 6.17$ is .1036.

p-value > .05; do not reject H_0. Conclude that the assumption of a binomial distribution cannot be rejected.