In this module, you will learn:

- How the empirical relationship between two variables is affected by random sampling error
- How to use an informal test in making inferences about relationships
- How to use formal statistical tests in making inferences about relationships
- How measures of association gauge the strength of an empirical relationship
- Which measure of association to use in performing and interpreting political analysis

Tests of significance and measures of association: What’s the difference?

- A **test of statistical significance** helps you decide whether an observed relationship between an independent variable and a dependent variable really exists in the population or whether it could have happened by chance when the sample was drawn.
- A **measure of association** tells the researcher how well the independent variable works in explaining the dependent variable.

Statistical Significance

- Suppose we calculate the mean Democratic rating among a female sample and a male sample
  - Female mean = 54
  - Male mean = 49
- Can we infer that, in the population, females do, in fact, rate the Democratic Party higher than do males?

Null hypothesis (H_0)

- No relationship exists in the population
  - Females and males give identical ratings to the Democratic Party. No gender difference exists.
- Any apparent relationship that we observe in a sample was produced by random sampling error
  - The 5-point gender difference is accounted for by random chance when the sample was drawn.

H_A

- "Our" hypothesis, which hypothesizes that women do, in fact, give the Democratic Party higher ratings, is a called the *alternative hypothesis* and labeled H_A.
- In order to have confidence in H_A, we must show that H_0's version of events—it all happened by chance—is sufficiently implausible.

Type I and Type II Error

- **Type I error** occurs when the researcher infers that there is a relationship in the population when, in fact, there is none.
- **Type II error** occurs when the researcher infers that there is no relationship in the population when, in fact, there is.

Table 7-1 Type I and Type II Error

	In the (unseen) population, is H_0 true or is H_0 false?	
Inferential decision, based on sample data	*H_0 is true*	*H_0 is false*
Do not reject H_0	Correct inference	Type II error
Reject H_0	Type I error	Correct inference

The .05 level of significance

- We assume the null hypothesis is true, and we set a high threshold for rejecting it.
- Ask this question: "If, in the *unobserved* population, H_0 is true, how often by chance will we obtain the relationship *observed* in the sample?"
- If the answer is, "more than 5 times out of 100," do not reject H_0.
- If the answer is, "5 times out of 100 or less," reject H_0.

Comparing two sample means: The confidence interval approach

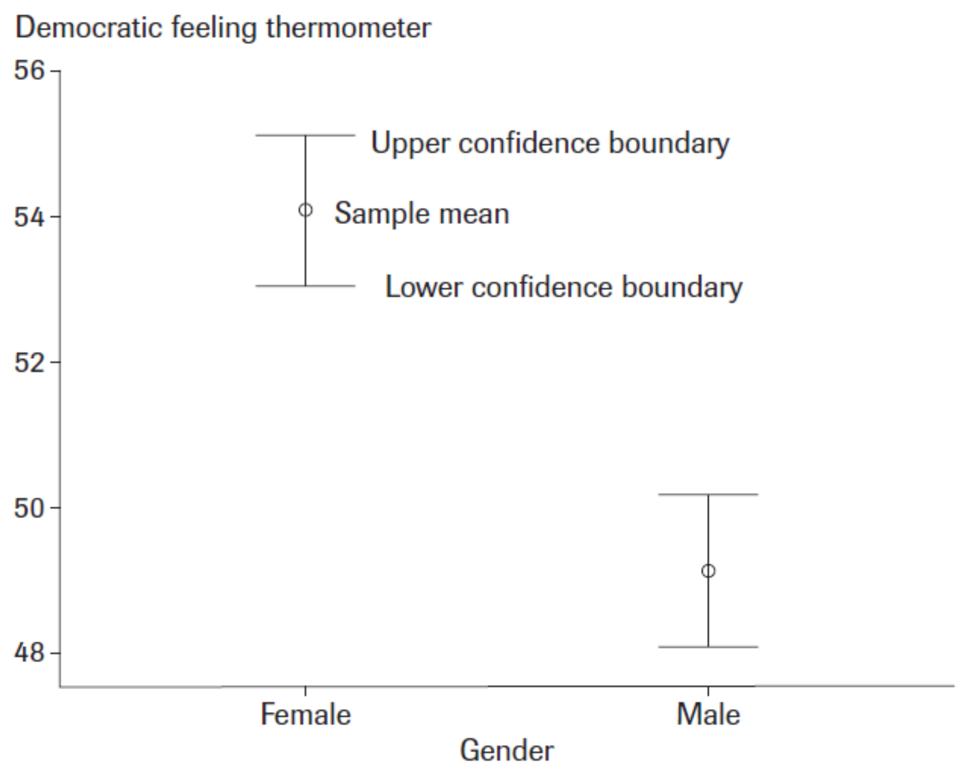
- Compare the 95% confidence intervals of two (or more) sample means.
- If the intervals overlap, then we cannot reject the null hypothesis.
- If they do not overlap, then we can reject the null hypothesis.

Example: Female and Male sample means

- Female mean = 54.1
  - Standard error = .52
  - Lower 95% CI boundary = 53.1
  - Upper 95% CI boundary = 55.1
- Male mean = 49.1
  - Standard error = .53
  - Lower 95% CI boundary = 48.1
  - Upper 95% CI boundary = 50.1

95% CIs do not overlap. Reject H_0

Figure 7-1 Error Bar Graph



Comparing two sample means: The P-value approach

- Determine the exact probability of obtaining the observed sample difference, under the assumption that the null hypothesis is correct.
- The *P*-value approach is based on the **standard error of the difference** between two sample means.

Standard error of a mean difference: 3 steps

- 1. Square each mean’s standard error. The square of the female mean’s standard error is equal to .27. The square of the male mean’s standard error is equal to . 28.
- 2. Sum the squared standard errors. In the example: .27 + .28 = .55.
- 3. Take the square root of the sum obtained in step 2: .55 = **.74**.

Obtain a test statistic

- $Z = (HA - H0)$ / Standard error of the difference
- $t = (HA - H0)$ / Standard error of the difference
- $Z = (5.0 - 0) / .74 =$ **6.8**
- $t = (5.0 - 0) / .74 =$ **6.8**

The inferential question

- "If, in the population, the mean difference between females and males is equal to 0, how often by chance will we obtain a mean that is 6.8 standard errors from 0?"
- Using normal estimation: about 400 times out of a billion.
- Reject the null hypothesis.

Chi-square Test of Significance

- The chi-square test is perhaps the most commonly used statistical test in cross-tabulation analysis.
- The chi-square test determines whether the observed dispersal of cases departs significantly from what we would expect to find if the null hypothesis were correct.

Table 7-6 Opinions on Diplomacy versus Military Force, by Gender

	Gender		
Favor diplomacy or force?	*Female*	*Male*	Total
Diplomacy	46.2%	34.4%	40.4%
	(245)	(176)	(421)
Middle	23.6%	27.8%	25.6%
	(125)	(142)	(267)
Force	30.2%	37.8%	33.9%
	(160)	(193)	(353)
Total	100.0%	100.0%	99.9%[a]
	(530)	(511)	(1,041)

Chi-square logic

- H_0 is based on the *expectation* that the distribution of cases down the Female column, and the distribution down the Male column, will be the same as the distribution down the Total column.
- So: 40.4% of Females and 40.4% of Males will say "Diplomacy," 25.6% will take the "Middle" position, and 33.9% will say "Force."
- Any *observed* departures from this *expected* pattern are accounted for by random sampling error.

Table 7-7 Chi-square for Opinions on Diplomacy versus Military Force, by Gender

Favor diplomacy or force?		Gender: *Female*	*Male*
Diplomacy	Observed frequency (f_o)	245	176
	Expected frequency (f_e)	214.3	206.7
	$f_o - f_e$	30.7	–30.7
	$(f_o - f_e)^2$	942.5	942.5
	$(f_o - f_e)^2 / f_e$	**4.4**	**4.6**
Middle	Observed frequency (f_o)	125	142
	Expected frequency (f_e)	135.9	131.1
	$f_o - f_e$	–10.9	10.9
	$(f_o - f_e)^2$	118.8	118.8
	$(f_o - f_e)^2 / f_e$	**.9**	**.9**
Force	Observed frequency (f_o)	160	193
	Expected frequency (f_e)	179.7	173.3
	$f_o - f_e$	–19.7	19.7
	$(f_o - f_e)^2$	388.1	388.1
	$(f_o - f_e)^2 / f_e$	**2.2**	**2.2**

Source: 2004 American National Election Study.

Chi-square calculations: $\chi^2 = \sum(fo - fe)^2 / fe$

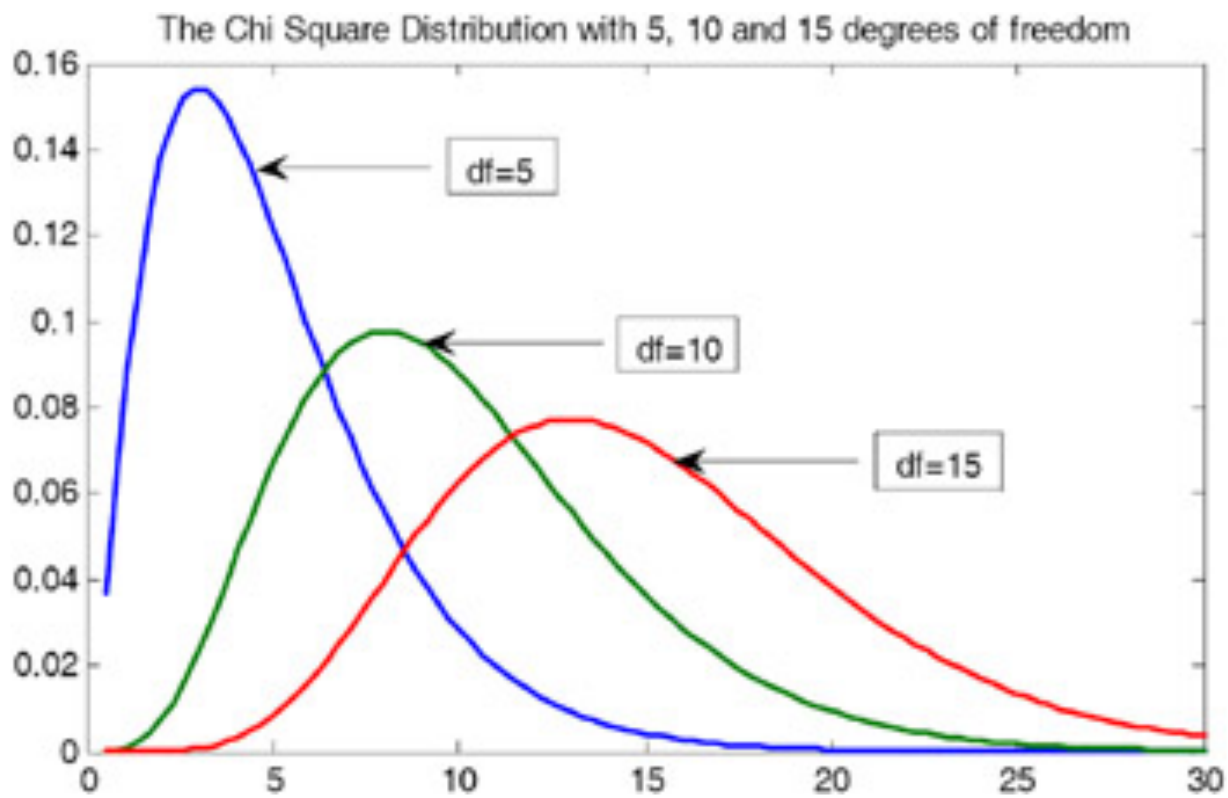
- 1. Find the expected frequency for each cell. Example: Since .339 of the total sample fall into the "force" category of the dependent variable, the expected frequency (*fe*) for the "female–force" cell would be .339 times 530, or about 180.9.
- 2. For each cell, subtract the expected frequency from the observed frequency. Example: *fo* for the "female–force" cell is equal to 160 and *fe* (from step 1) is equal to 180. Subtracting 180 from 160 yields −20.
- 3. Square this number for each cell. Example: Squaring −20 yields 400.
- 4. Divide this number by the expected frequency (*fe*). Example: 400 divided by 180 is equal to 2.2.
- 5. To arrive at the chi-square test statistic, add up all the cell-by-cell calculations.

Determining statistical significance: logic

- For the example, chi-square = 15.2
- The null hypothesis says it should have been 0, give or take random sampling error.
- How do we figure out the amount of random sampling error?

Degrees of freedom

- The size of the chi-square statistic that is needed to “beat” the null hypothesis depends on degrees of freedom.
- Degrees of freedom = (number of rows − 1)(number of columns − 1)
  - Example: (3-1)(2-1) = 2 degrees of freedom
- The number of degrees of freedom determines the shape of the distribution:


The Chi Square Distribution with 5, 10 and 15 degrees of freedom
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
5
10
15
20
25
30
df=5
df=10
df=15

Critical values

- A critical value marks the upper plausible boundary of random error and so defines H_0's limit.
- Example: Using .05 level and 2 d.f., critical value = 5.991.
- The calculated value of chi-square, 15.2, exceeds the critical value, so reject H_0.

H_0 owns the area between 0 and the critical value….

Table 7-8 Critical Values of χ^2

	Area to the right of critical value			
Degrees of freedom	*.10*	*.05*	*.025*	*.01*
1	2.706	3.841	5.024	6.635
2	4.605	5.991	7.378	9.210
3	6.251	7.815	9.348	11.345
4	7.779	9.488	11.143	13.277
5	9.236	11.070	12.833	15.086
6	10.645	12.592	14.449	16.812