JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

Section D

Two Sample t-test, Two Choices

FYI: Equal Variances Assumption

- The "traditional" t-test assumes equal variances in the two groups
  - This can be formally tested with another hypothesis test!
  - But why not just compare observed values of s1 to s2?

- There is a slight modification to allow for unequal variances—this modification adjusts the degrees of freedom for the test, using slightly different SE computation (the formula I give you)

- If you want to be truly "safe" (desert island choice of t-test)
  - More conservative to use test that allows for unequal variance

- Makes little to no difference in large sample

FYI: Equal Variances Assumption

- Actually, the following occurs:
  - If underlying population level standard deviations are equal:
    - Both approaches give valid confidence intervals but intervals by approach assuming unequal standard deviations slightly wider (and p-values slightly larger)
  - If underlying population level standard deviations are not equal:
    - The approach assuming equal variances does not give valid confidence intervals and can severely under-cover the goal of 95%

Unequal SD Approach: Diet Type/ Weight Loss Example

- Command syntax:
  - ttesti n_1 $\bar{x}_1$ s_1 n_2 $\bar{x}_2$ s_2, unequal

```
. ttesti 64 -5.7 8.6 68 -1.8 3.9, unequal

Two-sample t test with unequal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |      64        -5.7       1.075         8.6   -7.848216   -3.551784
       y |      68        -1.8    .4729445         3.9   -2.744001   -.8559989
---------+--------------------------------------------------------------------
combined |     132   -3.690909    .5978226    6.868458   -4.873545   -2.508273
---------+--------------------------------------------------------------------
    diff |                -3.9    1.174437               -6.234436   -1.565564
------------------------------------------------------------------------------
    diff = mean(x) - mean(y)                                      t =  -3.3207
Ho: diff = 0                     Satterthwaite's degrees of freedom =  86.6941

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0007         Pr(|T| > |t|) = 0.0013          Pr(T > t) = 0.9993
```

Equal SD Approach: Diet Type/ Weight Loss Example

- Command syntax:
  - ttesti $n_1\ \bar{x}_1\ s_1\ n_2\ \bar{x}_2\ s_2$

```
. ttesti 64 -5.7 8.6 68 -1.8 3.9

Two-sample t test with equal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |      64        -5.7       1.075         8.6   -7.848216   -3.551784
       y |      68        -1.8    .4729445         3.9   -2.744001   -.8559989
---------+--------------------------------------------------------------------
combined |     132   -3.690909    .5978226    6.868458   -4.873545   -2.508273
---------+--------------------------------------------------------------------
    diff |                -3.9    1.151038               -6.177191   -1.622809
------------------------------------------------------------------------------
    diff = mean(x) - mean(y)                                      t =  -3.3882
Ho: diff = 0                                     degrees of freedom =      130

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0005         Pr(|T| > |t|) = 0.0009          Pr(T > t) = 0.9995
```

Unequal SD Approach: LDL/ Treatment Example

- Command syntax:
  - ttesti $n_1\ \bar{x}_1\ s_1\ n_2\ \bar{x}_2\ s_2$, unequal

```
. ttesti 11 -1.41 .55 12 -.32 .65, unequal

Two-sample t test with unequal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |      11       -1.41    .1658312         .55   -1.779495   -1.040505
       y |      12        -.32    .1876388         .65   -.7329903    .0929903
---------+--------------------------------------------------------------------
combined |      23   -.8413043    .1692296    .8115967   -1.192265   -.4903436
---------+--------------------------------------------------------------------
    diff |               -1.09    .2504163                -1.61095   -.5690505
------------------------------------------------------------------------------
    diff = mean(x) - mean(y)                                      t =  -4.3528
Ho: diff = 0                     Satterthwaite's degrees of freedom =  20.8813

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0001         Pr(|T| > |t|) = 0.0003          Pr(T > t) = 0.9999
```

Equal SD Approach: LDL/Treatment Example

- Command syntax:
  - ttesti n_1 $\bar{x}_1$ s_1 n_2 $\bar{x}_2$ s_2, unequal

```
. ttesti 11 -1.41 .55 12 -.32 .65

Two-sample t test with equal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |      11       -1.41    .1658312         .55   -1.779495   -1.040505
       y |      12        -.32    .1876388         .65   -.7329903    .0929903
---------+--------------------------------------------------------------------
combined |      23   -.8413043    .1692296    .8115967   -1.192265   -.4903436
---------+--------------------------------------------------------------------
    diff |               -1.09    .2523107               -1.614709   -.5652911
------------------------------------------------------------------------------
    diff = mean(x) - mean(y)                                      t =  -4.3201
Ho: diff = 0                                     degrees of freedom =       21

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0002         Pr(|T| > |t|) = 0.0003          Pr(T > t) = 0.9998
```