Biostat 200
Lecture 6

Recap

- We calculate confidence intervals to give the most plausible values for the population mean or proportion – upon repeated sampling, 95% of the intervals contain the population mean

- We conduct hypothesis tests of a mean or a proportion to make a conclusion about how our sample mean or proportion compares with some hypothesized value for the population mean or proportion. P-values give you the probability of rejecting the null in the case that it is true.

Recap

- You can use 95% confidence intervals to reach the same conclusions as hypothesis tests
  - If the null value of the mean is outside of the 95% confidence interval, that would be the same as rejecting the null of a two-sided hypothesis test at significance level 0.05.

Recap

- We make these conclusions based on what we observed in our sample -- we will never know the true population mean or proportion
  - If the data are very different from the hypothesized mean or proportion, we reject the null
    - Example: Phase I vaccine trial – does the candidate vaccine meet minimum thresholds for safety and efficacy?
  - Statistical significance can be driven by n, and does not equal clinical or biological significance
  - On the other hand, you might have a suggestive result but not statistical significance with a small sample that deserves a larger follow up study

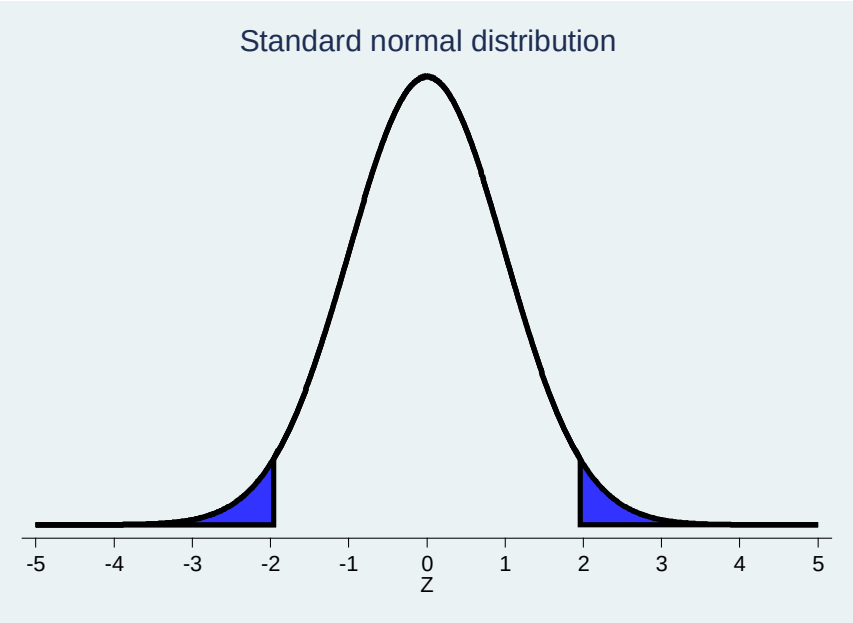
Types of error

- Type I error = significance level of the test

  α=P(reject H_0 | H_0 is true)
- Incorrectly reject the null
- We take a sample from the population, calculate the statistics, and make inference about the true population. If we did this repeatedly, we would incorrectly reject the null 5% of the time that it is true if α is set to 0.05.

How do we know we will sometimes mistakenly reject the null?

- For a 2-sided test, we will reject if our $2*P(Z>z_{stat}) < \alpha$
- So if $(\bar{X}-\mu)/(\sigma/\sqrt{n}) > 1.96$ (or < -1.96) we will reject (for $\alpha=0.05$).
- But if our hypothesized mean μ really is the population mean (i.e. the null is true), then we just got a $\bar{X}$ that looks very different from μ by chance.



Types of error

- Type II error –

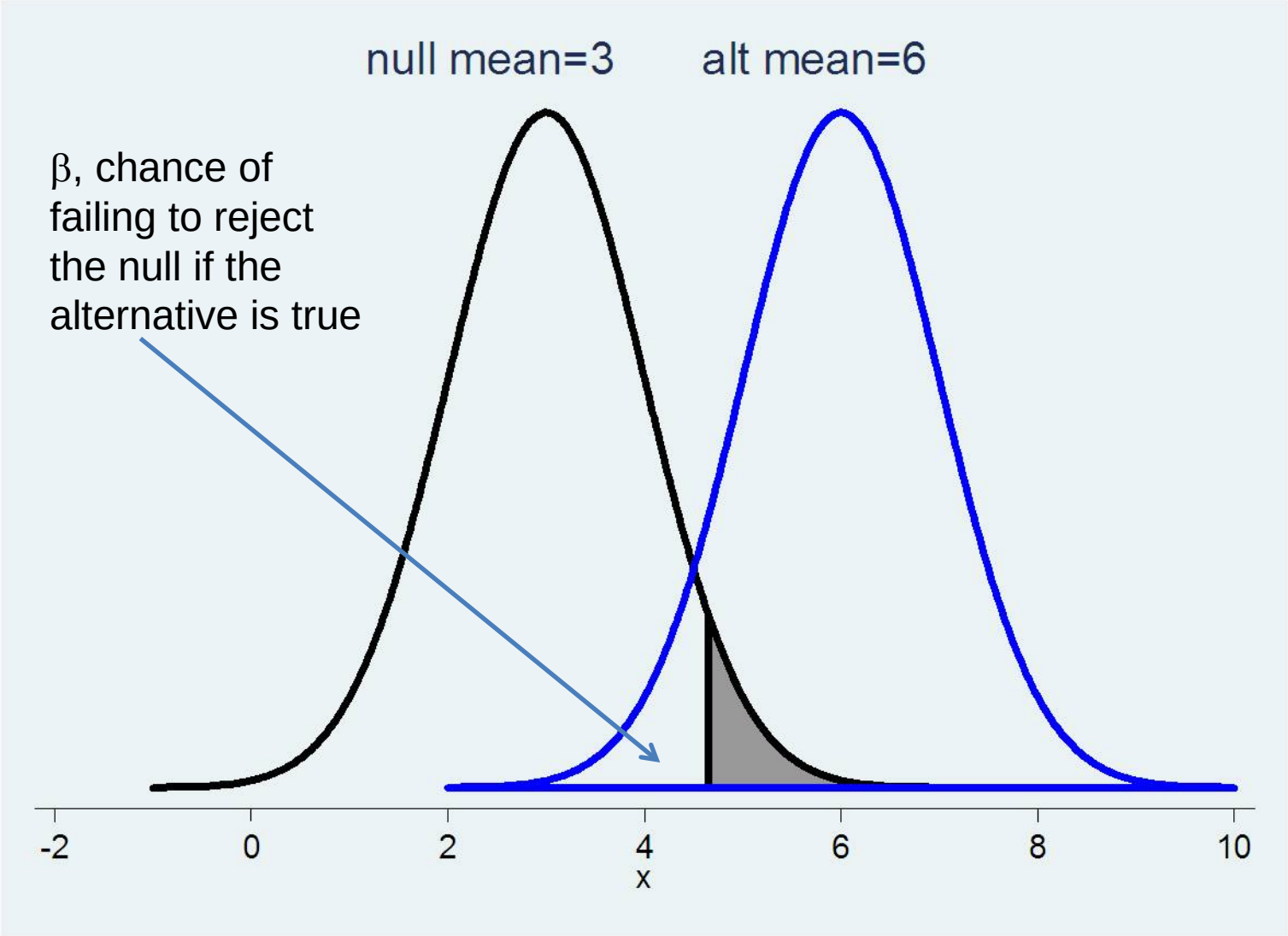
  β = P(do not reject H_0 | H_0 is false)
- Incorrectly fail to reject the null
- This happens when the test statistic (z_{stat} or t_{stat}) is not large enough, even though the null is false

Types of error

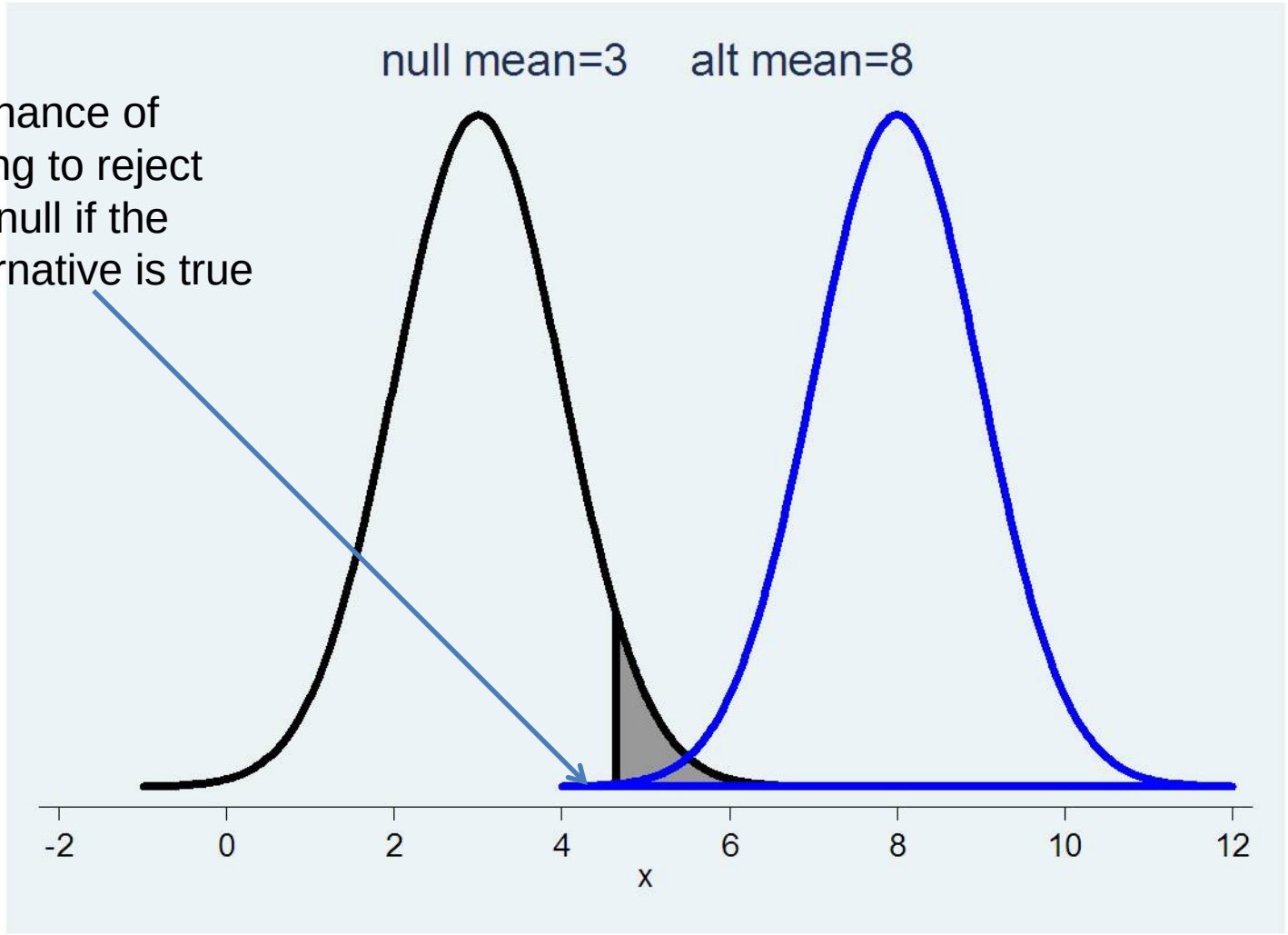
- Remember, H_0 is a statement about the population and is either true or false
- We take a sample and use the information in the sample to try to determine the answer
- Whether we make a Type I error or a Type II error depends on whether H_0 is true or false
- We set α, the chance of a Type I error, and we can design our study to minimize the chance of a Type II error

	True state	
Decision	**H_0 is true**	**H_0 is false**
Do not reject H_0	Correct	Type II error=β
Reject H_0	Type I error=α	Correct

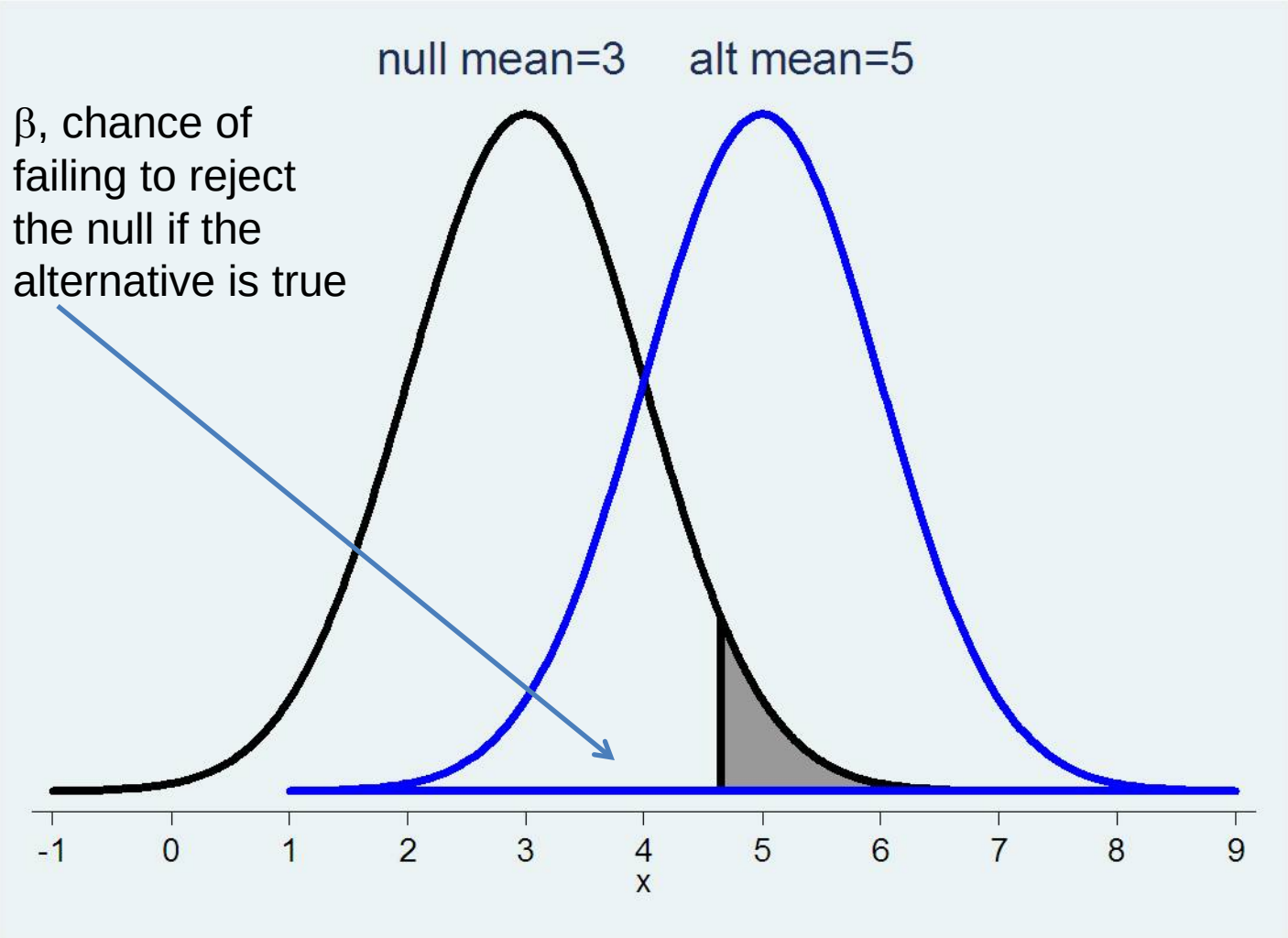
Chance of a type II error



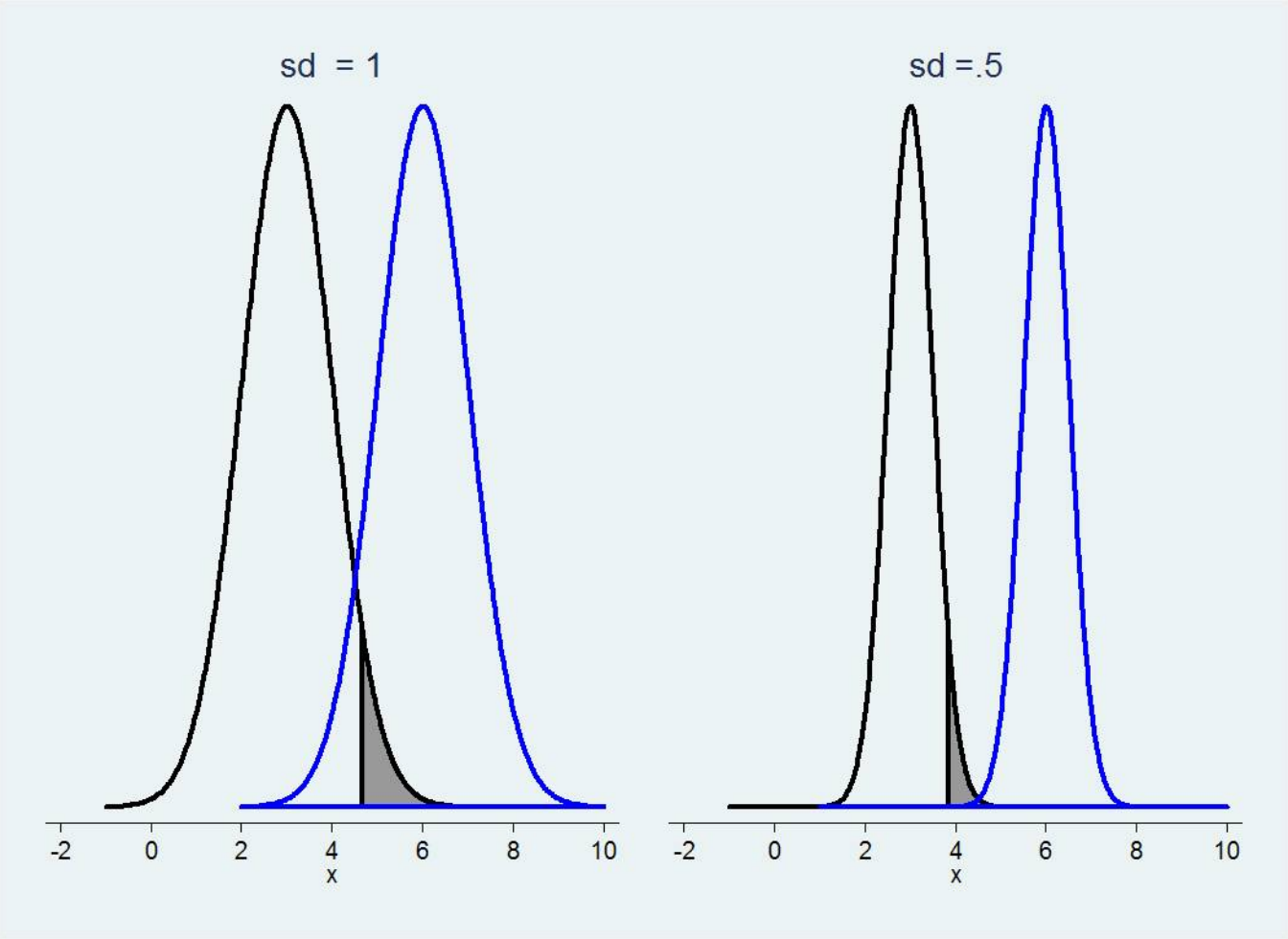
If the alternative is very different from the null, the chance of a Type II error is low



If the alternative is not very different from the null, the chance of a Type II error is high



Chance of a Type II error is lower if the SEM is smaller



This is relevant because the SD for the distribution of a sample mean is $\sigma/\sqrt{n}$ So increasing n decreases the SD of the mean. Also $\sigma/\sqrt{n}$ (or $s/\sqrt{n}$) is in the denominator of your test statistic, which if large means you will reject.

Criminal justice

- A person is accused of a crime, say stealing bread, and is on trial.
- The null hypothesis is:
- The alternative hypothesis is:
- The evidence against him/her is the data.
- If the data is very clear, then it is easy to convict. If the data are not very clear, then the jury needs to err on the side of caution.
  - The data may be made stronger by multiple witnesses saying the same thing (thus reducing σ)
  - The data may be made stronger by evidence that itself is strong (like a large effect size), even if not replicated (e.g. fingerprints in the flour)

Finding β, P(Type II error)

- Find the critical value for your test
  - At what value of $\bar{x}$ will your z_{stat} be greater than 1.96 (or 1.645 for a one-sided test)?
  - This depends on n, σ, α and the hypothesized alternative value for the mean
- Under the hypothesized alternative value, the probability of getting a sample mean less extreme as the critical value is β

Power

- The power (i.e. 1-β) of a statistical test is lower for alternative values that are closer to the null value (the chance of a Type II error is higher) and higher for more extreme alternative values.
  - The jury is more likely to convict with stronger evidence.
- It is standard to fix α=0.05 and β=0.20 (for 80% power) and determine n for various alternative hypotheses

Sample size calculation

- Mean age of walking
- $H_0 : \mu_0 < 11.4$ months
- Known $\sigma=2$
- Significance level =0.05

- For what value of $\bar{x}$ will we reject the null?

  $P(Z>z)=0.05$ for $z_{stat} > 1.645$

  So $(\bar{x} - 11.4)/(2/\sqrt{n}) > 1.645$

  $\bar{x} > 11.4 + 1.645*2/\sqrt{n}$

Sample size calculation, con't

- If the mean is actually 13, for what value of $\bar{x}$ will the probability of not rejecting be β, say 0.10
- P(Z<z) = 0.10, so $z_{stat} < -1.282$
- So $(\bar{x} - 13)/(2/\sqrt{n}) < -1.282$
- So $\bar{x} < -1.282*2/\sqrt{n} + 13$

- The $\bar{X}$ that meets these two conditions is found by setting both equations equal

$11.4 + 1.645*2/ \sqrt{n} = -1.282*2/ \sqrt{n} + 13$

$(1.645+1.282)*2/ \sqrt{n} = 13+11.4$

$\sqrt{n} = (1.645+1.282)*2/(13-11.4)$

```
di ((1.645+1.282)*2/(13-11.4))^2
13.386452
```

n=13.4

In Stata

```
. power onemean 11.4 13, power(0.9) sd(2) knownsd onesided

Estimated sample size for a one-sample mean test
z test
Ho: m = m0  versus  Ha: m > m0

Study parameters:

        alpha =    0.0500
        power =    0.9000
        delta =    0.8000
           m0 =   11.4000
           ma =   13.0000
           sd =    2.0000

Estimated sample size:

            N =        14
```

Sample size calculations

- In practice, you often have your n fixed by cost
- Then you can calculate how big the alternative has to be to reject the null with 80% probability assuming the alternative is true
- The difference between this alternative and the null is called the minimum detectable difference
  - In epidemiology when wanting to estimate an odds ratio it is called the minimum detectable odds ratio
- So if the minimum detectable difference is large, that is a bad thing – you will only have statistical significance if the alternative is very far from the null (very large effect sizes)

Comparison of two means: the *paired* t-test

- Paired samples, numerical variables
  - Two determinations on the same person (before and after) – e.g. before and after intervention
  - Matched samples – measurement on pairs of persons similar in some characteristics, i.e. identical twins (matching is on genetics)
  - Matching or pairing is performed to control for extraneous factors
- Each person or pair has 2 data points, and we calculate the difference for each
- Then we can use our one-sample methods to test hypotheses about the value of the difference

Comparison of two means: *paired* t-test

- Step 1: The hypotheses (two sided)
- Generically $H_0: \mu_1-\mu_2 = \delta$

  $H_A: \mu_1-\mu_2 \neq \delta$

- Often $\delta=0$, no difference

  So $H_0: \mu_1-\mu_2 = 0$, i.e. $H_0: \mu_1=\mu_2$

  $H_A: \mu_1-\mu_2 \neq 0$, i.e. $H_A: \mu_1\neq\mu_2$

Comparison of two means: *paired* t-test

- Step 1: The hypotheses (one sided)
- Generically $H_0: \mu_1-\mu_2 \geq \delta$ or $H_0: \mu_1-\mu_2 \leq \delta$

  $H_A: \mu_1-\mu_2 < \delta$ $H_0: \mu_1-\mu_2 < \delta$

- Often $\delta=0$, no difference

  So $H_0: \mu_1 \geq \mu_2$ or $H_0: \mu_1 \leq \mu_2$

  $H_A: \mu_1 < \mu_2$ $H_A: \mu_1 > \mu_2$

Comparison of two means: *paired* t-test

- Step 2: Calculate the test statistic

$$t_{stat} = \frac{\bar{d} - \delta}{s_d / \sqrt{n}} \text{ where}$$

$\bar{d}$ is the average of the differences d_i between the pairs

$$s_d = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \bar{d})^2}{n-1}}$$

δ is the hypothesized difference between the means

Comparison of two means: *paired* t-test

- If δ=0, a null of no difference, the formula for t_{stat} is

$$t_{stat} = \frac{\bar{d}}{s_d / \sqrt{n}} \text{ where}$$

$$s_d = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \bar{d})^2}{n-1}}$$

Comparison of two means: *paired* t-test

- Step 3: Reject or fail to reject the null
  - Is the p-value (the probability of observing a difference as large or larger, under the null hypothesis) greater than or less than the significance level, α?

Example: BREATH Study

- We have found that self-reported alcohol consumption decreases after ART initiation, however we suspect a high degree of under-report. We use the biomarker PEth to obtain a biological measure of alcohol use in a one-year prospective study. Looking at baseline and 3 months…

Example

- We think participants may be decreasing their alcohol use over time. The null hypothesis is that they are drinking the same amount.

  $H_0: \mu_2-\mu_1=0 \rightarrow \mu_2=\mu_1$

  $H_A: \mu_1-\mu_2 \neq 0 \rightarrow \mu_2 \neq \mu_1$

- Significance level=0.05

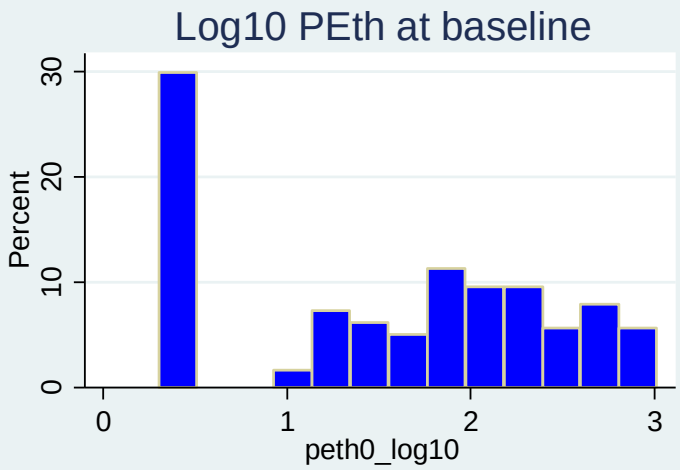
Paired data – PEth at 0 and 3 months

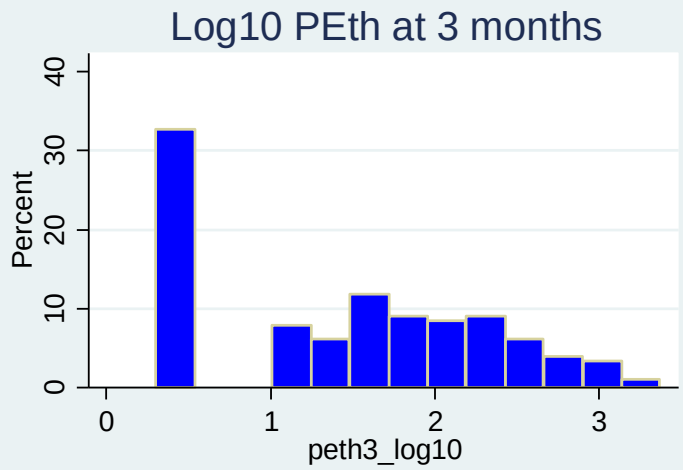
** print out 1st 10 observations

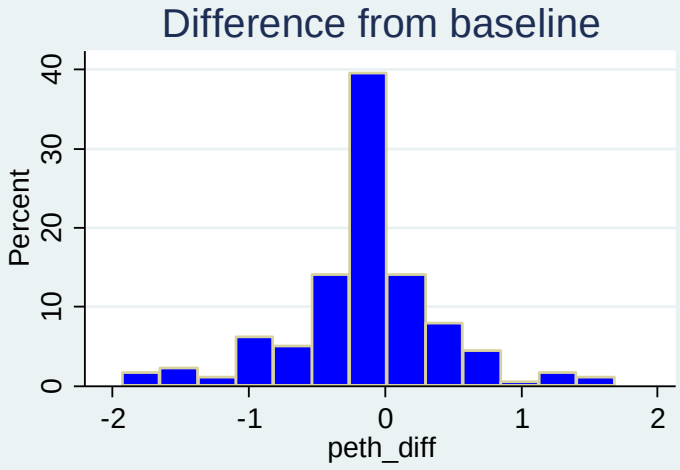
```
. list studyid peth0_log10 peth3_log10 peth_diff in 1/10
     +-----------------------------------------------+
     | studyid   peth0_log10   peth3_log10   peth_diff |
     |-----------------------------------------------|
  1. | MBB2001   2.054766   1.950681    -.104085 |
  2. | MBB2002     .30103     .30103           0 |
  3. | MBB2006     .30103     .30103           0 |
  4. | MBB2007     .30103     .30103           0 |
  5. | MBB2008   1.388012   1.575477    .1874642 |
     |-----------------------------------------------|
  6. | MBB2011     .30103     .30103           0 |
  7. | MBB2013     .30103     .30103           0 |
  8. | MBB2016   2.122216     .30103   -1.821186 |
  9. | MBB2021   1.341533   1.348207    .0066741 |
 10. | MBB2022     .30103   1.111766    .8107364 |
     +-----------------------------------------------+
```

These data from BREATH wide data.dta"

Graphs of the data







```
. summ peth_diff

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   peth_diff |       177    -.101553    .5782935  -1.929419   1.685742


*** calculate the t statistic
 di -.101553/.5782935*sqrt(177)
-2.3363133


*** calculate the p-value
. di 2*ttail(176,2.336133)
.02061082
```

→ So we reject the null

$$t_{stat} = \frac{\bar{d}}{s_d / \sqrt{n}} \text{ where}$$

$$s_d = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \bar{d})^2}{n-1}}$$

Using the ttest command

```
. . ttest peth_diff==0

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
peth_d~f |     177    -.101553    .0434672    .5782935    -.187337    -.015769
------------------------------------------------------------------------------
    mean = mean(peth_diff)                                        t =  -2.3363
Ho: mean = 0                                     degrees of freedom =      176

    Ha: mean < 0                 Ha: mean != 0                  Ha: mean > 0
 Pr(T < t) = 0.0103         Pr(|T| > |t|) = 0.0206          Pr(T > t) = 0.9897
```

Note that mean>0 here is *mean difference*

Another way without calculating the difference

The command is

```
ttest var1==var2
```

```
. . ttest peth3_log10==peth0_log10

Paired t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
peth3~10 |     177    1.426734    .0686534    .9133744    1.291244    1.562224
peth0~10 |     177    1.528287    .0690468    .9186077    1.392021    1.664553
---------+--------------------------------------------------------------------
    diff |     177    -.101553    .0434672    .5782935    -.187337    -.015769
------------------------------------------------------------------------------
     mean(diff) = mean(peth3_log10 - peth0_log10)                 t =  -2.3363
 Ho: mean(diff) = 0                          degrees of freedom =      176

 Ha: mean(diff) < 0           Ha: mean(diff) != 0           Ha: mean(diff) > 0
 Pr(T < t) = 0.0103         Pr(|T| > |t|) = 0.0206          Pr(T > t) = 0.9897
```

.

Another example

- Breast cancer tumor volume

Comparison of two means: t-test

- The goal is to compare means from two independent samples
- Two different populations
  - E.g. vaccine versus placebo group
  - E.g. women with adequate versus in adequate micronutrient levels

Comparison of two means: t-test

- Two sided hypothesis

$H_0: \mu_1 = \mu_2$

$H_A: \mu_1 \neq \mu_2$

- One sided hypothesis

$H_0: \mu_1 \geq \mu_2$

$H_A: \mu_1 < \mu_2$

- One sided hypothesis

$H_0: \mu_1 \leq \mu_2$

$H_A: \mu_1 > \mu_2$

Comparison of two means: t-test

- Even though the null and alternative hypotheses are the same as for the paired t-test, the test is different, it is wrong to use a paired t-test with independent samples and vice versa

Comparison of two means: t-test

- By the CLT, if X_1 and X_2 are normally distributed, then $\overline{X}_1 - \overline{X}_2$ is normally distributed with mean μ_1-μ_2 and standard deviation $\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}$

- In one version of the t-test, we assume that the population standard deviations are equal, so $\sigma_1 = \sigma_2 = \sigma$

- Substituting, the standard deviation for the distribution of the difference of two sample means is $\sqrt{\sigma^2 (1/n_1 + 1/n_2)}$

Comparison of two means: t-test

- So we can calculate a z-score for the difference in the means and compare it to the standard normal distribution. The test statistic is

$$z_{stat} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2 \left(1/n_1 + 1/n_2\right)}}$$

Comparison of two means: t-test when σ is unknown

- T-test test statistic

$$t_{stat} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(1/n_1 + 1/n_2\right)}} \text{ where}$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

  - The formula for the pooled SD is a weighted average of the individual sample SDs
  - The degrees of freedom for the test are n_1+n_2-2

Comparison of two means: t-test

- As in our other hypothesis tests, compare the t statistic to the t-distribution to determine the probability of obtaining a mean difference as large or larger as the observed difference
- Reject the null if the probability, the p-value, is less than α, the significance level
- Fail to reject the null if $p \geq \alpha$

Comparison of two means, example

- Study of non-pneumatic anti-shock garment (Miller et al)
- Two groups – pre-intervention received usual treatment, intervention group received NASG
- Comparison of hemorrhaging in the two groups
- Null hypothesis: The hemorrhaging is the same in the two groups $H_0: \mu_1=\mu_2$

  $H_A: \mu_1 \neq \mu_2$

- The data:
- External blood loss after entry:
  - Pre-intervention group (n=83) mean blood loss =340.4 SD=248.2
  - Intervention group (n=83) mean blood loss =73.5 SD=93.9

Calculating by hand

- External blood loss:
  - Pre-intervention group (n=83) mean=340.4 SD=248.2
  - Intervention group (n=83) mean=73.5 SD=93.9
- First calculate s_p^2

$$s_p^2 = (82*248.2^2 + 82*93.9^2)/(83+83-2)$$
$$= 35210.2$$
$$t_{stat} = (340.4-73.5)/\text{sqrt}(35210.2*(2/83))$$
$$= 9.16$$
$$df = 83+83-2=164$$

```
. di 2*ttail(164,9.16)
2.041e-16
```

$$t_{stat} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(1/n_1 + 1/n_2\right)}} \text{ where}$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Comparison of two means, example

*** ttesti n1 mean1 sd1 n2 mean2 sd2**

```
ttesti 83 340.4 248.2 83 73.5 93.9

Two-sample t test with equal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |      83       340.4    27.24349       248.2     286.204     394.596
       y |      83        73.5    10.30686        93.9    52.99636    94.00364
---------+--------------------------------------------------------------------
combined |     166      206.95    17.85377    230.0297    171.6987    242.2013
---------+--------------------------------------------------------------------
    diff |               266.9    29.12798                209.3858    324.4142
------------------------------------------------------------------------------
    diff = mean(x) - mean(y)                                      t =   9.1630
Ho: diff = 0                     degrees of freedom =      164

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

- Remember that when conducting a hypothesis test that a mean is equal to or less than or greater than some value (a one-sample t-test), use

```
.ttesti  n mean sd hypothesizedmean
```

- When testing the equality of <u>two</u> means, use

```
ttesti   n1 mean1   sd1   n2  mean2 sd2
```

- Stata will know which one based on how many numbers you enter (4 vs. 6)

- You can calculate a 95% confidence interval for the difference between the 2 means

$$\text{Lower bound}: (\bar{x}_1 - \bar{x}_2) - t_{df,.025}\sqrt{s_p^2\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}$$

$$\text{Upper bound}: (\bar{x}_1 - \bar{x}_2) + t_{df,.025}\sqrt{s_p^2\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}$$

- If the confidence interval for the difference does not include 0, then you can reject the null hypothesis of no difference

This is NOT equivalent to calculating separate confidence intervals for each mean and determining whether they overlap

Comparison of two means: t-test

- This t-test assumes equal variances in the two underlying populations
- If we do not assume equal variances we use a slightly different test statistic
    - Variances not assumed to be equal, so you do not use a pooled estimate
    - There is another formula for degrees of freedom
- Often the two different t-tests yield the same answer, but you should not assume equivalence unless you have a good reason for it
    - If the sample sizes are equal, you will get the same test statistic, just the df changes

The t statistic is

$$t_{stat} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{(s_1^2 / n_1) + (s_2^2 / n_2)}}$$

with degrees of freedom ν (pronounced "nu")

$$\nu = \frac{\left[(s_1^2 / n_1) + (s_2^2 / n_2)\right]^2}{\left[(s_1^2 / n_1)^2 / (n_1 - 1) + (s_2^2 / n_2)^2 / (n_2 - 1)\right]}$$

Round up to the nearest integer to get the degrees of freedom

Comparison of two means, example

```
ttesti 83 340.4 248.2 83 73.5 93.9, unequal

Two-sample t test with unequal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |      83       340.4    27.24349       248.2     286.204     394.596
       y |      83        73.5    10.30686        93.9    52.99636    94.00364
---------+--------------------------------------------------------------------
combined |     166      206.95    17.85377    230.0297    171.6987    242.2013
---------+--------------------------------------------------------------------
    diff |               266.9    29.12798                209.1446    324.6554
------------------------------------------------------------------------------
    diff = mean(x) - mean(y)                                      t =   9.1630
Ho: diff = 0                     Satterthwaite's degrees of freedom =  105.002

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

Test of the means of independent samples

- When you have the data in Stata, with the different groups in different columns, use

```
ttest  var1==var2, unpaired
```

or

```
ttest  var1==var2, unpaired unequal
```

Test of the means of independent samples

- More commonly you will have the data all in one variable, and the grouping in another variable. Then use

```
ttest  var, by(groupvar)
```

or

```
ttest var, by(groupvar) unequal
```

Alcohol study example

Testing whether log PEth at 6 months differs by study arm

Null hypothesis: Log PEth for cohort arm = Log PEth for comparison arm

```
. ttest peth_log10, by(studyarm)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
Cohort a |     181    1.406739    .0709392    .9543891     1.26676    1.546718
Min asse |     106    1.490502    .0897145    .9236671    1.312615    1.668389
---------+--------------------------------------------------------------------
combined |     287    1.437676    .0556285     .942407    1.328183    1.547169
---------+--------------------------------------------------------------------
    diff |           -.0837633    .1153577               -.3108244    .1432978
------------------------------------------------------------------------------
    diff = mean(Cohort a) - mean(Min asse)                        t =  -0.7261
Ho: diff = 0                                     degrees of freedom =      285

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.2342         Pr(|T| > |t|) = 0.4684          Pr(T > t) = 0.7658
```

```
. ttest peth_log10, by(studyarm) unequal

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
Cohort a |     181    1.406739    .0709392    .9543891     1.26676    1.546718
Min asse |     106    1.490502    .0897145    .9236671    1.312615    1.668389
---------+--------------------------------------------------------------------
combined |     287    1.437676    .0556285     .942407    1.328183    1.547169
---------+--------------------------------------------------------------------
    diff |           -.0837633    .1143724               -.3091369    .1416103
------------------------------------------------------------------------------
    diff = mean(Cohort a) - mean(Min asse)                        t =  -0.7324
Ho: diff = 0                     Satterthwaite's degrees of freedom =  225.846

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.2324         Pr(|T| > |t|) = 0.4647          Pr(T > t) = 0.7676
```

- Confidence interval for the difference of two means from independent samples, when unequal variances are assumed

$$\text{Lower bound} : (\bar{x}_1 - \bar{x}_2) - t_{\nu,.025}\sqrt{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]}$$

$$\text{Upper bound} : (\bar{x}_1 - \bar{x}_2) + t_{\nu,.025}\sqrt{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]}$$

Comparison of two proportions

- Similar to comparing two means
- Null hypothesis about two proportions, p_1 and p_2,

  H_0: $p_1 = p_2$

  H_A: $p_1 \neq p_2$
- If n_1 and n_2 are sufficiently large, the difference between the two proportions follows a normal distribution.

Comparison of two proportions

- So we can use the z statistic

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})[1/n_1 + 1/n_2]}}$$

$$\text{where } \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

to find the probability of observing a difference as large as we do, under the null hypothesis of no difference

Comparison of two proportions

- Example: Proportion with PEth>=50 ng/ml

```
. tab peth_50

   peth_50 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        155       54.01       54.01
          1 |        132       45.99      100.00
------------+-----------------------------------
      Total |        287      100.00

. bysort studyarm: tab peth_50
-------------------------------------------------------------------------------->
   studyarm = Cohort assessed

    peth_50 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         97       53.59       53.59
          1 |         84       46.41      100.00
------------+-----------------------------------
      Total |        181      100.00
-------------------------------------------------------------------------------
-> studyarm = Min assessed

    peth_50 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         58       54.72       54.72
          1 |         48       45.28      100.00
------------+-----------------------------------
      Total |        106      100.00


.
```

Comparison of two proportions

- Null hypothesis: The proportion with PEth>=50 is the same in both groups

  H_0: $p_1 = p_2$ (so $p_1 - p_2 = 0$)

- Z statistic is calculated:

  $\hat{p}$ = 0.460 (overall proportion)

  z_{stat} = (.464-.453)/sqrt(.460*(1-.460)*(1/181+1/106))

  =.18045468

```
. di 2*(1-normal(.18045468))
.85679563
```

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})[1/n_1 + 1/n_2]}}$$

where $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$

Comparison of two proportions

prtesti n1 p1 n2 p2

```
. . prtesti 106 .453 181 .464

Two-sample test of proportions                        x: Number of obs =      106
                                                      y: Number of obs =      181
------------------------------------------------------------------------------
    Variable |       Mean   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |       .453   .0483493                      .3582372    .5477628
           y |       .464   .0370683                      .3913476    .5366524
-------------+----------------------------------------------------------------
        diff |      -.011   .0609238                     -.1304084    .1084084
             |  under Ho:   .0609565    -0.18   0.857
------------------------------------------------------------------------------
        diff = prop(x) - prop(y)                                      z =  -0.1805
    Ho: diff = 0

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(Z < z) = 0.4284         Pr(|Z| < |z|) = 0.8568          Pr(Z > z) = 0.5716
```

Comparison of two proportions

```
. prtest peth_50, by(month)

Two-sample test of proportions                        6: Number of obs =      181
                                                     88: Number of obs =      106
------------------------------------------------------------------------------
    Variable |       Mean   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           6 |   .4640884   .0370687                       .391435    .5367418
          88 |   .4528302   .0483477                      .3580704    .5475899
-------------+----------------------------------------------------------------
        diff |   .0112582   .0609228                     -.1081483    .1306647
             |  under Ho:   .0609564     0.18   0.853
------------------------------------------------------------------------------
        diff = prop(6) - prop(88)                                     z =   0.1847
    Ho: diff = 0

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(Z < z) = 0.5733         Pr(|Z| < |z|) = 0.8535          Pr(Z > z) = 0.4267

.
```

Statistical hypothesis tests

Data and comparison type	Alternative hypotheses	Parametric test Stata command
Numerical; One mean	H_a: $\mu \neq \mu_a$ (two-sided) H_a: $\mu > \mu_a$ or $\mu < \mu_a$ (one-sided)	Z or t-test •ttest var1==hypoth val.*
Numerical; Two means, paired data	H_a: $\mu_1 \neq \mu_2$ (two-sided) H_a: $\mu_1 > \mu_2$ or $\mu < \mu_a$ (one-sided)	Paired t-test •ttest var1==var2*
Numerical; Two means, independent data	H_a: $\mu_1 \neq \mu_2$ (two-sided) H_a: $\mu_1 > \mu_2$ or $\mu < \mu_a$ (one-sided)	T-test (equal or unequal variance) •ttest var1, by(byvar) •ttest var1, by(byvar) unequal
Numerical, Two or more means, independent data		
Dichotomous; One proportion	H_a: $p \neq p_a$ (two-sided) H_a: $p > p_a$ or $p < p_a$ (one-sided)	Proportion test • prtest var1=hypoth value* • bitest var1=hypoth value
Dichotomous; two proportions	H_a: $p_1 \neq p_2$ (two-sided) H_a: $p_1 > p_2$ (one-sided)	Proportion test (z-test) • prtest var1, by(byvar)
Categorical by categorical (nxk)		

For next time

- Read Pagano and Gauvreau
  - Pagano and Gavreau Chapters 11, and 14 (pages 332-338)
  - Pagano and Gavreau Chapter 12-13