CHRISTOPHER PRENER, PH.D.
FALL, 2016

WEEK 10
LECTURE 11

QUANTITATIVE ANALYSIS

DIFFERENCE OF MEANS (2)

AGENDA

1. Follow-up
2. Macros
3. Difference of Means in Stata
4. Effect Sizes
5. Graphing Means
6. Power Analysis

1 FOLLOW-UP

2 MACROS

AUTOMATING YOUR WORK

- Automation helps
    - improve consistency,
    - manage repetitive tasks,
    - limit debugging,
    - and ultimately decrease time spent coding.

TYPES OF MACROS

LOCAL

STORE A STRING OR VALUE TEMPORARILY IN A SINGLE DO-FILE

GLOBAL

STORE A STRING OR VALUE THAT IS VALID FOR AN ENTIRE STATA SESSION

LOCAL MACROS

```
// storing lists or other strings:

local macroName "string"
. local varlist "var1 var2 var3"
```

DATASET EXPLORATION

```
. describe mpg weight length turn displacement gear_ratio
```

```
. summarize mpg weight length turn displacement gear_ratio
```

DATASET EXPLORATION

```
. describe mpg weight length turn displacement gear_ratio
```

```
. summarize mpg weight length turn displacement gear_ratio
```

DATASET EXPLORATION

```
. local coreVars "mpg weight length turn displacement gear_ratio"


. describe `coreVars'


. summarize `coreVars'
```

DATASET EXPLORATION

```
. local coreVars "mpg weight length turn displacement gear_ratio"


. describe `coreVars'


. summarize `coreVars'
```

SETTING GRAPH OPTIONS

```
. histogram mpg, frequency scheme(s2mono) title("Miles per
Gallon") subtitle("1978 Automobiles") note("Graph produced by
Christopher Prener, Ph.D.; Data via Stata")
```

```
. histogram weight, frequency scheme(s2mono) title("Vehicle
Weight") subtitle("1978 Automobiles") note("Graph produced by
Christopher Prener, Ph.D.; Data via Stata")
```

```
. histogram displacement, frequency scheme(s2mono) title("Vehicle
Displacement") subtitle("1978 Automobiles") note("Graph produced
by Christopher Prener, Ph.D.; Data via Stata")
```

SETTING GRAPH OPTIONS

```
. histogram mpg, frequency scheme(s2mono) title("Miles per Gallon") subtitle("1978 Automobiles") note("Graph produced by Christopher Prener, Ph.D.; Data via Stata")
```

```
. histogram weight, frequency scheme(s2mono) title("Vehicle Weight") subtitle("1978 Automobiles") note("Graph produced by Christopher Prener, Ph.D.; Data via Stata")
```

```
. histogram displacement, frequency scheme(s2mono) title("Vehicle Displacement") subtitle("1978 Automobiles") note("Graph produced by Christopher Prener, Ph.D.; Data via Stata")
```

SETTING GRAPH OPTIONS

```
. local graphOptions "frequency scheme(s2mono) subtitle("1978
Automobiles") note("Graph produced by Christopher Prener, Ph.D.;
Data via Stata")"
```

```
. histogram mpg, title("Miles per Gallon") `graphOptions'
```

```
. histogram weight, title("Vehicle Weight") `graphOptions'
```

```
. histogram displacement, title("Vehicle Displacement")
`graphOptions'
```

SETTING GRAPH OPTIONS

```
. local graphOptions "frequency scheme(s2mono) subtitle("1978
Automobiles") note("Graph produced by Christopher Prener, Ph.D.;
Data via Stata")"
```

```
. histogram mpg, title("Miles per Gallon") `graphOptions'
```

```
. histogram weight, title("Vehicle Weight") `graphOptions'
```

```
. histogram displacement, title("Vehicle Displacement")
`graphOptions'
```

LOCAL MACROS

```
// storing values:

local macroName = val
```

```
. local x = 1
```

WORKING WITH VALUES

```
. summarize mpg
```

```
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
         mpg |         74     21.2973    5.785503         12         41
```

```
. return list

scalars:
                  r(N) =  74
              r(sum_w) =  74
               r(mean) =  21.2972972972973
                r(Var) =  33.47204738985561
                 r(sd) =  5.785503209735141
                r(min) =  12
                r(max) =  41
                r(sum) =  1576
```

WORKING WITH VALUES

```
. return list

scalars:
                  r(N) =  74
              r(sum_w) =  74
               r(mean) =  21.2972972972973
                r(Var) =  33.47204738985561
                 r(sd) =  5.785503209735141
               r(min) =  12
               r(max) =  41
               r(sum) =  1576

. local mpgMean = `r(mean)'
```

WORKING WITH VALUES

```
. summarize mpg
```

```
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
         mpg |         74     21.2973    5.785503        12         41
```

```
. local mpgMean = `r(mean)'

. display "The average of the variable mpg is `mpgMean'."
The average of the variable mpg is 21.2972972972973.
```

WORKING WITH VALUES

```
. generate mpgHigh = mpg



. recode mpgHigh (1/21.297=0) (21.298/45=1)
(mpgHigh: 74 changes made)

. tabulate mpgHigh
    mpgHigh |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         43       58.11       58.11
          1 |         31       41.89      100.00
------------+-----------------------------------
      Total |         74      100.00
```

WORKING WITH VALUES

```
. generate mpgHigh = .

. replace mpgHigh = 0 if mpg <= `mpgMean'
(43 real changes made)

. replace mpgHigh = 1 if mpg > `mpgMean'
(31 real changes made)
```

```
. tabulate mpgHigh

    mpgHigh |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         43       58.11       58.11
          1 |         31       41.89      100.00
------------+-----------------------------------
      Total |         74      100.00
```

WORKING WITH VALUES

```
. summarize mpg
```

```
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
         mpg |         74     21.2973    5.785503        12         41
```

```
. local mpgMean = `r(mean)'

. txt "The mean of the miles per gallon variable is " ///
%10.3f `mpgMean' "."
```

WORKING WITH VALUES

surround static text with double quotes

```
. txt "The mean of the miles per gallon variable is " ///
%10.3f `mpgMean' "."
```

round values to 3 sig. digits

insert macro

more static text

trailing space

manage line length

WORKING WITH VALUES

```
. summarize mpg
```

```
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
         mpg |         74     21.2973    5.785503        12         41
```

```
. local mpgMean = `r(mean)'

. txt " The mean of the miles per gallon variable is " ///
%10.3f `mpgMean' "."
```

The mean of the miles per gallon variable is 21.297.

WORKING WITH VALUES

```
. summarize mpg
```

```
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
         mpg |         74     21.2973    5.785503        12         41
```

```
. // OFF

. local mpgMean = `r(mean)'

. // ON


. txt "**1.** The mean of the miles per gallon variable is " ///
%10.3f `mpgMean' "."
```

WORKING WITH VALUES

```
. quietly summarize mpg, detail


. return list

scalars:
                  r(N) =  74
              r(sum_w) =  74
               r(mean) =  21.2972972972973
                r(Var) =  33.47204738985561
                 r(sd) =  5.785503209735141
           r(skewness) =  .9487175964588155
           r(kurtosis) =  3.97500459645325
                r(sum) =  1576
                r(min) =  12
                r(max) =  41
                 r(p1) =  12
                 r(p5) =  14
                r(p10) =  14
                r(p25) =  18
                r(p50) =  20
                r(p75) =  25
                r(p90) =  29
                r(p95) =  34
                r(p99) =  41
```

WORKING WITH VALUES

```
. swilk mpg

                    Shapiro-Wilk W test for normal data

    Variable |        Obs       W           V         z       Prob>z
-------------+--------------------------------------------------------
         mpg |         74    0.94821      3.335     2.627    0.00430


. return list

scalars:
                  r(p) =  .0043009091676854
                  r(z) =  2.627486783178489
                  r(V) =  3.334909342486645
                  r(W) =  .9482147278617608
                  r(N) =  74
```

WORKING WITH VALUES

```
. sfrancia mpg

                  Shapiro-Francia W' test for normal data

    Variable |       Obs       W'           V'       z       Prob>z
-------------+--------------------------------------------------------
         mpg |        74    0.94872      3.650     2.510    0.00604


. return list

scalars:
                  r(z) =  2.50953087416346
                  r(V) =  3.649896467022918
                  r(W) =  .9487225945150459
                  r(N) =  74
                  r(p) =  .0060445823550852
```

SUBSAMPLES

```
// restrict analyses to domestic vehicles with below average mpg
. local insample foreign == 0 & mpg <= `mpgMean'
. summarize price if `insample'
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       price |         36    6839.778    3433.552        3291      15906
```

GLOBAL MACROS

```
// storing lists or other strings:

global macroName "string"
```

```
. global varlist "var1 var2 var3"
```

```
// storing values:

global macroName = val
```

```
. global x = 1
```

GLOBAL MACROS

```
// returning lists or other strings:

. summarize $varlist
```

```
// returning values:

. display $x
```

GLOBAL MACROS

```
// ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

// check to see if appropriate directories exist

global projName "mpgMeans"

capture mkdir $projName


// ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

// log process

log using "$projName/$projName.txt", text replace
```

3 DIFFERENCE OF MEANS IN STATA

ONE SAMPLE T-TEST

```
ttest var=mu
```

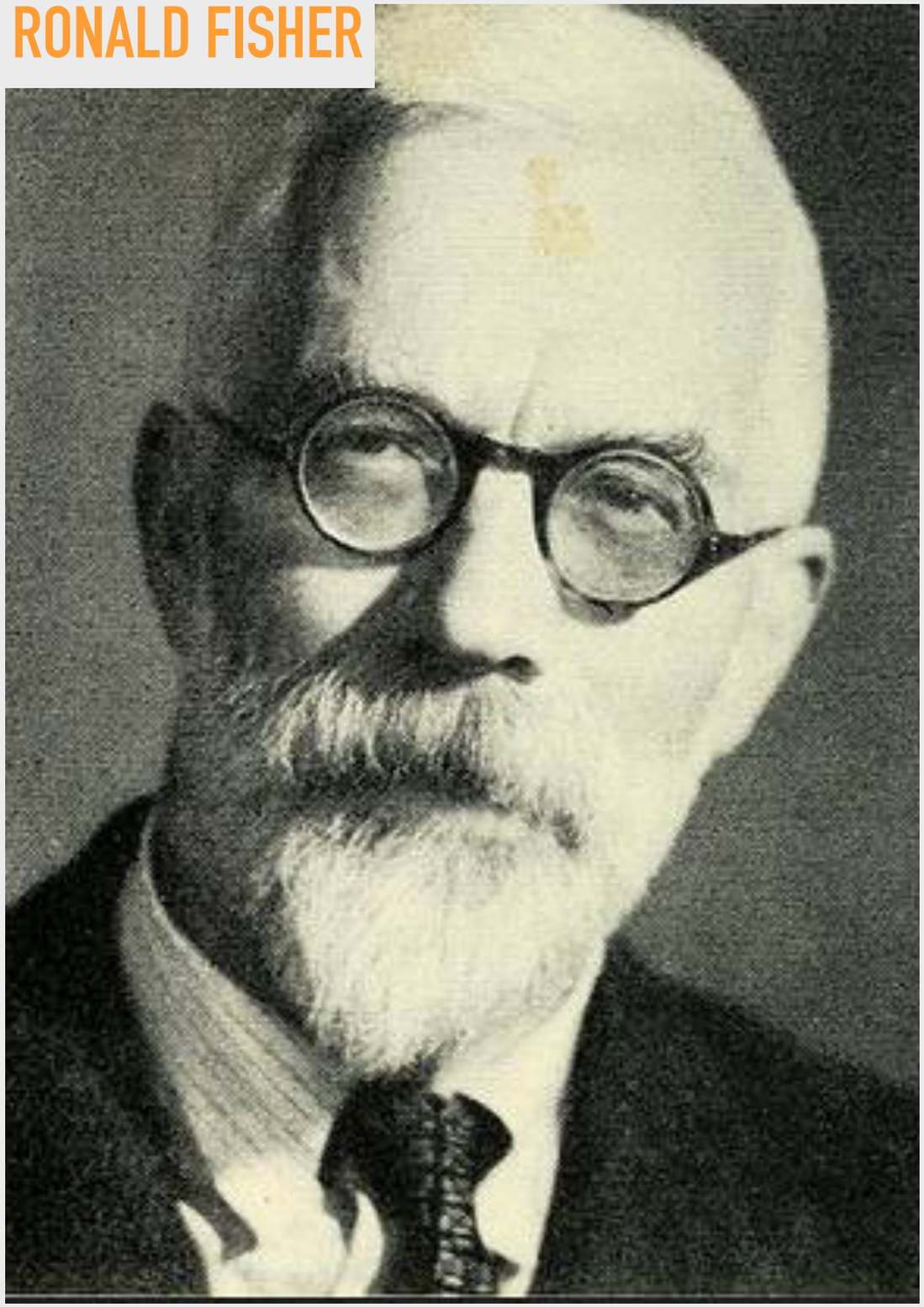
```
. ttest mpg=25

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
     mpg |      74     21.2973    .6725511    5.785503     19.9569    22.63769
------------------------------------------------------------------------------
    mean = mean(mpg)                                              t =  -5.5055
Ho: mean = 25                                  degrees of freedom =       73

    Ha: mean < 25               Ha: mean != 25                 Ha: mean > 25
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

LEVENE'S TEST

```
sdtest yVar, by(xVar)
```

```
. sdtest mpg, by(foreign)
Variance ratio test
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
Domestic |      52    19.82692     .657777    4.743297    18.50638    21.14747
 Foreign |      22    24.77273     1.40951    6.611187    21.84149    27.70396
---------+--------------------------------------------------------------------
combined |      74     21.2973    .6725511    5.785503     19.9569    22.63769
------------------------------------------------------------------------------
    ratio = sd(Domestic) / sd(Foreign)                             f =   0.5148
Ho: ratio = 1                                  degrees of freedom =   51, 21

    Ha: ratio < 1               Ha: ratio != 1                 Ha: ratio > 1
  Pr(F < f) = 0.0275         2*Pr(F < f) = 0.0549           Pr(F > f) = 0.9725
```

RONALD FISHER



F-DISTRIBUTION

- Named in honor of Ronald Fisher
- Models the distribution of the ratio between two groups based on their variance
- Used to test whether two estimates of variance can be assumed to come from the same population
- Not symmetrical like *t*, and its shape varies based on the given degrees of freedom

LEVENE'S TEST

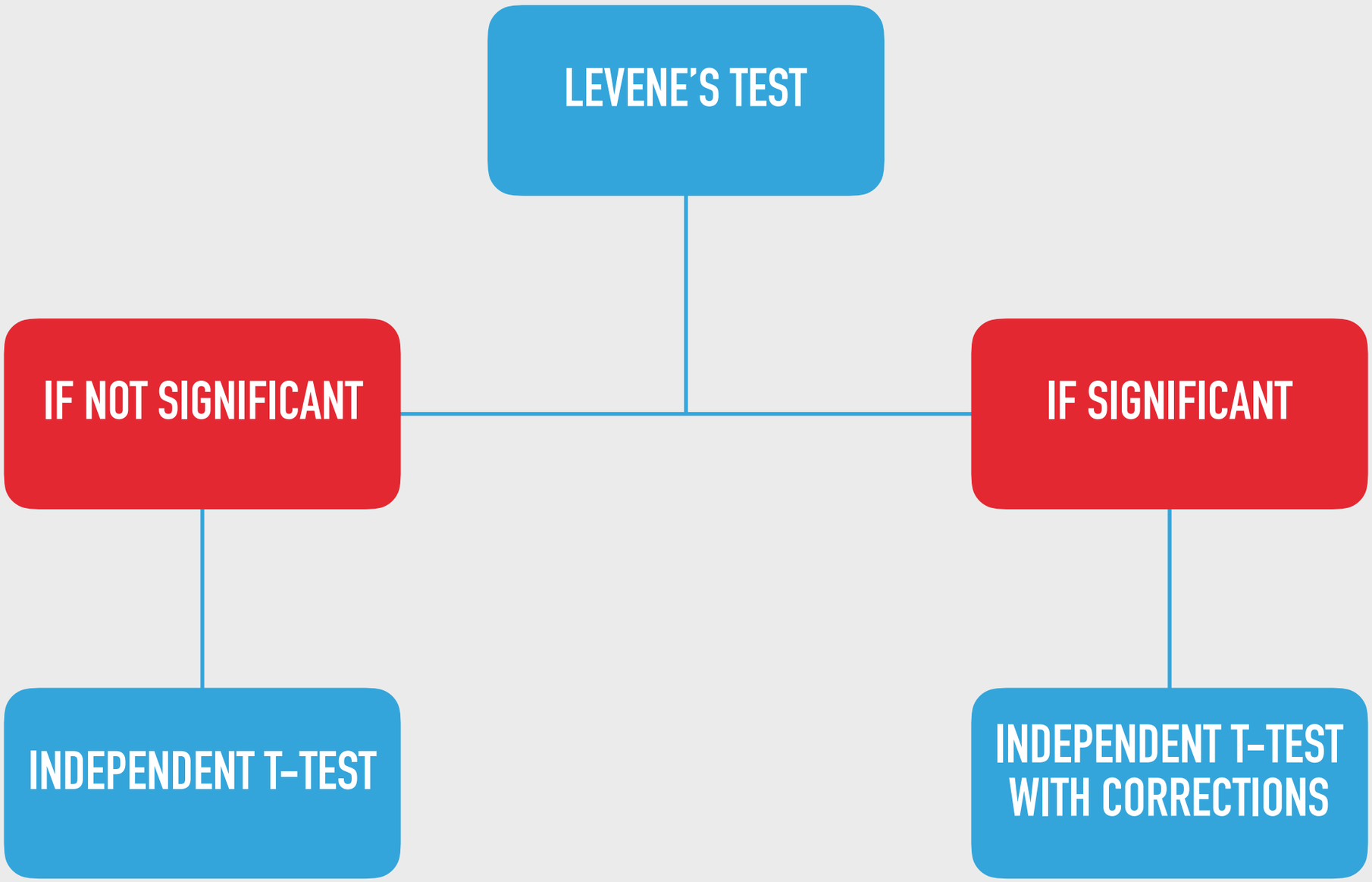
```
sdtest yVar, by(xVar)
```

```
. sdtest mpg, by(foreign)
Variance ratio test
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
Domestic |      52    19.82692     .657777    4.743297    18.50638    21.14747
 Foreign |      22    24.77273     1.40951    6.611187    21.84149    27.70396
---------+--------------------------------------------------------------------
combined |      74     21.2973    .6725511    5.785503     19.9569    22.63769
------------------------------------------------------------------------------
    ratio = sd(Domestic) / sd(Foreign)                             f =   0.5148
Ho: ratio = 1                                  degrees of freedom =   51, 21

    Ha: ratio < 1               Ha: ratio != 1                 Ha: ratio > 1
  Pr(F < f) = 0.0275         2*Pr(F < f) = 0.0549           Pr(F > f) = 0.9725
```

LEVENE'S TEST SCALARS

```
. quietly sdtest mpg, by(foreign)

. return list

scalars:
                  r(sd) =  5.785503209735139
                   r(p) =  .054948631637733
                 r(p_l) =  .027474315818867
                 r(p_u) =  .972525684181133
               r(sd_2) =  6.611186898567625
               r(sd_1) =  4.743297247514701
                   r(F) =  .514756468853841
               r(df_2) =  21
               r(df_1) =  51
                   r(N) =  74
```

LEVENE'S TEST



INDEPENDENT SAMPLES T-TEST

```
ttest yVar, by(xVar)
```

```
. ttest mpg, by(foreign)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
Domestic |      52    19.82692     .657777    4.743297    18.50638    21.14747
 Foreign |      22    24.77273     1.40951    6.611187    21.84149    27.70396
---------+--------------------------------------------------------------------
combined |      74     21.2973    .6725511    5.785503     19.9569    22.63769
---------+--------------------------------------------------------------------
    diff |           -4.945804    1.362162               -7.661225   -2.230384
------------------------------------------------------------------------------
    diff = mean(Domestic) - mean(Foreign)                         t =  -3.6308
Ho: diff = 0                                     degrees of freedom =       72

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0003         Pr(|T| > |t|) = 0.0005          Pr(T > t) = 0.9997
```

INDEPENDENT SAMPLES T-TEST

```
ttest yVar, by(xVar)
```

```
. ttest mpg, by(foreign) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
Domestic |      52    19.82692     .657777    4.743297    18.50638    21.14747
 Foreign |      22    24.77273     1.40951    6.611187    21.84149    27.70396
---------+--------------------------------------------------------------------
combined |      74     21.2973    .6725511    5.785503     19.9569    22.63769
---------+--------------------------------------------------------------------
    diff |           -4.945804    1.555438               -8.11642   -1.775188
------------------------------------------------------------------------------
    diff = mean(Domestic) - mean(Foreign)                         t =  -3.1797
Ho: diff = 0                     Welch's degrees of freedom =  31.4209

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0017         Pr(|T| > |t|) = 0.0033          Pr(T > t) = 0.9983
```

DEPENDENT SAMPLES T-TEST

```
ttest varG1==varG2
```

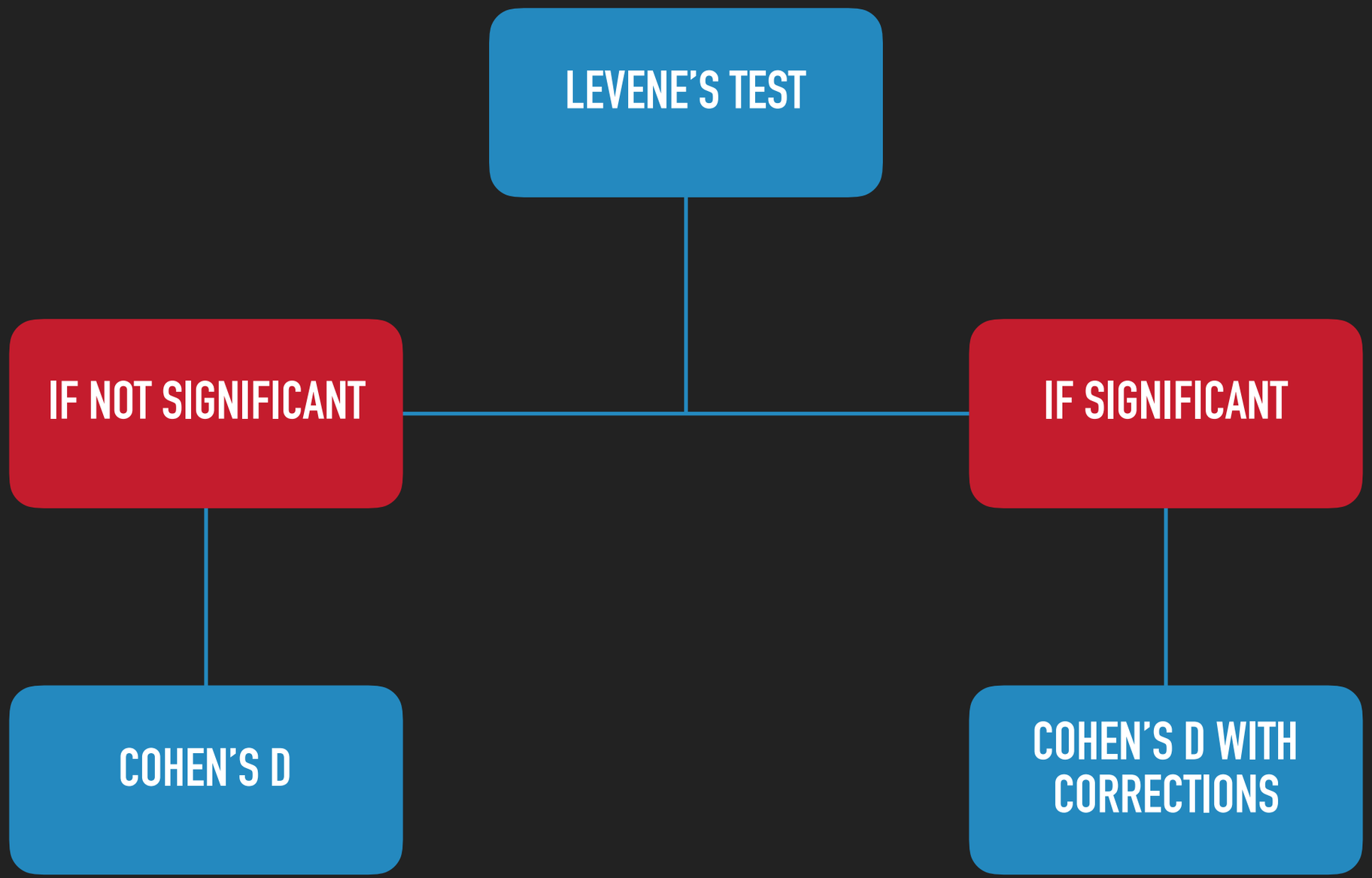
```
. use http://www.ats.ucla.edu/stat/stata/notes/hsb2, clear
. ttest read==write
Paired t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
    read |     200       52.23    .7249921    10.25294    50.80035    53.65965
   write |     200      52.775    .6702372    9.478586    51.45332    54.09668
---------+--------------------------------------------------------------------
    diff |     200       -.545    .6283822    8.886666   -1.784142    .6941424
------------------------------------------------------------------------------
     mean(diff) = mean(read - write)                               t =  -0.8673
 Ho: mean(diff) = 0                         degrees of freedom =      199

 Ha: mean(diff) < 0           Ha: mean(diff) != 0           Ha: mean(diff) > 0
 Pr(T < t) = 0.1934         Pr(|T| > |t|) = 0.3868          Pr(T > t) = 0.8066
```

T-TEST SCALARS

```
. use http://www.ats.ucla.edu/stat/stata/notes/hsb2, clear

. quietly ttest read==write

. return list

scalars:
                  r(level) =  95
                 r(sd_2) =  9.47858602138653
                 r(sd_1) =  10.25293682648241
                  r(se) =  .6283822053335848
                 r(p_u) =  .8065906589542506
                 r(p_l) =  .1934093410457493
                   r(p) =  .3868186820914985
                   r(t) =  -.8673065458794776
                r(df_t) =  199
                r(mu_2) =  52.775
                 r(N_2) =  200
                r(mu_1) =  52.23
                 r(N_1) =  200
```

4 EFFECT SIZES

COHEN'S D AFTER INDEPENDENT T-TEST



COHEN'S D AFTER INDEPENDENT T-TEST

```
. esize twosample mpg, by(foreign) cohensd

Effect size based on mean comparison

                                Obs per group:
                                 Domestic =         52
                                  Foreign =         22
------------------------------------------------------------------------------
        Effect Size |   Estimate     [95% Conf. Interval]
--------------------+---------------------------------------------------------
          Cohen's d |  -.9234449    -1.441225   -.3997744
------------------------------------------------------------------------------
```

COHEN'S D AFTER INDEPENDENT T-TEST

```
. esize twosample mpg, by(foreign) cohensd unequal welch

Effect size based on mean comparison, unequal variances

                                  Obs per group:
                                 Domestic =          52
                                  Foreign =          22
------------------------------------------------------------------------------
        Effect Size |   Estimate     [95% Conf. Interval]
--------------------+---------------------------------------------------------
          Cohen's d |  -.9234449    -1.465242   -.3694511
------------------------------------------------------------------------------
                     Welch's degrees of freedom =  31.4209
```

COHEN'S D SCALARS

```
. quietly esize twosample mpg, by(foreign) cohensd

. return list

scalars:
                  r(d) =  -.9234448623222905
               r(lb_d) =  -1.441225213139341
               r(ub_d) =  -.3997744112393538
                r(N_1) =  52
                r(N_2) =  22
              r(level) =  95
```

5 GRAPHING MEANS

ONE SAMPLE T-TEST

```
ttest var=mu
```

```
. ttest mpg=25

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
     mpg |      74     21.2973    .6725511    5.785503     19.9569    22.63769
------------------------------------------------------------------------------
    mean = mean(mpg)                                              t =  -5.5055
Ho: mean = 25                                  degrees of freedom =       73

    Ha: mean < 25               Ha: mean != 25                 Ha: mean > 25
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

EXTRACT KEY VALUES

```
// create dataset with key values

. collapse (mean) meanMpg = mpg (sd) sdMpg = mpg (count) n=mpg

// calculated standard error

. generate seMpg = sdMpg/sqrt(n)


// create category variable

. generate cat = 1
```

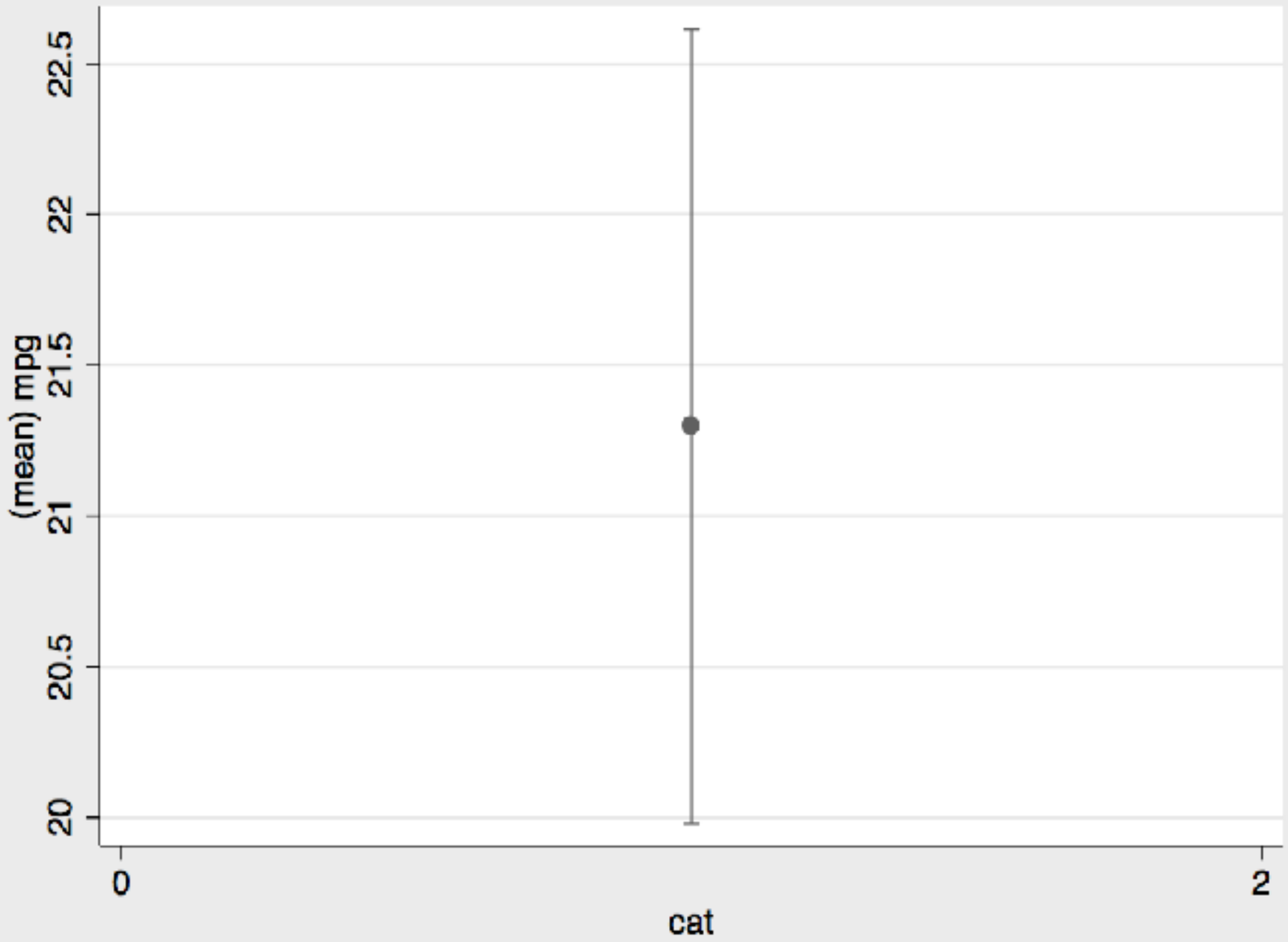
EXTRACT KEY VALUES

```
. ttest mpg=25

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
     mpg |      74     21.2973    .6725511    5.785503     19.9569    22.63769
------------------------------------------------------------------------------


. list n meanMpg seMpg sdMpg

     +-------------------------------------+
     |  n   meanMpg       seMpg    sdMpg  |
     |-------------------------------------|
  1. | 74   21.2973   .6725511   5.7855  |
     +-------------------------------------+
```

GRAPH MEAN AND 95% CONFIDENCE INTERVAL

```
serrbar mVar sVar xVar, scale(val)
```

```
. serrbar meanMpg seMpg cat, scale(1.96) scheme(s2mono)
```



One Sample T-Test: Miles per Gallon and μ=25

1978 Automobiles

Results: $t = -5.505$, $p < 0.001$

Produced by Christopher Prener, Ph.D.; Data via Stata

INDEPENDENT SAMPLES T-TEST

```
ttest yVar, by(xVar)
```

```
. ttest mpg, by(foreign)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
Domestic |      52    19.82692     .657777    4.743297    18.50638    21.14747
 Foreign |      22    24.77273     1.40951    6.611187    21.84149    27.70396
---------+--------------------------------------------------------------------
combined |      74     21.2973    .6725511    5.785503     19.9569    22.63769
---------+--------------------------------------------------------------------
    diff |           -4.945804    1.362162               -7.661225   -2.230384
------------------------------------------------------------------------------
    diff = mean(Domestic) - mean(Foreign)                         t =  -3.6308
Ho: diff = 0                                     degrees of freedom =       72

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0003         Pr(|T| > |t|) = 0.0005          Pr(T > t) = 0.9997
```

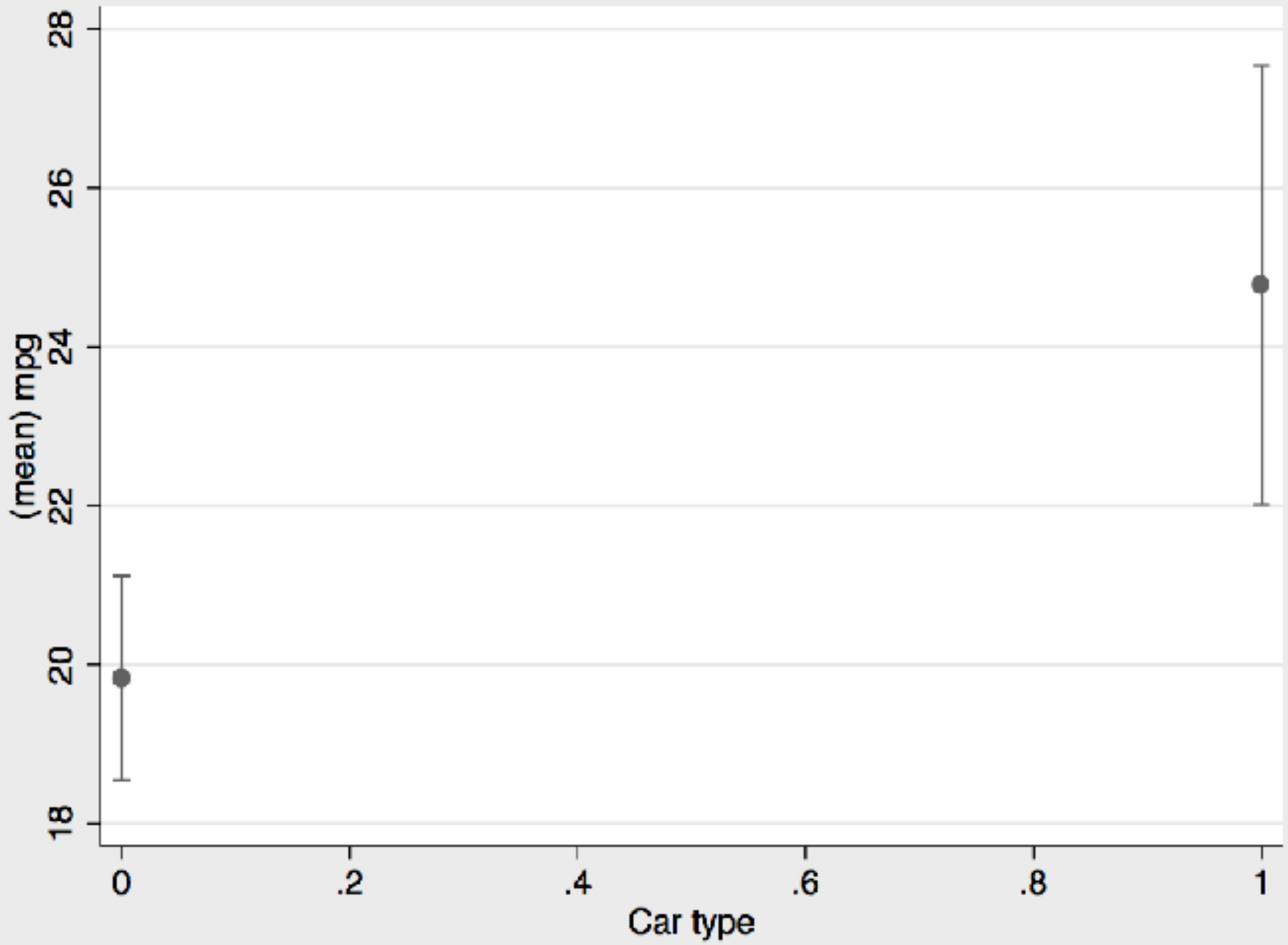
EXTRACT KEY VALUES

```
// create dataset with key values

. collapse (mean) meanMpg = mpg (sd) sdMpg = mpg (count) n=mpg, by(foreign)


// calculated standard error

. generate seMpg = sdMpg/sqrt(n)
```

GRAPH MEANS AND 95% CONFIDENCE INTERVAL

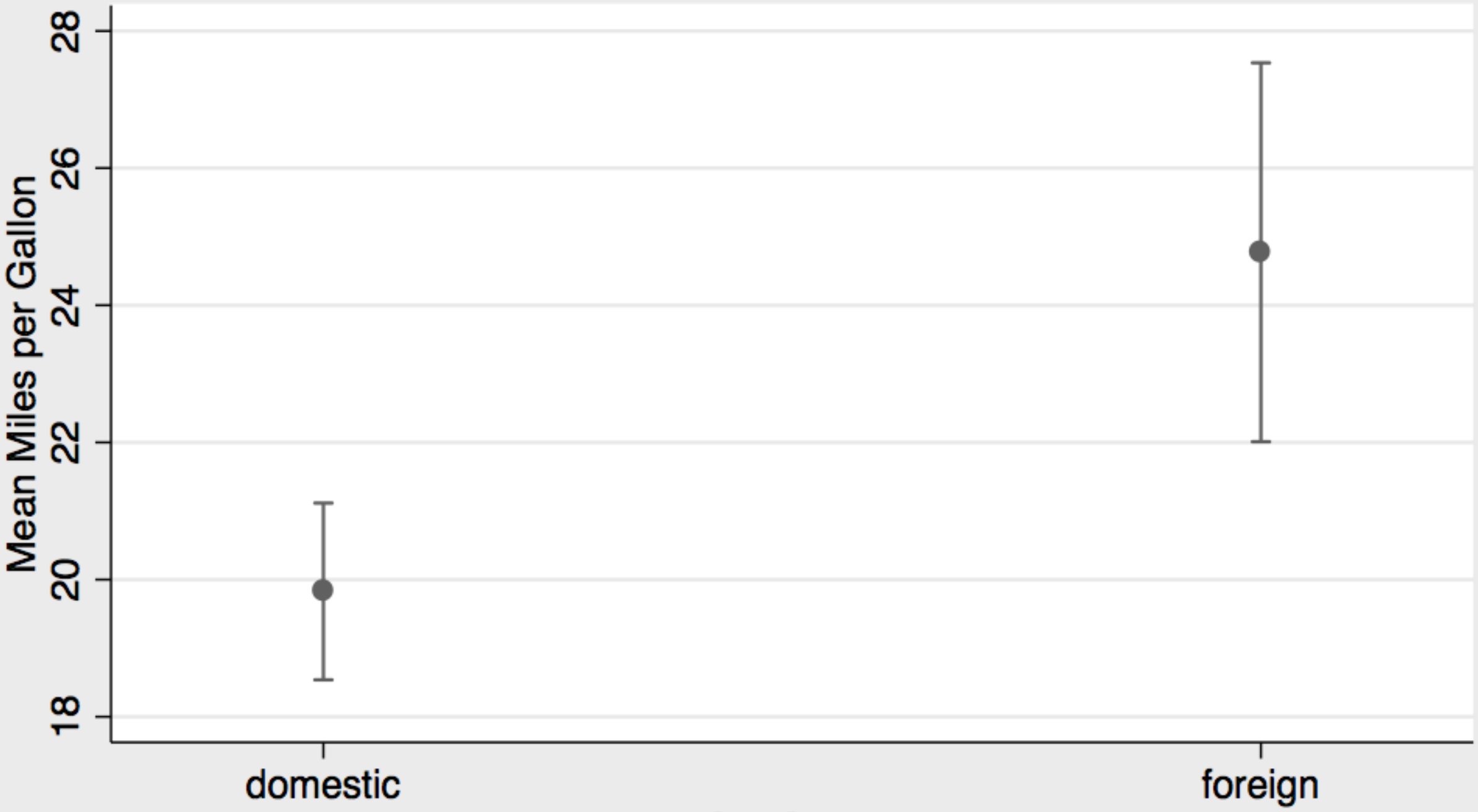
```
serrbar mVar sVar xVar, scale(val)
```

```
. serrbar meanMpg seMpg foreign, scale(1.96) scheme(s2mono)
```



Two Sample T-Test: Miles per Gallon by Car Type

1978 Automobiles



Results: t = -3.631, p = .0005, Cohen's d = -.923

Produced by Christopher Prener, Ph.D.; Data via Stata

DEPENDENT SAMPLES T-TEST

```
ttest varG1==varG2
```

```
. use http://www.ats.ucla.edu/stat/stata/notes/hsb2, clear
. ttest read==write
Paired t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
    read |     200       52.23    .7249921    10.25294    50.80035    53.65965
   write |     200      52.775    .6702372    9.478586    51.45332    54.09668
---------+--------------------------------------------------------------------
    diff |     200       -.545    .6283822    8.886666   -1.784142    .6941424
------------------------------------------------------------------------------
     mean(diff) = mean(read - write)                               t =  -0.8673
 Ho: mean(diff) = 0                         degrees of freedom =      199

 Ha: mean(diff) < 0           Ha: mean(diff) != 0           Ha: mean(diff) > 0
 Pr(T < t) = 0.1934         Pr(|T| > |t|) = 0.3868          Pr(T > t) = 0.8066
```

EXTRACT KEY VALUES

```
// save key values

. local meanRead = r(mu_1)
. local sdRead = r(sd_1)
. local nRead = r(N_1)

. local meanWrite = r(mu_2)
. local sdWrite = r(sd_2)
. local nWrite = r(N_2)
```

EXTRACT KEY VALUES

```
// create dataset with key values

clear

set obs 2

generate cat = .
generate mean = .
generate sd = .
generate n = .

replace cat = 0 in 1
replace mean = `meanRead' in 1
replace sd = `sdRead' in 1
replace n = `nRead' in 1

replace cat = 1 in 2
replace mean = `meanWrite' in 2
replace sd = `sdWrite' in 2
replace n = `nWrite' in 2
```

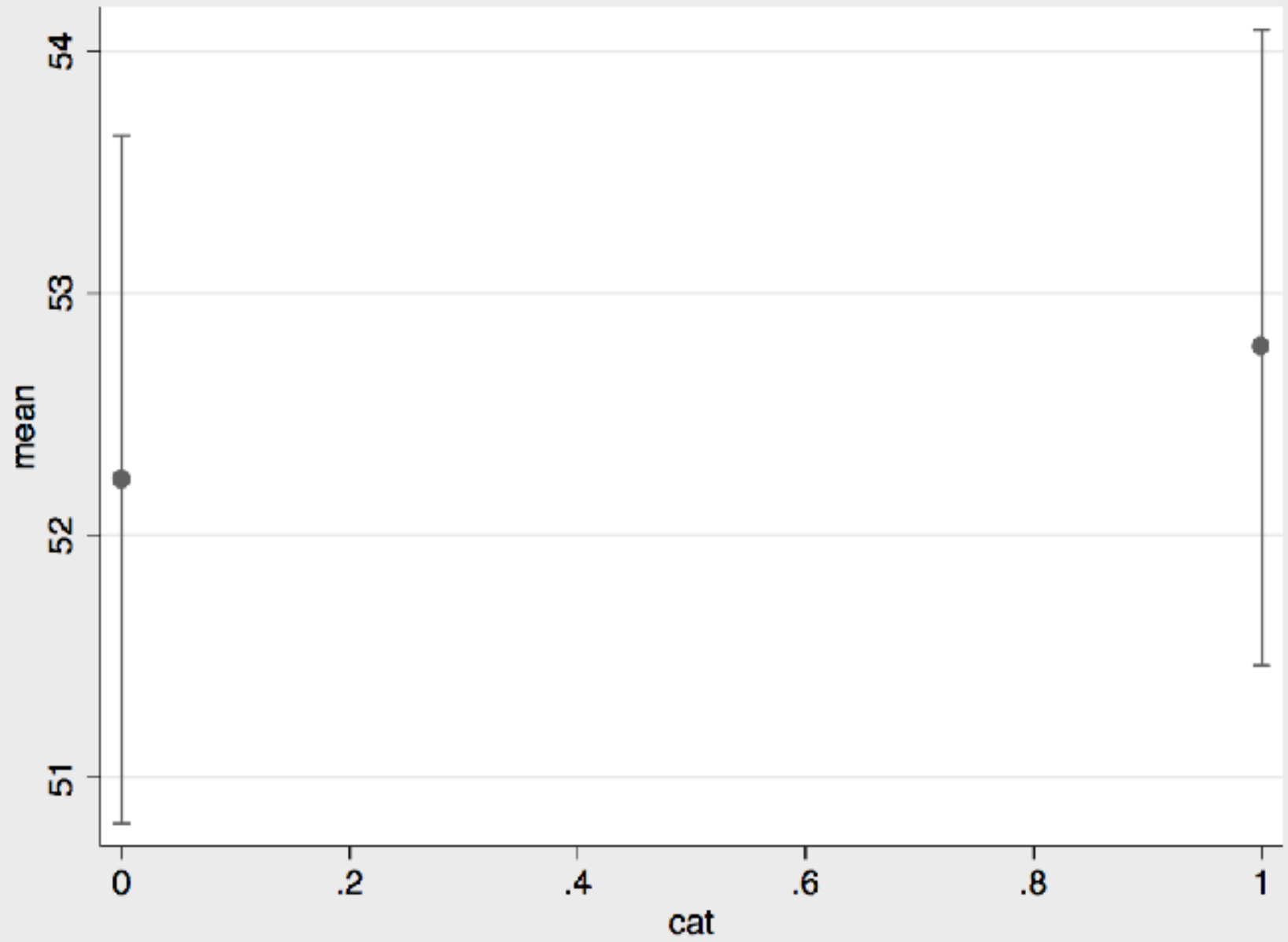
EXTRACT KEY VALUES

```
// calculated standard error

generate se = sd/sqrt(n)
```

GRAPH MEANS AND 95% CONFIDENCE INTERVAL

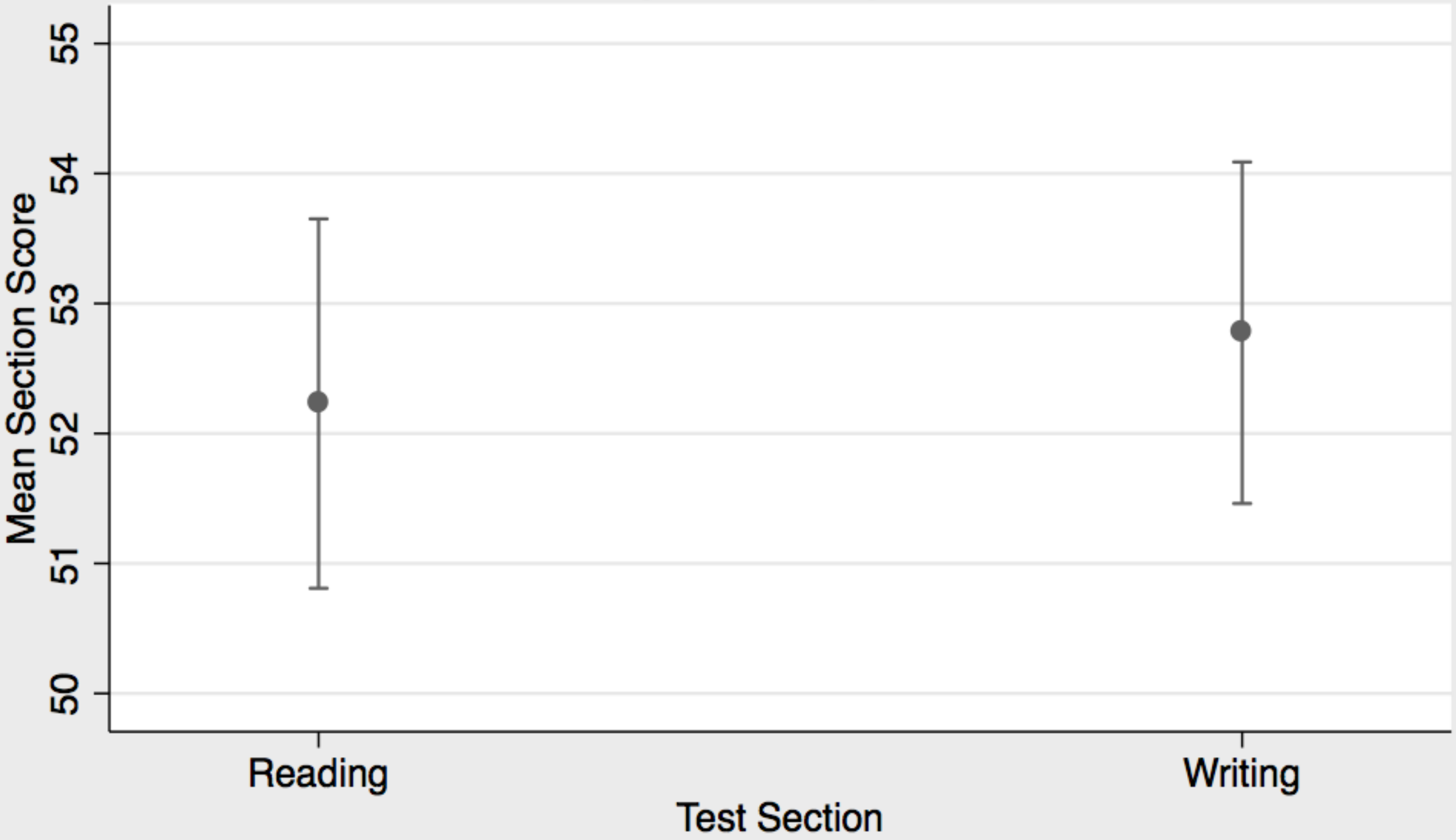
```
serrbar mVar sVar xVar, scale(val)
```

```
. serrbar mean se cat, scale(1.96) scheme(s2mono)
```



Dependent Sample T-Test: Standardized Testing

Reading and Writing Sections



Results: $t = -.867$, $p = .387$

Produced by Christopher Prener, Ph.D.; Data via Stata

6 POWER ANALYSIS

ERROR

Pr(Type II)=β

1-β=power

Sample	Population	
	$\mu = \mu_0$	$\mu \neq \mu_0$
Not Reject	yes	Type II
Reject	Type I	yes

*The null hypothesis is that $\mu=\mu_0$

Pr(Type I)=α

SCENARIO

- Scaled risk for foreclosure instrument with range 0 to 20

SCENARIO

- Scaled risk for foreclosure instrument with range 0 to 20
- Based on Acock (2016:178), we can estimate standard deviation by assuming that few people will be at either extreme of the distribution (< 2 or > 18):

$$s = \frac{18 - 2}{4} = 4$$

SCENARIO

- Scaled risk for foreclosure instrument with range 0 to 20
- Based on Acock (2016:178), we can estimate standard deviation by assuming that few people will be at either extreme of the distribution (< 2 or > 18):

$$s = \frac{18 - 2}{4} = 4$$

- If we want to detect a moderate size, we can multiply *s* by 0.5 (Cohen's *d*!), which equals 2.

SCENARIO

- Scaled risk for foreclosure instrument with range 0 to 20
- Based on Acock (2016:178), we can estimate standard deviation by assuming that few people will be at either extreme of the distribution (< 2 or > 18):

$$s = \frac{18 - 2}{4} = 4$$

- If we want to detect a moderate size, we can multiply *s* by 0.5 (Cohen's *d*!), which equals 2.
- We then pick two reasonable values that have a difference of 2 (12 and 14, 14 and 16, or 15 and 17, for example).

KEY ASSUMPTIONS

From Acock (2016:181):

- What is the clinically important difference?
- What is the standard deviation of each group?
- How much power do I want to have?
- What alpha level am I using?

KEY ASSUMPTIONS

From Acock (2016:181):

- What is the clinically important difference? 2; possible means 8 and 10
- What is the standard deviation of each group? 4
- How much power do I want to have? β=0.8 or β=0.9 are common
- What alpha level am I using? typically α=0.05

CALCULATING N

```
power twomeans mean1 mean2, sd(val) power(beta)
```

```
. power twomeans 8 10, sd(4) power(0.8)
{some output omitted}
Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    2.0000
           m1 =    8.0000
           m2 =   10.0000
           sd =    4.0000

Estimated sample sizes:

            N =       128
  N per group =        64
```

CALCULATING N

```
power twomeans mean1 mean2, sd(val) power(beta)
```

```
. power twomeans 8 10, sd(4) power(0.9)
{some output omitted}
Study parameters:

        alpha =    0.0500
        power =    0.9000
        delta =    2.0000
           m1 =    8.0000
           m2 =   10.0000
           sd =    4.0000

Estimated sample sizes:

            N =       172
  N per group =        86
```

CALCULATING N

```
power twomeans mean1 mean2, sd(val) power(beta)
```

```
. power twomeans 8 10, sd(4) power(0.95)
{some output omitted}
Study parameters:

        alpha =    0.0500
        power =    0.9500
        delta =    2.0000
           m1 =    8.0000
           m2 =   10.0000
           sd =    4.0000

Estimated sample sizes:

            N =       210
  N per group =       105
```

CALCULATING N

```
power twomeans mean1 mean2, sd(val) power(beta)
```

```
. power twomeans 7.8 11, sd(4) power(0.9)
{some output omitted}
Study parameters:

        alpha =    0.0500
        power =    0.9000
        delta =    3.2000
           m1 =    7.8000
           m2 =   11.0000
           sd =    4.0000

Estimated sample sizes:

            N =        68
  N per group =        34
```

DOCUMENT DETAILS

Document produced by Christopher Prener, Ph.D for the Saint Louis University course SOC 5050: QUANTITATIVE ANALYSIS - APPLIED INFERENTIAL STATISTICS. See the course wiki and the repository README.md file for additional details.