Confidence Intervals & Hypothesis Testing

- Paired t-test
- Types of error

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)
- 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) $H_0: \mu_1 - \mu_2 = \delta$ $H_A: \mu_1 - \mu_2 \neq \delta$

  Often $\delta=0$, no difference So $H_0: \mu_1 - \mu_2 = 0$ OR $H_0: \mu_1 = \mu_2$

  $H_A: \mu_1 - \mu_2 \neq 0$, OR $H_A: \mu_1 \neq \mu_2$

- 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

- 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?

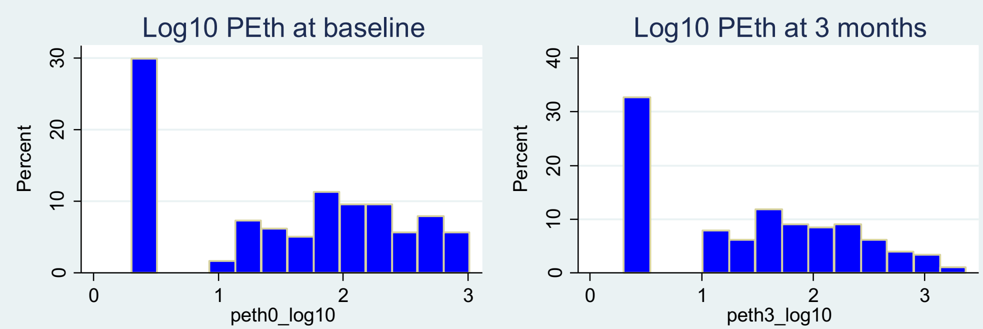
Comparison of two means: *paired* t-test
Example: BREATH Study

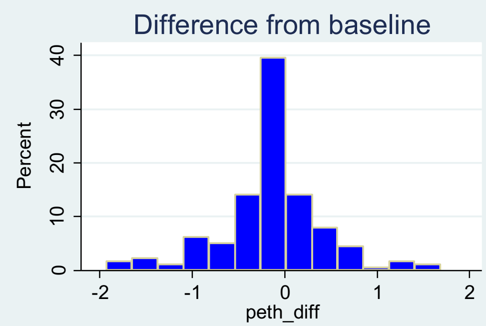
- We have found that self-reported alcohol consumption decreases after ART initiation, however we suspect a high degree of under-report. We use the biomarker PEth to obtain a biological measure of alcohol use in a one-year prospective study. Looking at baseline and 3 months…

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

  H_0: μ_2-μ_1 =0 → $\mu_2=\mu_1$ H_A: μ_1-$\mu_2 \neq 0$ → $\mu_2 \neq \mu_1$
- Significance level=0.05

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





The mean of the paired difference is a new random variable

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

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

```
. 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 hypothesis

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

Comparison of two means: *paired* t-test – another method

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

    .
```

Comparison of two *independent* means

- The goal is to compare means from two independent samples
- Two different populations
- 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
- So we can calculate a t or a z-score for the difference in the means and compare it to the standard normal distribution.

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

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

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

Comparison of two *independent* means

- 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

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
```

Comparison of two *independent* means

- 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

Comparison of two *independent* means

- 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

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
```

Comparison of two *independent* means

- 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
```

- 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
```

Comparison of two *independent* means

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
```

Comparison of two *independent* means

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) 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
```

Comparison of two proportions

- 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.
- So we can use the z statistic to find the probability of observing a difference as large as we do, under the null hypothesis of no difference

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

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

Comparison of two proportions

Example: Proportion with PEth>=50 ng/ml

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

.


.
```