Chapter 9

Analysis of Two-Way Tables

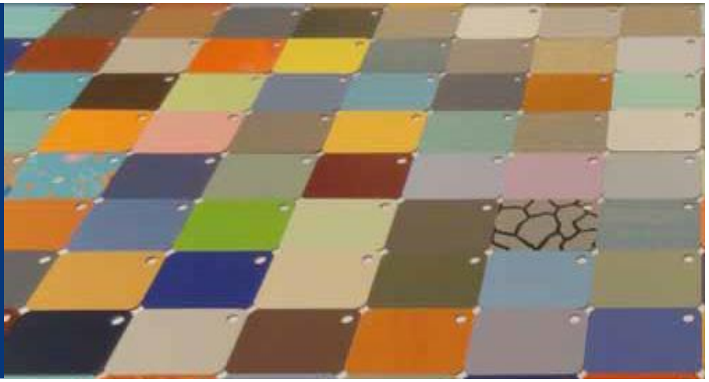
Introduction to the Practice of

STATISTICS EIGHTH EDITION

Moore / McCabe / Craig

Lecture Presentation Slides

Chapter 9
Analysis of Two-Way Tables



9.1 Inference for Two-Way Tables

9.2 Goodness of Fit

9.1 Inference for Two-Way Tables

- Two-way tables
- Expected cell counts
- The chi-square statistic
- The chi-square distributions
- The chi-square test
- Computations
- Computing conditional distributions
- Models for two-way tables

Two-Way Tables

Chapter 2 discussed how two-way tables can be used to describe the joint distribution of two categorical variables. Such tables can be used to describe the **relationship** between two categorical variables. A relationship exists when the distribution of one variable depends on the value of the other variable.

When the data are obtained from random sampling, two-way tables of counts can be used to formally test the hypothesis that **the two categorical variables are independent in the population from which the data were obtained.**

Expected Cell Counts

The rows of a two-way table are the values of one categorical variable and the columns are the values of the other variable. The count in any particular cell of the table equals the number of subjects who fall into that cell. We want to test the hypothesis (H_0) that there is no relationship between the two categorical variables.

To test this hypothesis, we compare **actual counts** from the sample data with **expected counts,** where the latter counts are those expected when there is no relationship between the two variables.

The expected count in any cell of a two-way table when H_0 is true is:

$$\boxed{\text{expected count} = \frac{\text{row total} \times \text{column total}}{n}}$$

The Chi-Square Statistic

To see if the data give convincing evidence against the null hypothesis, we compare the observed counts from our sample with the expected counts assuming H_0 is true.

Assume there are r rows in the two-way table and c columns, which means there are $r \times c$ cells.

The test statistic that makes the comparison is the **chi-square statistic.**

The **chi-square statistic** is a measure of how far the observed counts are from the expected counts. The formula for the statistic is:

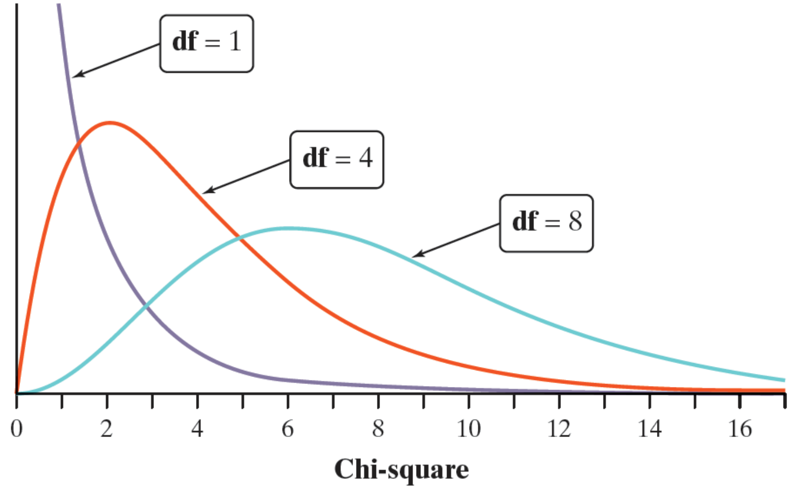
$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

where "observed" represents an observed cell count, "expected" represents the expected count for the same cell, and the sum is over all $r \times c$ cells in the table.

The Chi-Square Distributions

When the observed counts are very different from the expected counts, a large value of χ^2 will result, providing evidence against the null hypothesis. When the observed and expected counts are in close agreement, a small value of χ^2 will result.

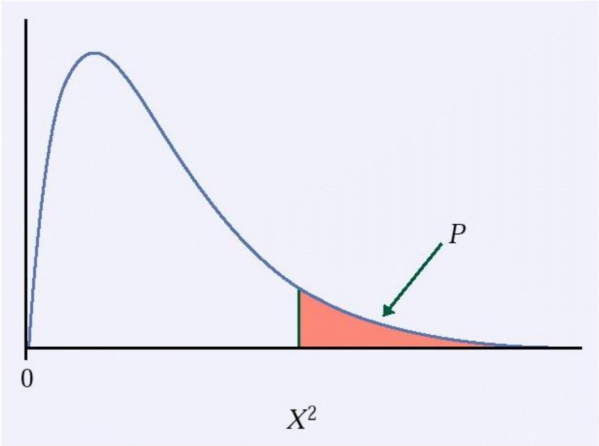
The *P*-value for a χ^2 test comes from comparing the value of the χ^2 statistic with critical values for a **chi-square distribution.**



The Chi-Square Distributions

The **chi-square distributions** are a family of distributions that take only positive values and are skewed to the right. A particular χ^2 distribution is specified by giving its **degrees of freedom.**

The χ^2 test for a two-way table with *r* rows and *c* columns uses critical values from the χ^2 distribution with $(r – 1)(c – 1)$ degrees of freedom. The *P*-value is the area under the density curve of this χ^2 distribution to the right of the value of the test statistic.



Cell Counts Required for the Chi-Square Test

The chi-square test is an approximate method that becomes more accurate as the counts in the cells of the table get larger. We must therefore check that the counts are large enough to allow us to trust the *P*-value. Fortunately, the chi-square approximation is accurate for quite modest counts.

Cell Counts Required for the Chi-Square Test

You can safely use the chi-square test with critical values from the chi-square distribution when the average of the expected counts is 5 or more and all individual expected counts are 1 or greater. In the special case of a 2×2 table, all four expected counts should be 5 or more.

The Chi-Square Test

The chi-square test is an overall test for detecting relationships between two categorical variables. If the test is significant, it is important to look at the data to learn the nature of the relationship. We have three ways to look at the data:

1) Compare conditional distributions: Compute the conditional distribution of the column variable for each row. The chi-square test is testing whether or not these conditional distributions are the same in the population.

2) Compare observed and expected cell counts: Which cells have more or fewer observations than we would expect if H_0 were true?

3) Look at the terms of the chi-square statistic: Which cells contribute the most to the value of χ^2?

The Chi-Square Test

In addition to the conditions on expected cell counts, there is another important assumption required for validity of the chi-square test.

It must be true that **we have a simple random sample of subjects from a population of subjects, and each subject in the population falls into one and only one cell of the two-way table.**

There is a *distinctly different* situation in which the chi-square test may also be used. Suppose we have c independent random samples from c different populations, and we classify the individuals within each sample according to a categorical variable that takes on r values.

We now have an $r \times c$ table that looks the same as before but is obtained via a different sampling scheme.

In this second situation, we may use the chi-square test to test the hypothesis

H_0: the distribution of the categorical variable is the same in each of the c populations

The Chi-Square Test

In summary, the chi-square test can be used when sampling is done in either of the following two ways:

- Independent SRSs from two or more populations, with each individual classified according to one categorical variable
- A single SRS, with each individual classified according to both of two categorical variables

In either scenario, in order for the chi-square test to be valid, the following should be true:

- ❑**The average of the expected cell counts should be at least 5.**
- ❑**All individual cell counts should be at least 1.**
- ❑**In a 2 × 2 table, all four expected cell counts should be at least 5.**

Computations for Two-Way Tables

The calculations required to analyze a two-way table are straightforward, but tedious. In practice, it is recommended to use software. However, it is possible to do the work with a calculator. When analyzing relationships between two categorical variables, follow this procedure:

1) Calculate descriptive statistics that convey the important information in the table. Usually, these will be column or row percents.

2) Find the expected counts and use them to compute the χ^2 statistic.

3) Compare your χ^2 statistic to the chi-square critical values from Table F to find the approximate *P*-value for your test.

4) Draw a conclusion about the association between the row and column variables.

Calculating Expected Cell Counts

Consider the expected count of French wine bought when no music was playing:

$$\frac{99}{243} \times 84 = 34.22$$

Observed Counts				
	Music			
Wine	**None**	**French**	**Italian**	**Total**
French	30	39	30	$\boxed{\textbf{99}}$
Italian	11	1	19	**31**
Other	43	35	35	**113**
Total	$\boxed{\textbf{84}}$	**75**	**84**	$\boxed{\textbf{243}}$

The values in the calculation are the row total for French wine, the column total for no music, and the table total. We can rewrite the original calculation as:

$$\frac{\quad \cdot \quad}{\qquad} = 34.22$$

Expected Counts

The **expected count** in any cell of a two-way table when H_0 is true is:

$$\text{expected count} = \frac{\text{row total} \times \text{column total}}{\text{table total}}$$

The Chi-Square Calculation

Observed Counts				
	Music			
Wine	**None**	**French**	**Italian**	**Total**
French	30	39	30	**99**
Italian	11	1	19	**31**
Other	43	35	35	**113**
Total	**84**	**75**	**84**	**243**

Expected Counts				
	Music			
Wine	**None**	**French**	**Italian**	**Total**
French	34.22	30.56	34.22	**99**
Italian	10.72	9.57	10.72	**31**
Other	39.06	34.88	39.06	**113**
Total	**84**	**75**	**84**	**243**

For the French wine with no music, the observed count is 30 bottles and the expected count is 34.22. The contribution to the χ^2 statistic for this cell is

$$\frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}} = \frac{(30 - 34.22)^2}{34.22} = 0.52$$

The χ^2 statistic is the sum of nine such terms:

$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}} = \frac{(30 - 34.22)^2}{34.22} + \frac{(39 - 30.56)^2}{30.56} + \ldots + \frac{(35 - 39.06)^2}{39.06}$$

$$= 0.52 + 2.33 + \ldots + 0.42 = 18.28$$

The χ^2 Statistic and its *P*-Value

H_0: There is no difference in the distributions of wine purchases at this store when no music, French accordion music, or Italian string music is played.
H_a: There is a difference in the distributions of wine purchases at this store when no music, French accordion music, or Italian string music is played.

	Music			
Wine	**None**	**French**	**Italian**	**Total**
French	30	39	30	**99**
Italian	11	1	19	**31**
Other	43	35	35	**113**
Total	**84**	**75**	**84**	**243**

Our calculated test statistic is $\chi^2 = 18.28$.

To find the *P*-value using a chi-square table look for df = (3-1)(3-1) = 4.

	P	
df	**.0025**	**.001**
4	16.42	18.47

The small *P*-value (between 0.001 and 0.0025) gives us convincing evidence to reject H_0 and conclude that there is a difference in the distributions of wine purchases at this store when no music, French accordion music, or Italian string music is played.

Models for Two-Way Tables

The chi-square test is a technique that may be used to compare the distributions of a categorical variable in several populations or to test for evidence of a relationship between two categorical variables. We can either:

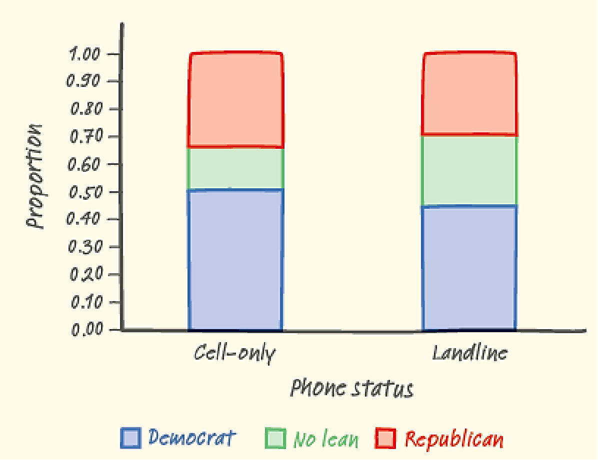
- **Compare several populations:** Randomly select several SRSs each from a different population (or from a population subjected to different treatments) → experimental study.
- **Test for independence:** Take one SRS and classify the individuals in the sample according to two categorical variables (attribute or condition) → observational study, historical design.

Both models use the χ^2 test to test the hypothesis of *no relationship*.

Comparing Several Populations

Random digit dialing telephone surveys used to exclude cell phone numbers. If the opinions of people who have only cell phones differ from those of people who have landline service, the poll results may not represent the entire adult population. The Pew Research Center interviewed separate random samples of cell-only and landline telephone users who were less than 30 years old. Here's what the Pew survey found about how these people describe their political party affiliation.

	Cell-only sample	Landline sample
Democrat or lean Democratic	49	47
Refuse to lean either way	15	27
Republican or lean Republican	32	30
Total	96	104



Testing for Independence

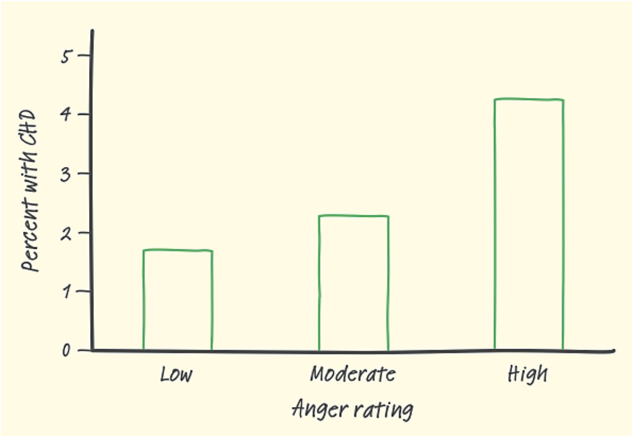
Suppose we have a *single* sample from a *single* population. For each individual in this SRS of size *n,* we measure two categorical variables. The results are then summarized in a two-way table.

The null hypothesis is that the row and column variables are independent. The alternative hypothesis is that the row and column variables are dependent.

Testing for Independence

We're interested in whether angrier people tend to get chronic heart disease (CHD) more often. We can compare the percents of people who did and did not get heart disease in each of the three anger categories:

	Low anger	Moderate anger	High anger	Total
CHD	53	110	27	190
No CHD	3057	4621	606	8284
Total	3110	4731	633	8474



There is a clear trend: As the anger score increases, so does the percent who suffer heart disease. A much higher percent of people in the high-anger category developed CHD (4.27%) than in the moderate (2.33%) and low (1.70%) anger categories.

9.2 Goodness of Fit

- The chi-square goodness of fit test

The Chi-Square Test for Goodness of Fit

Mars, Inc. makes milk chocolate candies. Here's what the company's Consumer Affairs Department says about the color distribution of its M&M's candies:

On average, the new mix of colors of M&M's milk chocolate candies will contain 13 percent of each of browns and reds, 14 percent yellows, 16 percent greens, 20 percent oranges, and 24 percent blues.

The **one-way table** below summarizes the data from a sample bag of M&M's. In general, one-way tables display the distribution of a categorical variable for the individuals in a sample.

Color	Blue	Orange	Green	Yellow	Red	Brown	**Total**
Count	9	8	12	15	10	6	**60**

The sample proportion of blue M&M'S is $\hat{p} = \frac{9}{60} = 0.15.$

The Chi-Square Test for Goodness of Fit

Since the company claims that 24% of all M&M's are blue, we might believe that something fishy is going on. We could use the *z* test for a proportion to test the hypotheses:

$$H_0: p = 0.24$$
$$H_a: p \neq 0.24$$

where *p* is the true population proportion of blue M&M'S. We could then perform additional significance tests for each of the remaining colors.

However, performing a one-sample *z* test for each proportion would be pretty inefficient and would lead to the problem of *multiple comparisons,* in which we have to adjust for the fact that several tests are done at the same time.

More important, performing one-sample *z* tests for each color wouldn't tell us how likely it is to get a random sample of 60 candies with a color distribution that differs as much from the one claimed by the company as this bag does (taking *all* the colors into consideration at one time).

For that, we need a new kind of significance test, called a **chi-square test for goodness of fit.**

The Chi-Square Test for Goodness of Fit

We can write the hypotheses in symbols as:

$$H_0: p_{blue} = 0.24,\ p_{orange} = 0.20,\ p_{green} = 0.16,$$
$$p_{yellow} = 0.14,\ p_{red} = 0.13,\ p_{brown} = 0.13,$$
$$H_a\text{: At least one of the proportions is different than claimed}$$

where p_{color} = the true population proportion of M&M's of that color.

The idea of the chi-square test for goodness of fit is this: We compare the **observed counts** from our sample with the counts that would be expected if H_0 is true. The more the observed counts differ from the **expected counts,** the more evidence we have against the null hypothesis.

In general, the expected counts can be obtained by multiplying the proportion of the population distribution in each category by the sample size.

The Chi-Square Test for Goodness of Fit

Assuming that the color distribution stated by Mars, Inc. is true, 24% of all M&M's produced are blue.

For random samples of 60 candies, the average number of blue M&M's should be (0.24)(60) = 14.40. This is our expected count of blue M&M's.

Using this same method, we can find the expected counts for the other color categories:

Orange:	**(0.20)(60) = 12.00**
Green:	**(0.16)(60) = 9.60**
Yellow:	**(0.14)(60) = 8.40**
Red:	**(0.13)(60) = 7.80**
Brown:	**(0.13)(60) = 7.80**

Color	Observed	Expected
Blue	9	14.40
Orange	8	12.00
Green	12	9.60
Yellow	15	8.40
Red	10	7.80
Brown	6	7.80

The Chi-Square Test for Goodness of Fit

To calculate the chi-square statistic, use the same formula as you did earlier in the chapter.

$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

Color	Observed	Expected
Blue	9	14.40
Orange	8	12.00
Green	12	9.60
Yellow	15	8.40
Red	10	7.80
Brown	6	7.80

$$\chi^2 = \boxed{\frac{(9 - 14.40)^2}{14.40}} + \boxed{\frac{(8 - 12.00)^2}{12.00}} + \boxed{\frac{(12 - 9.60)^2}{9.60}}$$

$$+ \boxed{\frac{(15 - 8.40)^2}{8.40}} + \boxed{\frac{(10 - 7.80)^2}{7.80}} + \boxed{\frac{(6 - 7.80)^2}{7.80}}$$

$$\begin{aligned}\chi^2 &= 2.025 + 1.333 + 0.600 + 5.186 + 0.621 + 0.415 \\ &= 10.180\end{aligned}$$

The Chi-Square Test for Goodness of Fit

The Chi-Square Test for Goodness of Fit

A categorical variable has k possible outcomes, with probabilities p_1, p_2, p_3, … , p_k. That is, p_i is the probability of the i^{th} outcome. We have n independent observations from this categorical variable.

To test the null hypothesis that the probabilities have specified values

$$H_0\text{: } p_1 = \_\_\_,\ p_2 = \_\_\_,\ \ldots,\ p_k = \_\_\_.$$

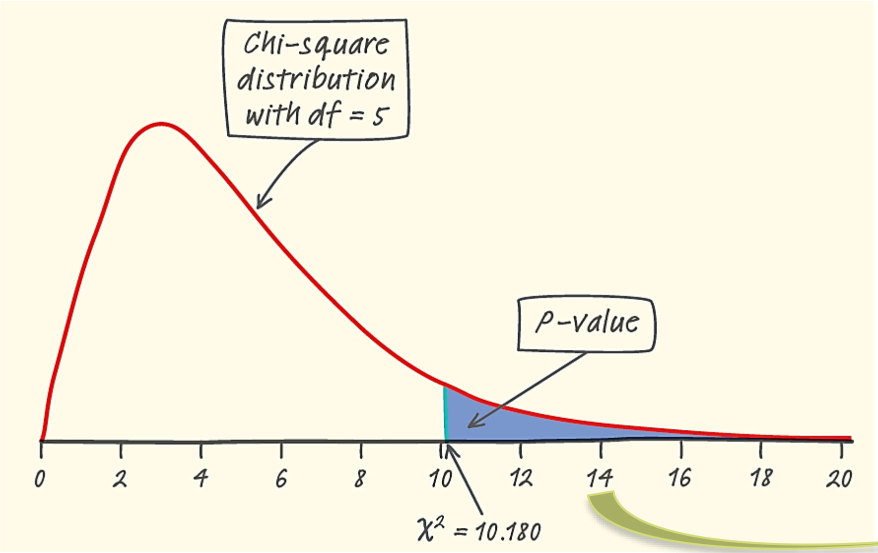
find the *expected count* for each category assuming that H_0 is true. Then calculate the chi-square statistic:

$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

where the sum is over the k different categories. The P-value is the area to the right of χ^2 under the density curve of the chi-square distribution with $k - 1$ degrees of freedom.

The Chi-Square Test for Goodness of Fit

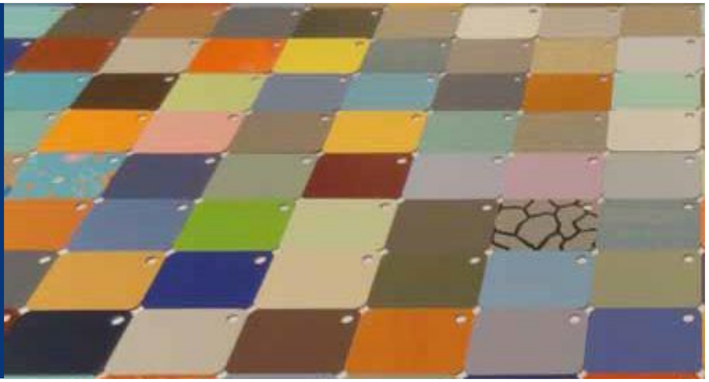
We computed the chi-square statistic for our sample of 60 M&M's to be $\chi^2 = 10.180$. Because all of the expected counts are at least 5, the χ^2 statistic will follow a chi-square distribution with df=6-1=5 reasonably well when H_0 is true.



		P	
df	**.15**	**.10**	**.05**
4	6.74	7.78	9.49
5	8.12	9.24	11.07
6	9.45	10.64	12.59

Since our *P*-value is between 0.05 and 0.10, it is greater than $\alpha = 0.05$. Therefore, we fail to reject H_0. We don't have sufficient evidence to conclude that the company's claimed color distribution is incorrect.

Chapter 9
Analysis of Two-Way Tables



9.1 Inference for Two-Way Tables

9.2 Goodness of Fit