**Chapter 11: Chi-Square and ANOVA Tests**

This chapter presents material on three more hypothesis tests. One is used to determine significant relationship between two qualitative variables, the second is used to determine if the sample data has a particular distribution, and the last is used to determine significant relationships between means of 3 or more samples.

**Section 11.1: Chi-Square Test for Independence**

Remember, qualitative data is where you collect data on individuals that are categories or names. Then you would count how many of the individuals had particular qualities. An example is that there is a theory that there is a relationship between breastfeeding and autism. To determine if there is a relationship, researchers could collect the time period that a mother breastfed her child and if that child was diagnosed with autism. Then you would have a table containing this information. Now you want to know if each cell is independent of each other cell. Remember, independence says that one event does not affect another event. Here it means that having autism is independent of being breastfed. What you really want is to see if they are not independent. In other words, does one affect the other? If you were to do a hypothesis test, this is your alternative hypothesis and the null hypothesis is that they are independent. There is a hypothesis test for this and it is called the **Chi-Square Test for Independence**. Technically it should be called the Chi-Square Test for Dependence, but for historical reasons it is known as the test for independence. Just as with previous hypothesis tests, all the steps are the same except for the assumptions and the test statistic.

**Hypothesis Test for Chi-Square Test**

1. State the null and alternative hypotheses and the level of significance

 the two variables are independent (this means that the one variable is not affected by the other)

 the two variables are dependent (this means that the one variable is affected by the other)

Also, state your  level here.

1. State and check the assumptions for the hypothesis test
2. A random sample is taken.
3. Expected frequencies for each cell are greater than or equal to 5 (The expected frequencies, *E*, will be calculated later, and this assumption means ).
4. Find the test statistic and p-value

Finding the test statistic involves several steps. First the data is collected and counted, and then it is organized into a table (in a table each entry is called a cell). These values are known as the observed frequencies, which the symbol for an observed frequency is *O*. Each table is made up of rows and columns. Then each row is totaled to give a row total and each column is totaled to give a column total.

The null hypothesis is that the variables are independent. Using the multiplication rule for independent events you can calculate the probability of being one value of the first variable, *A*, and one value of the second variable, *B* (the probability of a particular cell ). Remember in a hypothesis test, you assume that  is true, the two variables are assumed to be independent.



Now you want to find out how many individuals you expect to be in a certain cell. To find the expected frequencies, you just need to multiply the probability of that cell times the total number of individuals. Do not round the expected frequencies.



If the variables are independent the expected frequencies and the observed frequencies should be the same. The test statistic here will involve looking at the difference between the expected frequency and the observed frequency for each cell. Then you want to find the “total difference” of all of these differences. The larger the total, the smaller the chances that you could find that test statistic given that the assumption of independence is true. That means that the assumption of independence is not true. How do you find the test statistic? First find the differences between the observed and expected frequencies. Because some of these differences will be positive and some will be negative, you need to square these differences. These squares could be large just because the frequencies are large, you need to divide by the expected frequencies to scale them. Then finally add up all of these fractional values. This is the test statistic.

**Test Statistic:**

The symbol for Chi-Square is 

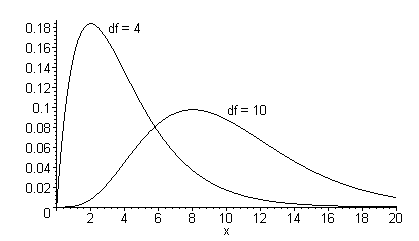


where *O* is the observed frequency and *E* is the expected frequency

**Distribution of Chi-Square**

 has different curves depending on the degrees of freedom. It is skewed to the right for small degrees of freedom and gets more symmetric as the degrees of freedom increases (see figure #11.1.1). Since the test statistic involves squaring the differences, the test statistics are all positive. A chi-squared test for independence is always right tailed.

**Figure #11.1.1: Chi-Square Distribution**



p-value:

Using the TI-83/84: 

Using R: 

Where the degrees of freedom is 

1. Conclusion

This is where you write reject  or fail to reject . The rule is: if the p-value < , then reject . If the p-value , then fail to reject 

1. Interpretation

This is where you interpret in real world terms the conclusion to the test. The conclusion for a hypothesis test is that you either have enough evidence to show  is true, or you do not have enough evidence to show  is true.

**Example #11.1.1: Hypothesis Test with Chi-Square Test Using Formula**

Is there a relationship between autism and breastfeeding? To determine if there is, a researcher asked mothers of autistic and non-autistic children to say what time period they breastfed their children. The data is in table #11.1.1 (Schultz, Klonoff-Cohen, Wingard, Askhoomoff, Macera, Ji & Bacher, 2006). Do the data provide enough evidence to show that that breastfeeding and autism are independent? Test at the1% level.

**Table #11.1.1: Autism Versus Breastfeeding**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Autism | Breast Feeding Timelines | | | | Row Total |
| None | Less than 2 months | 2 to 6 months | More than 6 months |
| Yes | 241 | 198 | 164 | 215 | 818 |
| No | 20 | 25 | 27 | 44 | 116 |
| Column Total | 261 | 223 | 191 | 259 | 934 |

**Solution:**

1. State the null and alternative hypotheses and the level of significance

 Breastfeeding and autism are independent

 Breastfeeding and autism are dependent



1. State and check the assumptions for the hypothesis test
2. A random sample of breastfeeding time frames and autism incidence was taken.
3. Expected frequencies for each cell are greater than or equal to 5 (ie. ). See step 3. All expected frequencies are more than 5.
4. Find the test statistic and p-value

Test statistic:

First find the expected frequencies for each cell.









Others are done similarly. It is easier to do the calculations for the test statistic with a table, the others are in table #11.1.2 along with the calculation for the test statistic. (Note: the column of  should add to 0 or close to 0.)

**Table #11.1.2: Calculations for Chi-Square Test Statistic**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *O* | *E* |  |  |  |
| 241 | 228.585 | 12.415 | 154.132225 | 0.674288448 |
| 198 | 195.304 | 2.696 | 7.268416 | 0.03721591 |
| 164 | 167.278 | -3.278 | 10.745284 | 0.064236086 |
| 215 | 226.833 | -11.833 | 140.019889 | 0.617281828 |
| 20 | 32.4154 | -12.4154 | 154.1421572 | 4.755213792 |
| 25 | 27.6959 | -2.6959 | 7.26787681 | 0.262417066 |
| 27 | 23.7216 | 3.2784 | 10.74790656 | 0.453085229 |
| 44 | 32.167 | 11.833 | 140.019889 | 4.352904809 |
| Total |  | 0.0001 |  | 11.2166432 |

The test statistic formula is , which is the total of the last column in table #11.1.2.

p-value:



Using TI-83/84: 

Using R: 

1. Conclusion

Fail to reject  since the p-value is more than 0.01.

1. Interpretation

There is not enough evidence to show that breastfeeding and autism are dependent. This means that you cannot say that the whether a child is breastfed or not will indicate if that the child will be diagnosed with autism.

**Example #11.1.2: Hypothesis Test with Chi-Square Test Using Technology**

Is there a relationship between autism and breastfeeding? To determine if there is, a researcher asked mothers of autistic and non-autistic children to say what time period they breastfed their children. The data is in table #11.1.1 (Schultz, Klonoff-Cohen, Wingard, Askhoomoff, Macera, Ji & Bacher, 2006). Do the data provide enough evidence to show that that breastfeeding and autism are independent? Test at the1% level.

**Solution:**

1. State the null and alternative hypotheses and the level of significance

 Breastfeeding and autism are independent

 Breastfeeding and autism are dependent

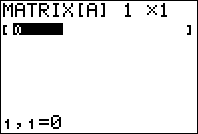


1. State and check the assumptions for the hypothesis test
2. A random sample of breastfeeding time frames and autism incidence was taken.
3. Expected frequencies for each cell are greater than or equal to 5 (ie. ). See step 3. All expected frequencies are more than 5.
4. Find the test statistic and p-value

Test statistic:

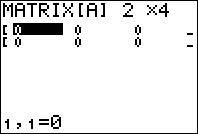
To use the TI-83/84 calculator to compute the test statistic, you must first put the data into the calculator. However, this process is different than for other hypothesis tests. You need to put the data in as a matrix instead of in the list. Go into the MATRX menu then move over to EDIT and choose 1:[A]. This will allow you to type the table into the calculator. Figure #11.1.2 shows what you will see on your calculator when you choose 1:[A] from the EDIT menu.

**Figure #11.1.2: Matrix Edit Menu on TI-83/84**



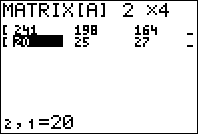
The table has 2 rows and 4 columns (don’t include the row total column and the column total row in your count). You need to tell the calculator that you have a 2 by 4. The 1 X1 (you might have another size in your matrix, but it doesn’t matter because you will change it) on the calculator is the size of the matrix. So type 2 ENTER and 4 ENTER and the calculator will make a matrix of the correct size. See figure #11.1.3.

**Figure #11.1.3: Matrix Setup for Table**



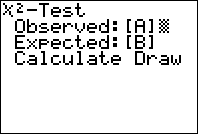
Now type the table in by pressing ENTER after each cell value. Figure #11.1.4 contains the complete table typed in. Once you have the data in, press QUIT.

**Figure #11.1.4: Data Typed into Matrix**



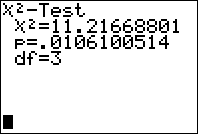
To run the test on the calculator, go into STAT, then move over to TEST and choose -Test from the list. The setup for the test is in figure #11.1.5.

**Figure #11.1.5: Setup for Chi-Square Test on TI-83/84**

****

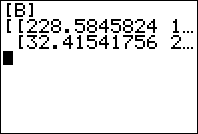
Once you press ENTER on Calculate you will see the results in figure #11.1.6.

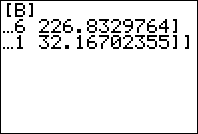
**Figure #11.1.6: Results for Chi-Square Test on TI-83/84**

****

The test statistic is  and the p-value is . Notice that the calculator calculates the expected values for you and places them in matrix B. To review the expected values, go into MATRX and choose 2:[B]. Figure #11.1.7 shows the output. Press the right arrows to see the entire matrix.

**Figure #11.1.7: Expected Frequency for Chi-Square Test on TI-83/84**





To compute the test statistic and p-value with R,

row1 = c(data from row 1 separated by commas)

row2 = c(data from row 2 separated by commas)

keep going until you have all of your rows typed in.

data.table = rbind(row1, row2, …) – makes the data into a table. You can call it what ever you want. It does not have to be data.table.

data.table – use if you want to look at the table

chisq.test(data.table) – calculates the chi-squared test for independence

chisq.test(data.table)$expected – let’s you see the expected values

For this example, the commands would be

row1 = c(241, 198, 164, 215)

row2 = c(20, 25, 27, 44)

data.table = rbind(row1, row2)

data.table

Output:

[,1] [,2] [,3] [,4]

row1 241 198 164 215

row2 20 25 27 44

chisq.test(data.table)

Output:

Pearson's Chi-squared test

data: data.table

X-squared = 11.217, df = 3, p-value = 0.01061

chisq.test(data.table)$expected

Output:

[,1] [,2] [,3] [,4]

row1 228.58458 195.30407 167.27837 226.83298

row2 32.41542 27.69593 23.72163 32.16702

The test statistic is  and the p-value is .

1. Conclusion

Fail to reject  since the p-value is more than 0.01.

1. Interpretation

There is not enough evidence to show that breastfeeding and autism are dependent. This means that you cannot say that the whether a child is breastfed or not will indicate if that the child will be diagnosed with autism.

**Example #11.1.3: Hypothesis Test with Chi-Square Test Using Formula**

The World Health Organization (WHO) keeps track of how many incidents of leprosy there are in the world. Using the WHO regions and the World Banks income groups, one can ask if an income level and a WHO region are dependent on each other in terms of predicting where the disease is. Data on leprosy cases in different countries was collected for the year 2011 and a summary is presented in table #11.1.3 ("Leprosy: Number of," 2013). Is there evidence to show that income level and WHO region are independent when dealing with the disease of leprosy? Test at the 5% level.

**Table #11.1.3: Number of Leprosy Cases**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| WHO Region | World Bank Income Group | | | | Row Total |
| High Income | Upper Middle Income | Lower Middle Income | Low Income |
| Americas | 174 | 36028 | 615 | 0 | 36817 |
| Eastern Mediterranean | 54 | 6 | 1883 | 604 | 2547 |
| Europe | 10 | 0 | 0 | 0 | 10 |
| Western Pacific | 26 | 216 | 3689 | 1155 | 5086 |
| Africa | 0 | 39 | 1986 | 15928 | 17953 |
| South-East Asia | 0 | 0 | 149896 | 10236 | 160132 |
| Column Total | 264 | 36289 | 158069 | 27923 | 222545 |

**Solution:**

1. State the null and alternative hypotheses and the level of significance

 WHO region and Income Level when dealing with the disease of leprosy are independent

 WHO region and Income Level when dealing with the disease of leprosy are dependent



1. State and check the assumptions for the hypothesis test
2. A random sample of incidence of leprosy was taken from different countries and the income level and WHO region was taken.
3. Expected frequencies for each cell are greater than or equal to 5 (ie. ). See step 3. There are actually 4 expected frequencies that are less than 5, and the results of the test may not be valid. If you look at the expected frequencies you will notice that they are all in Europe. This is because Europe didn’t have many cases in 2011.
4. Find the test statistic and p-value

Test statistic:

First find the expected frequencies for each cell.









Others are done similarly. It is easier to do the calculations for the test statistic with a table, and the others are in table #11.1.4 along with the calculation for the test statistic.

**Table #11.1.4: Calculations for Chi-Square Test Statistic**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *O* | *E* |  |  |  |
| 174 | 43.675 | 130.325 | 16984.564 | 388.8838719 |
| 54 | 3.021 | 50.979 | 2598.813 | 860.1218328 |
| 10 | 0.012 | 9.988 | 99.763 | 8409.746711 |
| 26 | 6.033 | 19.967 | 398.665 | 66.07628214 |
| 0 | 21.297 | -21.297 | 453.572 | 21.29722977 |
| 0 | 189.961 | -189.961 | 36085.143 | 189.9608978 |
| 36028 | 6003.514 | 30024.486 | 901469735.315 | 150157.0038 |
| 6 | 415.323 | -409.323 | 167545.414 | 403.4097962 |
| 0 | 1.631 | -1.631 | 2.659 | 1.6306365 |
| 216 | 829.342 | -613.342 | 376188.071 | 453.5983897 |
| 39 | 2927.482 | -2888.482 | 8343326.585 | 2850.001268 |
| 0 | 26111.708 | -26111.708 | 681821316.065 | 26111.70841 |
| 615 | 26150.335 | -25535.335 | 652053349.724 | 24934.7988 |
| 1883 | 1809.080 | 73.920 | 5464.144 | 3.020398811 |
| 0 | 7.103 | -7.103 | 50.450 | 7.1027882 |
| 3689 | 3612.478 | 76.522 | 5855.604 | 1.620938405 |
| 1986 | 12751.636 | -10765.636 | 115898911.071 | 9088.944681 |
| 149896 | 113738.368 | 36157.632 | 1307374351.380 | 11494.57632 |
| 0 | 4619.475 | -4619.475 | 21339550.402 | 4619.475122 |
| 604 | 319.575 | 284.425 | 80897.421 | 253.1404187 |
| 0 | 1.255 | -1.255 | 1.574 | 1.25471253 |
| 1155 | 638.147 | 516.853 | 267137.238 | 418.6140882 |
| 15928 | 2252.585 | 13675.415 | 187016964.340 | 83023.25138 |
| 10236 | 20091.963 | -9855.963 | 97140000.472 | 4834.769106 |
| Total |  | 0.000 |  | 328594.008 |

The test statistic formula is , which is the total of the last column in table #11.1.2.

p-value:



Using the TI-83/84: 

Using R: 

1. Conclusion

Reject  since the p-value is less than 0.05.

1. Interpretation

There is enough evidence to show that WHO region and income level are dependent when dealing with the disease of leprosy. WHO can decide how to focus their efforts based on region and income level. Do remember though that the results may not be valid due to the expected frequencies not all be more than 5.

**Example #11.1.4: Hypothesis Test with Chi-Square Test Using Technology**

The World Health Organization (WHO) keeps track of how many incidents of leprosy there are in the world. Using the WHO regions and the World Banks income groups, one can ask if an income level and a WHO region are dependent on each other in terms of predicting where the disease is. Data on leprosy cases in different countries was collected for the year 2011 and a summary is presented in table #11.1.3 ("Leprosy: Number of," 2013). Is there evidence to show that income level and WHO region are independent when dealing with the disease of leprosy? Test at the 5% level.

**Solution:**

1. State the null and alternative hypotheses and the level of significance

 WHO region and Income Level when dealing with the disease of leprosy are independent

 WHO region and Income Level when dealing with the disease of leprosy are dependent

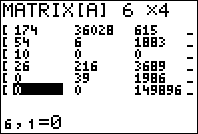


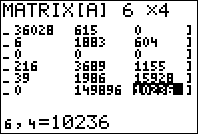
1. State and check the assumptions for the hypothesis test
2. A random sample of incidence of leprosy was taken from different countries and the income level and WHO region was taken.
3. Expected frequencies for each cell are greater than or equal to 5 (ie. ). See step 3. There are actually 4 expected frequencies that are less than 5, and the results of the test may not be valid. If you look at the expected frequencies you will notice that they are all in Europe. This is because Europe didn’t have many cases in 2011.
4. Find the test statistic and p-value

Test statistic:

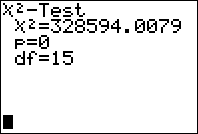
Using the TI-83/84. See example #11.1.2 for the process of doing the test on the calculator. Remember, you need to put the data in as a matrix instead of in the list.

**Figure #11.1.8: Setup for Matrix on TI-83/84**



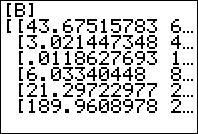


**Figure #11.1.9: Results for Chi-Square Test on TI-83/84**





**Figure #11.1.10: Expected Frequency for Chi-Square Test on TI-83/84**



Press the right arrow to look at the other expected frequencies.

p-value:



Using R:

row1=c(174, 36028, 615, 0)

row2=c(54, 6, 1883, 604)

row3=c(10, 0, 0, 0)

row4=c(26, 216, 3689, 1155)

row5=c(0, 39, 1986, 15928)

row6=c(0, 0, 149896, 10236)

chisq.test(data.table)

Pearson's Chi-squared test

data: data.table

X-squared = 328590, df = 15, p-value < 2.2e-16

Warning message:

In chisq.test(data.table) : Chi-squared approximation may be incorrect

chisq.test(data.table)$expected

[,1] [,2] [,3] [,4]

row1 43.67515783 6003.514404 2.615034e+04 4619.475122

row2 3.02144735 415.323117 1.809080e+03 319.575281

row3 0.01186277 1.630637 7.102788e+00 1.254713

row4 6.03340448 829.341724 3.612478e+03 638.146793

row5 21.29722977 2927.481709 1.275164e+04 2252.585405

row6 189.96089780 26111.708410 1.137384e+05 20091.962686

Warning message:

In chisq.test(data.table) : Chi-squared approximation may be incorrect

 and p-value = 

1. Conclusion

Reject  since the p-value is less than 0.05.

1. Interpretation

There is enough evidence to show that WHO region and income level are dependent when dealing with the disease of leprosy. WHO can decide how to focus their efforts based on region and income level. Do remember though that the results may not be valid due to the expected frequencies not all be more than 5.

**Section 11.1: Homework**

In each problem show all steps of the hypothesis test. If some of the assumptions are not met, note that the results of the test may not be correct and then continue the process of the hypothesis test.

1. The number of people who survived the Titanic based on class and sex is in table #11.1.5 ("Encyclopedia Titanica," 2013). Is there enough evidence to show that the class and the sex of a person who survived the Titanic are independent? Test at the 5% level.

**Table #11.1.5: Surviving the Titanic**

|  |  |  |  |
| --- | --- | --- | --- |
| Class | Sex | | Total |
| Female | Male |
| 1st | 134 | 59 | 193 |
| 2nd | 94 | 25 | 119 |
| 3rd | 80 | 58 | 138 |
| Total | 308 | 142 | 450 |

1. Researchers watched groups of dolphins off the coast of Ireland in 1998 to determine what activities the dolphins partake in at certain times of the day ("Activities of dolphin," 2013). The numbers in table #11.1.6 represent the number of groups of dolphins that were partaking in an activity at certain times of days. Is there enough evidence to show that the activity and the time period are independent for dolphins? Test at the 1% level.

**Table #11.1.6: Dolphin Activity**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Activity | Period | | | | Row  Total |
| Morning | Noon | Afternoon | Evening |
| Travel | 6 | 6 | 14 | 13 | 39 |
| Feed | 28 | 4 | 0 | 56 | 88 |
| Social | 38 | 5 | 9 | 10 | 62 |
| Column Total | 72 | 15 | 23 | 79 | 189 |

1. Is there a relationship between autism and what an infant is fed? To determine if there is, a researcher asked mothers of autistic and non-autistic children to say what they fed their infant. The data is in table #11.1.7 (Schultz, Klonoff-Cohen, Wingard, Askhoomoff, Macera, Ji & Bacher, 2006). Do the data provide enough evidence to show that that what an infant is fed and autism are independent? Test at the1% level.

**Table #11.1.7: Autism Versus Breastfeeding**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Autism | Feeding | | | Row Total |
| Brest-feeding | Formula with DHA/ARA | Formula without DHA/ARA |
| Yes | 12 | 39 | 65 | 116 |
| No | 6 | 22 | 10 | 38 |
| Column Total | 18 | 61 | 75 | 154 |

1. A person’s educational attainment and age group was collected by the U.S. Census Bureau in 1984 to see if age group and educational attainment are related. The counts in thousands are in table #11.1.8 ("Education by age," 2013). Do the data show that educational attainment and age are independent? Test at the 5% level.

**Table #11.1.8: Educational Attainment and Age Group**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Education | Age Group | | | | | Row Total |
| 25-34 | 35-44 | 45-54 | 55-64 | >64 |
| Did not complete HS | 5416 | 5030 | 5777 | 7606 | 13746 | 37575 |
| Competed HS | 16431 | 1855 | 9435 | 8795 | 7558 | 44074 |
| College 1-3 years | 8555 | 5576 | 3124 | 2524 | 2503 | 22282 |
| College 4 or more years | 9771 | 7596 | 3904 | 3109 | 2483 | 26863 |
| Column Total | 40173 | 20057 | 22240 | 22034 | 26290 | 130794 |

1. Students at multiple grade schools were asked what their personal goal (get good grades, be popular, be good at sports) was and how important good grades were to them (1 very important and 4 least important). The data is in table #11.1.9 ("Popular kids datafile," 2013). Do the data provide enough evidence to show that goal attainment and importance of grades are independent? Test at the 5% level.

**Table #11.1.9: Personal Goal and Importance of Grades**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Goal | Grades Importance Rating | | | | Row Total |
| 1 | 2 | 3 | 4 |
| Grades | 70 | 66 | 55 | 56 | 247 |
| Popular | 14 | 33 | 45 | 49 | 141 |
| Sports | 10 | 24 | 33 | 23 | 90 |
| Column Total | 94 | 123 | 133 | 128 | 478 |

1. Students at multiple grade schools were asked what their personal goal (get good grades, be popular, be good at sports) was and how important being good at sports were to them (1 very important and 4 least important). The data is in table #11.1.10 ("Popular kids datafile," 2013). Do the data provide enough evidence to show that goal attainment and importance of sports are independent? Test at the 5% level.

**Table #11.1.10: Personal Goal and Importance of Sports**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Goal | Sports Importance Rating | | | | Row Total |
| 1 | 2 | 3 | 4 |
| Grades | 83 | 81 | 55 | 28 | 247 |
| Popular | 32 | 49 | 43 | 17 | 141 |
| Sports | 50 | 24 | 14 | 2 | 90 |
| Column Total | 165 | 154 | 112 | 47 | 478 |

1. Students at multiple grade schools were asked what their personal goal (get good grades, be popular, be good at sports) was and how important having good looks were to them (1 very important and 4 least important). The data is in table #11.1.11 ("Popular kids datafile," 2013). Do the data provide enough evidence to show that goal attainment and importance of looks are independent? Test at the 5% level.

**Table #11.1.11: Personal Goal and Importance of Looks**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Goal | Looks Importance Rating | | | | Row Total |
| 1 | 2 | 3 | 4 |
| Grades | 80 | 66 | 66 | 35 | 247 |
| Popular | 81 | 30 | 18 | 12 | 141 |
| Sports | 24 | 30 | 17 | 19 | 90 |
| Column Total | 185 | 126 | 101 | 66 | 478 |

1. Students at multiple grade schools were asked what their personal goal (get good grades, be popular, be good at sports) was and how important having money were to them (1 very important and 4 least important). The data is in table #11.1.12 ("Popular kids datafile," 2013). Do the data provide enough evidence to show that goal attainment and importance of money are independent? Test at the 5% level.

**Table #11.1.12: Personal Goal and Importance of Money**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Goal | Money Importance Rating | | | | Row Total |
| 1 | 2 | 3 | 4 |
| Grades | 14 | 34 | 71 | 128 | 247 |
| Popular | 14 | 29 | 35 | 63 | 141 |
| Sports | 6 | 12 | 26 | 46 | 90 |
| Column Total | 34 | 75 | 132 | 237 | 478 |

**Section 11.2: Chi-Square Goodness of Fit**

In probability, you calculated probabilities using both experimental and theoretical methods. There are times when it is important to determine how well the experimental values match the theoretical values. An example of this is if you wish to verify if a die is fair. To determine if observed values fit the expected values, you want to see if the difference between observed values and expected values is large enough to say that the test statistic is unlikely to happen if you assume that the observed values fit the expected values. The test statistic in this case is also the chi-square. The process is the same as for the chi-square test for independence.

**Hypothesis Test for Goodness of Fit Test**

1. State the null and alternative hypotheses and the level of significance

 The data are consistent with a specific distribution

 The data are not consistent with a specific distribution

Also, state your  level here.

1. State and check the assumptions for the hypothesis test
2. A random sample is taken.
3. Expected frequencies for each cell are greater than or equal to 5 (The expected frequencies, *E*, will be calculated later, and this assumption means ).
4. Find the test statistic and p-value

Finding the test statistic involves several steps. First the data is collected and counted, and then it is organized into a table (in a table each entry is called a cell). These values are known as the observed frequencies, which the symbol for an observed frequency is *O*. The table is made up of *k* entries. The total number of observed frequencies is *n*. The expected frequencies are calculated by multiplying the probability of each entry, *p*, times *n*.



**Test Statistic:**



where *O* is the observed frequency and *E* is the expected frequency

Again, the test statistic involves squaring the differences, so the test statistics are all positive. Thus a chi-squared test for goodness of fit is always right tailed.

p-value:

Using the TI-83/84: 

Using R: 

Where the degrees of freedom is 

1. Conclusion

This is where you write reject  or fail to reject . The rule is: if the p-value < , then reject . If the p-value , then fail to reject 

1. Interpretation

This is where you interpret in real world terms the conclusion to the test. The conclusion for a hypothesis test is that you either have enough evidence to show  is true, or you do not have enough evidence to show  is true.

**Example #11.2.1: Goodness of Fit Test Using the Formula**

Suppose you have a die that you are curious if it is fair or not. If it is fair then the proportion for each value should be the same. You need to find the observed frequencies and to accomplish this you roll the die 500 times and count how often each side comes up. The data is in table #11.2.1. Do the data show that the die is fair? Test at the 5% level.

**Table #11.2.1: Observed Frequencies of Die**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Die values | 1 | 2 | 3 | 4 | 5 | 6 | Total |
| Observed Frequency | 78 | 87 | 87 | 76 | 85 | 87 | 100 |

**Solution:**

1. State the null and alternative hypotheses and the level of significance

 The observed frequencies are consistent with the distribution for fair die (the die is fair)

 The observed frequencies are not consistent with the distribution for fair die (the die is not fair)



1. State and check the assumptions for the hypothesis test
2. A random sample is taken since each throw of a die is a random event.
3. Expected frequencies for each cell are greater than or equal to 5. See step 3.
4. Find the test statistic and p-value

First you need to find the probability of rolling each side of the die. The sample space for rolling a die is {1, 2, 3, 4, 5, 6}. Since you are assuming that the die is fair, then .

Now you can find the expected frequency for each side of the die. Since all the probabilities are the same, then each expected frequency is the same.



Test Statistic:

It is easier to calculate the test statistic using a table.

**Table #11.2.2: Calculation of the Chi-Square Test Statistic**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *O* | *E* |  |  |  |
| 78 | 83.33 | -5.33 | 28.4089 | 0.340920437 |
| 87 | 83.33 | 3.67 | 13.4689 | 0.161633265 |
| 87 | 83.33 | 3.67 | 13.4689 | 0.161633265 |
| 76 | 83.33 | -7.33 | 53.7289 | 0.644772591 |
| 85 | 83.33 | 1.67 | 2.7889 | 0.033468139 |
| 87 | 83.33 | 3.67 | 13.4689 | 0.161633265 |
| Total |  | 0.02 |  | 1.504060962 |

The test statistic is  1.504060962

The degrees of freedom are 

Using TI-83/84: 

Using R: 

1. Conclusion

Fail to reject  since the p-value is greater than 0.05.

1. Interpretation

There is not enough evidence to show that the die is not consistent with the distribution for a fair die. There is not enough evidence to show that the die is not fair.

**Example #11.2.2: Goodness of Fit Test Using Technology**

Suppose you have a die that you are curious if it is fair or not. If it is fair then the proportion for each value should be the same. You need to find the observed frequencies and to accomplish this you roll the die 500 times and count how often each side comes up. The data is in table #11.2.1. Do the data show that the die is fair? Test at the 5% level.

**Solution:**

1. State the null and alternative hypotheses and the level of significance

 The observed frequencies are consistent with the distribution for fair die (the die is fair)

 The observed frequencies are not consistent with the distribution for fair die (the die is not fair)



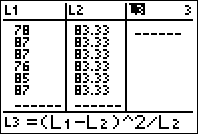
1. State and check the assumptions for the hypothesis test
2. A random sample is taken since each throw of a die is a random event.
3. Expected frequencies for each cell are greater than or equal to 5. See step 3.
4. Find the test statistic and p-value

Using the TI-83/84 calculator:

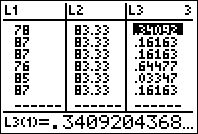
**Using the TI-83:**

To use the TI-83 calculator to compute the test statistic, you must first put the data into the calculator. Type the observed frequencies into L1 and the expected frequencies into L2. Then you will need to go to L3, arrow up onto the name, and type in . Now you use 1-Var Stats L3 to find the total. See figure #11.2.1 for the initial setup, figure #11.2.2 for the results of that calculation, and figure #11.2.3 for the result of the 1-Var Stats L3.

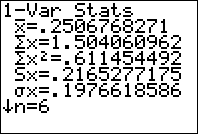
**Figure #11.2.1: Input into TI-83**

****

**Figure #11.2.2: Result for L3 on TI-83**



**Figure #11.2.3: 1-Var Stats L3 Result on TI-83**

****

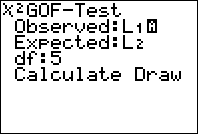
The total is the chi-square value, .

The p-value is found using , where the degrees of freedom is .

**Using the TI-84:**

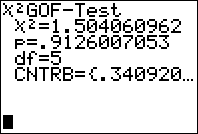
To run the test on the TI-84, type the observed frequencies into L1 and the expected frequencies into L2, then go into STAT, move over to TEST and choose GOF-Test from the list. The setup for the test is in figure #11.2.4.

**Figure #11.2.4: Setup for Chi-Square Goodness of Fit Test on TI-84**

****

Once you press ENTER on Calculate you will see the results in figure #11.2.5.

**Figure #11.2.5: Results for Chi-Square Test on TI-83/84**

****

The test statistic is  1.504060962

The 

The CNTRB represent the  for each die value. You can see the values by pressing the right arrow.

Using R:

Type in the observed frequencies. Call it something like observed.

observed<- c(type in data with commas in between)

Type in the probabilities that you are comparing to the observed frequencies. Call it something like null.probs.

null.probs <- c(type in probabilities with commas in between)

chisq.test(observed, p=null.probs) – the command for the hypothesis test

For this example (Note since you are looking to see if the die is fair, then the probability of each side of a fair die coming up is 1/6.)

observed<-c(78, 87, 87, 76, 85, 87)

null.probs<-c(1/6, 1/6, 1/6, 1/6, 1/6, 1/6)

chisq.test(observed, p=null.probs)

Output:

Chi-squared test for given probabilities

data: observed

X-squared = 1.504, df = 5, p-value = 0.9126

The test statistic is  and the p-value = 0.9126.

1. Conclusion

Fail to reject  since the p-value is greater than 0.05.

1. Interpretation

There is not enough evidence to show that the die is not consistent with the distribution for a fair die. There is not enough evidence to show that the die is not fair.

**Section 11.2: Homework**

In each problem show all steps of the hypothesis test. If some of the assumptions are not met, note that the results of the test may not be correct and then continue the process of the hypothesis test.

1. According to the M&M candy company, the expected proportion can be found in Table #11.2.3. In addition, the table contains the number of M&M’s of each color that were found in a case of candy (Madison, 2013). At the 5% level, do the observed frequencies support the claim of M&M?

**Table #11.2.3: M&M Observed and Proportions**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Blue | Brown | Green | Orange | Red | Yellow | Total |
| Observed Frequencies | 481 | 371 | 483 | 544 | 372 | 369 | 2620 |
| Expected Proportion | 0.24 | 0.13 | 0.16 | 0.20 | 0.13 | 0.14 |  |

1. Eyeglassomatic manufactures eyeglasses for different retailers. They test to see how many defective lenses they made the time period of January 1 to March 31. Table #11.2.4 gives the defect and the number of defects.

**Table #11.2.4: Number of Defective Lenses**

|  |  |
| --- | --- |
| Defect type | Number of defects |
| Scratch | 5865 |
| Right shaped – small | 4613 |
| Flaked | 1992 |
| Wrong axis | 1838 |
| Chamfer wrong | 1596 |
| Crazing, cracks | 1546 |
| Wrong shape | 1485 |
| Wrong PD | 1398 |
| Spots and bubbles | 1371 |
| Wrong height | 1130 |
| Right shape – big | 1105 |
| Lost in lab | 976 |
| Spots/bubble – intern | 976 |

Do the data support the notion that each defect type occurs in the same proportion? Test at the 10% level.

1. On occasion, medical studies need to model the proportion of the population that has a disease and compare that to observed frequencies of the disease actually occurring. Suppose the end-stage renal failure in south-west Wales was collected for different age groups. Do the data in table 11.2.5 show that the observed frequencies are in agreement with proportion of people in each age group (Boyle, Flowerdew & Williams, 1997)? Test at the 1% level.

**Table #11.2.5: Renal Failure Frequencies**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Age Group | 16-29 | 30-44 | 45-59 | 60-75 | 75+ | Total |
| Observed Frequency | 32 | 66 | 132 | 218 | 91 | 539 |
| Expected Proportion | 0.23 | 0.25 | 0.22 | 0.21 | 0.09 |  |

1. In Africa in 2011, the number of deaths of a female from cardiovascular disease for different age groups are in table #11.2.6 ("Global health observatory," 2013). In addition, the proportion of deaths of females from all causes for the same age groups are also in table #11.2.6. Do the data show that the death from cardiovascular disease are in the same proportion as all deaths for the different age groups? Test at the 5% level.

**Table #11.2.6: Deaths of Females for Different Age Groups**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Age | 5-14 | 15-29 | 30-49 | 50-69 | Total |
| Cardiovascular Frequency | 8 | 16 | 56 | 433 | 513 |
| All Cause Proportion | 0.10 | 0.12 | 0.26 | 0.52 |  |

1. In Australia in 1995, there was a question of whether indigenous people are more likely to die in prison than non-indigenous people. To figure out, the data in table 11.2.7 was collected. ("Aboriginal deaths in," 2013). Do the data show that indigenous people die in the same proportion as non-indigenous people? Test at the 1% level.

**Table #11.2.7: Death of Prisoners**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Prisoner Dies | Prisoner Did Not Die | Total |
| Indigenous Prisoner Frequency | 17 | 2890 | 2907 |
| Frequency of Non-Indigenous Prisoner | 42 | 14459 | 14501 |

1. A project conducted by the Australian Federal Office of Road Safety asked people many questions about their cars. One question was the reason that a person chooses a given car, and that data is in table #11.2.8 ("Car preferences," 2013).

**Table #11.2.8: Reason for Choosing a Car**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Safety | Reliability | Cost | Performance | Comfort | Looks |
| 84 | 62 | 46 | 34 | 47 | 27 |

Do the data show that the frequencies observed substantiate the claim that the reasons for choosing a car are equally likely? Test at the 5% level.

**Section 11.3: Analysis of Variance (ANOVA)**

There are times where you want to compare three or more population means. One idea is to just test different combinations of two means. The problem with that is that your chance for a type I error increases. Instead you need a process for analyzing all of them at the same time. This process is known as **analysis of variance (ANOVA)**. The test statistic for the ANOVA is fairly complicated, you will want to use technology to find the test statistic and p-value. The test statistic is distributed as an F-distribution, which is skewed right and depends on degrees of freedom. Since you will use technology to find these, the distribution and the test statistic will not be presented. Remember, all hypothesis tests are the same process. Note that to obtain a statistically significant result there need only be a difference between any two of the *k* means.

Before conducting the hypothesis test, it is helpful to look at the means and standard deviations for each data set. If the sample means with consideration of the sample standard deviations are different, it may mean that some of the population means are different. However, do realize that if they are different, it doesn’t provide enough evidence to show the population means are different. Calculating the sample statistics just gives you an idea that conducting the hypothesis test is a good idea.

**Hypothesis test using ANOVA to compare *k* means**

1. State the random variables and the parameters in words



1. State the null and alternative hypotheses and the level of significance





Also, state your  level here.

1. State and check the assumptions for the hypothesis test
2. A random sample of size  is taken from each population.
3. All the samples are independent of each other.
4. Each population is normally distributed. The ANOVA test is fairly robust to the assumption especially if the sample sizes are fairly close to each other. Unless the populations are really not normally distributed and the sample sizes are close to each other, then this is a loose assumption.
5. The population variances are all equal. If the sample sizes are close to each other, then this is a loose assumption.
6. Find the test statistic and p-value

The test statistic is , where  is the mean square between the groups (or factors), and  is the mean square within the groups. The degrees of freedom between the groups is  and the degrees of freedom within the groups is . To find all of the values, use technology such as the TI-83/84 calculator or R.

The test statistic, *F*, is distributed as an F-distribution, where both degrees of freedom are needed in this distribution. The p-value is also calculated by the calculator or R.

1. Conclusion

This is where you write reject  or fail to reject . The rule is: if the p-value < , then reject . If the p-value , then fail to reject 

1. Interpretation

This is where you interpret in real world terms the conclusion to the test. The conclusion for a hypothesis test is that you either have enough evidence to show  is true, or you do not have enough evidence to show  is true.

If you do in fact reject , then you know that at least two of the means are different. The next question you might ask is which are different? You can look at the sample means, but realize that these only give a preliminary result. To actually determine which means are different, you need to conduct other tests. Some of these tests are the range test, multiple comparison tests, Duncan test, Student-Newman-Keuls test, Tukey test, Scheffé test, Dunnett test, least significant different test, and the Bonferroni test. There is no consensus on which test to use. These tests are available in statistical computer packages such as Minitab and SPSS.

**Example #11.3.1: Hypothesis Test Involving Several Means**

Cancer is a terrible disease. Surviving may depend on the type of cancer the person has. To see if the mean survival time for several types of cancer are different, data was collected on the survival time in days of patients with one of these cancer in advanced stage. The data is in table #11.3.1 ("Cancer survival story," 2013). (Please realize that this data is from 1978. There have been many advances in cancer treatment, so do not use this data as an indication of survival rates from these cancers.) Do the data indicate that at least two of the mean survival time for these types of cancer are not all equal? Test at the 1% level.

**Table #11.3.1: Survival Times in Days of Five Cancer Types**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Stomach | Bronchus | Colon | Ovary | Breast |
| 124 | 81 | 248 | 1234 | 1235 |
| 42 | 461 | 377 | 89 | 24 |
| 25 | 20 | 189 | 201 | 1581 |
| 45 | 450 | 1843 | 356 | 1166 |
| 412 | 246 | 180 | 2970 | 40 |
| 51 | 166 | 537 | 456 | 727 |
| 1112 | 63 | 519 |  | 3808 |
| 46 | 64 | 455 |  | 791 |
| 103 | 155 | 406 |  | 1804 |
| 876 | 859 | 365 |  | 3460 |
| 146 | 151 | 942 |  | 719 |
| 340 | 166 | 776 |  |  |
| 396 | 37 | 372 |  |  |
|  | 223 | 163 |  |  |
|  | 138 | 101 |  |  |
|  | 72 | 20 |  |  |
|  | 245 | 283 |  |  |

**Solution:**

1. State the random variables and the parameters in words





















Now before conducting the hypothesis test, look at the means and standard deviations.



There appears to be a difference between at least two of the means, but realize that the standard deviations are very different. The difference you see may not be significant.

Notice the sample sizes are not the same. The sample sizes are 

1. State the null and alternative hypotheses and the level of significance



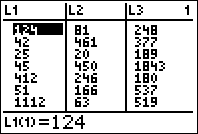


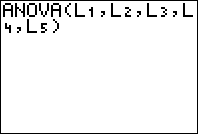


1. State and check the assumptions for the hypothesis test
2. A random sample of 13 survival times from stomach cancer was taken. A random sample of 17 survival times from bronchus cancer was taken. A random sample of 17 survival times from colon cancer was taken. A random sample of 6 survival times from ovarian cancer was taken. A random sample of 11 survival times from breast cancer was taken. These statements may not be true. This information was not shared as to whether the samples were random or not but it may be safe to assume that.
3. Since the individuals have different cancers, then the samples are independent.
4. Population of all survival times from stomach cancer is normally distributed. Population of all survival times from bronchus cancer is normally distributed. Population of all survival times from colon cancer is normally distributed. Population of all survival times from ovarian cancer is normally distributed. Population of all survival times from breast cancer is normally distributed. Looking at the histograms, box plots and normal quantile plots for each sample, it appears that none of the populations are normally distributed. The sample sizes are somewhat different for the problem. This assumption may not be true.
5. The population variances are all equal. The sample standard deviations are approximately 346.3, 209.9, 427.2, 1098.6, and 1239.0 respectively. This assumption does not appear to be met, since the sample standard deviations are very different. The sample sizes are somewhat different for the problem. This assumption may not be true.
6. Find the test statistic and p-value

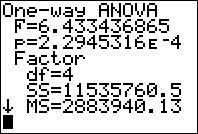
To find the test statistic and p-value using the TI-83/84, type each data set into L1 through L5. Then go into STAT and over to TESTS and choose ANOVA(. Then type in L1,L2,L3,L4,L5 and press enter. You will get the results of the ANOVA test.

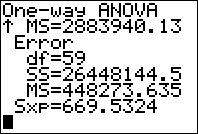
**Figure #11.3.1: Setup for ANOVA on TI-83/84**

****

****

**Figure #11.3.2: Results of ANOVA on TI-83/84**

****

****

The test statistic is  and .

Just so you know, the Factor information is between the groups and the Error is within the groups. So  and .

To find the test statistic and p-value on R:

The commands would be:

variable=c(type in all data values with commas in between) – this is the response variable

factor=c(rep("factor 1", number of data values for factor 1), rep("factor 2", number of data values for factor 2), etc) – this separates the data into the different factors that the measurements were based on.

data\_name = data.frame(variable, factor) – this puts the data into one variable. data\_name is the name you give this variable

aov(variable ~ factor, data = data name) – runs the ANOVA analysis

For this example, the commands would be:

time=c(124, 42, 25, 45, 412, 51, 1112, 46, 103, 876, 146, 340, 396, 81, 461, 20, 450, 246, 166, 63, 64, 155, 859, 151, 166, 37, 223, 138, 72, 245, 248, 377, 189, 1843, 180, 537, 519, 455, 406, 365, 942, 776, 372, 163, 101, 20, 283, 1234, 89, 201, 356, 2970, 456, 1235, 24, 1581, 1166, 40, 727, 3808, 791, 1804, 3460, 719)

factor=c(rep("Stomach", 13), rep("Bronchus", 17), rep("Colon", 17), rep("Ovary", 6), rep("Breast", 11))

survival=data.frame(time, factor)

results=aov(time~factor, data=survival)

summary(results)

Df Sum Sq Mean Sq F value Pr(>F)

factor 4 11535761 2883940 6.433 0.000229 \*\*\*

Residuals 59 26448144 448274

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

The test statistic is F = 6.433 and the p-value = 0.000229.

1. Conclusion

Reject  since the p-value is less than 0.01.

1. Interpretation

There is evidence to show that at least two of the mean survival times from different cancers are not equal.

By examination of the means, it appears that the mean survival time for breast cancer is different from the mean survival times for both stomach and bronchus cancers. It may also be different for the mean survival time for colon cancer. The others may not be different enough to actually say for sure.

**Section 11.3: Homework**

In each problem show all steps of the hypothesis test. If some of the assumptions are not met, note that the results of the test may not be correct and then continue the process of the hypothesis test.

1. Cuckoo birds are in the habit of laying their eggs in other birds’ nest. The other birds adopt and hatch the eggs. The lengths (in cm) of cuckoo birds’ eggs in the other species nests were measured and are in table #11.3.2 ("Cuckoo eggs in," 2013). Do the data show that the mean length of cuckoo bird’s eggs is not all the same when put into different nests? Test at the 5% level.

**Table #11.3.2: Lengths of Cuckoo Bird Eggs in Different Species Nests**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Meadow Pipit | | Tree Pipit | Hedge Sparrow | Robin | Pied Wagtail | Wren |
| 19.65 | 22.25 | 21.05 | 20.85 | 21.05 | 21.05 | 19.85 |
| 20.05 | 22.45 | 21.85 | 21.65 | 21.85 | 21.85 | 20.05 |
| 20.65 | 22.45 | 22.05 | 22.05 | 22.05 | 21.85 | 20.25 |
| 20.85 | 22.45 | 22.45 | 22.85 | 22.05 | 21.85 | 20.85 |
| 21.65 | 22.65 | 22.65 | 23.05 | 22.05 | 22.05 | 20.85 |
| 21.65 | 22.65 | 23.25 | 23.05 | 22.25 | 22.45 | 20.85 |
| 21.65 | 22.85 | 23.25 | 23.05 | 22.45 | 22.65 | 21.05 |
| 21.85 | 22.85 | 23.25 | 23.05 | 22.45 | 23.05 | 21.05 |
| 21.85 | 22.85 | 23.45 | 23.45 | 22.65 | 23.05 | 21.05 |
| 21.85 | 22.85 | 23.45 | 23.85 | 23.05 | 23.25 | 21.25 |
| 22.05 | 23.05 | 23.65 | 23.85 | 23.05 | 23.45 | 21.45 |
| 22.05 | 23.25 | 23.85 | 23.85 | 23.05 | 24.05 | 22.05 |
| 22.05 | 23.25 | 24.05 | 24.05 | 23.05 | 24.05 | 22.05 |
| 22.05 | 23.45 | 24.05 | 25.05 | 23.05 | 24.05 | 22.05 |
| 22.05 | 23.65 | 24.05 |  | 23.25 | 24.85 | 22.25 |
| 22.05 | 23.85 |  |  | 23.85 |  |  |
| 22.05 | 24.25 |  |  |  |  |  |
| 22.05 | 24.45 |  |  |  |  |  |
| 22.05 | 22.25 |  |  |  |  |  |
| 22.05 | 22.25 |  |  |  |  |  |
| 22.25 | 22.25 |  |  |  |  |  |
| 22.25 | 22.25 |  |  |  |  |  |
| 22.25 |  |  |  |  |  |  |

1. Levi-Strauss Co manufactures clothing. The quality control department measures weekly values of different suppliers for the percentage difference of waste between the layout on the computer and the actual waste when the clothing is made (called run-up). The data is in table #11.3.3, and there are some negative values because sometimes the supplier is able to layout the pattern better than the computer ("Waste run up," 2013). Do the data show that there is a difference between some of the suppliers? Test at the 1% level.

**Table #11.3.3: Run-ups for Different Plants Making Levi Strauss Clothing**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Plant 1 | Plant 2 | Plant 3 | Plant 4 | Plant 5 |
| 1.2 | 16.4 | 12.1 | 11.5 | 24 |
| 10.1 | -6 | 9.7 | 10.2 | -3.7 |
| -2 | -11.6 | 7.4 | 3.8 | 8.2 |
| 1.5 | -1.3 | -2.1 | 8.3 | 9.2 |
| -3 | 4 | 10.1 | 6.6 | -9.3 |
| -0.7 | 17 | 4.7 | 10.2 | 8 |
| 3.2 | 3.8 | 4.6 | 8.8 | 15.8 |
| 2.7 | 4.3 | 3.9 | 2.7 | 22.3 |
| -3.2 | 10.4 | 3.6 | 5.1 | 3.1 |
| -1.7 | 4.2 | 9.6 | 11.2 | 16.8 |
| 2.4 | 8.5 | 9.8 | 5.9 | 11.3 |
| 0.3 | 6.3 | 6.5 | 13 | 12.3 |
| 3.5 | 9 | 5.7 | 6.8 | 16.9 |
| -0.8 | 7.1 | 5.1 | 14.5 |  |
| 19.4 | 4.3 | 3.4 | 5.2 |  |
| 2.8 | 19.7 | -0.8 | 7.3 |  |
| 13 | 3 | -3.9 | 7.1 |  |
| 42.7 | 7.6 | 0.9 | 3.4 |  |
| 1.4 | 70.2 | 1.5 | 0.7 |  |
| 3 | 8.5 |  |  |  |
| 2.4 | 6 |  |  |  |
| 1.3 | 2.9 |  |  |  |

1. Several magazines were grouped into three categories based on what level of education of their readers the magazines are geared towards: high, medium, or low level. Then random samples of the magazines were selected to determine the number of three-plus-syllable words were in the advertising copy, and the data is in table #11.3.4 ("Magazine ads readability," 2013). Is there enough evidence to show that the mean number of three-plus-syllable words in advertising copy is different for at least two of the education levels? Test at the 5% level.

**Table #11.3.4: Number of Three Plus Syllable Words in Advertising Copy**

|  |  |  |
| --- | --- | --- |
| High Education | Medium Education | Low Education |
| 34 | 13 | 7 |
| 21 | 22 | 7 |
| 37 | 25 | 7 |
| 31 | 3 | 7 |
| 10 | 5 | 7 |
| 24 | 2 | 7 |
| 39 | 9 | 8 |
| 10 | 3 | 8 |
| 17 | 0 | 8 |
| 18 | 4 | 8 |
| 32 | 29 | 8 |
| 17 | 26 | 8 |
| 3 | 5 | 9 |
| 10 | 5 | 9 |
| 6 | 24 | 9 |
| 5 | 15 | 9 |
| 6 | 3 | 9 |
| 6 | 8 | 9 |

1. A study was undertaken to see how accurate food labeling for calories on food that is considered reduced calorie. The group measured the amount of calories for each item of food and then found the percent difference between measured and labeled food, . The group also looked at food that was nationally advertised, regionally distributed, or locally prepared. The data is in table #11.3.5 ("Calories datafile," 2013). Do the data indicate that at least two of the mean percent differences between the three groups are different? Test at the 10% level.

**Table #11.3.5: Percent Differences Between Measured and Labeled Food**

|  |  |  |
| --- | --- | --- |
| National Advertised | Regionally Distributed | Locally Prepared |
| 2 | 41 | 15 |
| -28 | 46 | 60 |
| -6 | 2 | 250 |
| 8 | 25 | 145 |
| 6 | 39 | 6 |
| -1 | 16.5 | 80 |
| 10 | 17 | 95 |
| 13 | 28 | 3 |
| 15 | -3 |  |
| -4 | 14 |  |
| -4 | 34 |  |
| -18 | 42 |  |
| 10 |  |  |
| 5 |  |  |
| 3 |  |  |
| -7 |  |  |
| 3 |  |  |
| -0.5 |  |  |
| -10 |  |  |
| 6 |  |  |

1. The amount of sodium (in mg) in different types of hotdogs is in table #11.3.6 ("Hot dogs story," 2013). Is there sufficient evidence to show that the mean amount of sodium in the types of hotdogs are not all equal? Test at the 5% level.

**Table #11.3.6: Amount of Sodium (in mg) in Beef, Meat, and Poultry Hotdogs**

|  |  |  |
| --- | --- | --- |
| Beef | Meat | Poultry |
| 495 | 458 | 430 |
| 477 | 506 | 375 |
| 425 | 473 | 396 |
| 322 | 545 | 383 |
| 482 | 496 | 387 |
| 587 | 360 | 542 |
| 370 | 387 | 359 |
| 322 | 386 | 357 |
| 479 | 507 | 528 |
| 375 | 393 | 513 |
| 330 | 405 | 426 |
| 300 | 372 | 513 |
| 386 | 144 | 358 |
| 401 | 511 | 581 |
| 645 | 405 | 588 |
| 440 | 428 | 522 |
| 317 | 339 | 545 |
| 319 |  |  |
| 298 |  |  |
| 253 |  |  |

Data Source:

*Aboriginal deaths in custody*. (2013, September 26). Retrieved from <http://www.statsci.org/data/oz/custody.html>

*Activities of dolphin groups*. (2013, September 26). Retrieved from <http://www.statsci.org/data/general/dolpacti.html>

Boyle, P., Flowerdew, R., & Williams, A. (1997). Evaluating the goodness of fit in models of sparse medical data: A simulation approach. *International Journal of Epidemiology*, *26*(3), 651-656. Retrieved from http://ije.oxfordjournals.org/content/26/3/651.full.pdf html

*Calories datafile*. (2013, December 07). Retrieved from <http://lib.stat.cmu.edu/DASL/Datafiles/Calories.html>

*Cancer survival story*. (2013, December 04). Retrieved from <http://lib.stat.cmu.edu/DASL/Stories/CancerSurvival.html>

*Car preferences*. (2013, September 26). Retrieved from <http://www.statsci.org/data/oz/carprefs.html>

*Cuckoo eggs in nest of other birds*. (2013, December 04). Retrieved from <http://lib.stat.cmu.edu/DASL/Stories/cuckoo.html>

*Education by age datafile*. (2013, December 05). Retrieved from <http://lib.stat.cmu.edu/DASL/Datafiles/Educationbyage.html>

*Encyclopedia Titanica*. (2013, November 09). Retrieved from <http://www.encyclopedia-titanica.org/>

*Global health observatory data respository*. (2013, October 09). Retrieved from http://apps.who.int/gho/athena/data/download.xsl?format=xml&target=GHO/MORT\_400&profile=excel&filter=AGEGROUP:YEARS05-14;AGEGROUP:YEARS15-29;AGEGROUP:YEARS30-49;AGEGROUP:YEARS50-69;AGEGROUP:YEARS70 ;MGHEREG:REG6\_AFR;GHECAUSES:\*;SEX:\*

*Hot dogs story*. (2013, November 16). Retrieved from <http://lib.stat.cmu.edu/DASL/Stories/Hotdogs.html>

*Leprosy: Number of reported cases by country*. (2013, September 04). Retrieved from <http://apps.who.int/gho/data/node.main.A1639>

Madison, J. (2013, October 15). *M&M's color distribution analysis*. Retrieved from <http://joshmadison.com/2007/12/02/mms-color-distribution-analysis/>

*Magazine ads readability*. (2013, December 04). Retrieved from <http://lib.stat.cmu.edu/DASL/Datafiles/magadsdat.html>

*Popular kids datafile*. (2013, December 05). Retrieved from <http://lib.stat.cmu.edu/DASL/Datafiles/PopularKids.html>

Schultz, S. T., Klonoff-Cohen, H. S., Wingard, D. L., Askhoomoff, N. A., Macera, C. A., Ji, M., & Bacher, C. (2006). Breastfeeding, infant formula supplementation, and autistic disorder: the results of a parent survey. *International Breastfeeding Journal*, *1*(16), doi: 10.1186/1746-4358-1-16

*Waste run up*. (2013, December 04). Retrieved from <http://lib.stat.cmu.edu/DASL/Stories/wasterunup.html>