Lecture 11

Statistical Significance for 2 x 2 Tables

Thought Question 1:

Suppose a sample of 400 people included 100 under age thirty and 300 aged thirty and over. Each person asked whether or not they supported requiring public school children to wear uniforms. Fill in the table below, if there is **no relationship** between age and opinion on this question. Explain your reasoning. (*Hint*: Notice that overall, 30% favor uniforms.)

	Yes, favor uniforms	No, don't favor uniforms	Total
Under 30			100
30 and over			300
Total	120	280	400

Thought Question 2:

Suppose in a random sample of 10 males and 10 females, 7 of the males (70%) and 4 of the females (40%) admitted that they have fallen asleep at least once while driving. Would these numbers convince you that there is a difference in the proportions of males and females in the *population* who have fallen asleep while driving?

Now suppose sample consisted of 1000 of each sex, with 700 males and 400 females admitting they had fallen asleep. Would these numbers convince you that there is a difference in the population proportions who have fallen asleep? Explain the difference in the two scenarios.

Thought Question 3:

Based on the data from Example 1 in Chapter 12, we can conclude that there is a statistically significant relationship between taking aspirin or not and having a heart attack or not.

What do you think it means to say that the relationship is "***statistically significant***"?

Thought Question 4:

Refer to the previous question.

Do you think that a statistically significant relationship is the same thing as an important and sizeable relationship? Explain.

13.1 Measuring the Strength of the Relationship

The Meaning of Statistical Significance

If a relationship between two categorical variables is ***statistically significant*** it means that the relationship observed in the sample was ***unlikely*** to have occurred unless there really is a relationship in the population.

However, even if the relationship in the population is real, it may be so small as to be of little practical importance. Statistical significance doesn't necessarily mean a relationship is meaningful!

Measuring the Relationship in a 2 x 2 Contingency Table

Strength of the relationship measured by ***difference* in the percentages of outcomes for the two categories of the explanatory variable**. *Sample size* is important too.

Example 1: Aspirin and Heart Attacks

Aspirin Group: Percentage who had heart attacks = 0.94%

Placebo Group: Percentage who had heart attacks = 1.71%

Difference: only 1.71% – 0.94% = 0.77%

Are we convinced by the data that there is a ***real*** relationship in the population between taking aspirin and risk of heart attack? Need to assess if the relationship is statistically significant.

Measuring the Relationship in a 2 x 2 Contingency Table

Example 2: Young Drivers, Gender and Alcohol

Males: Percentage drinking in past 2 hours = 16%

Females: Percentage drinking in past 2 hours = 11.6%

Difference: 16% – 11.6% = 4.4%

If population percents were equal, how likely would we be to observed a sample with a difference as large as 4.4% or larger?

Example 3: Ease of Pregnancy and Smoking

Nonsmokers: Percentage pregnant in 1st cycle = 41%

Smokers: Percentage pregnant in 1st cycle = 29%

Difference: 41% – 29% = 12%

Is a difference of 12% large enough to rule out chance?

Strength of the Relationship versus Size of the Study

Example 1: Aspirin and Heart Attacks

Difference: only 1.71% – 0.94% = 0.77%

Experiment included ***over 22,000 men***, so small difference should convince us aspirin works for represented population.

Example 2: Young Drivers, Gender and Alcohol

Difference: 16% – 11.6% = 4.4%

Rather small, based on ***only 619 respondents*** and was not convincing to the Supreme Court.

Example 3: Ease of Pregnancy and Smoking

Difference: 41% – 29% = 12%

Larger, but again ***only based on 586 subjects***. Convincing?

13.2 Steps for Assessing Statistical Significance

Basic steps for *hypothesis testing*:

1. Determine the null hypothesis and the alternative hypothesis.
2. Collect the data and summarize them with a single number called a 'test statistic'.
3. Determine how unlikely the test statistic would be if the null hypothesis were true.
4. Make a decision.

Step 1: Determine the null hypothesis and the alternative hypothesis.

Alternative Hypothesis: what researchers interested in showing to be true (*aka* 'research hypothesis').

Null Hypothesis: some form of '***nothing happening***'

Q: Are two categorical variables related? …

Null Hypothesis: There is no relationship between the two variables in the population.

Alternative Hypothesis: There is a relationship between the two variables in the population.

Note: *Hypotheses established before collecting any data.*
Not acceptable to use same data to determine and test hypotheses.

Stating The Hypotheses

Example 1: Aspirin and Heart Attacks

Null Hypothesis: There is ***no relationship*** between taking aspirin and risk of heart attack in the population.

Alternative Hypothesis: There ***is a relationship*** between taking aspirin and risk of heart attack in the population.

Example 2: Young Drivers, Gender and Alcohol

Null Hypothesis: Males and females in the population ***are equally likely*** to drive within two hours of drinking alcohol.

Alternative Hypothesis: Males and females in the population ***are not equally likely*** to drive within two hours of drinking alcohol.

Example 3: Ease of Pregnancy and Smoking

Null Hypothesis: Smokers and nonsmokers ***are equally likely*** to get pregnant in 1^{st} cycle in population of women trying to get pregnant.

Alternative Hypothesis: Smokers and nonsmokers ***are not equally likely*** to get pregnant in 1^{st} cycle in population of women trying to get pregnant.

13.3 The Chi-Square Test

Step 2: Collect data and summarize with a 'test statistic'.

Chi-square statistic: compares data in *sample* to what would be expected if no relationship between variables in the population*population*.

Step 3: Determine how unlikely test statistic would be if the null hypothesis were true.

***p*-value:** probability of observing a test statistic as extreme as the one observed or more so, if the null hypothesis is really true. (For chi-square: more extreme = larger value of chi-square statistic.)

Step 4: Make a decision.

If ***chi-square statistic is at least 3.84***, the ***p-value is 0.05 or less***, so conclude relationship in population is real. That is, we reject the null hypothesis and conclude the relationship is ***statistically significant***.

Computing the Chi-Square Statistic

1. **Compute the expected number in each cell, assuming the null hypothesis is true.**
2. **Compare the observed and expected numbers.**
3. **Compute the chi-squared statistic.**

Expected number = $\dfrac{\text{(row total)(column total)}}{\text{(table total)}}$

Compare observed and expected in each cell by:

$$\frac{(\text{observed number} - \text{expected number})^2}{\text{expected number}}$$

Chi-square statistic = sum of comparison values over all cells.

Note: Method valid only if no empty cells and all expected counts ≥ 5.

Example 3: Ease of Pregnancy and Smoking

	Pregnancy Occurred After			
	First Cycle	**Two or More Cycles**	**Total**	**Percentage in First**
Smoker	29	71	100	29%
Nonsmoker	198	288	486	41%
Total	227	359	586	38.7%

1. **Compute the expected numbers.**

   Expected number of smokers pregnant after 1st cycle:

   (100)(227)/586 = 38.74

   Can find the remaining expected numbers by subtraction.

Expected Numbers:

	Pregnancy Occurred After		
	First Cycle	**Two or More Cycles**	**Total**
Smoker	38.74	100 – 38.74 = 61.26	100
Nonsmoker	227 – 38.74 = 188.26	486 – 188.26 = 297.74	486
Total	227	359	586

Example 3: Ease of Pregnancy and Smoking

Here are the Observed (and Expected) numbers:

	Pregnancy Occurred After		
	First Cycle	**Two or More Cycles**	**Total**
Smoker	29 (38.74)	71 (61.26)	100
Nonsmoker	198 (188.26)	288 (297.74)	486
Total	227	359	586

2. Compare Observed and Expected numbers.

(observed number – expected number)^2/(expected number)

First cell: $(29 - 38.74)^2/(38.74) = 2.45$

Remaining cells shown in table below.

	Pregnancy Occurred After	
	First Cycle	**Two or More Cycles**
Smoker	2.45	1.55
Nonsmoker	0.50	0.32

3. Compute the chi-squared statistic.

chi-square statistic = 2.45 + 1.55 + 0.50 + 0.32 = **4.82**

Making the Decision

Relationship ***statistically significant*** if the chi-square statistic is at least 3.84.

Example 3: Ease of Pregnancy and Smoking

The chi-square statistic of 4.82 is larger than 3.84.

- There is a ***statistically significant relationship*** between smoking and time to pregnancy.
- The difference observed in time to pregnancy between smokers and nonsmokers in the sample indicates a ***real difference for the population*** of all similar women.

Computers, Calculators and Chi-Square Tests

Minitab Results for Example 3: Ease of Pregnancy and Smoking

```
Expected counts are printed below observed counts

              C1              C2              Total
1             29              71               100
              38.74           61.26

2            198             288               486
             188.26          297.74

Total        227             359               586

ChiSq = 2.448 + 1.548 +
      0.504 + 0.318 = 4.817
```

DF = 1, P-Value = 0.028

Example 4: Age at Birth of First Child and Breast Cancer

First Child at Age 25 or older?	Breast Cancer	No Breast Cancer	Total
Yes	31	1597	1628
No	65	4475	4540
Total	96	6072	6168

Step 1: Determine null and alternative hypotheses.

Null Hypothesis: There is ***no relationship*** between age at birth of first child and breast cancer in the population of women who have had children.

Alternative Hypothesis: There is a ***relationship*** between age at birth of first child and breast cancer in the population of women who have had children.

Example 4: Age at Birth of First Child and Breast Cancer

Step 2: Collect data and compute test statistic.

Expected count for "Yes and Breast Cancer":

(1628)(96)/6168 = 25.34

Can find the remaining expected numbers by subtraction.

Here are the Observed (and Expected) numbers:

First Child at Age 25 or older?	Breast Cancer	No Breast Cancer	Total
Yes	31 (25.84)	1597 (1602.66)	1628
No	65 (70.66)	4475 (4469.34)	4540
Total	96	6072	6168

Chi-square statistic = 1.75

Example 4: Age at Birth of First Child and Breast Cancer

Step 3: Determine how unlikely the test statistic would be if null hypothesis were true.

Using Excel: *p*-value = CHIDIST(1.75,1) = 0.186

Step 4: Make a decision.

- The chi-square statistic of 1.75 *is less than* 3.84 (and *p*-value of 0.186 *is greater than* 0.05).
- Therefore, the relationship is ***not statistically significant***.
- We ***cannot*** conclude that the increased risk observed in the *sample* would hold for the *population* of women.

Note: Possible *confounding variable* is use of oral contraceptives at a young age, that may be related to age at birth of first child and may have effect on likelihood of breast cancer.

13.4 Practical Versus Statistical Significance

Statistical significance *does not mean* two variables have a relationship that is of **practical importance**.

- Table based on very large number of observations will have little trouble achieving statistical significance, even if relationship between two variables is only minor.
- An interesting relationship in a population may fail to achieve statistical significance in sample if there are too few observations.

It is difficult to rule out chance unless you have either a strong relationship or a sufficiently large sample.

Example 2, continued: Drinking and Driving

Case Study 6.5: Court case challenging law that differentiated the ages at which young men and women could buy 3.2% beer.

Results of Roadside Survey for Young Drivers

	Drank Alcohol in Last 2 Hours?			
	Yes	No	Total	Percentage Who Drank
Males	77	404	481	16.0%
Females	16	122	138	11.6%
Total	93	526	619	15.0%

Difference: 16% – 11.6% = 4.4%

Might think there is a real relationship in the population, but chi-square statistic = 1.637 => a difference of 4.4% in a sample of ***only 619*** would not be surprising if equal proportions in the population.

If sample were three times larger but same difference, chi-square statistic would be 3(1.637) = 5.01=> it ***would*** be ***statistically significant***.

Case Study 13.1: ESP and Movies

Study Details:

- Subjects (“receivers”) asked to select (from four choices), what another person (“sender”) watched on TV in another room (the “target”).
- Actual target randomly selected from the four choices => if no ESP, guesses successful by chance 25% of the time.
- Surprisingly, were successful 34% of the time.

Researchers **hypothesized** videos (dynamic pictures) would be received with better success than photographs (static pictures).

Case Study 13.1: ESP and Movies

Results of ESP Study

	Successful ESP Guess?			
	Yes	No	Total	% Success
Static picture	45	119	164	27
Dynamic picture	77	113	190	41
Total	122	232	354	34

- Chi-square statistic = 6.675 is larger than 3.84.
- Results are ***statistically significant***.
- It does appear that success in ESP guessing depends on the type of picture used as the target.

Source: Bem and Honorton, 1994.

For Those Who Like Formulas

To represent the *observed numbers* in a 2 × 2 contingency table, we use the notation:

	Variable 2		
Variable 1	Yes	No	Total
Yes	a	b	$a + b$
No	c	d	$c + d$
Total	$a + c$	$b + d$	n

Therefore, the *expected* numbers are computed as follows:

	Variable 2		
Variable 1	Yes	No	Total
Yes	$(a + b)(a + c)/n$	$(a + b)(b + d)/n$	$a + b$
No	$(c + d)(a + c)/n$	$(c + d)(b + d)/n$	$c + d$
Total	$a + c$	$b + d$	n

Computing the Chi-Squared Statistic, χ^2, for an $r \times c$ Contingency Table

Let O_i = observed number in cell *i*, and E_i = expected number in cell *i*. Then:

$$\chi^2 = \sum_{i=1}^{rc} \frac{(O_i - E_i)^2}{E_i}$$

χ^2 Test Basic Idea

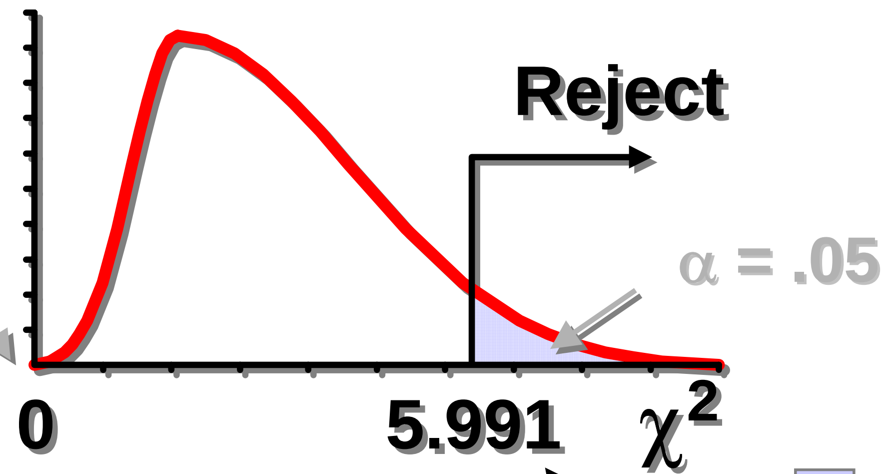
- 1. Compares Observed Count to Expected Count If Null Hypothesis Is True
- 2. Closer Observed Count to Expected Count, the More Likely the H_0 Is True
  - – Measured by Squared Difference Relative to Expected Count
    - Reject Large Values

Finding Critical Value Example

What is the critical χ^2 value if k = 3, & α =.05?

If $n_i = E(n_i)$, $\chi^2 = 0$.
Do not reject H_0

df = k - 1 = 2



χ^2 **Table (Portion)**

	Upper Tail Area				
DF	**.995**	**...**	**.95**	**...**	**.05**
1	...	...	0.004	...	3.841
2	0.010	...	0.103	...	5.991

Expected Count Example

House Style	Location Urban Obs.	 Rural Obs.	Total
Split-Level	63	49	112
Ranch	15	33	48
Total	78	82	160

Expected Count Example

Marginal probability = $\frac{112}{160}$

	Location		
	Urban	Rural	
House Style	Obs.	Obs.	Total
Split-Level	63	49	112
Ranch	15	33	48
Total	78	82	160

Expected Count Example

Marginal probability = $\frac{112}{160}$

House Style	Location Urban Obs.	 Rural Obs.	Total
Split-Level	63	49	112
Ranch	15	33	48
Total	78	82	160

Marginal probability = $\frac{78}{160}$

Expected Count Example

Joint probability = $\frac{112}{160}\frac{78}{160}$

Marginal probability = $\frac{112}{160}$

	Location		
House Style	**Urban Obs.**	**Rural Obs.**	**Total**
Split-Level	63	49	112
Ranch	15	33	48
Total	78	82	160

Marginal probability = $\frac{78}{160}$

Expected Count Example

Joint probability = $\frac{112}{160} \cdot \frac{78}{160}$

Marginal probability = $\frac{112}{160}$

House Style	Location: Urban Obs.	Location: Rural Obs.	Total
Split-Level	63	49	112
Ranch	15	33	48
Total	78	82	160

Marginal probability = $\frac{78}{160}$

Expected count = $160 \cdot \frac{112}{160} \cdot \frac{78}{160}$

= 54.6

Expected Count Calculation

Expected Count Calculation

$$\text{Expected count} = \frac{(\text{Row total}) \cdot (\text{Column total})}{\text{Sample size}}$$

Expected Count Calculation

$$\textbf{Expected count} = \frac{(\text{Row total}) \cdot (\text{Column total})}{\text{Sample size}}$$

$\frac{112 \cdot 78}{160}$ $\frac{112 \cdot 82}{160}$

	House Location				
	Urban		Rural		
House Style	Obs.	Exp.	Obs.	Exp.	Total
Split-Level	63	54.6	49	57.4	112
Ranch	15	23.4	33	24.6	48
Total	78	78	82	82	160

$\frac{48 \cdot 78}{160}$ $\frac{48 \cdot 82}{160}$

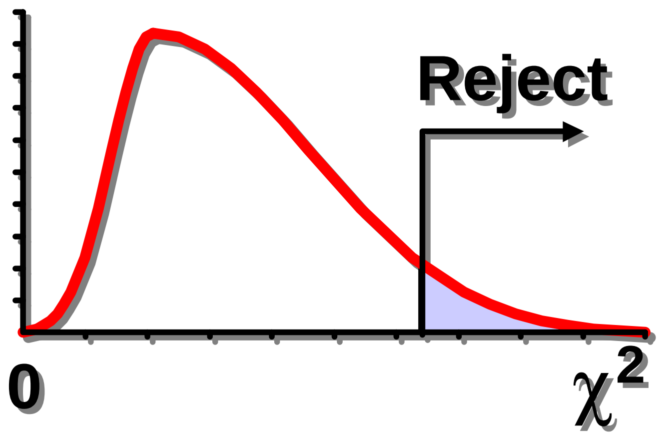
χ^2 Test of Independence Example

- You're a marketing research analyst. You ask a random sample of **286** consumers if they purchase Diet Pepsi or Diet Coke. At the **.05** level, is there evidence of a **relationship**?

	Diet Pepsi		
Diet Coke	**No**	**Yes**	**Total**
No	**84**	**32**	**116**
Yes	**48**	**122**	**170**
Total	**132**	**154**	**286**

χ^2 Test of Independence Solution

- **H_0:**
- **H_a:**
- α **=**
- **df =**
- **Critical Value(s):**



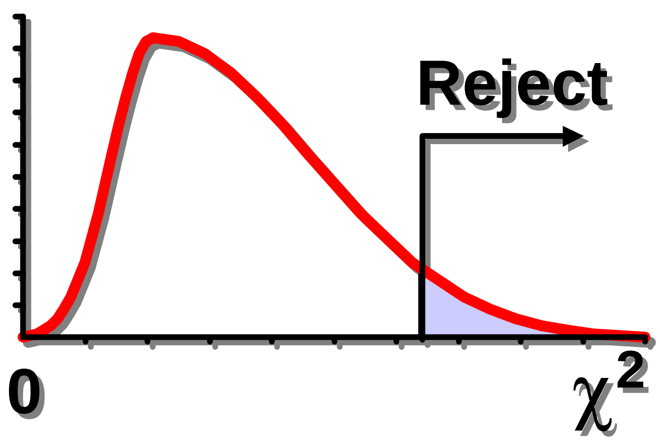
Test Statistic:

Decision:

Conclusion:

χ^2 Test of Independence Solution

- H_0: No Relationship
- H_a: Relationship
- α =
- df =
- Critical Value(s):



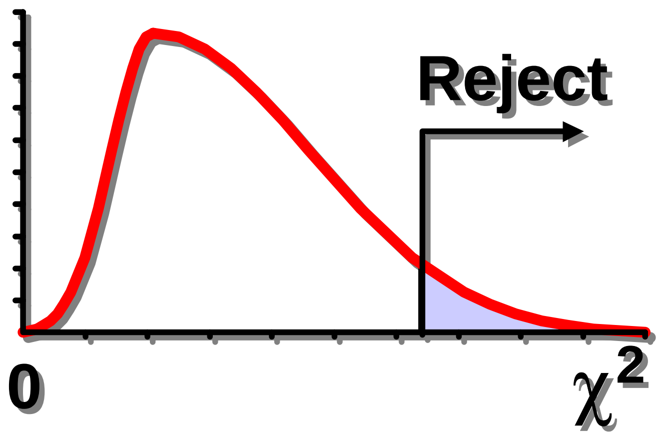
Test Statistic:

Decision:

Conclusion:

χ^2 Test of Independence Solution

- **H_0: No Relationship**
- **H_a: Relationship**
- α **= .05**
- **df = (2 - 1)(2 - 1) = 1**
- **Critical Value(s):**



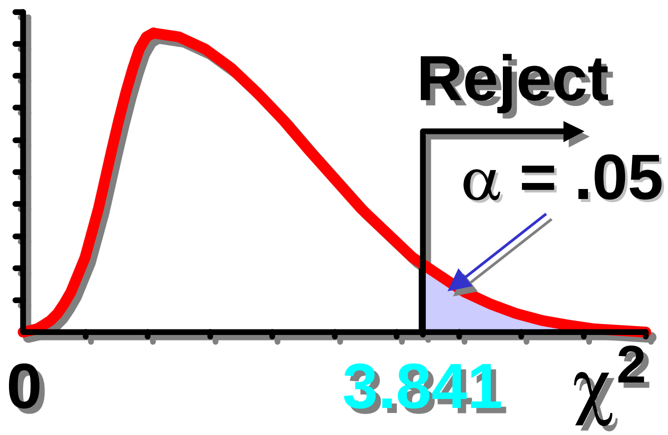
Test Statistic:

Decision:

Conclusion:

χ^2 Test of Independence Solution

- H_0: No Relationship
- H_a: Relationship
- $\alpha = .05$
- df = (2 - 1)(2 - 1) = 1
- Critical Value(s):



Test Statistic:

Decision:

Conclusion:

χ^2 Test of Independence Solution

- $E(n_{ij}) \geq 5$ in all cells

$$\frac{116 \cdot 132}{286} \qquad \frac{154 \cdot 132}{286}$$

	Diet Pepsi				
	No		Yes		
Diet Coke	Obs.	Exp.	Obs.	Exp.	Total
No	84	53.5	32	62.5	116
Yes	48	78.5	122	91.5	170
Total	132	132	154	154	286

$$\frac{170 \cdot 132}{286} \qquad \frac{170 \cdot 154}{286}$$

χ^2 Test of Independence Solution

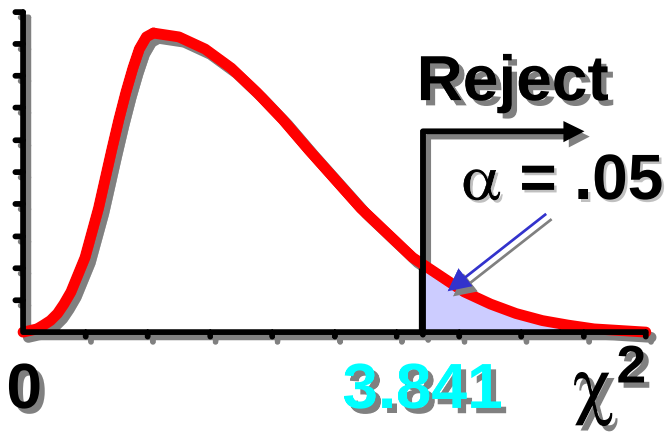
$$\chi^2 = \sum_{\text{all cells}} \frac{\left[n_{ij} - E(n_{ij})\right]^2}{E(n_{ij})}$$

$$= \frac{\left[n_{11} - E(n_{11})\right]^2}{E(n_{11})} + \frac{\left[n_{12} - E(n_{12})\right]^2}{E(n_{12})} + \cdots + \frac{\left[n_{22} - E(n_{22})\right]^2}{E(n_{22})}$$

$$= \frac{\left[84 - 53.5\right]^2}{53.5} + \frac{\left[32 - 62.5\right]^2}{62.5} + \cdots + \frac{\left[122 - 91.5\right]^2}{91.5} = 54.29$$

χ^2 Test of Independence Solution

- **H_0: No Relationship**
- **H_a: Relationship**
- **α = .05**
- **df = (2 - 1)(2 - 1) = 1**
- **Critical Value(s):**



Test Statistic:

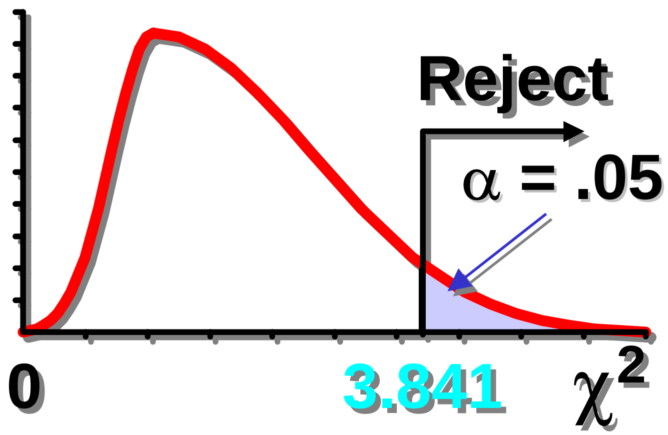
$\chi^2 = 54.29$

Decision:

Conclusion:

χ^2 Test of Independence Solution

- H_0: No Relationship
- H_a: Relationship
- $\alpha = .05$
- df = (2 - 1)(2 - 1) = 1
- Critical Value(s):



Test Statistic:

$$\chi^2 = 54.29$$

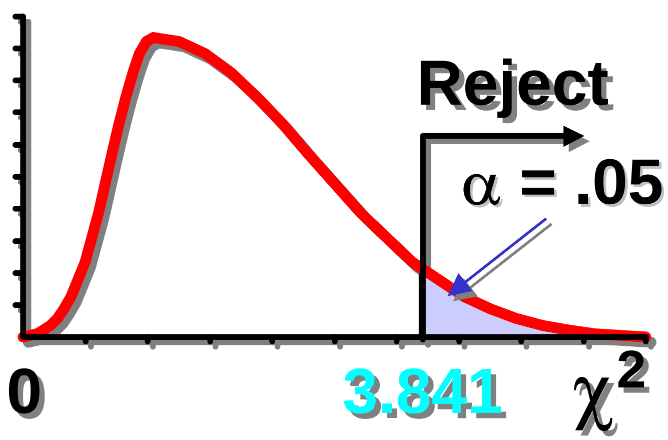
Decision:

Reject at $\alpha = .05$

Conclusion:

χ^2 Test of Independence Solution

- **H_0: No Relationship**
- **H_a: Relationship**
- **α = .05**
- **df = (2 - 1)(2 - 1) = 1**
- **Critical Value(s):**



Test Statistic:

$$\chi^2 = 54.29$$

Decision:

Reject at α = .05

Conclusion:

There is evidence of a relationship