

City University of New York (CUNY)

CUNY Academic Works

Open Educational Resources

Queensborough Community College

2020

Clear-Sighted Statistics: Module 17: Chi-Square Tests

Edward Volchok

CUNY Queensborough Community College

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/qb_oers/131

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).

Contact: AcademicWorks@cuny.edu

Clear-Sighted Statistics: An OER Textbook

Module 17: Chi-Square Tests

“This (Chi-Square), I believe is the great contribution to statistical methodology which the unsurpassed energy of Professor [Karl] Pearson’s work will be remembered....”¹

R. A. Fisher
Scientific Methods for Research Workers, 13th Edition
1958

I. Introduction

While R. A. Fisher and Karl Pearson had an [acrimonious](#) relationship, Fisher was right about the importance of Chi-Square (χ^2) techniques. Chi-square tests first became popular due to an article written by Karl Pearson in 1900.² In his honor, Chi-Square techniques are sometimes called “Pearson’s Chi-Squared tests.” One hundred and twenty years after Pearson first started using these techniques, they are still widely used. Fisher’s muted praise for Pearson here and elsewhere is evidence of Fisher’s lingering [animus](#).³ History, however, has been kinder to Pearson than Fisher’s lukewarm praise, made 22 years after Pearson’s death, would suggest. In this module we will introduce Chi-Square tests. In Module 18, we shall examine another of Pearson’s major contributions: The coefficient of correlation or Pearson’s r , which measures the strength of a linear correlation.

Chi-Square is the most commonly used nonparametric technique, a technique that makes no assumptions about how the population is distributed. So far we have dealt with parametric statistics: Techniques that include population parameters like the mean (μ), variance (σ^2), standard deviation (σ), and proportion (π). With parametric distributions, we assume that the data fit a normal or symmetrical distribution like the one shown in Figure 1. The unknown population parameters can be estimated using sample statistics.

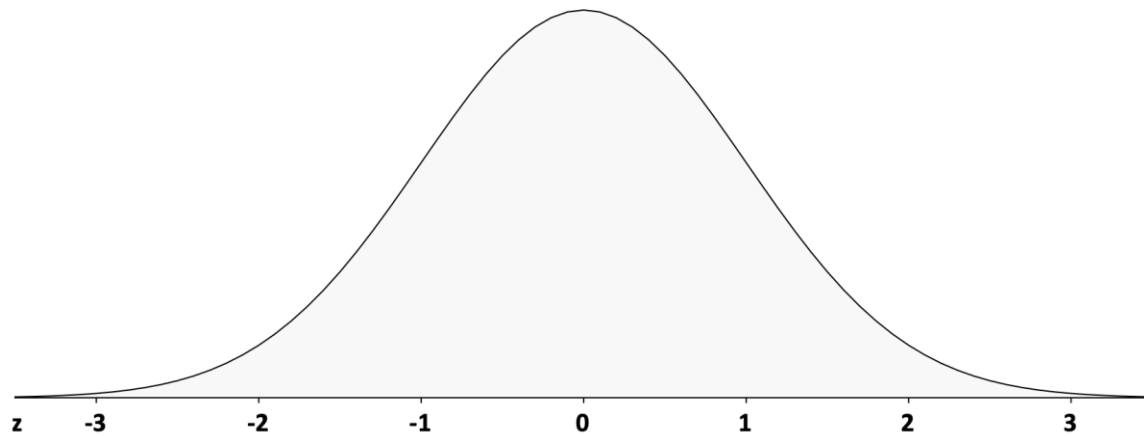


Figure 1: The Normal Distribution

Nonparametric statistics refers to statistical techniques that do not require that the data fit a normal distribution. We can use qualitative data—nominal and ordinal data—and frequency counts or proportions, which are quantitative.⁴ Chi-Square tests focus on the relationships among the nominal or ordinal classifications of categories or treatments, not interval or ratio measures like mean, proportion, variance, or standard deviation.

Chi-Square stands for the Greek letter χ , and is pronounced “kai.” It has nothing to do with the Chinese principle of *ch’i*, 氣, which is pronounced “chee,” and stands for the vital energy of any living being. It is also not related to the Hebrew word *chai*, חי, which is pronounced as “hi.” This word has symbolic and numerical meaning. It means “life,” but more deeply connotes a life dedicated to kindness, thoughtfulness, and selflessness. It also signifies the number 18.

After completing this module, you will be able to:

- List the characteristics of the Chi-Square distribution.
- Understand the restrictions of Chi-Square tests.
- Conduct a Chi-Square goodness of fit test when the expected frequencies are equal.

- Conduct a Chi-Square goodness of fit test when the expected frequencies are *not* equal.
- Conduct a Chi-Square test for data organized into contingency tables (Also known as a χ^2 test of independence).
- Conduct a Chi-Square goodness of fit test to confirm normality.
- Calculate Effect Size using ϕ , which is the simplest technique for determining Chi-Square effect size.
- Determine the power of a Chi-Square test using G*Power.
- Use charts of the chi-square distribution to draw conclusions about the results of your null hypothesis significance tests.

You can download the following files for this module:

- The Chi-Square critical values table: ChiSq_Distribution.xlsx or ChiSq_Distribution.pdf.
- The p-values calculator: p-value_Calculator.xlsx.
- Data and answers for the module examples: 17_ChiSq_Examples.xlsx.
- Data for the end-of-module exercises: 17_ChiSq_Exercises.xlsx.

II. Characteristics of a Chi-Square Distribution

The Chi-Square distribution has five major characteristics:

1. **Chi-Square is always a non-negative number.** This is because Chi-Square values are computed using the squared differences between observed frequencies, O , and expected frequencies, E . Whenever you square a negative or positive number, the result is always a positive number. The smallest possible value for Chi-Square, therefore, is zero, which indicates that the observed and expected frequencies are identical. As the value of Chi-Square increases there are more differences between the observed and expected frequencies.

2. **Chi-Square distributions are defined by the number of degrees of freedom.** How degrees of freedom are calculated depends on the type of Chi-Square test. As illustrated by Figure 2, which shows the Chi-Square distribution curves with 1, 4, 8, and 12 degrees of freedom. The distribution moves closer to a symmetrical or normal distribution as the number of degrees of freedom increases.

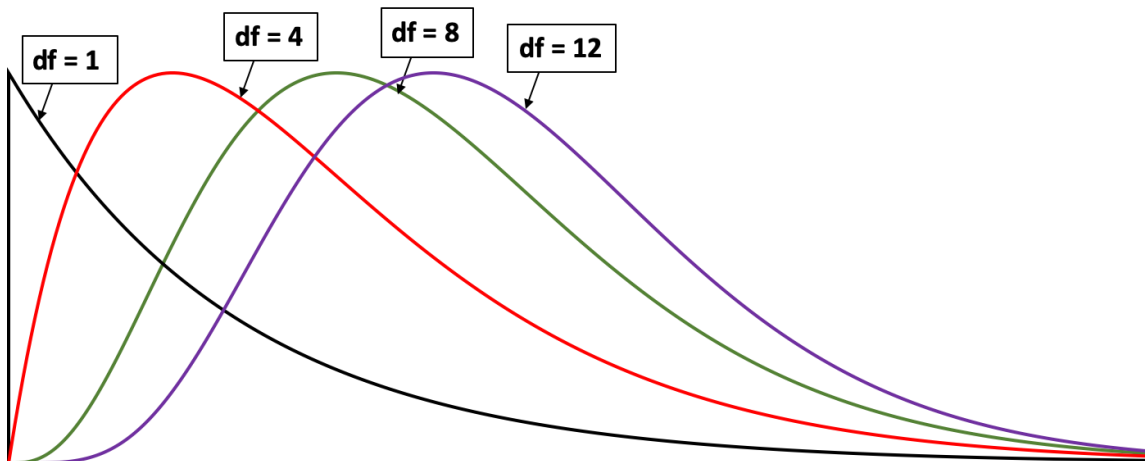


Figure 2: Chi-Square Distributions with 1, 4, 8, and 12 Degrees of Freedom

3. Because Chi-Square must be a non-negative number (zero or a positive number), **Chi-Square distributions are always positively skewed (right skewed).** Chi-Square tests, therefore, will always be right-tailed tests.
4. Like ANOVA, Chi-Square tests are **omnibus** or **global tests**. When there are more than two treatments, the test cannot determine which treatment pairs differ.
5. You should also be aware of three requirements for a Chi-Square tests:
- The categories or treatments are independent and mutually exclusive.
 - The data for the observed frequencies are the result of counts of the qualitative data, which will become evident in the examples presented below.
 - Chi-square is sensitive to low sample size and low category frequencies. The expected frequencies for each category must be at least five. If this requirement

cannot be met, we may be able to combine categories. To do this, of course, there would have to be more than two categories.

III. Types of Chi-Square (χ^2) Tests

In this module we will cover three broad categories of Chi-Square tests.⁵ These are basic Chi-Square tests.

1) Chi-Square Goodness of Fit Test

This test is sometimes called a one-way Chi-Square test because we use it to test one population and one qualitative (nominal or ordinal) variable. This test calculates how closely the data fit the expected distribution of the data. The Null Hypothesis states that the observed distribution “fits,” or matches the expected distribution. The alternate hypothesis declares that the observed distribution does not fit with the expected distribution; or, there is a statistically significant difference between the observed and expected frequencies.

Degrees of freedom for Chi-Square goodness of fit tests are defined as the number of categories, k , minus one, $k - 1$.

2) Chi-Square Contingency Table Tests

These tests are sometimes called a Chi-Square test of independence, or a two-way Chi-Square test, because we test multiple qualitative variables. The Null Hypothesis for this test declares that the observed distribution matches the expected distribution or the categories are independent. The alternate hypothesis states that the observed distribution does not match the expected distribution or the categories are dependent. You will see examples of Chi-Square contingency tables shortly. Degrees of freedom for these tests are found by multiplying the number of rows in the contingency table minus 1 by the number of columns minus 1:

$$(r - 1)(c - 1) = \text{degrees of freedom}$$

Equation 1: Degrees of Freedom for A Chi-Square Contingency Table Test

3) Chi-Square Test for Normality

We will also cover an extension of the Chi-Square goodness of fit test, the **Chi-Square test for normality**. This test allows us to determine whether the data follow a normal distribution. The Null Hypothesis states that the data are normally distributed while the alternate hypothesis states that the data are not normally distributed. This test is considered a quick way of testing whether the data are normally distributed. The expected frequencies are calculated based on the mean and standard deviation of sampled data and how the data fit with a normal distribution. This test, however, is not considered as precise as other tests for normality like the Shapiro-Wilk or Kolmogorov-Smirnov tests, which are not typically covered in an introductory statistics class. Dedicated statistical software like SPSS can perform these tests in seconds with a few clicks of the mouse.

With Chi-Square tests for normality, Chi-Square degrees of freedom depend on whether we used population parameters (μ and σ) or sample statistics (\bar{X} and s). If we are using population parameters, degrees of freedom are defined as $k - 1$. But when we use sample statistics, degrees of freedom are defined as $k - 1 - 2$. In essence, we lose two degrees of freedom when we estimate μ and σ using \bar{X} and s .

All Chi-Square tests use the following formula:

$$\chi^2 = \Sigma \left[\frac{(O - E)^2}{E} \right]$$

Where: O are the observed frequencies and
E are the expected frequencies

Equation 2: Formula for the Chi-Square Test Statistic

With all Chi-Square techniques, the observed frequencies are determined by counting the variables in the data we collect. The techniques used to calculate the expected frequencies depend on the type of Chi-Square test, which we will review shortly.

IV. The Chi-Square (χ^2) Goodness of fit Test

The Chi-Square goodness of fit test compares the distribution of observed frequencies to the distribution of expected frequencies to determine how closely our observations match or “fit” our expectations.

The Null Hypothesis for this test states that there is no statistically significant difference between the observed frequencies and the expected frequencies. The alternate hypothesis states that there is a statistically significant difference between the observed and expected frequencies. Both the null and alternate hypothesis are written in a short sentence without mathematical notation.

The critical value of Chi-Square is found using a critical values table or statistical software. Figure 3 shows a section of the critical values table for Chi-Square. The left-most column shows the degrees of freedom. There are five columns for the level of significance, α : 0.10, 0.05, 0.02, 0.01, and 0.001. The critical value for Chi-Square is found at the intersection of the row for degrees of freedom and the significance level column. For example, if we were using a 0.05 significance level with 5 degrees of freedom, the critical value would be 11.070. For a test using a 0.01 significance level, the critical value would be 15.086. **Please note:** The critical value is always expressed with three digits past the decimal point. This is the thousandths column.

df	Right-Tail Area				
	Level of Significance, α				
	0.10	0.05	0.02	0.01	0.001
1	2.706	3.841	5.412	6.635	10.828
2	4.605	5.991	7.824	9.210	13.816
3	6.251	7.815	9.837	11.345	16.266
4	7.779	9.488	11.668	13.277	18.467
5	9.236	11.070	13.388	15.086	20.515
6	10.645	12.592	15.033	16.812	22.458
7	12.017	14.067	16.622	18.475	24.322
8	13.362	15.507	18.168	20.090	26.124
9	14.684	16.919	19.679	21.666	27.877
10	15.987	18.307	21.161	23.209	29.588

Figure 3: Abbreviated Chi-Square Critical Values Table

Microsoft Excel can also find the critical value of Chi-Square using the following function:

$$=CHISQ.INV.RT(\alpha,df)$$

Equation 3: Excel's Function for Finding the Critical Value for Chi-Square

The decision rule for the example cited above, α equals 5 percent with 5 degrees of freedom is: Reject the Null Hypothesis if Chi-Square is greater than 11.070. You can estimate the p-value using the Chi-Square critical value table in the same manner as the Student-t table can be used to estimate p-values. For example, if the value of Chi-Square were 10.052, that p-value would be greater than 5 percent and less than 10 percent, and we would fail to reject the Null Hypothesis. Excel can calculate a precise p-value using this function:

$$=CHISQ.DIST.RT(\text{Chi-Square},df)$$

Equation 4: Excel's Function for Finding the p-Value of Chi-Square

The exact p-value for a Chi-Square of 10.052 with 5 degrees of freedom is 0.0738 or 7.38 percent.

A. Example 1: Chi-Square Goodness of Fit Test with Equal Expected Frequencies

Your friend loves to bake cookies and she is very good at it. So good, in fact, people tell her that they want to buy her cookies. She has decided to start selling her cookies at street fairs

and farmer's markets. Her cookies are large, about six inches in diameter. She intends to sell cookies made from her six favorite recipes: Almond Sugar, Brown Sugar Shortbread, Chocolate Chip, Gingerbread, Oatmeal, and Peanut Butter.

After her first street fair, she asks for your help. She wants you to determine whether her sales were equally distributed among the six recipes. You agree and tell her she needs this information to manage how many of each type of cookie to bake. Table 1 shows how many cookies she sold broken down by recipe.

Table 1: Cookies Sold – Observed Frequencies

Cookie	Sales
Almond Sugar	24
Brown Sugar Shortbread	22
Chocolate Chip	44
Gingerbread	25
Oatmeal Raisin	25
Peanut Butter	40
Total	180

1: Test Set-Up

You conduct a Chi-Square goodness of fit test to compare actual or observed sales with expected sales. This test will tell you whether the actual sales match the expected sales, or if there is a statistically significant difference between the observed sales and the expected sales.

The first question: What are the expected sales per cookie recipe? After all, your friend has never before sold her cookies. You have no experience to guide you in estimating the sales for each recipe. The solution is to assume that sales for each type of cookie will be equal. You can take your friend's total cookie sales and divide by the number of types of cookies to get the expected frequencies: $E = 180 \text{ cookies sold} / 6 \text{ cookie recipes} = 30 \text{ cookies per recipe}$. You expect to sell 30 of each kind of cookie.

2: Select the Significance Level, α

You decide to conduct this test at a 5 percent significance level.

3: State the Null and Alternate Hypotheses

- H_0 : The six cookie recipes have equal sales
- H_1 : The six cookie recipes have significantly different sales

4: Compose the Decision Rule

To write the decision rule, we must find the critical value for Chi-Square at a 5 percent significance level with $k - 1$ or 5 degrees of freedom. Figure 4 shows the Chi-Square critical value table with 11.070 as the critical value.

df	Right-Tail Area				
	Level of Significance, α				
	0.10	0.05	0.02	0.01	0.001
1	2.706	3.841	5.412	6.635	10.828
2	4.605	5.991	7.824	9.210	13.816
3	6.251	7.815	9.837	11.345	16.266
4	7.779	9.488	11.668	13.277	18.467
5	9.236	11.070	13.388	15.086	20.515
6	10.645	12.592	15.033	16.812	22.458
7	12.017	14.067	16.622	18.475	24.322
8	13.362	15.507	18.168	20.090	26.124
9	14.684	16.919	19.679	21.666	27.877
10	15.987	18.307	21.161	23.209	29.588

Figure 4: Critical Value with $df = 5$ and $\alpha = 0.05$ is 11.070

The critical value for Chi-Square can also be calculated using Excel:

$$=CHISQ.INV.RT(0.05,5)$$

Equation 5: The Excel function for critical values

The decision rule: Reject the Null Hypothesis if Chi-Square is greater than 11.070.

Figure 5 shows the Chi-Square distribution for the 5 percent rejection region starting at 11.070. The area in black is the rejection region.

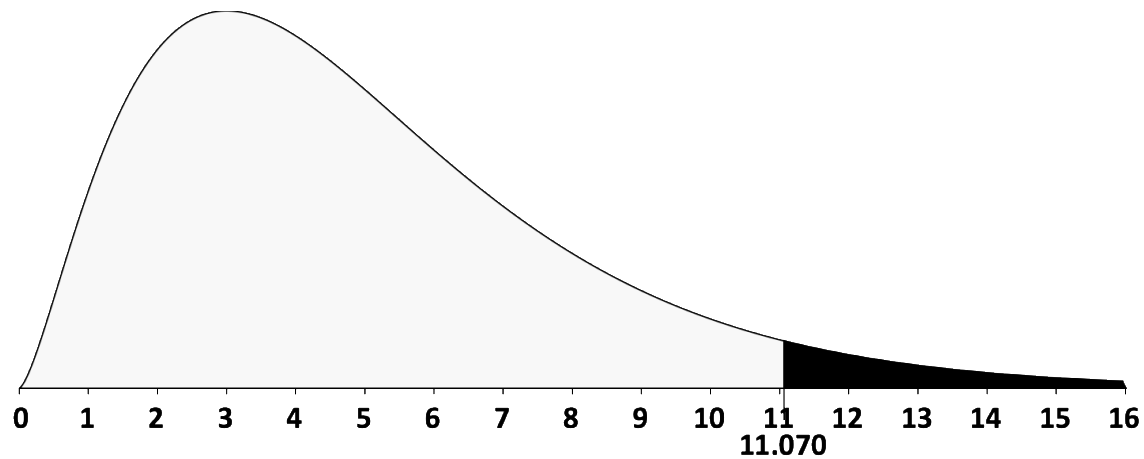


Figure 5: Chi-Square Distribution with 5 df and a 5% α , $\chi^2_{5,0.05}$

5: Calculate the Test Statistic and p-Value

Excel does not have a built-in Chi-Square goodness of fit function. The closest function Excel has is the CHISQ.TEST, =CHISQ.TEST(observed_range,expected_range). This function returns the p-value for a contingency table test. Fortunately, we can set up a worksheet to calculate the Chi-Square value, the p-value for a goodness of fit test, and effect size measured with Phi. We can also use these measures to estimate statistical power and the probability of a Type II error using G*Power. Figure 6 shows the Chi-Square goodness of fit calculations performed in Excel.

	A	B	C	D	E	F
1		Cookie	O	E	(O - E)²/E	
2	1	Almond Sugar	24	30	1.200	=(C2-D2)^2/D2
3	2	Brown Sugar Shortbread	22	30	2.133	=(C3-D3)^2/D3
4	3	Chocolate Chip	44	30	6.533	=(C4-D4)^2/D4
5	4	Gingerbread	25	30	0.833	=(C5-D5)^2/D5
6	5	Oatmeal Raisin	25	30	0.833	=(C6-D6)^2/D6
7	6	Peanut Butter	40	30	3.333	=(C7-D7)^2/D7
8		Total	180	180	14.867	=SUM(E2:E7)
9		Significance Level, α	0.05 Given by user			
10		k	6 =COUNTA(B2:B7)			
11		df (k - 1)	5 =COUNTA(B2:B7)-1			
12		Critical Value	11.070 =CHISQ.INV.RT(C15,5)			
13		Chi-Square	14.867 Σ(O - E)²/E			
14		p-value	0.0109 =CHISQ.DIST.RT(E8,C11)			
15		Effect Size, φ	0.2874 =SQRT(C13/C8)			
16		Interpretation of ES	Moderate See A17			
17	=IF(C15<0.1,I17,IF(C15<0.2,I18,IF(C15<0.4,I19,IF(C15<0.6,I20,IF(C15<0.8,I21,I22))))))					

Figure 6: Chi-Square Calculation Performed in Excel

The Chi-Square value equals 14.867, found by $\Sigma[(O - E)^2 / E]$. The p-value equals 0.0109, or 1.09 percent. Figure 7 shows a chart with the rejection region and the p-value.

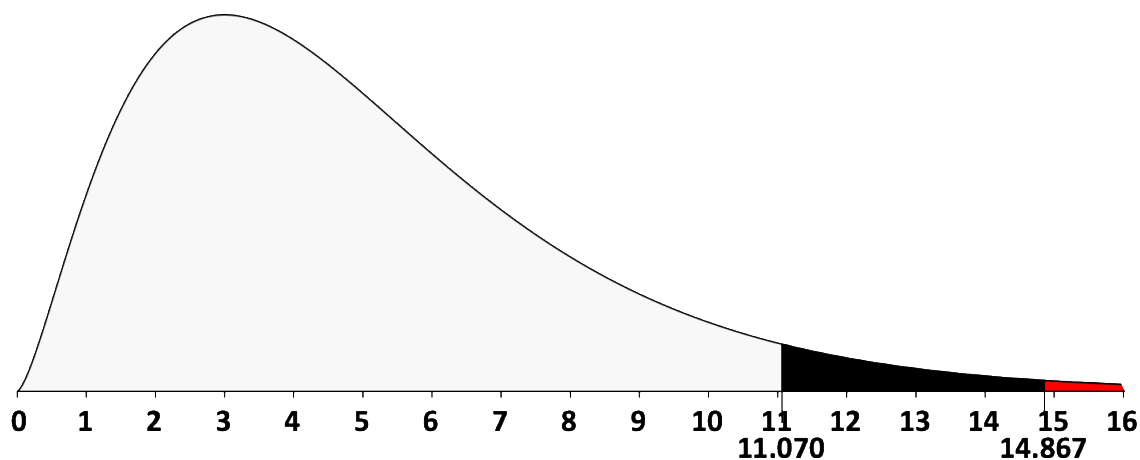


Figure 7: Chi-Square Distribution with a p-Value of 1.09%

We will measure Effect Size using Phi, ϕ :

$$\varphi = \sqrt{\frac{\chi^2}{n}} = \sqrt{\frac{14.867}{180}} = 0.2874$$

Equation 6: Effect Size

The effect size of 0.2874 indicates that the type of cookie has a moderate effect on cookie sales. Table 2 shows the guidelines for interpreting effect size for Chi-Square.⁶

Table 2: Effect Size Thresholds

w	Effect Size
0.00 < 0.10	Negligible
0.10 < 0.20	Weak
0.20 < 0.40	Moderate
0.40 < 0.60	Relatively Strong
0.60 < 0.80	Strong
0.80 < 1.00	Very Strong

Given the fact that we will reject the null hypothesis, there is no need for a power analysis. But, we should illustrate this analysis by conducting a *post hoc* analysis using G*Power, a free program that runs on Windows and Macintosh computers. [Here is the link for G*Power.](#) To run this analysis in G*Power, we select χ^2 tests from the family of tests, Generic χ^2 test from the Statistical test box, and Post Hoc: Computer power... from the type of power analysis. There are three inputs:

- 1) Noncentrality parameter lambda, λ , which is the same as Chi-Square, 14.867
- 2) α , err probability, or the level of significance set at 0.05
- 3) Degrees of freedom defined as the number of categories minus one. Degrees of freedom is 5, found by 6 minus

G*Power provides three outputs:

- 1) The critical value for Chi-Square, 11.070
- 2) The probability of a Type II error, 0.137

3) The statistical power, 0.863. With power of 86.3 percent, this test has sufficient power

See Figure 8:

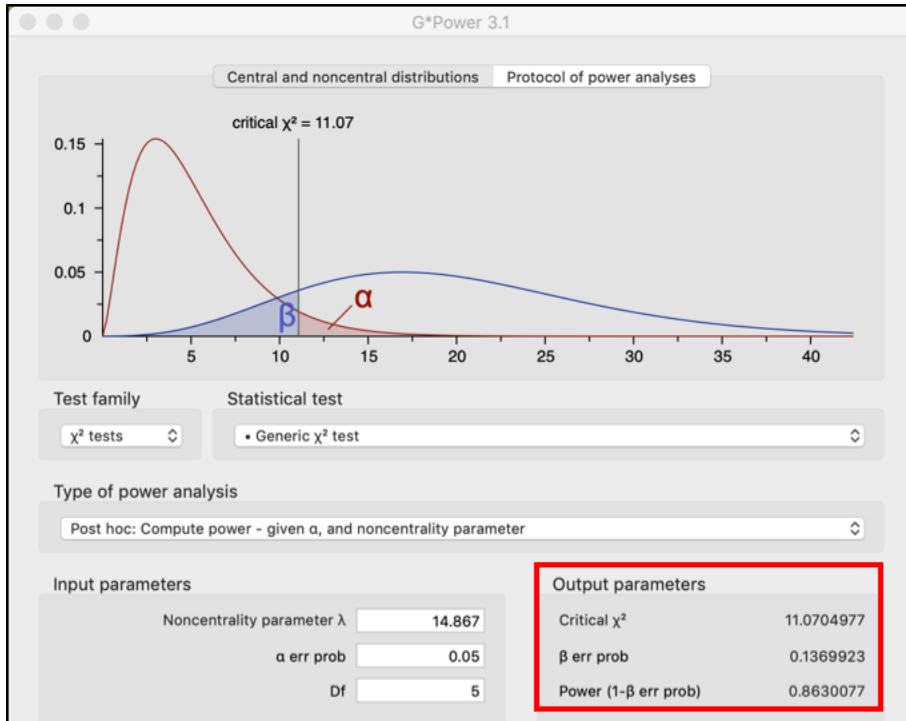


Figure 8: Statistical Power and the $P(\text{Type II})$ Found Using G*Power

6: Decide and Report

Given the fact that the Chi-Square value of 14.867 is above the critical value of 11.070 and in the rejection region. The decision to reject the Null Hypothesis can also be made using the p-value of 1.09 percent, which is below the 5 percent significance level. There is a statistically significant difference between the observed sales and the expected sales.

Conclusion: The different cookie recipes do not sell equally. **Please note:** At a 1 percent significance level, we would fail to reject the Null Hypothesis because the p-value is greater than the significance level.

Like ANOVA tests, Chi-Square goodness of fit tests are *omnibus* tests. We do not know which of the 15 different two-cookie pairs of cookies have different sales. To answer

this question we would have to conduct a *post hoc* analysis. There are several *post hoc* analyses for Chi-Square, but we will skip them because they take us beyond the scope of an introductory statistics class.

Conclusion: The six types of cookies do not sell equally. This result also has practical significance because this information can be used to manage production and inventory.

B. Example 2: Chi-Square Goodness of Fit Test with Unequal Expected Frequencies

One week after you completed this analysis, your friend again calls you for help. Over the weekend she sold her cookies at another street fair. She asks if you can repeat your analysis. You agree, but you tell her that this time it will be slightly different because we do not expect each type of cookie to sell equally. She asks, “How do we determine the expected sales for each cookie?” Here is the answer: “We take the relative frequency of the sales per cookie for week 1, and use those proportions to estimate the expected sales for each of the six cookie types. Table 3 shows the relative frequencies for each cookie:

Table 3: Week 1 Sales and Relative Frequencies

Cookie	Week 1 Sales	Relative Frequency	Found By
Almond Sugar	24	0.13	$24/180 = 0.13$
Brown Sugar Shortbread	22	0.12	$22/180 = 0.12$
Chocolate Chip	44	0.24	$44/180 = 0.24$
Gingerbread	25	0.14	$25/180 = 0.14$
Oatmeal Raisin	25	0.14	$25/180 = 0.14$
Peanut Butter	40	0.22	$40/180 = 0.22$
Total	180	1.00	

You tell your friend that you will find the expected frequencies by multiplying the relative frequencies by the sales for all six cookie types, which are the observed frequencies for the second sales event.

1: Test Set-Up

Your friend gives you a breakout of sales by cookie for the second street fair. See

Table 4:

Table 4: Week 2 Sales per Cookie

Cookie	Week 2 Sales
Almond Sugar	34
Brown Sugar Shortbread	45
Chocolate Chip	100
Gingerbread	39
Oatmeal Raisin	62
Peanut Butter	80
Total	360

You will then use the relative frequencies from Week 1 to calculate the expected frequencies for Week 2. See Table 5:

Table 5: Expected Frequency Calculations for Week 2

Cookie	O	E*	Found By
Almond Sugar	34	48	= 360 * 0.13
Brown Sugar Shortbread	45	44	= 360 * 0.12
Chocolate Chip	100	88	= 360 * 0.24
Gingerbread	39	50	= 360 * 0.14
Oatmeal Raisin	62	50	= 360 * 0.14
Peanut Butter	80	80	= 360 * 0.22
Total	360	360	= 360 * 1.00

***Please note:** The expected frequencies for this problem have been rounded off to a whole number.

2: Select the Significance Level, α

Like last week, you select a 0.05 significance level.

3: State the Null and Alternate Hypotheses

We revise the null and alternate hypothesis because we now know that each of the cookie recipes do not have equal sales. **Our research question now is:** Do the actual (observed) sales match our expected sales? Here are the null and alternate hypotheses:

- H_0 : The observed sales per cookie match expected sales
- H_1 : The observed sales per cookie do not match expected sales

4: Compose the Decision Rule

Given that the degrees of freedom is still 5 and the level of significance stays at 0.05, the critical value for Chi-Square remains 11.070. The decision rule does not change: Reject the Null Hypothesis if Chi-Square is greater than 11.070.

5: Calculate the Test Statistic and p-Value

Using the observed and expected frequencies found in Table 5, you are now ready to conduct the Chi-Square goodness of fit test for unequal expected frequencies.

Figure 9 shows the complete test results with the p-value:

	A	B	C	D	E	F
1		Cookie	O	E	(O - E)²/E	
2	1	Almond Sugar	34	48	4.083	=(C2-D2)^2/D2
3	2	Brown Sugar Shortbread	45	44	0.023	=(C3-D3)^2/D3
4	3	Chocolate Chip	100	88	1.636	=(C4-D4)^2/D4
5	4	Gingerbread	39	50	2.420	=(C5-D5)^2/D5
6	5	Oatmeal Raisin	62	50	2.880	=(C6-D6)^2/D6
7	6	Peanut Butter	80	80	0.000	=(C7-D7)^2/D7
8		Total	360	360	11.042	=SUM(E2:E7)
9		Significance Level, α	0.05 Given by user			
10		k	6 =COUNTA(B2:B7)			
11		df (k - 1)	5 =COUNTA(B2:B7)-1			
12		Critical Value	11.070 =CHISQ.INV.RT(C15,5)			
13		Chi-Square	11.042 Σ[O - E]²/E]			
14		p-value	0.0505 =CHISQ.DIST.RT(E8,C11)			
15		Effect Size, ϕ	0.1751 =SQRT(C13/C8)			
16		Interpretation of ES	Weak See A17			
17		=IF(C15<0.1,I17,IF(C15<0.2,I18,IF(C15<0.4,I19,IF(C15<0.6,I20,IF(C15<0.8,I21,I22))))))				

Figure 9: Chi-Square Goodness of fit Test Calculations

The value of Chi-Square is 11.042, just below the critical value of 11.070. The p-value is 0.0505, just above the significance level of 0.05. Figure 10 shows the Chi-Square distribution with a 5.05 percent p-value.

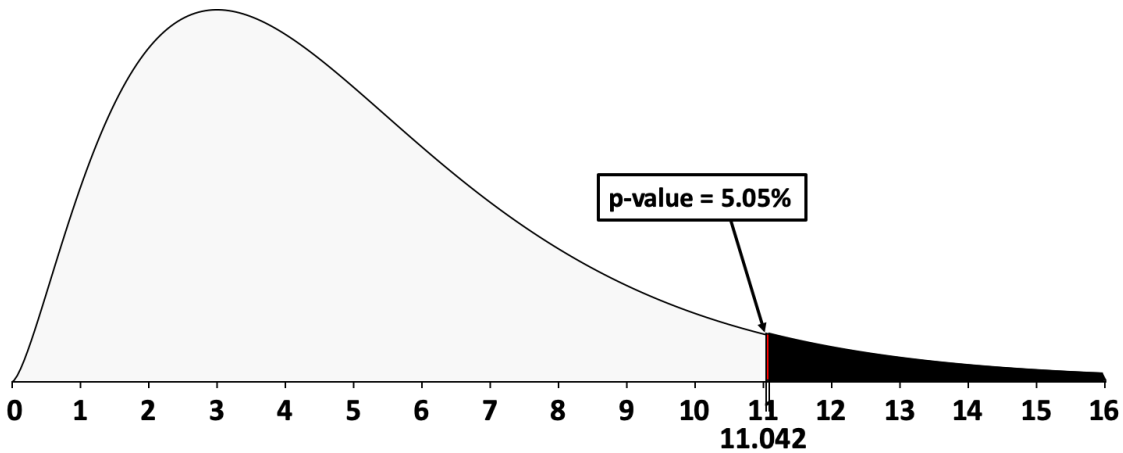


Figure 10: Chi-Square Distribution with a p-Value of 1.09%

Effect size fell from 0.2874 in the previous example to 0.1751. This is a weak effect size that suggests that the recipe counts for 17.51 percent of the cookie sales.

$$\varphi, \text{ Effect Size} = \sqrt{\frac{11.042}{360}} = 0.1751$$

Equation 7: Effect Size

Given that we will fail to reject the null hypothesis given the p-value of 5.05 percent, we need to conduct a power analysis. The smaller effect size reduces statistical power. Here are the estimates of statistical power and the probability of a Type II error performed using G*Power. **Please note:** We must change the noncentrality lambda parameter to match the Chi-Square value of 11.042.

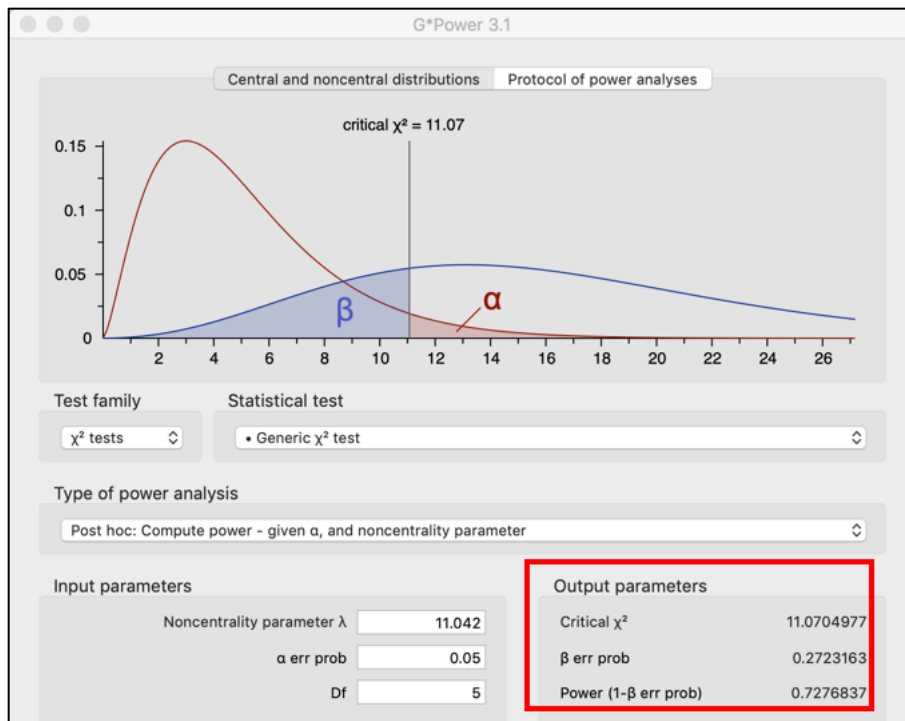


Figure 11: Statistical Power = 72.77%, $P(\text{Type II}) = 27.23\%$

Statistical power is low, 72.77 percent. It is, however, above the bare minimum of 70 percent. We conclude that we have sufficient power for this test, with the probability of a Type II error—a false negative—of 27.23 percent.

6: Decide and Report

We fail to reject the null hypothesis based on the Chi-Square value and p-value. We have insufficient evidence to conclude that the observed sales do not match the expected sales. The difference between the observed and expected frequencies is most likely random sampling error.

This test lacks statistical significance, but has practical significance. The information uncovered provides the baker with useful information for managing her inventory. She knows her cookies do not sell equally and has figured each recipe's portion to total sales. The data suggest that she should prepare more chocolate chip cookies than almond sugar, gingerbread, or brown sugar shortbread cookies. **Given**

that the Null Hypothesis was nearly rejected, the baker should continue to monitor the observed and expected sales of her cookies.

V. Chi-Square (χ^2) Contingency Table Test

Since 1969, the Gallup organization has been polling Americans regarding their attitudes toward the legalization of marijuana. In 1969, only 12 percent of respondents favored legalizing marijuana. On October 23, 2019, Gallup published the results of their latest survey: 66 percent of Americans favored the legalization of marijuana, the same as the 2018 survey. Table 6 shows the results for Republicans, Independents, and Democrats:

Table 6: 2019 Gallup Poll on the Legalization of Marijuana

Party	n	Yes, Legal	No, DK, NA
Republican	393	199, 51%	194, 49%
Independent	655	442, 67%	213, 33%
Democrat	450	352, 78%	98, 22%
Total	1,498		

Research Question: Does a person’s position on the legalization of marijuana depend on his or her political identification? We have addressed this question with confidence intervals and two-sample z-tests for proportion. We can also answer this question using a Chi-Square contingency table test. This test has big advantages over two-sample z-tests for proportion because the analysis can be conducted with a single test without increasing the probability of a Type I error. The down-side is that the Chi-Square tests have lower statistical power than parametric tests like the z-test, t-test, or ANOVA.

1: Test Set-Up - 2 x 3 Contingency Table

Table 6, shown above, details the proportion of people who do and do not favor the legalization of marijuana by political affiliation. **Please note:** Gallup has three responses: 1) Yes/legal, 2) No/not legal, and 3) Don’t Know/Refused to Answer. Because the third

response has low frequencies (less than 5 for Republicans and Democrats and only 8 for independents), the “no, not legal” and the “don’t know/refused” responses were combined to meet the requirements of Chi-Square tests. This is legitimate because all of these respondents do not favor the legalization of marijuana.

Remember: Small expected frequencies can lead to erroneous conclusions. Chi-Square is very sensitive to small sample sizes and tiny Expected Frequencies. The reason for combining the “don’t know” and “no answer” responses with the “do not legalize” responses was to avoid having a cell in the Chi-Square contingency table with unacceptably low Expected Frequencies.

2: Select the Significance Level, α

A 5 percent level of significance was selected.

3: State the Null and Alternate Hypotheses

- H_0 : A person’s attitude toward the legalization of marijuana is independent of his or her political identification. (No relationship between a person’s attitude towards legalization of marijuana and his or her political identification).
- H_1 : A person’s attitude toward the legalization of marijuana is dependent on his or her political identification. (There is a relationship between a person’s attitude towards legalization of marijuana and his or her political identification).

4: Compose the Decision Rule

Here is the formula to determine degrees of freedom for a contingency table test and the results for this problem:

$$df = (\# \text{ of rows} - 1)(\# \text{ of columns} - 1) = (3 - 1)(2 - 1) = 2$$

Equation 8: Formula for Degrees of Freedom For Contingency Tables

The three rows are Republicans, Independents, and Democrats. The two columns are “favors the legalization of marijuana,” and “does not favor the legalization of marijuana.”

Please note: We do not count the “total” columns or rows in the calculation of degrees of freedom.

For this problem with three rows and three columns there are 2 degrees of freedom.

As shown in Figure 12, the critical value for Chi-Square is 5.991.

		Right-Tail Area				
		Level of Significance, α				
df	0.10	0.05	0.02	0.01	0.001	
1	2.706	3.841	5.412	6.635	10.828	
2	4.605	5.991	7.824	9.210	13.816	
3	6.251	7.815	9.837	11.345	16.266	
4	7.779	9.488	11.668	13.277	18.467	
5	9.236	11.070	13.388	15.086	20.515	
6	10.645	12.592	15.033	16.812	22.458	
7	12.017	14.067	16.622	18.475	24.322	
8	13.362	15.507	18.168	20.090	26.124	
9	14.684	16.919	19.679	21.666	27.877	
10	15.987	18.307	21.161	23.209	29.588	

Figure 12: Critical value for Chi-Square with 4 degrees of freedom at a 5% significance level = 9.488

The critical value of Chi-Square can be found using Excel: =CHISQ.INV.RT(alpha,df).

The decision rule: Reject the Null Hypothesis if Chi-Square is greater than 5.991. Figure

13 shows a chart of this Chi-Square distribution with its rejection region:

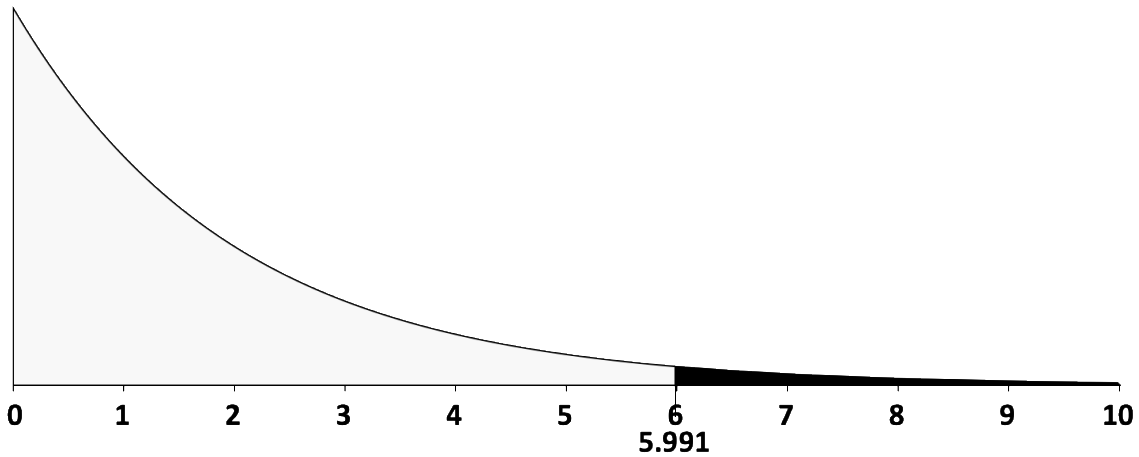


Figure 13: Chi-Square Distribution with 2 degrees of freedom at a 5% significance level, $\chi^2_{2,0.05}$

5: Calculate the Test Statistic and p-Value

Figure 14 shows the observed frequencies based on the data reported by Gallup:

Party	<u>Yes, Legal</u>		<u>No/DK/NA</u>		<u>Total</u>	
	O	E	O	E	O	E
Republican	199		194		393	
Independent	442		213		655	
Democrat	352		98		450	
Total	993		505		1,498	

Figure 14: Chi-Square Contingency Table with Observed Frequencies

Equation 9 shows the formula we use to calculate the expected frequencies:

$$E = \frac{(\text{Row Total})(\text{Column Total})}{\text{Grand Total}}$$

Equation 9: Formula for Calculating Expected Frequencies

Please note: The row and column totals for the observed frequencies must equal those for the expected frequencies.

Figure 15 shows the contingency table with the expected frequencies.

Party	Yes, Legal		No/DK/NA		Total	
	O	E	O	E	O	E
Republican	199	261	194	132	393	393
Independent	442	434	213	221	655	655
Democrat	352	298	98	152	450	450
Total	993	993	505	505	1,498	1,498

Figure 15: Chi-Square Contingency Table with Expected Frequencies

The next step is to calculate the Chi-Square test statistic. Figure 16 shows the result of this calculation:

	A	B	C	D	E	F	G	H	I	
1	Gallup, October 23, 2019									
2	https://news.gallup.com/poll/267698/support-legal-marijuana-steady-past-year.aspx									
3	Interviews conducted 10/1-13/2019									
4										
5		Yes, Legal		No/DK/NA		Total				
6	Party	O	E	O	E	O	E			
7	Republican	199	261	194	132	393	393			
8	Independent	442	434	213	221	655	655			
9	Democrat	352	298	98	152	450	450			
10	Total	993	993	505	505	1,498	1,498			
11										
12										
13		Yes, Legal		No/DK/NA		(O - E) ² /E				
14	Party	O	E	O	E	Yes	No			
15	Republican	199	261	194	132	14.525	28.561	=((D15-E15)^2/E15)		
16	Independent	442	434	213	221	0.141	0.276	=((D16-E16)^2/E16)		
17	Democrat	352	298	98	152	9.668	19.010	19.010		
18	Total	993	993	505	505	$\chi^2 =$	72.181	=SUM(F15:G17)		
19	Critical Value =						5.991	=CHISQ.INV.RT(0.05,2)		
20	df =						2	(# rows -1)(# columns - 1)		
21	Alpha, $\alpha =$						0.05	CV = 5.991 =CHISQ.INV.RT(D21,G20)		
22	p-value =						>0.001	=CHISQ.DIST.RT(G18,df)		
23	Effect Size, $\phi =$						0.27	=SQRT(G19/B18)		
24	Interpretation of ES						Moderate	See A25		
25	=IF(G23<0.1,L17,IF(G23<0.2,L18,IF(G23<0.4,L19,IF(G23<0.6,L20,IF(G23<0.8,L21,L22))))))									

Figure 16: Completed Chi-Square Contingency Table

The Chi-Square value of 72.181 is very large and the p-value is extremely small, well below 0.001.

As shown in Figure 17, statistical power is 100%. This raises the issue of whether this test is over-powered; that there is no chance of failing to reject the null hypothesis. But

the effect size is 0.27, which is a moderate effect. The test, therefore, is not over-powered. The difference between the approval rates for the legalization of marijuana—51 percent for Republicans, 67 percent of independents, and 78 percent for Democrats—has both statistical and practical significance.

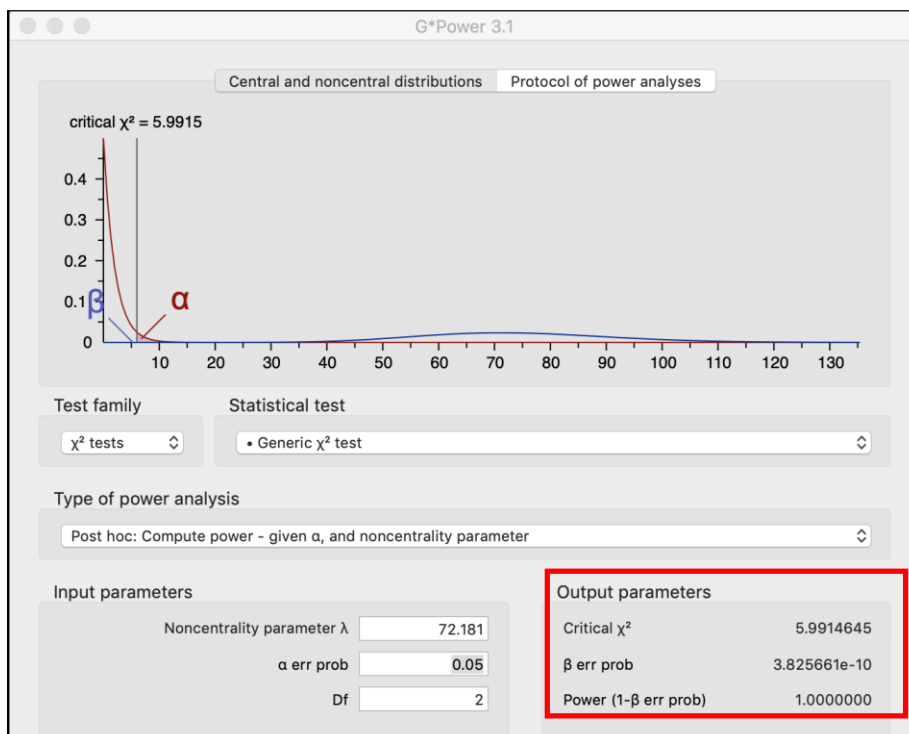


Figure 17: Power Analysis

6: Decide and Report

With Chi-Square as high as 72.181 and a p-value far less than 0.001, we reject the Null Hypothesis. **Conclusion:** There is a relationship between a person's attitude toward the legalization of marijuana and his or her political identification. Republicans are less likely to think marijuana should be legal than either Independents or Democrats. But, as previously stated, Chi-Square is an omnibus test and we would have to conduct a *post hoc* analysis to determine which of the three of the pairs are unequal: 1) Republicans vs. Independents, 2) Republicans vs. Democrats, and 3) Independents vs. Democrats.

V. Chi-Square (χ^2) Normality Test

Dr. V is studying his students' test results. He took a random sample of 144 exams. Before he can proceed in his analysis, he needs to determine whether the data are normally distributed. If so, he will conduct his analysis using parametric techniques. If the data are not normally distributed, he must employ nonparametric techniques. Table 7 shows Dr. V's data:

Table 7: 144 Test Results

94	99	73	62	55	78	92	98	97	50	79	70
98	98	58	75	40	33	55	75	96	98	19	67
74	98	30	46	46	99	76	99	87	68	75	35
99	85	85	69	72	56	65	94	38	75	79	32
91	92	90	88	99	72	60	83	70	30	51	20
96	98	96	35	97	96	98	66	88	63	38	18
97	98	94	97	97	56	33	67	62	72	78	88
98	97	66	19	99	96	62	84	73	81	42	19
80	86	32	90	72	72	82	56	49	95	97	15
84	87	97	73	25	61	72	72	74	43	66	78
96	57	58	25	98	60	72	78	94	60	56	68
92	67	71	74	71	72	50	60	25	57	32	72

1: Test Set-Up

Here are the seven steps to conduct a Chi-Square goodness of fit test for normality. We will do this using Microsoft Excel because is much faster and easier than doing this by hand:

- 1) Calculate the sample mean, \bar{X} , and the sample standard deviation, s . (**Please note:** if possible calculate the population mean, μ , and population standard deviation, σ).
- 2) Find the z-values for each random variable using Excel's STANDARDIZE function.
- 3) Create a frequency table for the z-values from -3.50 to +3.50 in 0.50 increments.

This table should have four columns: 1) The bins or categories, 2) The observed frequencies, 3) The probability of the expected frequencies, and 4) The expected frequencies.

- 4) Enter the observed frequencies in the second column of the frequency distribution using Excel's FREQUENCY array function. **Please Note:** You must press Control+Shift+Enter when entering an array function.
- 5) In the third column, enter the probability of the expected frequencies using Excel's NORM.S. DIST function.
- 6) Enter the expected frequencies by multiplying the probability of the expected frequencies by the total number of observations.
- 7) To conform to the Chi-Square requirement that the expected frequencies must be at least 5, combine adjacent categories when their expected frequencies are less than 5.

Here are the seven steps to solve this problem:

1: Calculate the sample mean and standard deviation. The sample mean, \bar{X} , and the sample standard deviation, s , are stand-ins for the unknown population parameters, μ and σ . The mean, \bar{X} , for our data is 70.65 and the standard deviation, s , is 23.19.

2: Find the z-value for each random value using Excel's STANDARDIZE function:

$$= \text{STANDARDIZE}(X, \text{Mean}, \text{Standard Deviation})$$

Where: X is the random value

Mean is the sample mean

Standard Deviation is the sample standard deviation

Equation 10: Microsoft Excel's STANDARDIZE Function

3: Create a frequency table with 17 rows and four columns. In the first row, place the names of the four columns. The first column will be for the bins or categories for a normal distribution. We will have 15 categories for the z-values ranging from -3.50 to +3.50. These categories represent the distance from the mean in z-values. See Table 8.

Table 8: Chi-Square Normalcy Table

Bins	O	P(E)	E
-3.50			
-3.00			
-2.50			
-2.00			
-1.50			
-1.00			
-0.50			
0.00			
0.50			
1.00			
1.50			
2.00			
2.50			
3.00			
3.50			

Total

4: Column O is for the observed frequencies using the FREQUENCY function. Here is

the syntax for this function:

$$\{=FREQUENCY(\text{data array}, \text{bin array})\}$$

Equation 11: FREQUENCY Function

Where: data array is the range of cells with the data, and bin array is the cell range within the bins. The curly bracket symbols “{}” mean that this function is an array function. Enter array functions by holding down the Control, Shift, and Enter (or RETURN) keys.

Table 9: Observed Frequencies Found Using the FREQUENCY Function

Bins	O	P(E)	E
-3.50	0		
-3.00	0		
-2.50	0		
-2.00	6		
-1.50	12		
-1.00	7		
-0.50	14		
0.00	21		
0.50	32		
1.00	35		
1.50	0		
2.00	0		
2.50	0		
3.00	0		
3.50	0		
Total	144		

Please Note: The first three bins and the last five bins have no values. We may eliminate these cells after we find the expected frequencies.

5: The third column is for the probability of the expected frequency. The first cell in this column has a different formula than the other cells.

$$= \text{NORM. S. DIST}(F2, \text{True})$$

Where: F2 is the first cell in the Bins column.

Equation 12: Excel Formula for the first cell in P(E) column

The second cell gets this function, which can then be pasted into the remaining cells

in this column:

$$= \text{NORM. S. DIST}(F3, \text{True}) - \text{NORM. S. DIST}(F2, \text{True})$$

Equation 13: Excel Function for the Second Cell in the P(E) Column

Table 10: $P(E)$ Column Values

Bins	O	P(E)	E
-3.50	0	0.0002	
-3.00	0	0.0011	
-2.50	0	0.0049	
-2.00	6	0.0165	
-1.50	12	0.0441	
-1.00	7	0.0918	
-0.50	14	0.1499	
0.00	21	0.1915	
0.50	32	0.1915	
1.00	35	0.1499	
1.50	0	0.0918	
2.00	0	0.0441	
2.50	0	0.0165	
3.00	0	0.0049	
3.50	0	0.0011	
Total	144		

#6: To find the expected frequencies, we multiply $P(E)$ by the total number of observations, n , 144. See Table 11 for the expected frequencies.

Table 11: Expected Frequencies

Bins	O	P(E)	E
-3.50	0	0.0002	0.03
-3.00	0	0.0011	0.16
-2.50	0	0.0049	0.70
-2.00	6	0.0165	2.38
-1.50	12	0.0441	6.34
-1.00	7	0.0918	13.23
-0.50	14	0.1499	21.58
0.00	21	0.1915	27.57
0.50	32	0.1915	27.57
1.00	35	0.1499	21.58
1.50	0	0.0918	13.23
2.00	0	0.0441	6.34
2.50	0	0.0165	2.38
3.00	0	0.0049	0.70
3.50	0	0.0011	0.16
Total	144	0.9998	144

We now have our expected frequencies. **But, we also have a problem.** We cannot conduct a Chi-Square goodness of fit test when some cells have expected frequencies of less than five. In addition, observed frequencies in the first or last cell must be greater than zero. We must combine cells.

7: Combine the cells, when necessary. Table 12, shows the contingency table with combined bins.

Table 12: Combined Bins

Bins	O	P(E)	E
-2.00	6	0.023	3.28
-1.50	12	0.044	6.34
-1.00	7	0.092	13.23
-0.50	14	0.150	21.58
0.00	21	0.191	27.57
0.50	32	0.191	27.57
1.00	17	0.150	21.58
3.50	35	0.158	22.81
Total	144	1.000	144

2: Select the Significance Level, α

We will use a 0.05 significance.

3: State the Null and Alternate Hypotheses

- H_0 : The distribution follows a normal probability distribution
- H_1 : The distribution does **not** follow a normal probability distribution

4: Compose the Decision Rule

Degrees of freedom for most Chi-Square goodness of fit tests are found by using the formula $k - 1$. But, with tests for normality that use the sample mean and the sample standard deviation to estimate μ and σ , two degrees of freedom are lost. The formula for degrees of freedom, therefore, changes to $k - 3$. With eight classes, we have five degrees of

freedom. The critical value for Chi-Square with 5 degrees of freedom and a 5 percent significance level is 11.070. Figure 18 shows a Chi-Square distribution with 5 degrees of freedom and a 5 percent significance level:

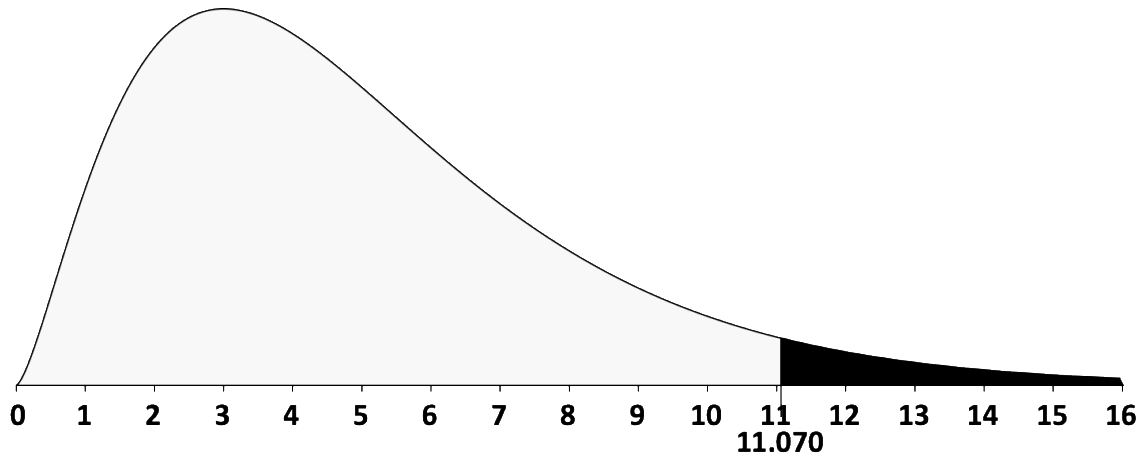


Figure 18: Chi- Chi-Square Distribution with 5 df and a 5% α , $\chi^2_{5,0.05}$

5: Calculate the Test Statistic, p-Value, Effect Size, and Power

Table 13 shows the calculation for Chi-Square and the p-value:

Table 13: Chi-Square Calculation

Bins	O	P(E)	E	$((O-E)^2)/E$
-2.00	6	0.023	3.28	2.265
-1.50	12	0.044	6.34	5.042
-1.00	7	0.092	13.23	2.931
-0.50	14	0.150	21.58	2.664
0.00	21	0.191	27.57	1.566
0.50	32	0.191	27.57	0.712
1.00	17	0.150	21.58	0.973
3.50	35	0.158	22.81	6.511
Total	144	1.000	144	22.665
			df	5
			CV	11.070
			p-value	0.0004

Effect size for this problem is 0.3697, found by:

$$\varphi \text{ Effect Size} = \sqrt{\frac{\chi^2}{n}} = \frac{22.664}{144} = 0.3697$$

Equation 14: Effect Size

This is a moderate effect based on the interpretation of Phi Effect Size found in Table 2.

As shown in Figure 19, this test has ample statistical power: 97.4 percent with a 2.6 percent probability of a Type II error. Given the moderate effect size of 0.3697, we need not worry that this test is over-powered.

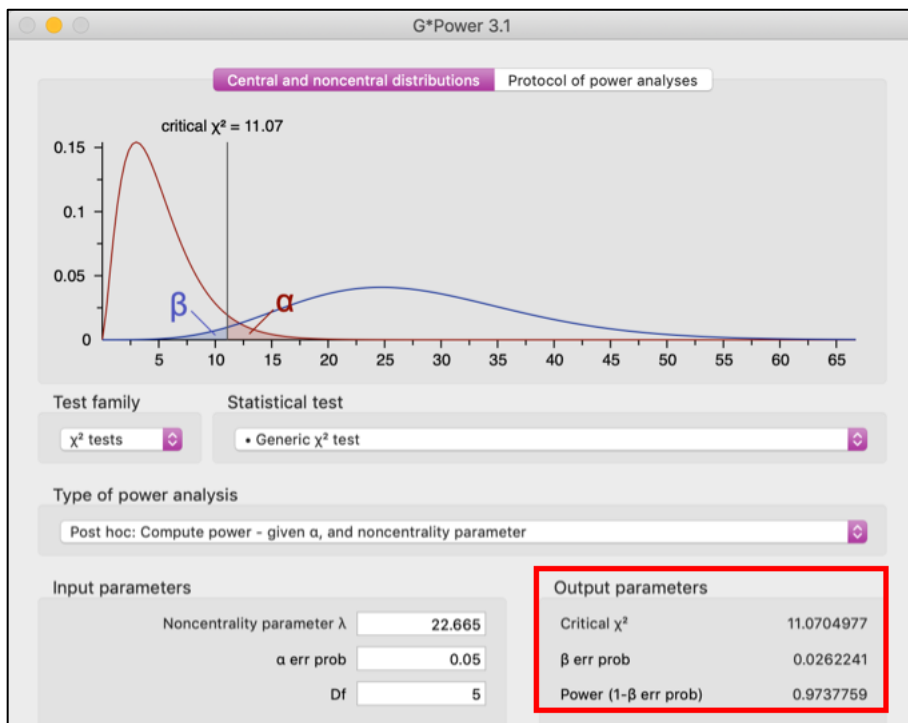


Figure 19: G*Power Output - Power = 97.4%, P(Type II) = 2.6%

6: Decide and Report

The Chi-Square value of 22.665 is in the rejection region and the p-value of 0.04 percent is well below the 5 percent significance level. We reject the Null Hypothesis. There is very strong evidence that Dr. V's data are not normally distributed. **Conclusion:** Dr. V cannot use parametric tests to analyze this data. This conclusion has practical significance because

any analysis using parametric techniques would be seriously flawed and Dr. V would look foolish if he were to present his analysis using parametric techniques.

VI. Summary

We have demonstrated that Chi-Square is an easy-to-use technique to determine whether the relationship of the data we observed fits the expected distribution. We looked at three types of Chi-Square tests: 1) The Chi-Square goodness of fit or one-way test, 2) The Chi-Square contingency or two-way test, and 3) The Chi-Square goodness of fit test for normality. All three test use the same equation:

$$\chi^2 = \Sigma \left[\frac{(O - E)^2}{E} \right]$$

Equation 15: The Chi-Square Test Statistic Equation

The most challenging aspect of Chi-Square is determining the expected frequencies. We examined several examples of how to find the expected frequencies.

VII. Exercises

Data for these exercises can be found in 16_Exercises.xlsx.

Exercise 1 – Chi-Square Goodness of Fit Test

A recent study in a national business journal asked consumers the following question: “In general, how would you rate the level of service provided by your supermarket?” A district manager of Piggly Wiggly supermarkets asked the same question about Piggly Wiggly supermarkets in her district. She took a random sample of 525 Piggly Wiggly shoppers. The question she seeks to answer is: Do the results of the Piggly Wiggly survey contradict the national study using a 5 percent significance level?

1: Test Set-Up

Table 14 shows the result of the Piggly Wiggly survey along with consumers' ratings of supermarkets from the national study.

Table 14: Piggly Wiggly Supermarket Survey

National			
Rating	Study	O	E
Excellent	30%	225	?
Good	30%	160	?
Fair	30%	100	?
Poor	10%	40	?
Total	100%	525	525

Find the expected frequencies

Table 15: Expected Frequencies (Note: Expected frequencies have been rounded off to a whole number.)

National				
Rating	Study	O	E	Found by
Excellent	30%	225		
Good	30%	160		
Fair	30%	100		
Poor	10%	40		
Total	100%	525	525	

2: Select the Significance Level, α

A 5 percent significance level has been selected

3: State the Null and Alternate Hypotheses

- H_0 :
- H_1 :

4: Compose the Decision Rule

5: Calculate the Test Statistic, ES, and p-Value

6: Decide, Report, and Conduct a Power Analysis Using G*Power

Exercise 2 – Chi-Square Contingency Table (Test of Independence)

1: Test Set-Up

A major retailer is looking at an old study on shoppers' favorite person to shop for during the holiday season. They have data for baby boomers (people born between 1946 and 1964), Gen-Xers (people born between 1965 and 1980), and Millennials (people born between 1981 and 1996). Results of the study are shown in Table 16.

Table 16: Observed Frequencies Source – Russell Research

Observed Frequencies

Favorite person to shop for during the holiday season	Baby Boomers	Gen-Xers	Millennials	Total
Spouse/Partner	106	76	50	232
Boyfriend/Girlfriend	10	13	60	83
Parent	21	37	50	108
Child(ren)	137	169	94	400
Extended Family Member	31	17	19	67
Friend	17	13	35	65
Someone Else	17	3	6	26
Total	339	328	314	981

Find the expected frequencies:

Table 17: Expected Frequencies

Favorite person to shop for during the holiday season	Baby Boomers	Gen-Xers	Millennials	Total
Spouse/Partner				
Boyfriend/Girlfriend				
Parent				
Child(ren)				
Extended Family Member				
Friend				
Someone Else				
Total				

2: Select the Significance Level, α

A 5 percent significance level has been selected.

3: State the Null and Alternate Hypotheses

- H_0 :

- H_1 :

4: Compose the Decision Rule

5: Calculate the Test Statistic, ES, and p-Value

6: Decide and Report

Exercise 3 – Chi-Square Goodness of Fit Test for Normality

Dr. Siegfried Zuckerkrank is a young endocrinologist who works at one of the world's [foremost](#) diabetes clinics. He is especially interested in how well newly-diagnosed insulin-dependent patients manage the difficult task of regulating their blood sugar levels. For a person without diabetes, blood sugar typically ranges from a low of 70 mg/dL to a high of 130 md/dL. An insulin-dependent diabetic is prone to both hypoglycemia (blood sugar <70 mg/dL) and hyperglycemia (blood sugar > 180 mg/dL). Insulin-dependent diabetics struggle to avoid both hypoglycemia and hyperglycemia.

One measure Dr. Zuckerkrank wants to study is the blood sugar levels of patients when they visit the clinic for their first examination after their initial diagnosis. Table 18 shows the blood sugar levels for a random sample of 121 patients:

Table 18: Blood Sugar Levels mg/dL

248	148	188	255	206	130	175	231	218	65	203
56	186	257	169	96	173	163	128	235	137	185
260	133	211	247	213	257	146	141	224	245	109
130	150	74	197	56	184	132	220	102	180	86
259	214	157	199	241	193	158	132	205	186	231
76	115	103	240	244	61	257	195	97	77	212
115	199	69	75	261	111	143	179	150	67	94
258	68	237	208	119	163	152	211	162	239	69
68	256	175	163	113	186	247	121	155	196	249
154	65	158	248	180	215	170	142	165	164	82
188	85	114	153	145	207	56	76	242	89	250

Dr. Zuckerkrank's first question: Are the blood sugar levels normally distributed? To answer this question, he conducts a Chi-Square goodness of fit test for normality. If the data are normally distributed, parametric tests may be used. Should the data not be normally distributed, nonparametric techniques must be employed.

1: Test Set-Up

Set up the test using the following steps:

- Calculate the sample mean, \bar{X} , and the sample standard deviation, s .
- Find the z-values for each random value using Excel's STANDARDIZE function.
- Create a frequency table for the z-values from -3.50 to +3.50 in 0.50 increments. The table should have four columns: 1) the bins, or categories, 2) the observed frequencies, 3) The probability of the expected frequencies, and 4) the expected frequencies. Table 19 shows the format of this table:

Table 19: Frequency Distribution Format

Bins	O	P(E)	E
-3.50			
-3.00			
-2.50			
-2.00			
-1.50			
-1.00			
-0.50			
0.00			
0.50			
1.00			
1.50			
2.00			
2.50			
3.00			
3.50			

Total

- d) Enter the observed frequencies in the second column of the frequency distribution using Excel's FREQUENCY array function.
- e) In the third column enter the probability of the expected frequencies using Excel's NORM.S. DIST function.
- f) Enter the expected frequencies by multiplying the probability of the expected frequencies by the total number of observations.
- g) Because the expected frequencies in the tails are small, combine cells to conform to the Chi-Square requirement that the expected frequencies must be at least 5. Also, combine the observed frequencies cells at the extremes, if these cells have a frequency of zero. The condensed frequency distribution meets the requirements for a Chi-Square test because none of the expected frequency cells have a frequency less than 5.

2: Select the Significance Level, α

Dr. Zuckerkrank intends on using a 0.05 significance level.

3: State the Null and Alternate Hypotheses

- H_0 :
- H_1 :

4: Compose the Decision Rule

5: Calculate the Test Statistic, ES, and p-Value

6: Decide and Report

¹ R. A. Fisher, *Scientific Methods for Research Workers*, 13th Edition, (New York: Harper, 1958), p. 22.

² Karl Pearson, "On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is Such that it Can Be Reasonably Supposed to Have Arisen from Random Sampling," *Philosophical Magazine*. Series 5. 50 (302), 1900, pp. 157–175.

³ A. W. F. Edwards, "R. A. Fisher on Karl Pearson," *Notes and Record of the Royal Society of London*, Volume 48, No. 1, January 1994. pp. 97-106.

⁴ Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences*, Second Edition, (New York: Psychology Press, 1988), p. 215.

⁵ Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences*, Second Edition, (New York: Psychology Press, 1988), p. 215.

⁶ Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences*, Second Edition, (New York: Psychology Press, 1988), p. 227.