**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:

*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 |



Degrees of freedom = *k* – 1 = (3 – 1) = 2

Using the  table with *df* = 2,= 7.99 shows the *p*–value is between .025 and .01

Using Excel or Minitab, the *p*–value corresponding to = 7.99 is .0184

*p*–value .05, reject H0. Conclude not all population proportions are equal.

2. a. 





b. Multiple comparisons

For 1 vs. 2



*df* = *k* –1 = 3 – 1 = *z*  = 5.991

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  | Critical | Significant |
| Comparison | *p*i | *p*j | Difference | *n*i | *n*j | Value | 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 vs. 3. The others are not significant. We can conclude that the population proportions differ for populations 1 and 3.

3. a. *H*0:

*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 |



Degrees of freedom = *k* – 1 = (3 – 1) = 2

Using the  table with *df* = 2,= 2.12 shows the *p*–value is greater than .10

Using Excel or Minitab, the *p*–value corresponding to = 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. 





Overall 

4. a. *H*0:

*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 |



Degrees of freedom = *k* – 1 = (3 – 1) = 2

Using the  table with *df* = 2,= 14.74 shows the *p*–value is less than .01

Using Excel or Minitab, the *p*–value corresponding to = 14.74 is .0006

*p*–value < .05, reject *H*0. Conclude that the three suppliers do not provide equal proportions of defective components.

c. 





Multiple comparisons

For Supplier A vs. Supplier B

*df* = *k* –1 = 3 – 1 = *z*  = 5.991



|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  | Critical | Significant |
| Comparison | *p*i | *p*j | Difference | *n*i | *n*j | Value | 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 of these suppliers should remain candidates for use by Benson.

5. a. *H*0:

*H*a: Not all population proportions are equal

Observed Frequencies (*f*ij)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Gender | A | B | C | D | Total |
| Male | 49 | 44 | 49 | 39 | 181 |
| Female | 41 | 46 | 36 | 44 | 167 |
|  | 90 | 90 | 85 | 83 | 348 |

Expected Frequencies (*e*ij)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Gender | A | B | C | D | Total |
| Male | 46.81 | 46.81 | 44.21 | 43.17 | 181 |
| Female | 43.19 | 43.19 | 40.79 | 39.83 | 167 |
|  | 90 | 90 | 85 | 83 | 348 |

Chi Square Calculations (*f*ij – *e*ij)2 / *e*ij

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Gender | A | B | C | D | Total |
| Male | .10 | .17 | .52 | .40 | 1.19 |
| Female | .11 | .18 | .56 | .44 | 1.29 |



Degrees of freedom = *k* – 1 = (4 – 1) = 3

Using the  table with *df* = 3,= 2.49 shows the *p*–value is greater than .10

Using Excel or Minitab, the *p*–value corresponding to = 2.49 is .4771

*p*–value > .05, do not reject *H*0. Conclude that we are unable to reject the hypothesis that the population proportion of male fish are equal in all four locations.

b. No. There is no evidence that differences in agricultural contaminants found at the four locations have altered the gender proportions of the fish populations.

6. a.  14% error rate

 9% error rate

b. *H*0:

*H*a:

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 |



*df* = *k* – 1 = (2 – 1) = 1

Using the  table with *df =* 1,= 3.41 shows the *p*–value is between .10 and .05

Using Excel or Minitab, the *p*–value corresponding to = 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 1 degree of freedom. In this case the test statistic  is always equal to *z2*. 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 a two–tailed hypothesis tests about the equality of the two population proportions.

7. a. *H*0:

*H*a: Not all population proportions are equal

Observed Frequencies (*f*ij)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Social Net | Great Britain | Israel | Russia | USA | Total |
| Yes | 344 | 265 | 301 | 500 | 1410 |
| No | 456 | 235 | 399 | 500 | 1590 |
|  | 800 | 500 | 700 | 1000 | 3000 |

Expected Frequencies (*e*ij)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Social Net | Great Britain | Israel | Russia | USA | Total |
| Yes | 376 | 235 | 329 | 470 | 1410 |
| No | 424 | 265 | 371 | 530 | 1590 |
|  | 800 | 500 | 700 | 1000 | 3000 |

Chi Square Calculations (*f*ij – *e*ij)2 / *e*ij

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Social Net | Great Britain | Israel | Russia | USA | Total |
| Yes | 2.72 | 3.83 | 2.38 | 1.91 | 10.85 |
| No | 2.42 | 3.40 | 2.11 | 1.70 | 9.62 |



Degrees of freedom = *df* = *k* – 1 = (4 – 1) = 3

Using the  table with *df* = 3,= 20.47 shows the *p*–value is less than .01

Using Excel or Minitab, the *p*–value corresponding to = 20.47 is .0001

*p*–value .05, reject *H*0. Conclude the population proportions are not all equal.

b. Great Britain 344/800 = .43

Israel 265/500 = .53 (Largest with 53% of adults)

Russia 301/700 = .43

United States 500/1000 = .50

c. Multiple pairwise comparisons



where *df* = *k* –1 = 4 – 1 = 3 and  = 7.815

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| Comparison | *p*i | *p*j | Difference | *n*i | *n*j | *CVij* | Diff > *CVij* |
| GB vs I | 0.43 | 0.53 | 0.10 | 800 | 500 | 0.0793 | Yes |
| GB v R | 0.43 | 0.43 | 0.00 | 800 | 700 | 0.0716 |  |
| GB vs USA | 0.43 | 0.50 | 0.07 | 800 | 1000 | 0.0659 | Yes |
| I vs R | 0.53 | 0.43 | 0.10 | 500 | 700 | 0.0814 | Yes |
| I vs USA | 0.53 | 0.50 | 0.03 | 500 | 1000 | 0.0765 |  |
| R vs USA | 0.43 | 0.50 | 0.07 | 700 | 1000 | 0.0685 | Yes |

Only two comparisons are not significant: Great Britain and Russia and then Israel and United States. 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 |



Degrees of freedom = (*r* – 1)(*k* – 1) = (3 – 1)(3 – 1) = 4

Using the  table with *df* = 4,= 5.70 shows the *p*–value is greater than .10

Using Excel or Minitab, the *p*–value corresponding to = 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 |

= 7.86

Degrees of freedom = (2–1)(3–1) = 2

Using the  table with *df* = 2,= 7.86 shows the *p*–value is between .01 and .025.

Using Excel or Minitab, the *p*–value corresponding to = 7.86 is .0196.

*p*–value .05, reject H0. 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 |

= 19.77

Degrees of freedom = (*r –* 1)(*c* – 1) = (3 *–* 1)(3– 1) = 4

Using the  table with *df* = 4,= 19.77 shows the *p*–value is less than .005.

Using Excel or Minitab, the *p*–value corresponding to = 19.77 is .0006.

*p*–value .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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | Observed | Expected |  |
|  |  | Frequency | Frequency | Chi–square |
| Ticket | Flight | (*f*i) | (*e*i) | (*f*i – *e*i)2 / *e*i |
| First | Domestic | 29 | 35.59 | 1.22 |
| First | International | 22 | 15.41 | 2.82 |
| Business | Domestic | 95 | 150.73 | 20.61 |
| Business | International | 121 | 65.27 | 47.59 |
| Full Fare | Domestic | 518 | 455.68 | 8.52 |
| Full Fare | International | 135 | 197.32 | 19.68 |
|  | Totals: | 920 |  | 100.43 |

Degrees of freedom = (*r* – 1)(*c* – 1) = (3 – 1)(2 – 1) = 2

Using the  table with *df* = 2,= 100.43 shows the *p*–value is less than .005.

Using Excel or Minitab, the *p*–value corresponding to = 100.43 is .0000.

*p*–value .05, reject H0. 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 |
|  |  | = | 9.44 |
|  |  |  |  |

Degrees of freedom = (*r* – 1)(*c* – 1) = (3 – 1)(2 – 1) = 2

Using the  table with *df* = 2,= 9.44 shows the *p*–value is less than .01

Using Excel or Minitab, the *p*–value corresponding to = 9.44 is .0089

*p*–value .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. a. *H*0: Having health insurance is independent of the size of the company

*H*a: Having health insurance is not independent of the size of the company

Observed Frequencies (*f*ij)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Health Insurance | Small | Medium | Large | Total |
| Yes | 36 | 65 | 88 | 189 |
| No | 14 | 10 | 12 | 36 |
| Total | 50 | 75 | 100 | 225 |

Expected Frequencies (*e*ij)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Health Insurance | Small | Medium | Large | Total |
| Yes | 42 | 63 | 84 | 189 |
| No | 8 | 12 | 16 | 36 |
| Total | 50 | 75 | 100 | 225 |

Chi–Square Calculations (*f*ij – *e*ij)2 / *e*ij

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Health Insurance | Small | Medium | Large | Total |
| Yes | .86 | .06 | .19 | 1.11 |
| No | 4.50 | .33 | 1.00 | 5.83 |

= 6.94

Degrees of freedom = (*r –* 1)(*c* – 1)= (2 – 1)(3 – 1) = 2

Using the  table with *df* = 2,= 6.94 shows the *p*–value is between .025 and .05.

Using Excel or Minitab, the *p*–value corresponding to = 6.94 is .0311.

*p*–value .05, reject *H*0. Conclude health insurance coverage is not independent of the size of

the company. Health coverage is expected to vary depending on the size of the company.

b. Percentage of no coverage by company size

Small 14/50 = 28%

Medium 10/75 = 13%

Large 12/100 = 12%

More than twice as many small companies do not provide health insurance coverage when compared to medium and large companies.

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 |



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

Using the  table with *df* = 6,= 6.57 shows the *p*–value is greater than .10

Using Excel or Minitab, the *p*–value corresponding to = 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 |



Degrees of freedom = (*r* – 1)(*c* – 1) = (3– 1)(3 – 1) = 4

Using the  table with *df* = 4,= 17.03 shows the *p*–value is less than .005

Using Excel or Minitab, the *p*–value corresponding to = 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 rating are associated.

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

b. Observed Frequency (*fij*)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Country | | | | | |  |
| Response | G.B. | France | Italy | Spain | Ger. | U.S. | Total |
| Strongly favor | 141 | 161 | 298 | 133 | 128 | 204 | 1065 |
| Favor | 348 | 366 | 309 | 222 | 272 | 326 | 1843 |
| Oppose | 381 | 334 | 219 | 311 | 322 | 316 | 1883 |
| Strongly Oppose | 217 | 215 | 219 | 443 | 389 | 174 | 1657 |
| Total | 1087 | 1076 | 1045 | 1109 | 1111 | 1020 | 6448 |

Expected Frequency (*eij*)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Country | | | | | |  |
| Response | G.B. | France | Italy | Spain | Ger. | U.S. | Total |
| Strongly favor | 180 | 178 | 173 | 183 | 183 | 168 | 1065 |
| Favor | 311 | 307 | 299 | 317 | 318 | 291 | 1843 |
| Oppose | 317 | 315 | 305 | 324 | 324 | 298 | 1883 |
| Strongly Oppose | 279 | 276 | 268 | 285 | 286 | 263 | 1657 |
| Total | 1087 | 1076 | 1045 | 1109 | 1111 | 1020 | 6448 |



Degrees of freedom = (*r* – 1)(*c* – 1) = (4– 1)(6 – 1) = 15

The *p*–value is approximately 0.

*p*–value .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 | G.B. | France | Italy | Spain | Ger. | U.S. |  |
| 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.

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 |

= 4.01

Degrees of freedom = (*r* – 1)(*c* – 1) = (4– 1)(2 – 1) = 3

Using the  table with *df* = 3,= 4.01 shows the *p*–value is greater than .10.

Using Excel or Minitab, the *p*–value corresponding to = 4.01 is .2604.

*p*–value > .05, do not reject *H*0. Cannot reject the assumption that age and hours of sleep are independent.

1. Since 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 | Expected |  |
|  |  | Frequency | 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 |

= 45.36

Degrees of freedom = (*r* – 1)(*c* – 1) = (3– 1)(3 – 1) = 4

Using the  table with *df* = 2,= 45.36 shows the *p*–value is less than .005.

Using Excel or Minitab, the *p*–value corresponding to = 45.36 is .0000.

*p*–value .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, *e*2 = 200 (.40) = 80

*e*3 = 200 (.20) = 40

Observed frequencies: *f*1 = 60, *f*2 = 120, *f*3 = 20



*k* – 1 = 2 degrees of freedom

Using the  table with *df* = 2,= 35 shows the *p*–value is less than .005.

Using Excel or Minitab, the *p*–value corresponding to = 35 is approximately 0.

*p*–value  .01, reject. Conclude the proportions differ from .40, .40, and .20.

b.  = 9.210

Reject *H*0 if9.210

= 35, reject . Conclude the proportions differ from .40, .40, and .20.

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

 = 22.8 *s* = 6.27

The *z* values that create 6 intervals, each with probability .1667 are –.98, –.43, 0, .43, .98

|  |  |
| --- | --- |
| *z* | Cut off value of *x* |
| –.98 | 22.8 – .98 (6.27) = 16.66 |
| –.43 | 22.8 – .43 (6.27) = 20.11 |
| 0 | 22.8 + 0 (6.27) = 22.80 |
| .43 | 22.8 + .43 (6.27) = 25.49 |
| .98 | 22.8 + .98 (6.27) = 28.94 |

|  |  |  |  |
| --- | --- | --- | --- |
| Interval | Observed Frequency | Expected Frequency | Difference |
| less than 16.66 | 3 | 5 | –2 |
| 16.66 – 20.11 | 7 | 5 | 2 |
| 20.11 – 22.80 | 5 | 5 | 0 |
| 22.80 – 25.49 | 7 | 5 | 2 |
| 25.49– 28.94 | 3 | 5 | –2 |
| 28.94 and up | 5 | 5 | 0 |



Degrees of freedom = *k* – *p* – 1 = 6 – 2 – 1 = 3

Using the  table with *df* = 3,= 3.20 shows the *p*–value is greater than .10.

Using Excel or Minitab, the *p*–value corresponding to = 3.20 is .3618.

*p*–value > .05, do not reject. The claim that the data comes 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



*k* – 1 = 3 degrees of freedom

Using the  table with *df* = 3,= 6.87 shows the *p*–value is between .05 and .10.

Using Excel or Minitab, the *p*–value corresponding to = 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.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | Observed | Expected |  |
|  | Hypothesized | Frequency | Frequency | Chi Square |
| Category | Proportion | (*f*i) | (*e*i) | (*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 |  | = 5.85 |

*k* – 1 = 6 – 1 = 5 degrees of freedom

Using the  table with *df* = 5,= 5.85 shows the *p*–value is greater than .10

Using Excel or Minitab, the *p*–value corresponding to = 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



*k* – 1 = 4 degrees of freedom

Using the  table with *df* = 4,= 11.50 shows the *p*–value is between .01 and .025.

Using Excel or Minitab, the *p*–value corresponding to = 11.50 is .0215.

*p*–value < .05; reject . 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:

*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

 = 14.33

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

Using the  table with *df* = 6,= 14.33 shows the *p*–value is between .05 and .025.

Using Excel or Minitab, the *p*–value corresponding to = 14.33 is .0262

*p*–value .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. = 71 *s* = 17 *n* = 25 Use 5 classes

|  |  |  |  |
| --- | --- | --- | --- |
| Percentage | z | Data Value | |
| 20.00% | –.84 | 71–.84(17) = | 56.72 |
| 40.00% | –.25 | 71–.84(17) = | 66.75 |
| 60.00% | .25 | 71–.84(17) = | 75.25 |
| 80.00% | .84 | 71–.84(17) = | 85.28 |

|  |  |  |
| --- | --- | --- |
| Interval | Observed Frequency | Expected Frequency |
| less than 56.72 | 7 | 5 |
| 56.72 – 66.75 | 7 | 5 |
| 66.75 – 75.25 | 1 | 5 |
| 75.25 – 85.28 | 1 | 5 |
| 85.28 up | 9 | 5 |

= 11.20

Degrees of freedom = *k* – *p* – 1 = 5 – 2 – 1 = 2

Using the  table with *df* = 2,= 11.20 shows the *p*–value is less than .005.

Using Excel or Minitab, the *p*–value corresponding to = 11.20 is .0037.

*p*–value .01, reject H0. Conclude the distribution does not have a normal probability distribution.

26.  = 24.5 *s* = 3 *n* = 30 Use 6 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 |
| 25.79 – 27.41 | 7 | 5 |
| 27.41 up | 4 | 5 |

 = 2.80

Degrees of freedom = (*k – p* – 1) = 6 – 2 – 1 = 3

Using the  table with *df* = 3,= 2.80 shows the *p*–value is greater than .10.

Using Excel or Minitab, the *p*–value corresponding to = 2.80 is .4235.

*p*–value > .10, do not reject H0. The assumption of a normal distribution cannot be rejected.

27. a. *H*0: 

*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 |



Degrees of freedom = *k* – 1 = (3 – 1) = 2

Using the  table with *df* = 2,= 8.10 shows the *p*–value is between .025 and .01.

Using Excel or Minitab, the *p*–value corresponding to = 8.10 is .0174

*p*–value  .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. 





*df* = *k* –1 = 3 – 1 = 2 

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  | Critical | Significant |
| Comparison | *p*i | *p*j | Difference | *n*i | *n*j | Value | 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.

28. a. 







Bridgeport 8.8%, Los Alamos 11.7%, Naples 9%, Washington DC 8.5%

b. *H*0:

*H*a: Not all population proportions are equal

Observed Frequencies (*f*ij)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Millionaire | Bridgeport | Los' Alamos | Naples | Washington | Total |
| Yes | 44 | 35 | 36 | 34 | 149 |
| No | 456 | 265 | 364 | 366 | 1451 |
| Total | 500 | 300 | 400 | 400 | 1600 |

Expected Frequencies (*e*ij)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Millionaire | Bridgeport | Los' Alamos | Naples | Washington | Total |
| Yes | 46.56 | 27.94 | 37.25 | 37.25 | 149 |
| No | 453.44 | 272.06 | 362.75 | 362.75 | 1451 |
| Total | 500 | 300 | 400 | 400 | 1600 |

Chi Square Calculations (*f*ij – *e*ij)2 / *e*ij

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Millionaire | Bridgeport | Los' Alamos | Naples | Washington | Total |
| Yes | .14 | 1.79 | .04 | .28 | 2.25 |
| No | .01 | .18 | .00 | .03 | .23 |



Degrees of freedom = *k* – 1 = (4 – 1) = 3

Using the  table with *df* = 3,= 2.48 shows the *p*–value is greater than .10

Using Excel or Minitab, the *p*–value corresponding to = 2.48 is .4789

*p*–value > .05, do not reject H0. Cannot conclude that there is a difference among the population proportion of millionaires for these four cities.

29. a. *H*0:

*H*a: Not all population proportions are equal

Observed Frequency (*f*ij)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Work | Anchorage | Atlanta | Minneapolis | Total |
| Both | 57 | 70 | 63 | 190 |
| Only One | 33 | 50 | 27 | 110 |
| Total | 90 | 120 | 90 | 300 |

Expected Frequency (*e*ij)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Work | Anchorage | Atlanta | Minneapolis | Total |
| Both | 57 | 76 | 57 | 190 |
| Only One | 33 | 44 | 33 | 110 |
| Total | 90 | 120 | 90 | 300 |

Chi Square (*f*ij – *e*ij)2 / *e*ij

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Work | Anchorage | Atlanta | Minneapolis | Total |
| Both | .00 | .47 | .63 | 1.11 |
| Only One | .00 | .82 | 1.09 | 1.91 |

= 3.01

Degrees of freedom = *k* – 1 = 3 – 1 = 2

Using the  table with *df* = 2,= 3.01 shows the *p*–value is greater than .10.

Using Excel or Minitab, the *p*–value corresponding to = 3.01 is .2220.

*p*–value > .05, do not reject *H*0. We cannot conclude that the population proportion with both husband and wife in the workforce differs for these three cities.

b. The overall proportion of married couples with both husband and wife in the workforce is 190/300 = .633, or 63.3%.

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 (*fij*)

|  |  |  |  |
| --- | --- | --- | --- |
| Preferred | Gender | |  |
| 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 (*eij*)

|  |  |  |  |
| --- | --- | --- | --- |
| Preferred | Gender | |  |
| 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 (*fij* – *eij*)2/ *eij*

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



Degrees of freedom = (*r* – 1)(*c* – 1) = (3 – 1)(2 – 1) = 2

Using the  table with *df* = 2,= 9.56 shows the *p*–value is less than .01.

Using Excel or Minitab, the *p*–value corresponding to = 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

|  |  |  |
| --- | --- | --- |
| Preferred | Gender | |
| 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)

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

Expected Frequencies (*e*ij)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Church | Age | | | |  |
| Attendance | 20 to 29 | 30 to 39 | 40 to 49 | 50 to 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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Church | Age | | | |  |
| Attendance | 20 to 29 | 30 to 39 | 40 to 49 | 50 to 59 | Total |
| Yes | 3.51 | .06 | .62 | .75 | 4.94 |
| No | 2.68 | .05 | .47 | .58 | 3.78 |



Degrees of freedom = (*r –* 1)(*c –* 1) = (2 *–* 1)(4 *–* 1) = 3

Using the  table with *df* = 3,= 8.72 shows the *p*–value is between .025 and .05.

Using Excel or Minitab, the *p*–value corresponding to = 8.72 is .0333.

*p*–value .05, reject. Conclude church attendance is not independent of age.

Church attendance by age group:

20 – 29 31/100  31%

30 – 39 63/150  42%

40 – 49 94/200  47%

50 – 59 72/150  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 |

= 6.17

Degrees of freedom = (*r –* 1)(*c –* 1) = (2 *–* 1)(7 *–* 1) =6

Using the  table with *df* = 6,= 6.17 shows the *p*–value is greater than .10.

Using Excel or Minitab, the *p*–value corresponding to = 6.17 is .4044.

*p*–value > .05, do not reject H0. 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. *H*0: The market shares for the five automobiles in Chicago are .24, .21, .19, .18, .17

*H*a: The market shares for the five automobiles in Chicago differ from the above shares

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Hypothesized | Observed | Expected | Chi Square |
| Compact Car | Market Share | Frequency | Frequency | (*f*i – *e*i)2 / *e*i |
| Chevy Cruze | .24 | 108 | 96 | 1.50 |
| Ford Focus | .21 | 92 | 84 | 0.76 |
| Hyundai Elantra | .19 | 64 | 76 | 1.89 |
| Honda Civic | .18 | 84 | 72 | 2.00 |
| Toyota Corolla | .17 | 52 | 68 | 3.76 |



Degrees of Freedom = *k* – 1 = 5 – 1 = 4

Using the  table with *df* = 4,= 9.92 shows the *p*–value is between .05 and .025.

Using Excel or Minitab, the *p*–value corresponding to = 9.92 is .0418

*p*–value < .05, reject. Conclude that the markets shares for the five compact cars in Chicago differ from the market shares reported by *Motor Trend*.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Hypothesized | Sample |  |
| Compact Car | Market Share | Market Share | Difference |
| Chevy Cruze | .24 | 108/400 = .27 | .03 |
| Ford Focus | .21 | 92/400 = .23 | .02 |
| Hyundai Elantra | .19 | 64/400 = .16 | –.03 |
| Honda Civic | .18 | 84/400 = .21 | .03 |
| Toyota Corolla | .17 | 52/400 = .13 | –.04 |

Chevy Cruze (.03) Honda Civic (.03) and Ford Focus (.02) show higher market shares in Chicago. Toyota Corolla (–.04) and Hyundai (–.03) show lower market shares in Chicago.

34.  = 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 |

= 2

Degrees of freedom = *k* – *p* – 1 = 8 – 2 – 1 = 5

Using the  table with *df* = 5,= 2.00 shows the *p*–value is greater than .10.

Using Excel or Minitab, the *p*–value corresponding to = 2.00 is .8491.

*p*–value > .05, do not reject. The assumption of a normal distribution cannot be rejected.

35. a.

|  |  |  |  |
| --- | --- | --- | --- |
| *x* | Observed Frequencies | Binomial Prob.  *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. = 6.17

Degrees of freedom = *k* – 1 = 4 – 1 = 3

Using the  table with *df* = 3,= 6.17 shows the *p*–value is greater than .10.

Using Excel or Minitab, the *p*–value corresponding to = 6.17 is .1036.

*p*–value > .05, do not reject H0. Conclude that the assumption of a binomial distribution cannot be rejected.