Day 3, Morning: Bivariate Tests of Statistical Inference

Instructor: Marshall A. Taylor
844 Flanner Hall
mtaylo15@nd.edu
marshalltaylor.net

Recap

- Yesterday we talked about the asymptotic theory of probability distributions, in particular the Central Limit Theorem (CLT), normal distributions, standard errors, and confidence intervals around population means.

- We then used our knowledge of the CLT to talk about how we can use what we know about our sample statistics to draw inferences about population parameters. We focused on univariate tests of statistical inference, in particular the one sample t-test (for means), the one sample differences-of-proportions test (for proportions), and the one sample sign test (for medians).

Recap

- Before moving forward with bivariate tests, let’s review how basic statistical inference works:
  - (1) We usually don’t know how the population is distributed along a certain variable. However, thanks to the CLT, we do know that if we take an infinite number of random samples of the variable (and of sufficient size) from the population, the resulting sampling distribution is approximately normal.

Recap

- (2) But we cannot possible take an infinite number of random samples. We generally only have one. However, we can quantify the standard deviation of the theoretical sampling distribution (standard error) if we assuming it is normal. The bigger the sample size and the smaller the statistic standard deviation, the more reliable (smaller) the standard error.
- (3) Assuming we have a single sample statistic and that it is a mean, we can test whether or not this number is significantly different from another hypothesized value (perhaps the known population mean) by subtracting the latter from the former and then dividing by the standard error for that mean. This gives us a t-statistic, which tells us how many standard deviations the observed difference between these two numbers is if we assume that the real difference is 0.

Recap

- (4) We also know, per the CLT, that approximately 95% of sample estimates from the sampling distribution fall within ±1.96 standard deviations of the population mean, 99% within ±2.58, and so on. These serve as our critical values.*
- (5) We then compare our test statistic to the critical value that corresponds to our desired significance level. If the absolute value of our test statistic is greater than the absolute value of the critical value, then we reject our null hypothesis that there is no difference and instead find support that there is a difference—and that the difference is not due to random chance. We can then generalize to the population.

*Technically, the t-distribution that we use to get a t-statistic is slightly different from the normal distribution, but the difference gets smaller with larger sample sizes.

Recap

- For example, say we were only willing to be wrong 95 out of 100 times. This is equivalent to saying that we have a .05 significance level (α = .05). Our test statistic is 2.33, which is larger than 1.96. Assuming a two-tailed test, we can that there is less than a 5% chance that we would observe the difference between our sample mean and the hypothesized mean if the real difference was, in fact, 0.

Any questions on that?

On to bivariate tests

- Luckily, moving up to **bivariate tests of statistical inference** is quite easy.

- Some slight mathematical adjustments notwithstanding, the difference between univariate and bivariate inferential tests basically amounts to just substituting the hypothesized values from the univariate tests with a second observed sample value!

On to bivariate tests

- We’re going to talk about a sampling of 6 tests today:
  - Independent samples *t*-test
  - Paired samples *t*-test
  - Independent samples difference-of-proportions test
  - McNemar test
  - Chi-square test
  - Gamma/ASE test

Independent Samples *t*-test

- What if we want to know whether or not two groups differ on some continuous trait in the population?

- We can use an independent samples *t*-test to answer these questions.

- By **independent samples**, we mean (1) that there are **two** different groups, and (2) they are **mutually exclusive** of one another (no co-membership).

Independent Samples *t*-test

- The formula is very similar to its one sample counterpart:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- Now we are just focusing on the difference in sample means between two groups (numerator) and taking into account the standard errors for both groups (denominator).

An Example: Independent Samples *t*-test

A manager is interested in testing whether or not lunch break lengths vary between their accounting and marketing departments. S/he takes a random sample from each department (n = 25 for both) and asks each respondent to indicate the number of minutes they spend at lunch during the typical work day. The accountants report an average of 55 minutes with a standard deviation of 5 minutes; the marketers report an average of 64 minutes with a standard deviation of 3 minutes. Do we have enough information to conclude that the accounting department tends to take shorter lunch breaks when α = .05?

An Example: Independent Samples *t*-test

- We first state our null hypothesis:

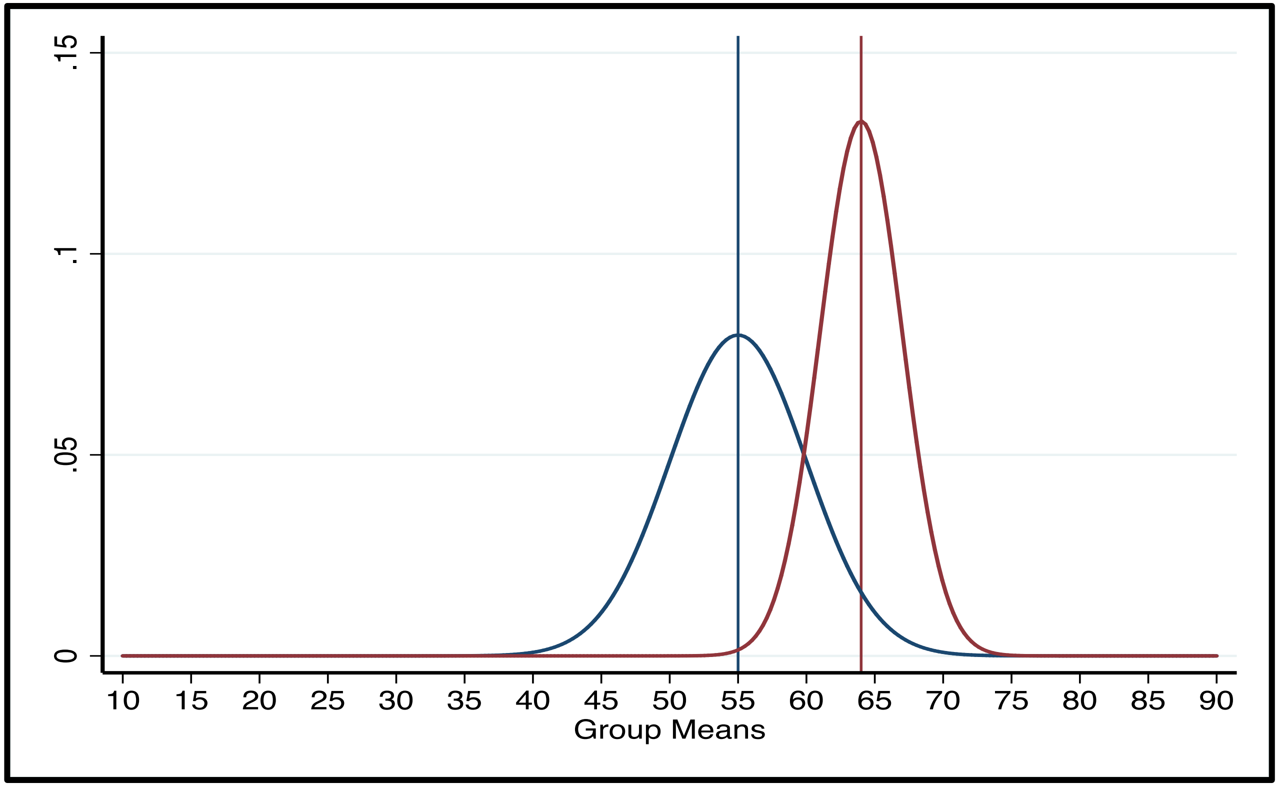
$$\mathrm{H}_0 : \mu_1 = \mu_2$$

- And our alternative hypothesis, which is one-tailed:

$$\mathrm{H}_\mathrm{A} : \mu_1 < \mu_2$$

An Example: Independent Samples *t*-test

- We want to examine how likely it would be to observe the difference between these two lines if the population difference is 0. Are these statistics due to random chance or not?



An Example: Independent Samples *t*-test

- Let’s plug in the numbers:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{55 - 64}{\sqrt{\frac{25}{25} + \frac{9}{25}}} = -7.717$$

An Example: Independent Samples *t*-test

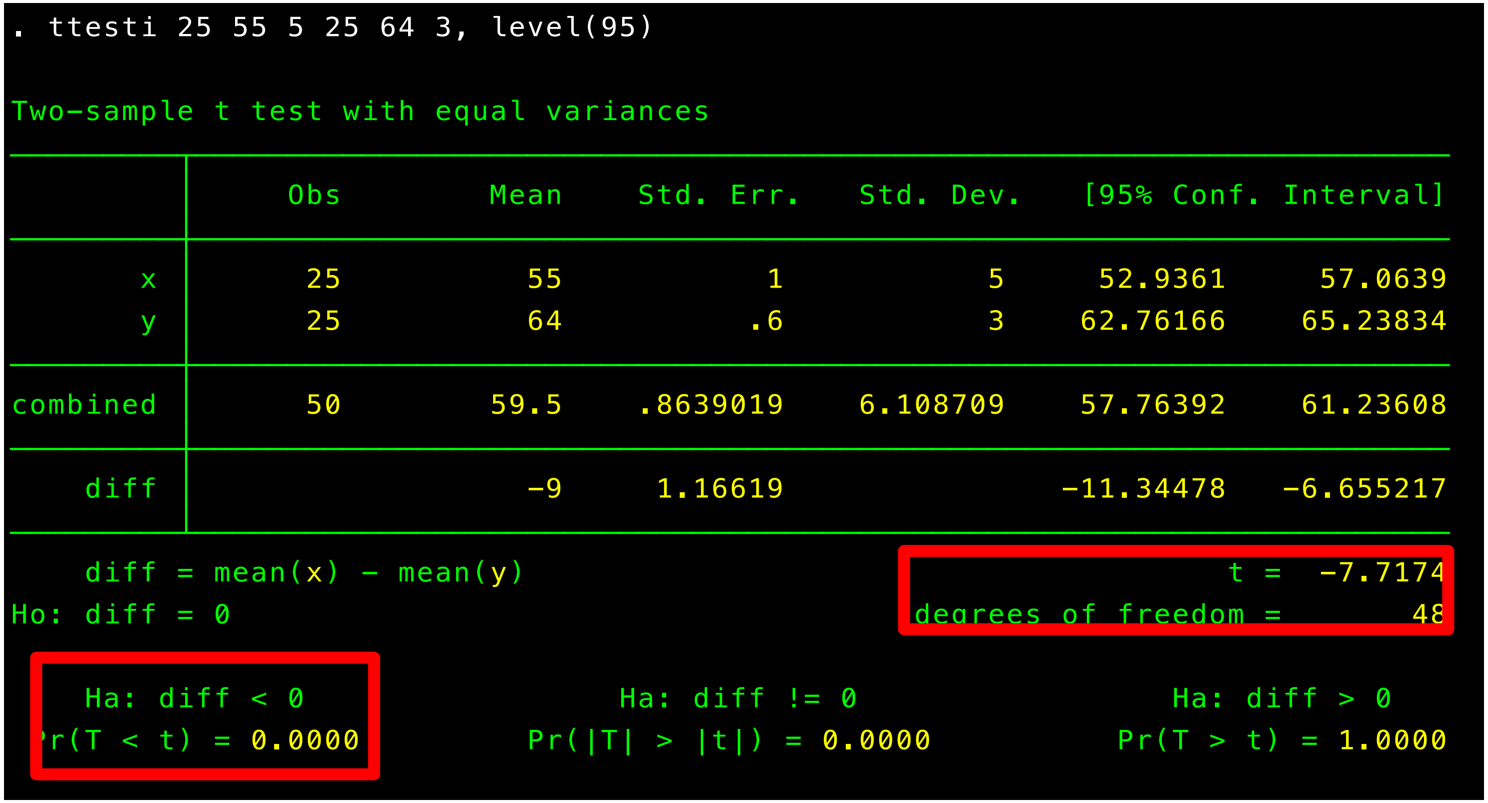
- We are interested if the accounting department takes a shorter break, so we are performing a one-tailed test.

- Therefore, the critical value is **–1.677**. Less than 5% of the *t*-distribution is less than 1.677 standard deviations away from 0.

- –7.717 is much smaller than –1.677. We can reject the null hypothesis that there is no difference and instead infer that there is less than a 5% chance that we would observe a difference of at least this magnitude if the real difference between the two groups was 0. As a whole, the accounting department likely takes shorter lunch breaks than the marketing department.

An Example: Independent Samples *t*-test

- Confirm with Stata:



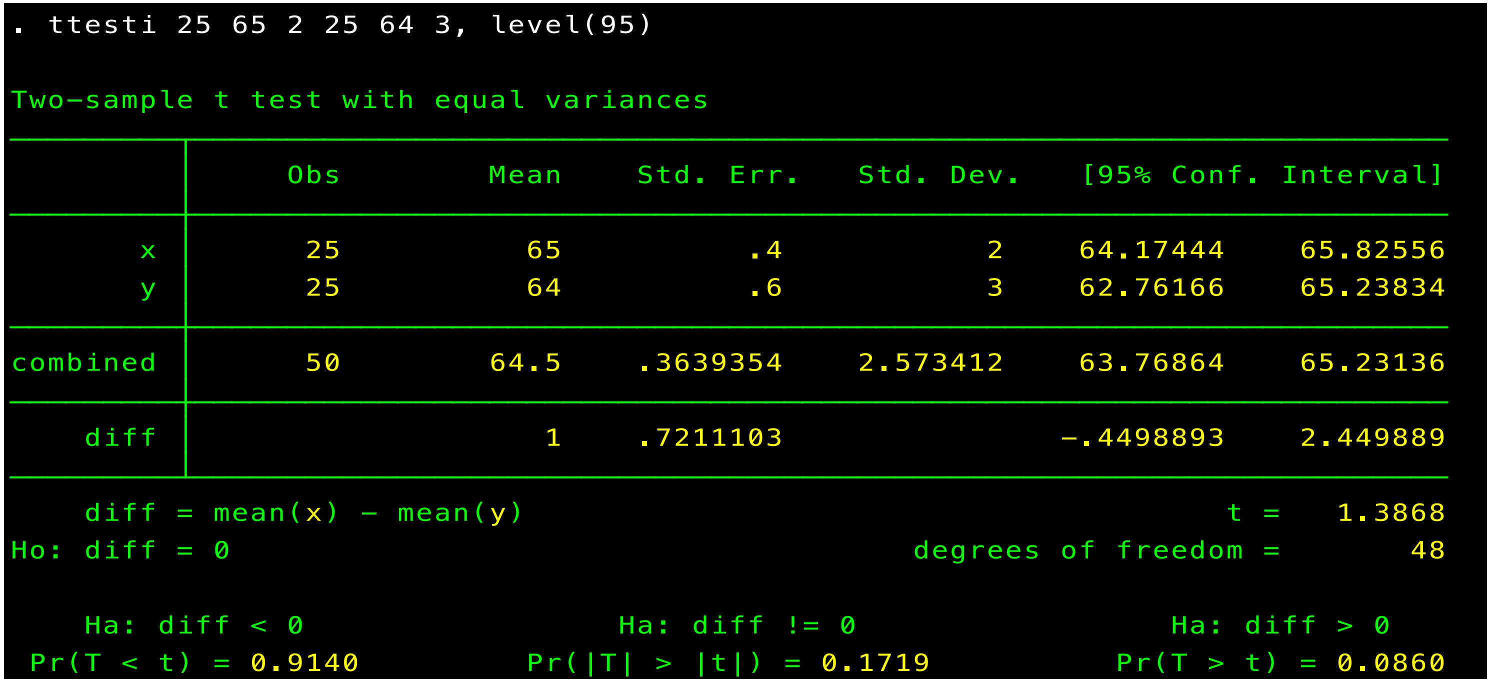
```
. ttesti 25 55 5 25 64 3, level(95)

Two-sample t test with equal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |      25          55           1           5     52.9361     57.0639
       y |      25          64          .6           3    62.76166    65.23834
---------+--------------------------------------------------------------------
combined |      50        59.5    .8639019    6.108709    57.76392    61.23608
---------+--------------------------------------------------------------------
    diff |                  -9     1.16619               -11.34478   -6.655217
------------------------------------------------------------------------------
    diff = mean(x) - mean(y)                                      t =  -7.7174
Ho: diff = 0                                     degrees of freedom =       48

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

An Example: Independent Samples *t*-test

- Now let’s say that the mean lunch break for accountants was 65 minutes—not 55—and that the standard deviation was 2 minutes. In that case there is no statistically significant difference.



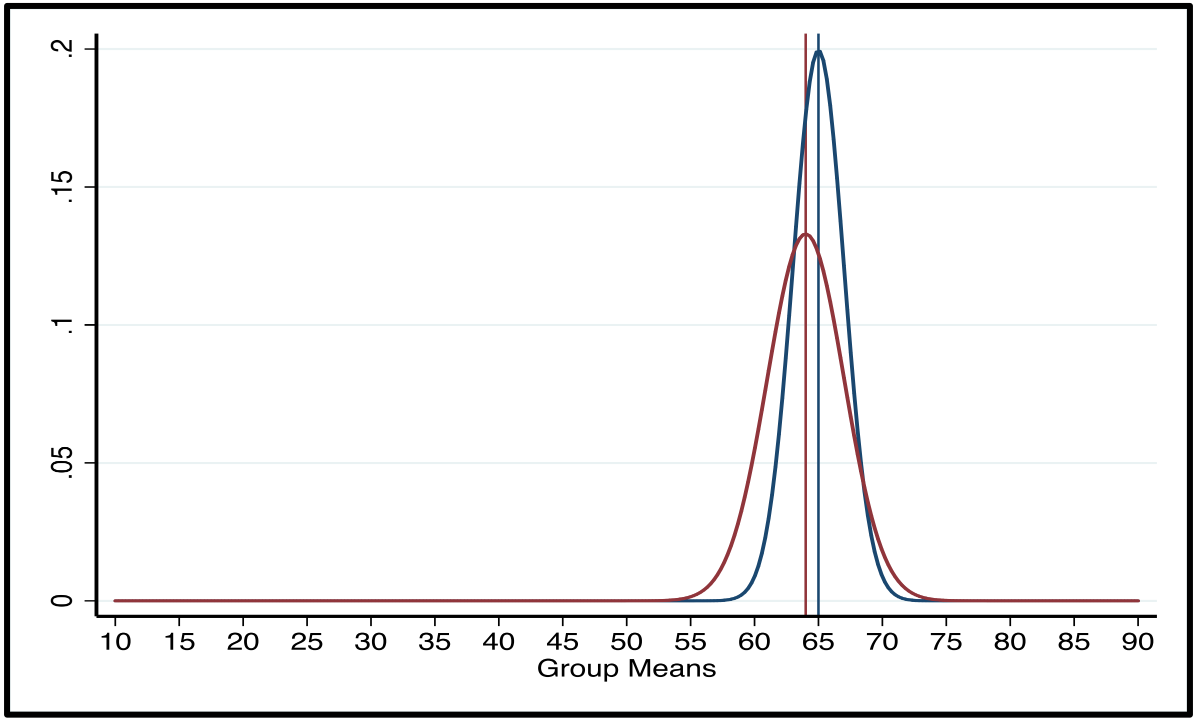
```
. ttesti 25 65 2 25 64 3, level(95)

Two-sample t test with equal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |      25          65          .4           2    64.17444    65.82556
       y |      25          64          .6           3    62.76166    65.23834
---------+--------------------------------------------------------------------
combined |      50        64.5    .3639354    2.573412    63.76864    65.23136
---------+--------------------------------------------------------------------
    diff |                   1    .7211103               -.4498893    2.449889
------------------------------------------------------------------------------
    diff = mean(x) - mean(y)                                      t =   1.3868
Ho: diff = 0                                     degrees of freedom =       48

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9140         Pr(|T| > |t|) = 0.1719          Pr(T > t) = 0.0860
```

An Example: Independent Samples *t*-test

- Look at how close the means are now. We don't have enough information to reject our H_0 that there is no statistically significant difference between the two means in the population.



Paired Samples *t*-test

- But what if the groups aren’t independent of one another?

- If you asked a sample of respondents what their mother’s educational expectations of them were and then asked them the same question about their father, you do not have independent samples. You have **paired** (or **dependent**) samples.

- If you want to examine the difference between these two means, you now have to model “within-unit” variation since you are focusing on different responses from the **same person**!

Paired Samples *t*-test

- The paired samples *t*-test is calculated with $t = \bar{d}/(s_d/\sqrt{n})$.

- Where the numerator is the mean difference between x_{1i} (the first value for individual i) and x_{2i} (the second value for individual i), and s_d is the standard deviation of the differences.

An Example: Paired Samples *t*-test

A school district superintendent is curious if adding five extra minutes to recess will lead to improved evening test scores. The superintendent randomly selects eight elementary schools and calculates the average evening math class test score for each school. They then increase recess time by five minutes in these schools. They again calculate the average evening test score for each school. Did the recess “treatment” account for a statistically significant change in evening test scores at the .05 level?*

*This is often referred to as a pretest-posttest design.

An Example: Paired Samples *t*-test

- State the null:

$$\mathrm{H}_0 : \mu_1 = \mu_2$$

- And the alternative (let’s go two-tailed):

$$\mathrm{H}_\mathrm{A} : \mu_1 \neq \mu_2$$

An Example: Paired Samples *t*-test

	x_{1i}	x_{2i}	**Difference**	**Mean Difference**	**Standard Deviation**
School 1	76	77	-1	-1.75	5.75
School 2	70	80	-10		
School 3	72	77	-5		
School 4	79	75	4		
School 5	77	83	-6		
School 6	64	62	2		
School 7	79	72	7		
School 8	80	85	-5		

An Example: Paired Samples *t*-test

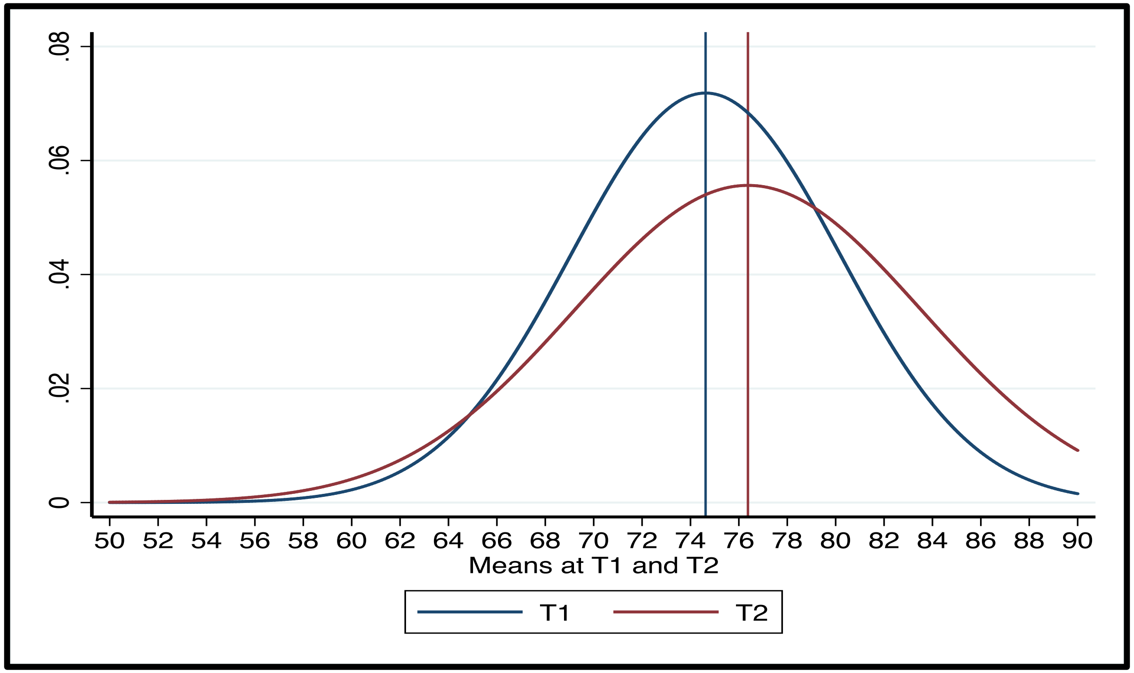
- Plug in the numbers:

$$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} = \frac{-1.75}{\frac{5.75}{\sqrt{8}}} = -.861$$

- Assuming a two-tailed test, our critical value with 7 (n – 1) degrees of freedom is ±2.365.

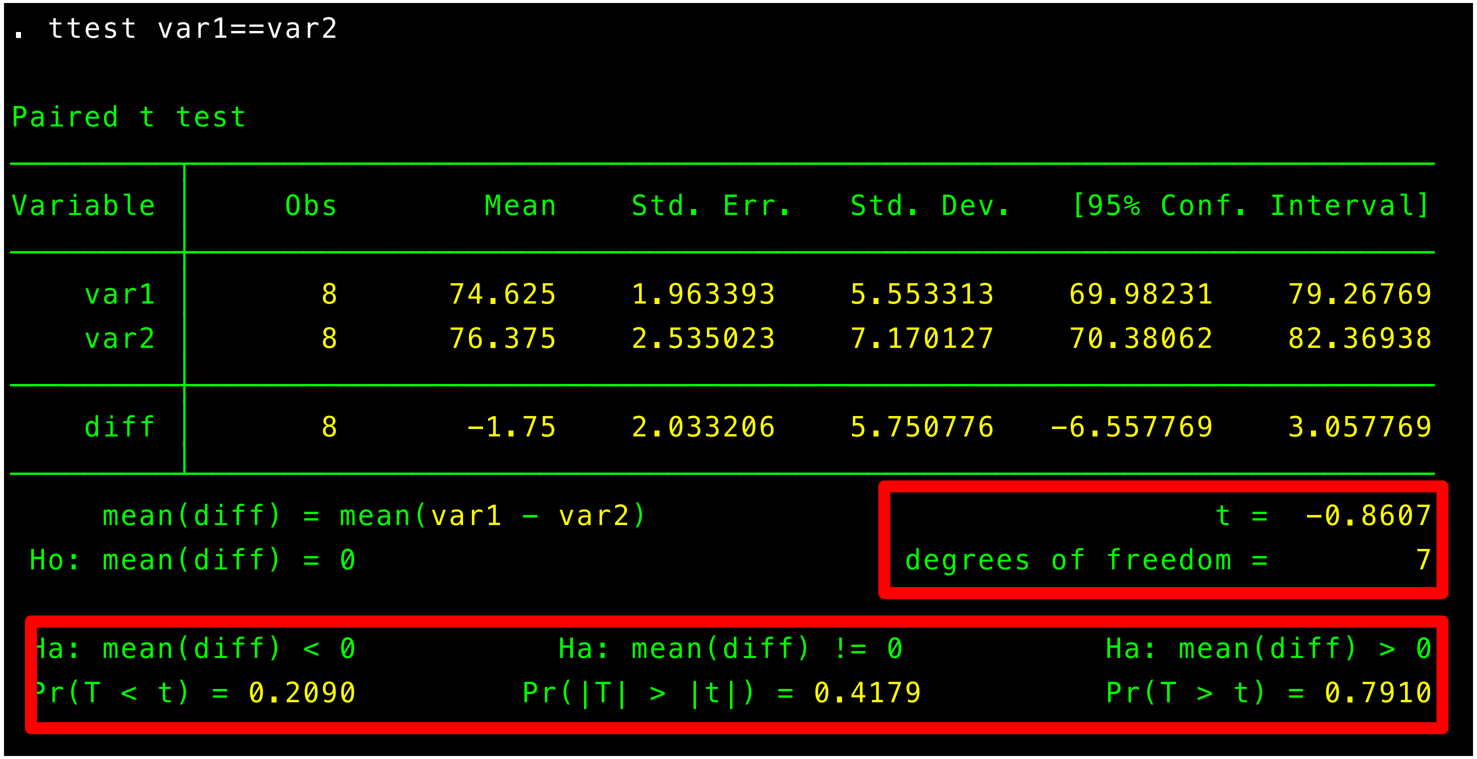
An Example: Paired Samples *t*-test

- Our test statistical is smaller than the lower bound of our critical value, so we fail to reject the null hypothesis that recess would make no difference in evening test scores in the district. This seems reasonable given how close the means are below.



An Example: Paired Samples *t*-test

- Check with Stata:



```
. ttest var1==var2

Paired t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
    var1 |       8      74.625    1.963393    5.553313    69.98231    79.26769
    var2 |       8      76.375    2.535023    7.170127    70.38062    82.36938
---------+--------------------------------------------------------------------
    diff |       8       -1.75    2.033206    5.750776   -6.557769    3.057769
------------------------------------------------------------------------------
     mean(diff) = mean(var1 - var2)                               t =  -0.8607
 Ho: mean(diff) = 0                              degrees of freedom =        7

 Ha: mean(diff) < 0           Ha: mean(diff) != 0           Ha: mean(diff) > 0
 Pr(T < t) = 0.2090         Pr(|T| > |t|) = 0.4179          Pr(T > t) = 0.7910
```

Independent Samples Difference-of-Proportions Test

- Sometimes we want to compare the distribution across a dichotomous variable from two separate groups.

- “Are women or men more likely to be diagnosed with lung cancer?”

- “Are those in the upper-middle class more likely to say they listen to jazz music than those in the lower-middle class?”

Independent Samples Difference-of-Proportions Test

- This is a simple extension of the one sample difference-of-proportions test!

- The test statistic can be found with:

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

  - Where p_1 and p_2 are the probabilities of “success” for the first and second groups, respectively, and n_1 and n_2 are their sample sizes.

Independent Samples Difference-of-Proportions Test

- Note that this formula assumes unequal variances between the two groups. If we assume equal variances (the default in many software programs), then the formula is as follows:

$$z = \frac{p_1 - p_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- Where $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$ and each *x* is the number of "successes" (usually a 1) for each respective variable.

An Example: Independent Samples Difference-of-Proportions Test

You have heard on the news that men are more likely than women to support the Tea Party. Looking at a nationally representative 2010 survey by ANES (American National Election Studies), you see that 28.5% of male respondents and 17.4% of female respondents explicitly support the Tea Party. There are 578 males and 633 females in the sample. Do we have enough information to say that, on average, males are more likely to support the Tea Party in the U.S. population?

An Example: Independent Samples Difference-of-Proportions Test

- The null hypothesis:

$$\mathrm{H}_0 : P_M = P_F$$

- The alternative hypothesis (one-tailed):

$$\mathrm{H}_\mathrm{A} : P_M > P_F$$

An Example: Independent Samples Difference-of-Proportions Test

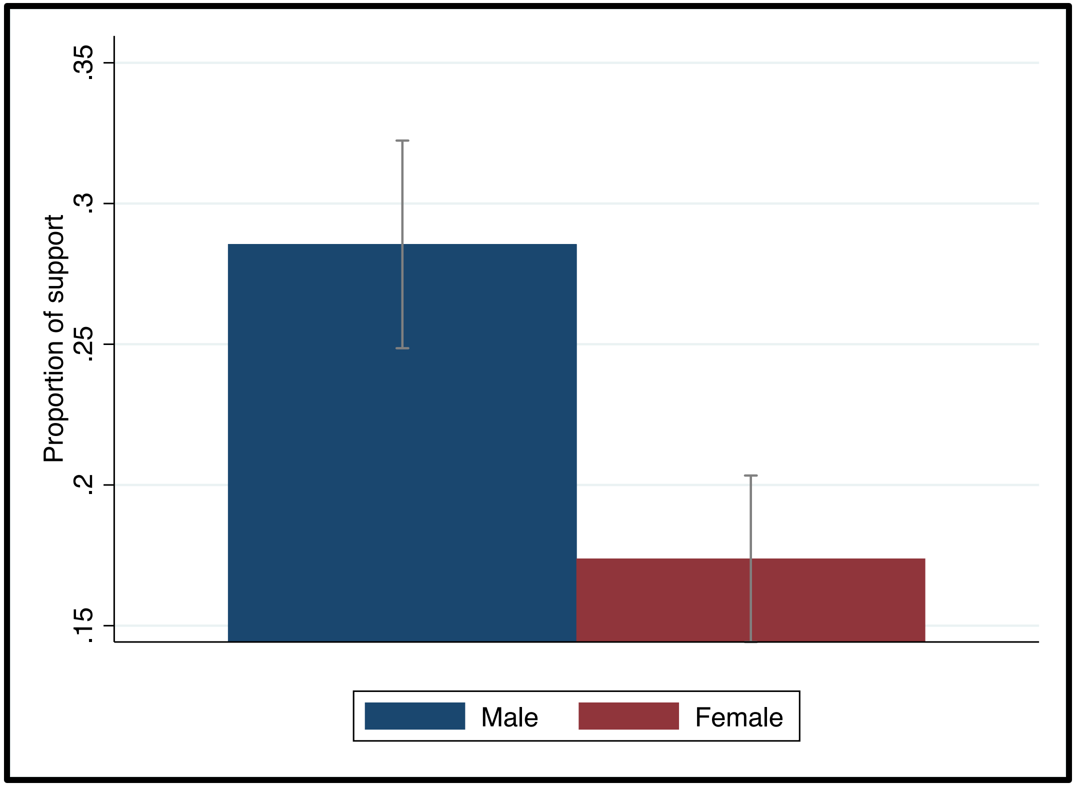
- Let's check it out:

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} = \frac{.285 - .174}{\sqrt{\frac{.285(1-.285)}{578} + \frac{.174(1-.174))}{633}}} = 4.612$$

- 4.612 is larger than our critical value of 1.64, so we reject the null hypothesis that there is no difference between men and women when it comes to support for the Tea Party.

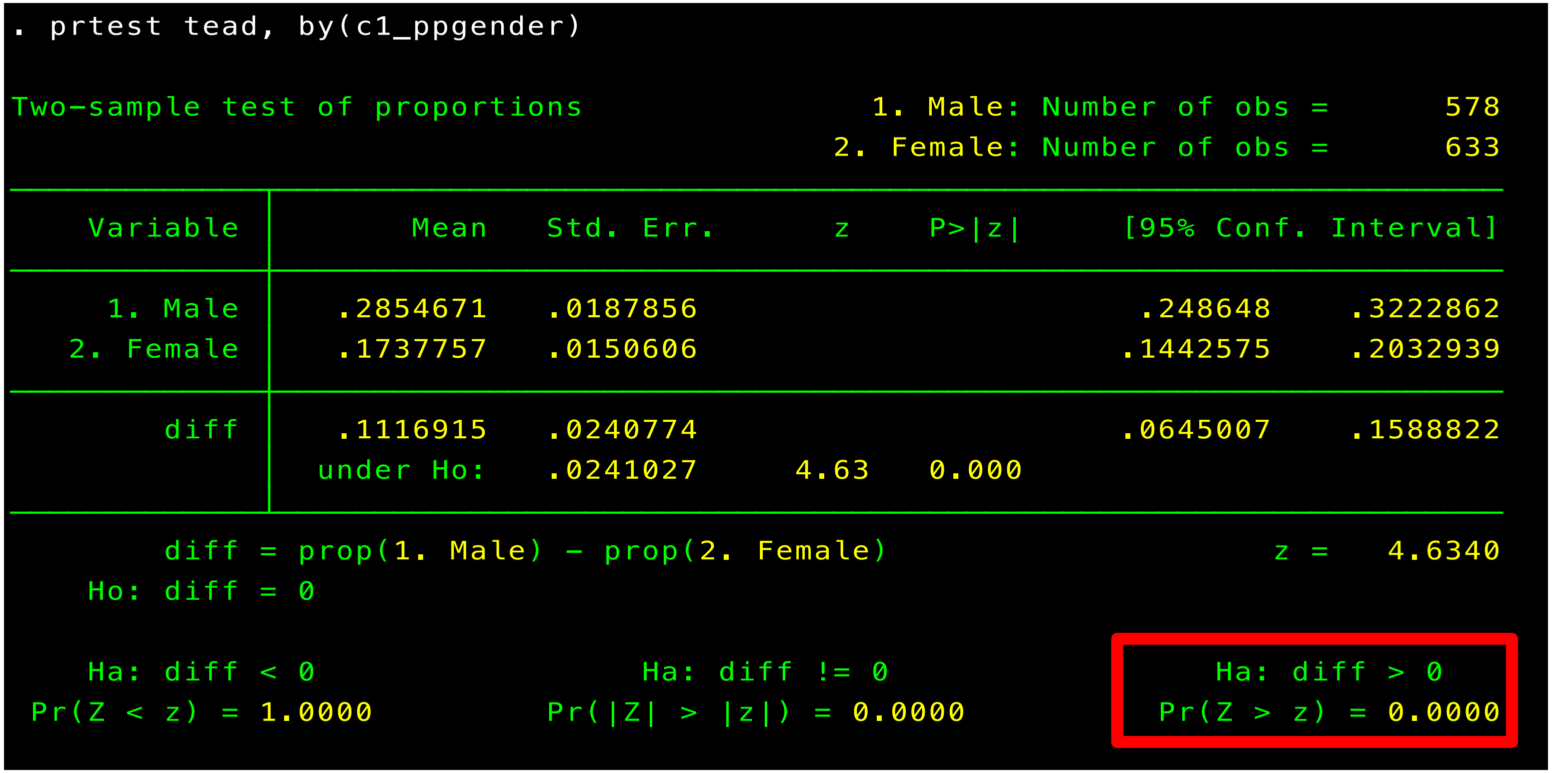
An Example: Independent Samples Difference-of-Proportions Test

- We can generalize to the population that, at least according to 2010 data, men are more likely to support the Tea Party. The bar graph below with 95% confidence intervals tells us the same story. Notice that the CIs don’t overlap (another way to determine statistical significance).



An Example: Independent Samples Difference-of-Proportions Test

- Confirm with Stata:



```
. prtest tead, by(c1_ppgender)

Two-sample test of proportions                     1. Male: Number of obs =       578
                                                 2. Female: Number of obs =       633
------------------------------------------------------------------------------
    Variable |       Mean   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     1. Male |   .2854671   .0187856                       .248648    .3222862
   2. Female |   .1737757   .0150606                      .1442575    .2032939
-------------+----------------------------------------------------------------
        diff |   .1116915   .0240774                      .0645007    .1588822
             |  under Ho:   .0241027     4.63   0.000
------------------------------------------------------------------------------
        diff = prop(1. Male) - prop(2. Female)                    z =   4.6340
    Ho: diff = 0

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(Z < z) = 1.0000         Pr(|Z| > |z|) = 0.0000          Pr(Z > z) = 0.0000
```

McNemar Test

- What if we want to assess the difference of proportions with paired (dependent) samples?

- What if we wanted to assess proportion differences "within" an individual—e.g., comparing their proportion at time 1 with their proportion on that same variable at time 2?

McNemar Test

- This paired samples difference-of-proportions test is referred to as the **McNemar test**.

- The test statistic (χ^2, or "chi-square") is found by taking the ratio of the squared difference between discordant pairs (respondents with a 0 on one variable and a 1 on the other) to the total number of discordant pairs:

$$\chi^2 = \frac{(b-c)^2}{b+c}$$

Quick tangent…

- χ^2?
- The "chi-square" distribution is a probabilistic frequency distribution of the differences between **observed** and **expected** cell counts in a cross-tabulation. A statistically significant χ^2-value is one where, given our available degrees of freedom and a set significance level, it is **not** likely that we would have at least this difference between observed and expected values under the assumption that the "real" difference in the population is zero.

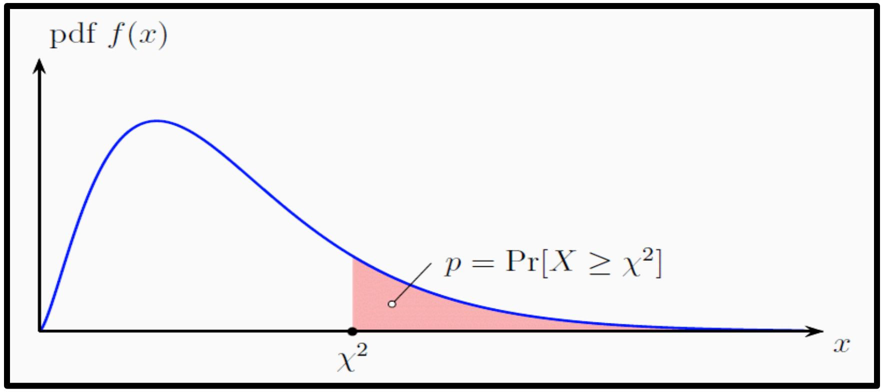


*Figure from DI Management (http://www.di-mgt.com.au/chisquare-calculator.html).

An Example: McNemar Test

- A researcher is studying how UK citizens' opinions on cuts to public spending change over short periods of time. Using Waves 2 (May-June of 2014) and 3 (September-October of 2014), they perform a McNemar test.

	Cuts are not too much	**Cuts are too much**
Cuts are not too much	6,677	1,441
Cuts are too much	1,477	9,702

An Example: McNemar Test

- A researcher is studying how UK citizens' opinions on cuts to public spending change over short periods of time. Using Waves 2 (May-June of 2014) and 3 (September-October of 2014), they perform a McNemar test.

	Cuts are not too much	**Cuts are too much**
Cuts are not too much	6,677	**1,441***
Cuts are too much	**1,477***	9,702

***These are our discordant pairs.**

An Example: McNemar Test

- The null hypothesis here would be that opinions on cuts to public spending did not change in the population during this time span ($\mathrm{H}_0 : P_{1i} = P_{2i}$).

- The alternative—assuming a two-tailed test—would be that opinions did change ($\mathrm{H}_\mathrm{A} : P_{1i} \neq P_{2i}$).

An Example: McNemar Test

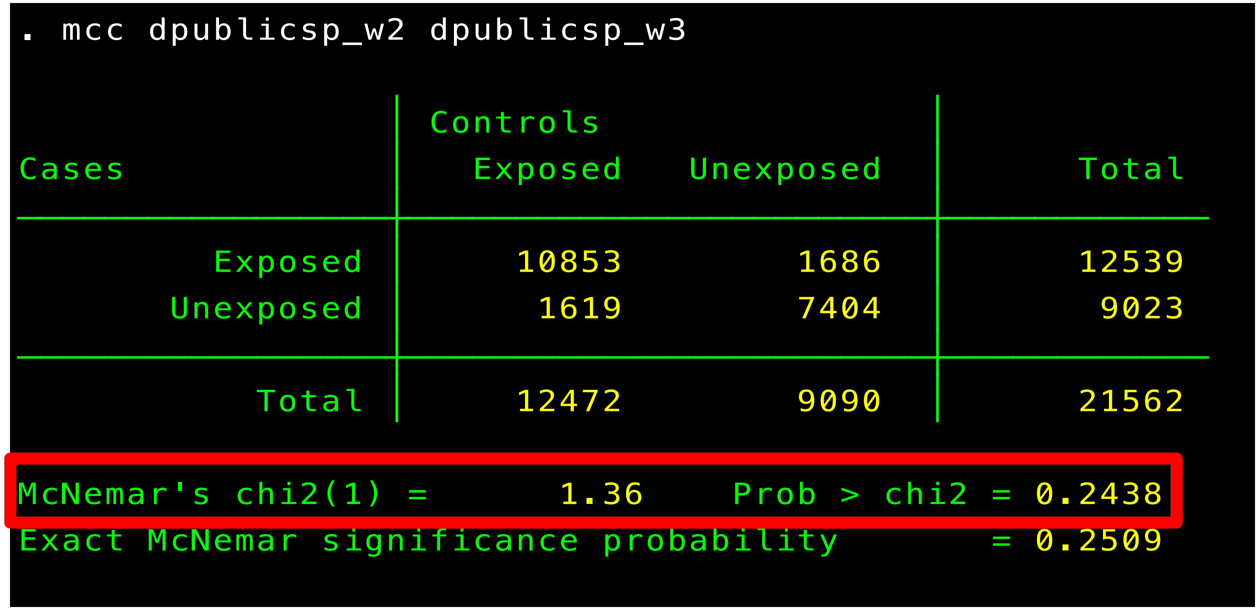
- Plug in the numbers:

$$\chi^2 = \frac{(b-c)^2}{b+c} = \frac{(1,619-1,686)^2}{1,619+1,686} = 1.36$$

- This test statistic is smaller than our critical χ^2-value of 3.84. We cannot reject our null hypothesis. It seems that opinions on this issue did not change during this time span.

An Example: McNemar Test

- Confirm with Stata:



```
. mcc dpublicsp_w2 dpublicsp_w3

                 | Controls               |
Cases            |   Exposed   Unexposed  |      Total
-----------------+------------------------+------------
         Exposed |     10853        1686  |      12539
       Unexposed |      1619        7404  |       9023
-----------------+------------------------+------------
           Total |     12472        9090  |      21562

McNemar's chi2(1) =      1.36    Prob > chi2 = 0.2438
Exact McNemar significance probability       = 0.2509
```

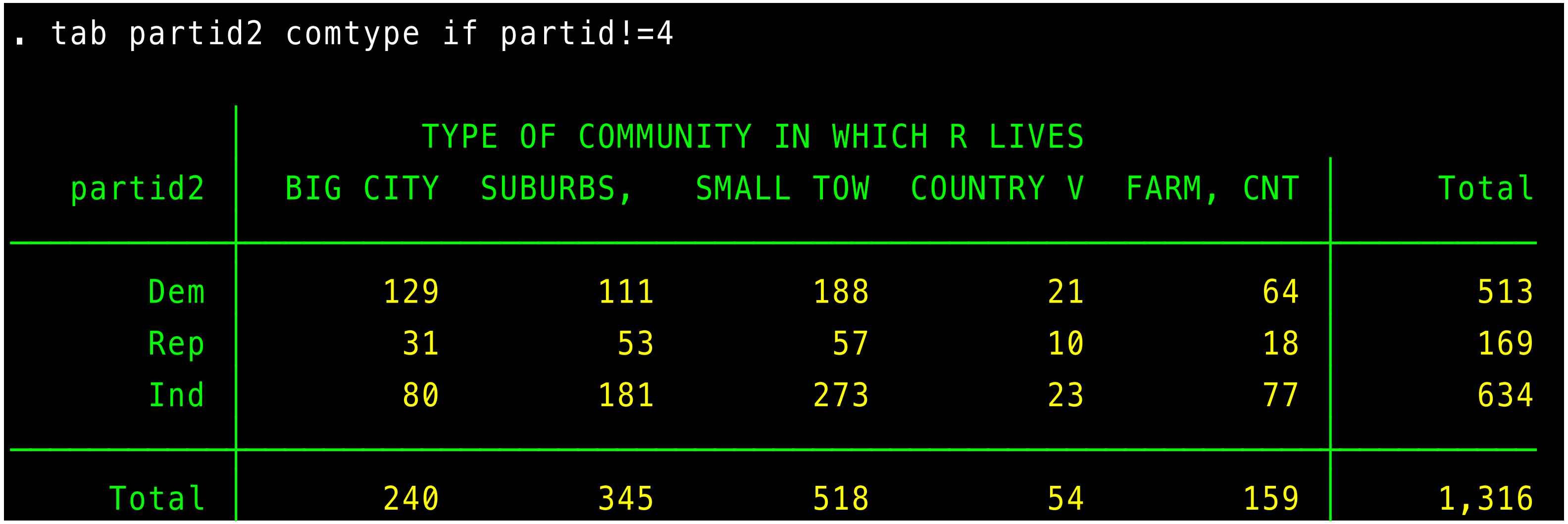
Chi-Square Test

- Speaking of χ^2…

- Since the chi-square distribution is one based on differences between observed and expected cell frequencies, we can also use it to draw inferences between nominal variables with two **or more** categories.

Chi-Square Test

- We start by first arranging our two nominal variables into a cross-tabulation, where the observed values for each category of each variable are distributed through the categories of the other variable. Below is an example, where we cross-tabulated respondents’ party identification with the type of community within which they live.



```
. tab partid2 comtype if partid!=4

           |          TYPE OF COMMUNITY IN WHICH R LIVES
   partid2 |  BIG CITY   SUBURBS,   SMALL TOW  COUNTRY V  FARM, CNT |     Total
-----------+-------------------------------------------------------+----------
       Dem |       129        111        188         21         64 |       513 
       Rep |        31         53         57         10         18 |       169 
       Ind |        80        181        273         23         77 |       634 
-----------+-------------------------------------------------------+----------
     Total |       240        345        518         54        159 |     1,316 
```

Chi-Square Test

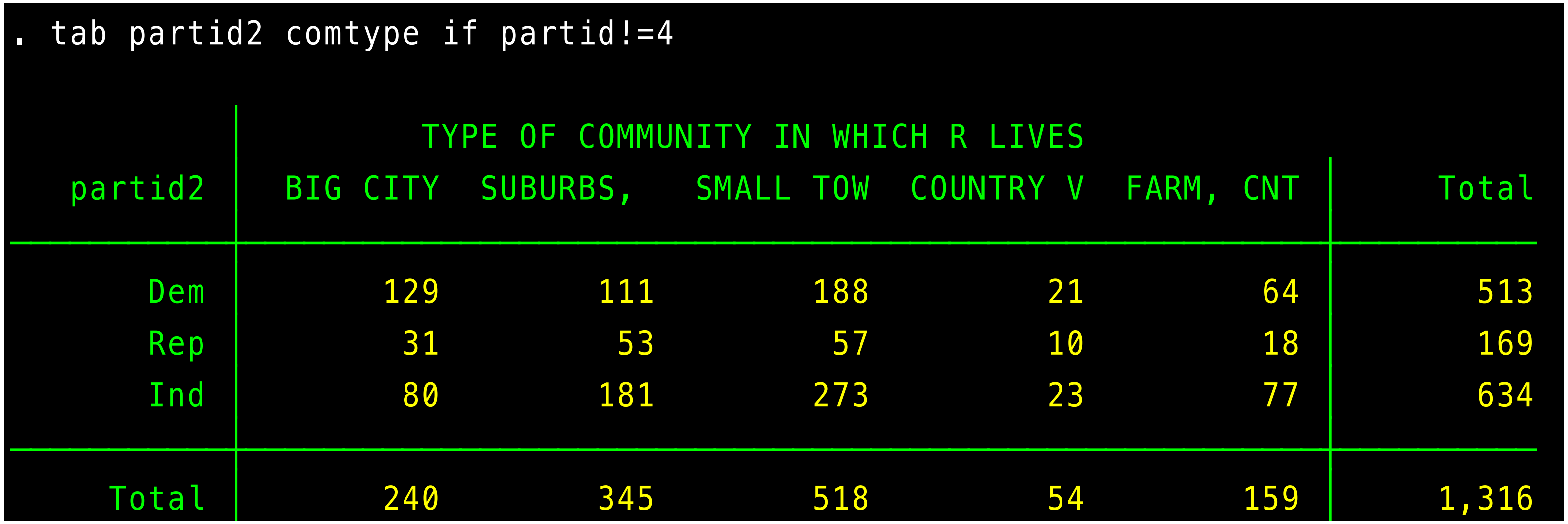
- Then we find the χ^2 test statistic by taking, for each cell, the squared difference between observed and expected frequencies and then dividing by the expected frequency for that cell. Then you simply sum them up. Formally:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O - E)^2}{E}$$

  - Where i is an individual cell in the total set of k cells.

Chi-Square Test

- By “expected frequency,” we mean $E_i = (r_i \times c_i)/n_i$ —or, the value in the *i*th cell of the $m \times n$ cross-tabulation that would be there if there was no relationship between the variables. For example, for Democrats in big cities:



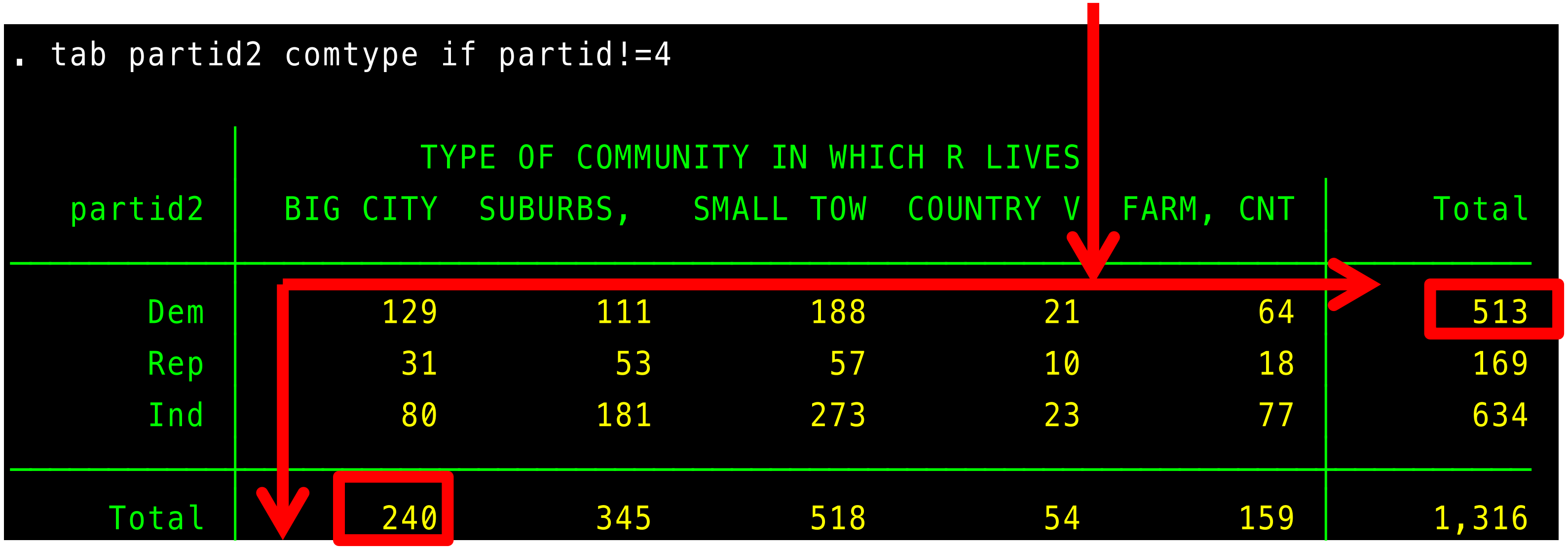
```
. tab partid2 comtype if partid!=4

           |          TYPE OF COMMUNITY IN WHICH R LIVES
   partid2 |  BIG CITY   SUBURBS,   SMALL TOW  COUNTRY V  FARM, CNT |     Total
-----------+-------------------------------------------------------+----------
       Dem |       129        111        188         21         64 |       513
       Rep |        31         53         57         10         18 |       169
       Ind |        80        181        273         23         77 |       634
-----------+-------------------------------------------------------+----------
     Total |       240        345        518         54        159 |     1,316
```

Chi-Square Test

- By “expected frequency,” we mean $E_i = (r_i \times c_i)/n_i$ —or, the value in the *i*th cell of the $m \times n$ cross-tabulation that would be there if there was no relationship between the variables. For example, for Democrats in big cities:

Multiply these numbers..



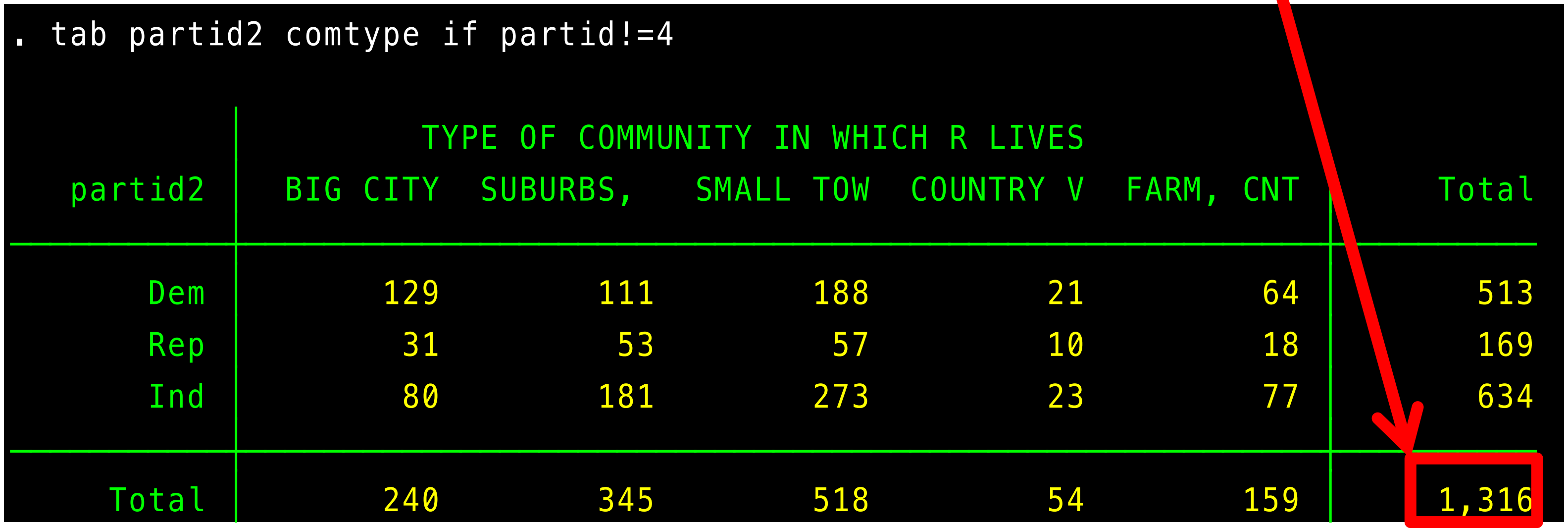
```
. tab partid2 comtype if partid!=4

           |            TYPE OF COMMUNITY IN WHICH R LIVES
   partid2 |  BIG CITY   SUBURBS,   SMALL TOW  COUNTRY V  FARM, CNT |     Total
-----------+-------------------------------------------------------+----------
       Dem |       129        111        188         21         64 |       513
       Rep |        31         53         57         10         18 |       169
       Ind |        80        181        273         23         77 |       634
-----------+-------------------------------------------------------+----------
     Total |       240        345        518         54        159 |     1,316
```

Chi-Square Test

- By “expected frequency,” we mean $E_i = (r_i \times c_i)/n_i$ —or, the value in the ith cell of the m×n cross-tabulation that would be there if there was no relationship between the variables. For example, for Democrats in big cities:

And divide by this number.



```
. tab partid2 comtype if partid!=4

           |              TYPE OF COMMUNITY IN WHICH R LIVES
   partid2 |  BIG CITY   SUBURBS,   SMALL TOW  COUNTRY V  FARM, CNT |     Total
-----------+-------------------------------------------------------+----------
       Dem |       129        111        188         21         64 |       513
       Rep |        31         53         57         10         18 |       169
       Ind |        80        181        273         23         77 |       634
-----------+-------------------------------------------------------+----------
     Total |       240        345        518         54        159 |     1,316
```

Chi-Square Test

- We then compare our test statistic to the critical χ^2-value associated with the degrees of freedom at the specified significance level.
  - FYI: In this case, the degrees of freedom are $(r - 1) \times (c - 1)$, where r is the total number of categories for the row variable and c is the total number of categories for the column variable.

An Example: Chi-Square Test

So is it likely that there is a relationship between a person’s political party identification and the type of community within which they live?

An Example: Chi-Square Test

- In the case of a chi-square test, the null hypothesis is that the two variables are independent of one another in the population—they are not related.

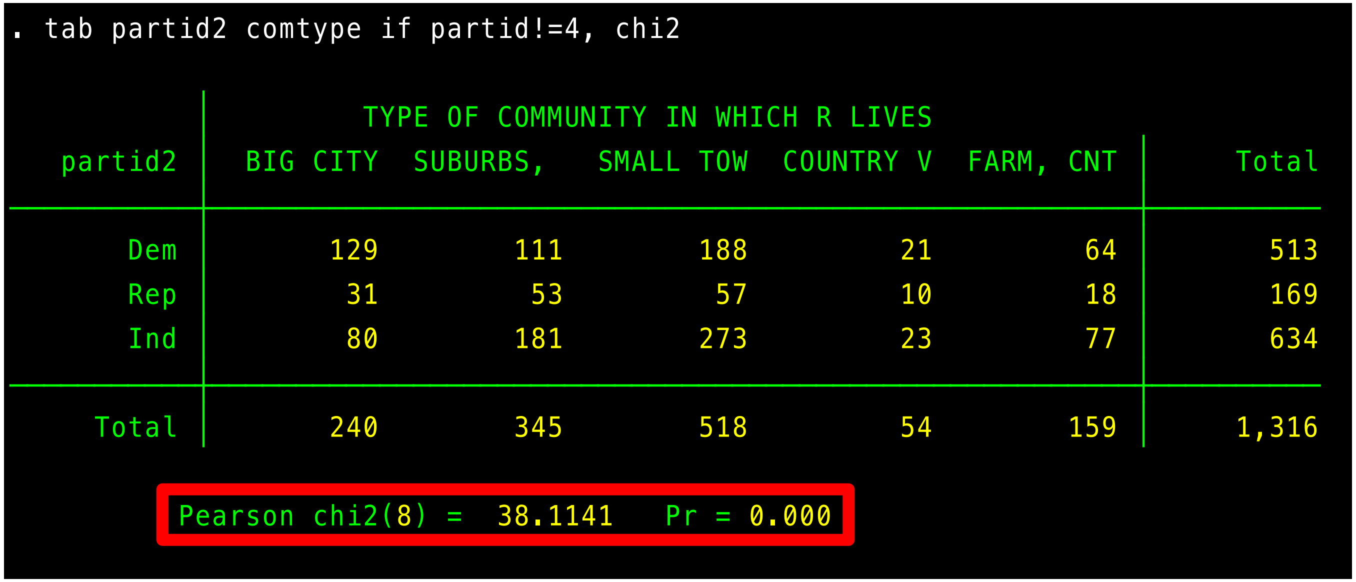
$$\mathrm{H}_0 : r \perp c$$

- And the alternative hypothesis is that they *are* related.

$$H_{\mathrm{A}} : r \text{ and } c \text{ are not } \perp$$

An Example: Chi-Square Test

- The result is statistically significant at α = .001. There is less than a 0.1% chance that we would observe differences of at least this magnitude between observed and expected frequencies if the difference in the population was, in fact, zero.



```
. tab partid2 comtype if partid!=4, chi2

           |           TYPE OF COMMUNITY IN WHICH R LIVES
   partid2 |  BIG CITY  SUBURBS,   SMALL TOW  COUNTRY V  FARM, CNT |     Total
-----------+-------------------------------------------------------+----------
       Dem |       129        111        188         21         64 |       513 
       Rep |        31         53         57         10         18 |       169 
       Ind |        80        181        273         23         77 |       634 
-----------+-------------------------------------------------------+----------
     Total |       240        345        518         54        159 |     1,316 

          Pearson chi2(8) =  38.1141   Pr = 0.000
```

Gamma/ASE Test

- The χ^2-statistic only takes into account the difference between observed and expected frequencies—in that sense each cell is treated as an independent entity.

- It does not take into account ordering. So, while you could perform a chi-square test with two ordinal variables, you are ignoring a lot of potentially important information.

Gamma/ASE Test

- This is where the **gamma/ASE test** comes in.

- We have not talked about gamma yet, since it is a **measure of association** between ordinal variables and not a measure of statistical significance. (We'll talk more about this distinction tomorrow.)

-

Gamma/ASE Test

- For now it is enough to note that gamma involves (1) cross-tabulating the two ordinal variables, and (2) taking the ratio of the difference between concordant and discordant pairs to the total number of pairs (minus ties). Closer to –1, the more negative the relationship (one goes up, the other goes down). Closer to +1, the more positive the relationship. Closer to 0, the less likely there is a relationship.

Gamma/ASE Test

- Though gamma is not for assessing statistical significance, we *can* take the ***z*-score** of gamma to assess statistical significance between two ordinal variables.

- This involves dividing gamma by its **asymptotic standard error** (ASE). Since our H_0 assumes the "real" gamma is 0 ($r \perp c$), then our z-score is $z = (\gamma - 0)/ASE$.*

*The γ symbol is the ancient Greek lowercase gamma.

An Example: Gamma/ASE Test

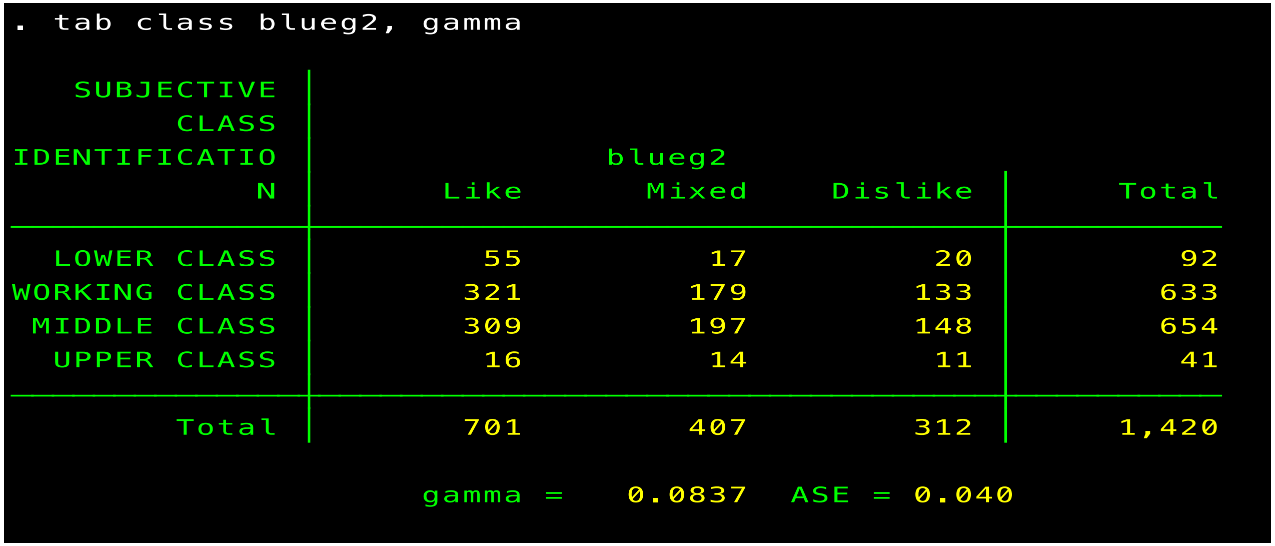
- We then have a good old-fashioned z-statistic which we can use to compare to our critical values.

- Let’s use an example.

- Is there a statistically significant relationship between a person’s self-perceived social class and the degree to which they like bluegrass music?

An Example: Gamma/ASE Test

- Dividing gamma by the ASE gives us 2.093. If we assume a two-tailed test with α = .05, we can reject the null hypothesis (because 2.093 > 1.96) and instead assert with some degree of confidence that there is a statistically significant relationship between one's self-perceived social class and their preference for bluegrass music.



```
. tab class blueg2, gamma

   SUBJECTIVE |
        CLASS |
IDENTIFICATIO |              blueg2
            N |      Like      Mixed    Dislike |     Total
--------------+---------------------------------+----------
  LOWER CLASS |        55         17         20 |        92
WORKING CLASS |       321        179        133 |       633
 MIDDLE CLASS |       309        197        148 |       654
  UPPER CLASS |        16         14         11 |        41
--------------+---------------------------------+----------
        Total |       701        407        312 |     1,420

                    gamma =   0.0837  ASE = 0.040
```

Conclusion

- This concludes our discussion on univariate and bivariate statistical inference.

- A critical take-home point is that you need to think about the assumptions of the type of inferential test you are using and assess whether or not it is appropriate given your variables.
  - Does the test involve standard deviations?
  - Are the samples independent or paired?
  - Does the ordering of your categorical variables matter?
  - Are you dealing with means or proportions?

Conclusion

- Statistical significance, however, should never be confused with **substantive significance**.

- Even if a relationship is not likely due to random chance, is it a **strong** relationship? Or is it **weak**? Somewhere in the middle?

- This is the function of **measures of association**, which we turn to tomorrow.

Datasets Used

- ANES 2010-2012 EGSS, first release (http://www.electionstudies.org/studypages/2010_2012EGSS/2010_2012EGSS.htm).
- BES 2015 Panel, Waves 1, 2, and 3 (http://www.britishelectionstudy.com/data-objects/panel-study-data/page/2/).
- GSS 1993 (http://gss.norc.org/get-the-data/stata).