One sample hypothesis testing about the mean

One sample hypothesis testing about the mean

1. Set up the null hypothesis
$(H_0 : \mu = \mu_0)$
2. Set up the alternative
$(H_A : \mu \neq \mu_0)$

One sample hypothesis testing about the mean

1. Set up the null hypothesis

$(H_0 : \mu = \mu_0)$

2. Set up the alternative

$(H_A : \mu \neq \mu_0)$

3. Choose α-level

$(\alpha = 0.05)$

One sample hypothesis testing about the mean

1. Set up the null hypothesis

$(H_0 : \mu = \mu_0)$

2. Set up the alternative

$(H_A : \mu \neq \mu_0)$

3. Choose α-level

$(\alpha = 0.05)$

4. Take a sample and calculate

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \quad \text{or} \quad t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

5. Calculate appropriate p-value.

5. Calculate appropriate p-value.

6. Compare p-value to α.

7. Either reject H_0 , or not.

5. Calculate appropriate p-value.

6. Compare p-value to α.

7. Either reject H_0 , or not.

Alternatively,

5. Find cutoff (±1.96)

6. Compare z or t to cutoff

7. Either reject H_0 , or not.

Comparative Situations – Two samples

- Before and After
- Treatment and Control
- Two groups

$$H_0 : \mu_1 - \mu_2 = \Delta$$

Comparative Situations – Two samples

- Before and After ✓ on same person
- Treatment and Control
- Two groups

$$H_0 : \mu_1 - \mu_2 = \Delta$$

Comparative Situations – Two samples

- Before and After
- Treatment and Control
- Two groups

$$H_0 : \mu_1 - \mu_2 = \Delta$$

Comparative Situations – Two samples

- Before and After
- Treatment and Control
- Two groups

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

$$H_0 : \mu_1 - \mu_2 = \Delta \quad ?$$

$$t = \frac{\bar{X}_1 - \bar{X}_2 - \Delta}{\text{standard deviation}}$$

Question: Are two samples independent?

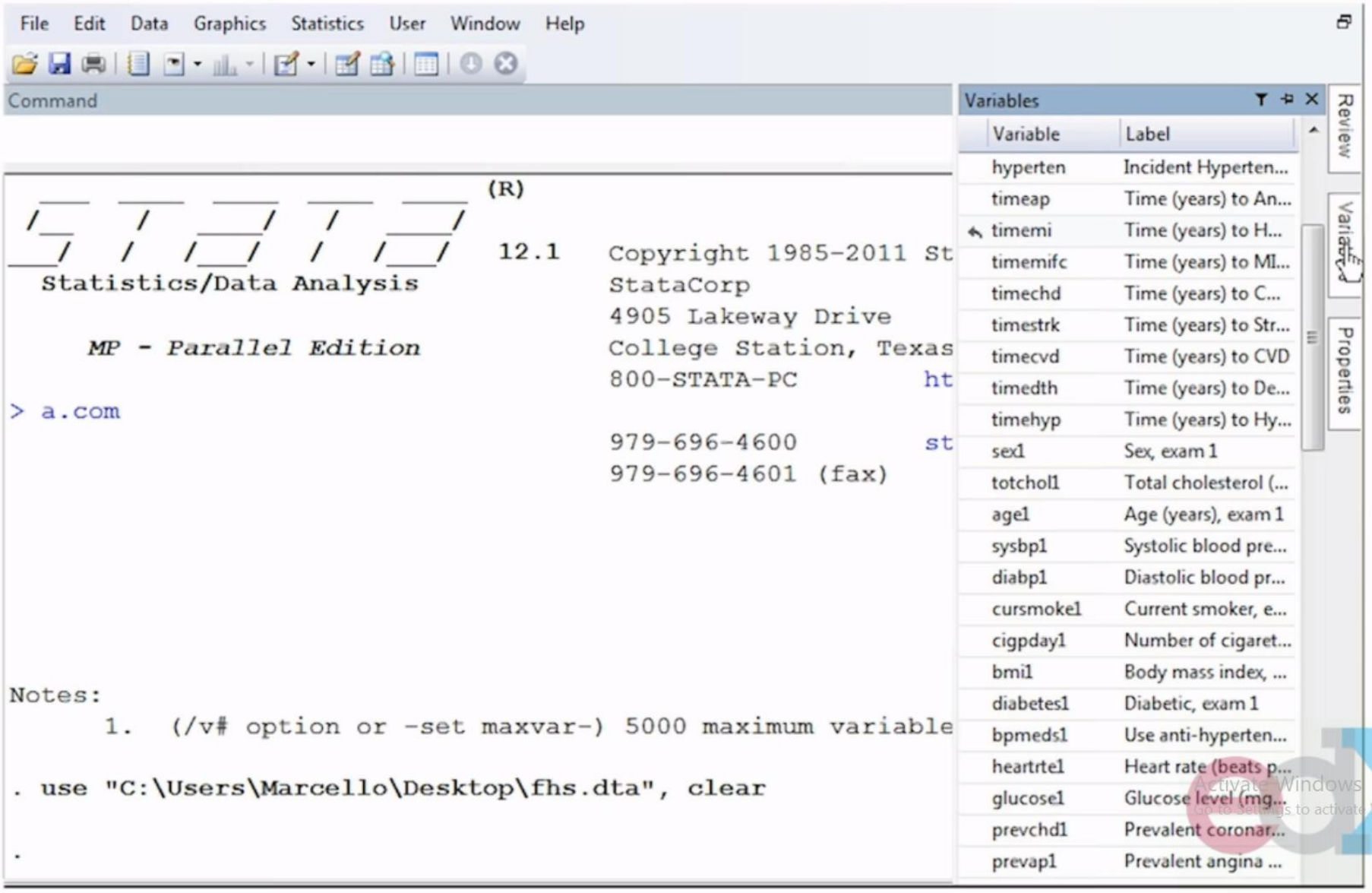
1. No (dependent, before/after)
2. Yes (different people)

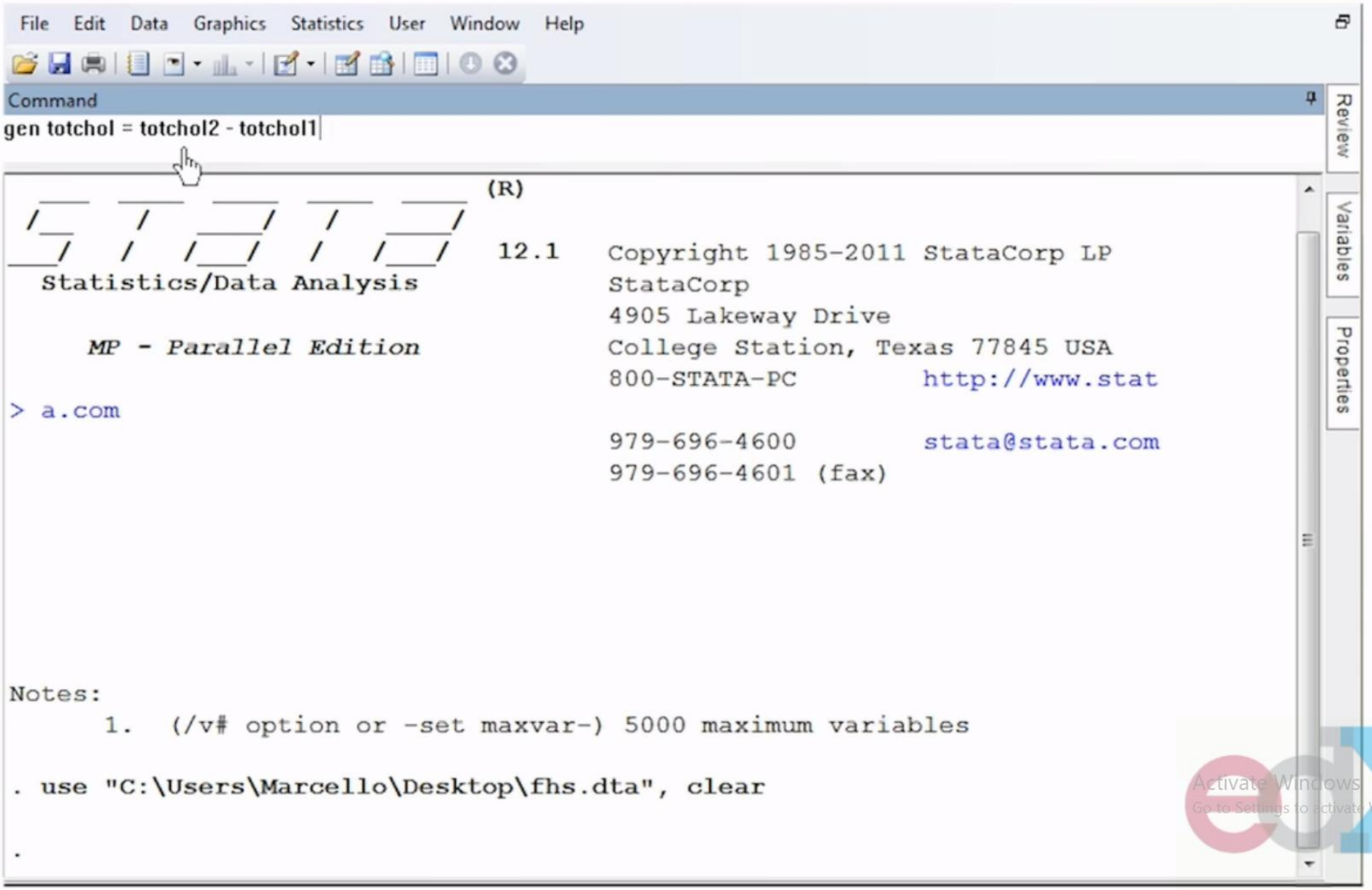
$$H_0 : \mu_1 - \mu_2 = \Delta \ ?$$

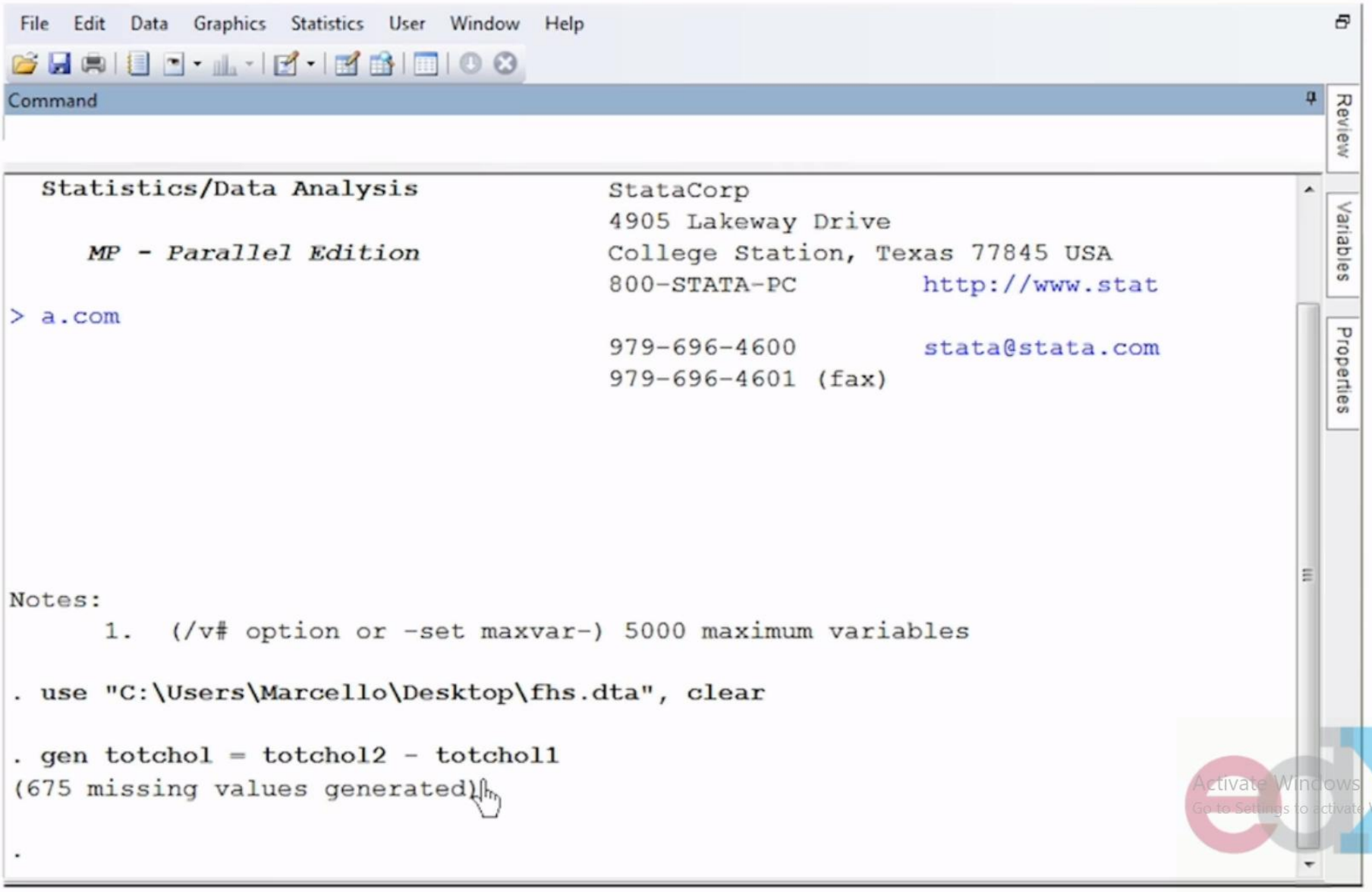
$$t = \frac{\bar{X}_1 - \bar{X}_2 - \Delta}{\text{standard deviation}}$$

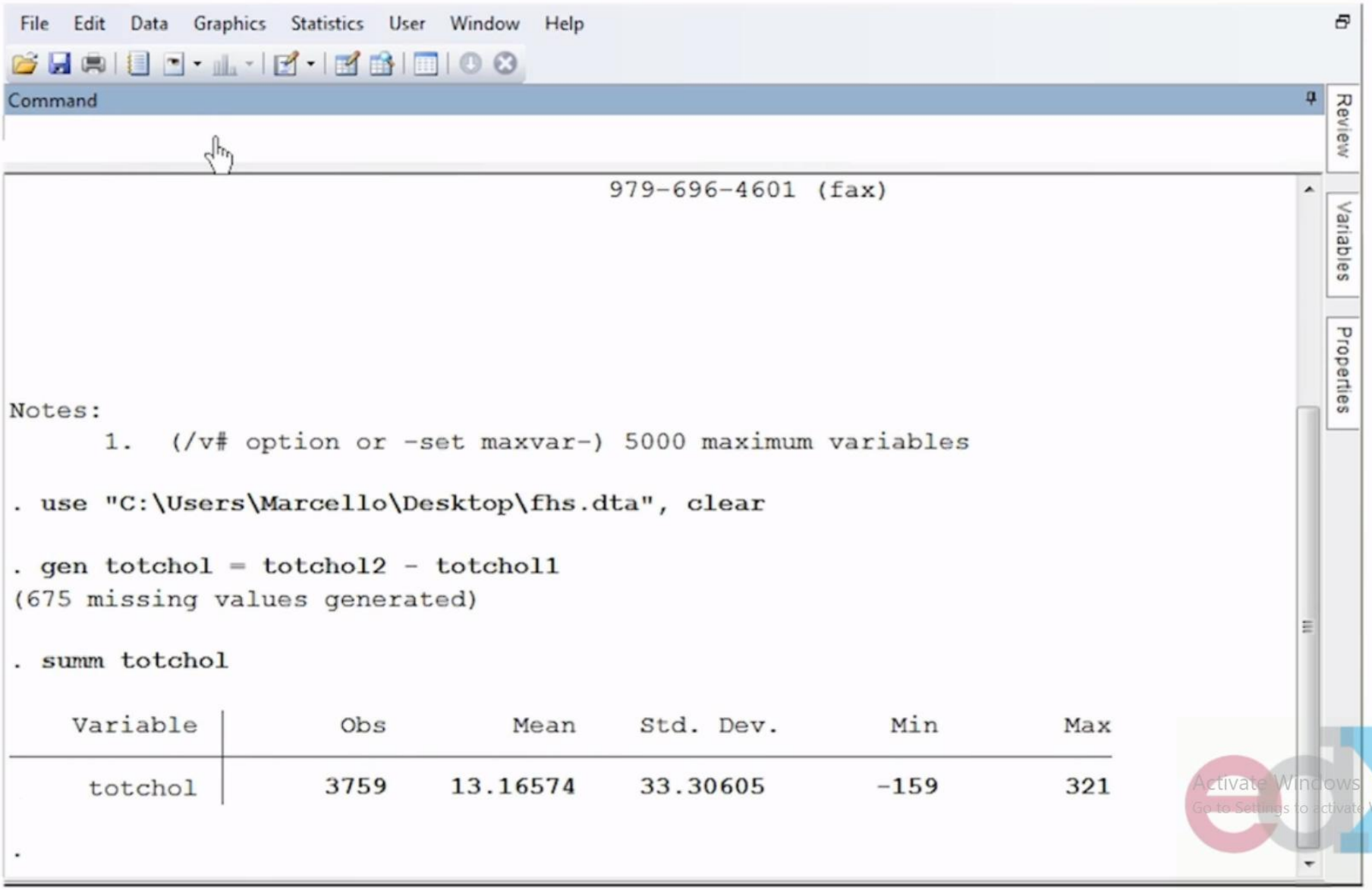
Question: Are two samples independent?

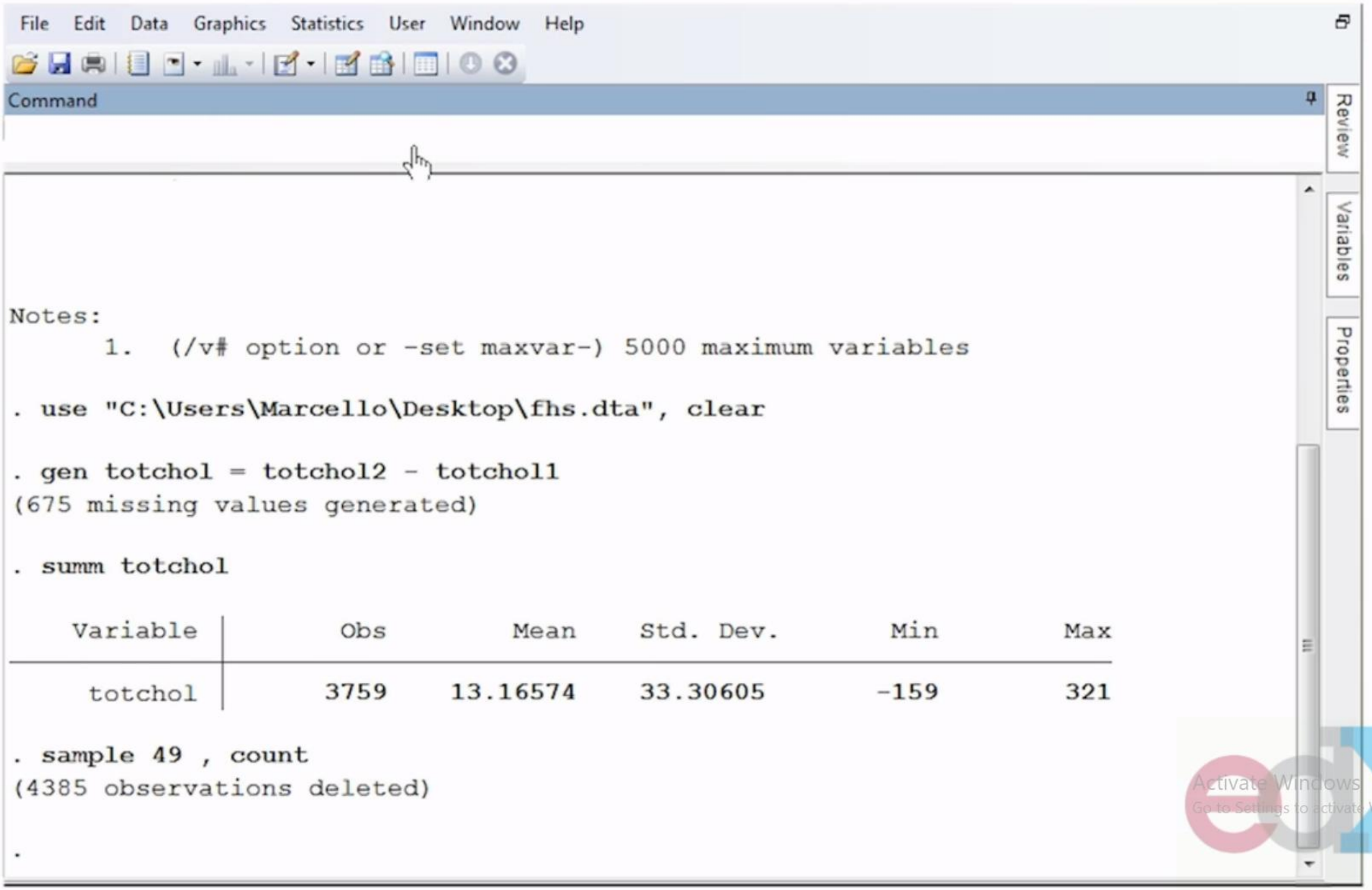
1. No (dependent, before/after)
2. Yes (different people)

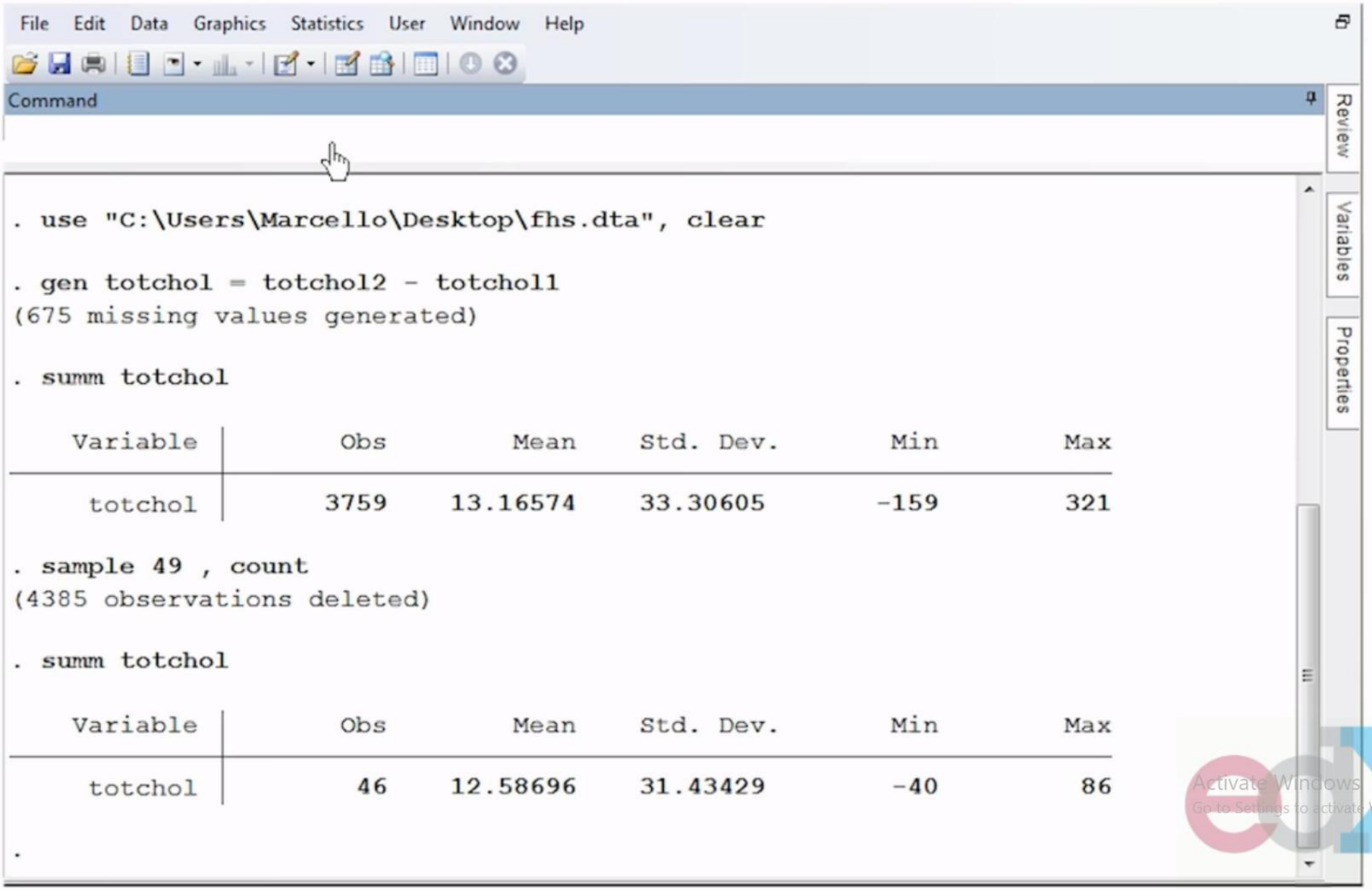

File Edit Data Graphics Statistics User Window Help
Command
(R)
12.1
Copyright 1985-2011 St
StataCorp
4905 Lakeway Drive
Statistics/Data Analysis
College Station, Texas
MP - Parallel Edition
800-STATA-PC
ht
> a.com
979-696-4600
st
979-696-4601 (fax)
Notes:
1. (/v# option or -set maxvar-) 5000 maximum variable
. use "C:\Users\Marcello\Desktop\fhs.dta", clear
.
Variables
Variable
Label
hyperten
Incident Hyperten...
timeap
Time (years) to An...
timemi
Time (years) to H...
timemifc
Time (years) to MI...
timechd
Time (years) to C...
timestrk
Time (years) to Str...
timecvd
Time (years) to CVD
timedth
Time (years) to De...
timehyp
Time (years) to Hy...
sex1
Sex, exam 1
totchol1
Total cholesterol (...
age1
Age (years), exam 1
sysbp1
Systolic blood pre...
diabp1
Diastolic blood pr...
cursmoke1
Current smoker, e...
cigpday1
Number of cigaret...
bmi1
Body mass index, ...
diabetes1
Diabetic, exam 1
bpmeds1
Use anti-hyperten...
heartrte1
Heart rate (beats p...
glucose1
Glucose level (mg...
prevchd1
Prevalent coronar...
prevap1
Prevalent angina ...
Review
Variables
Properties

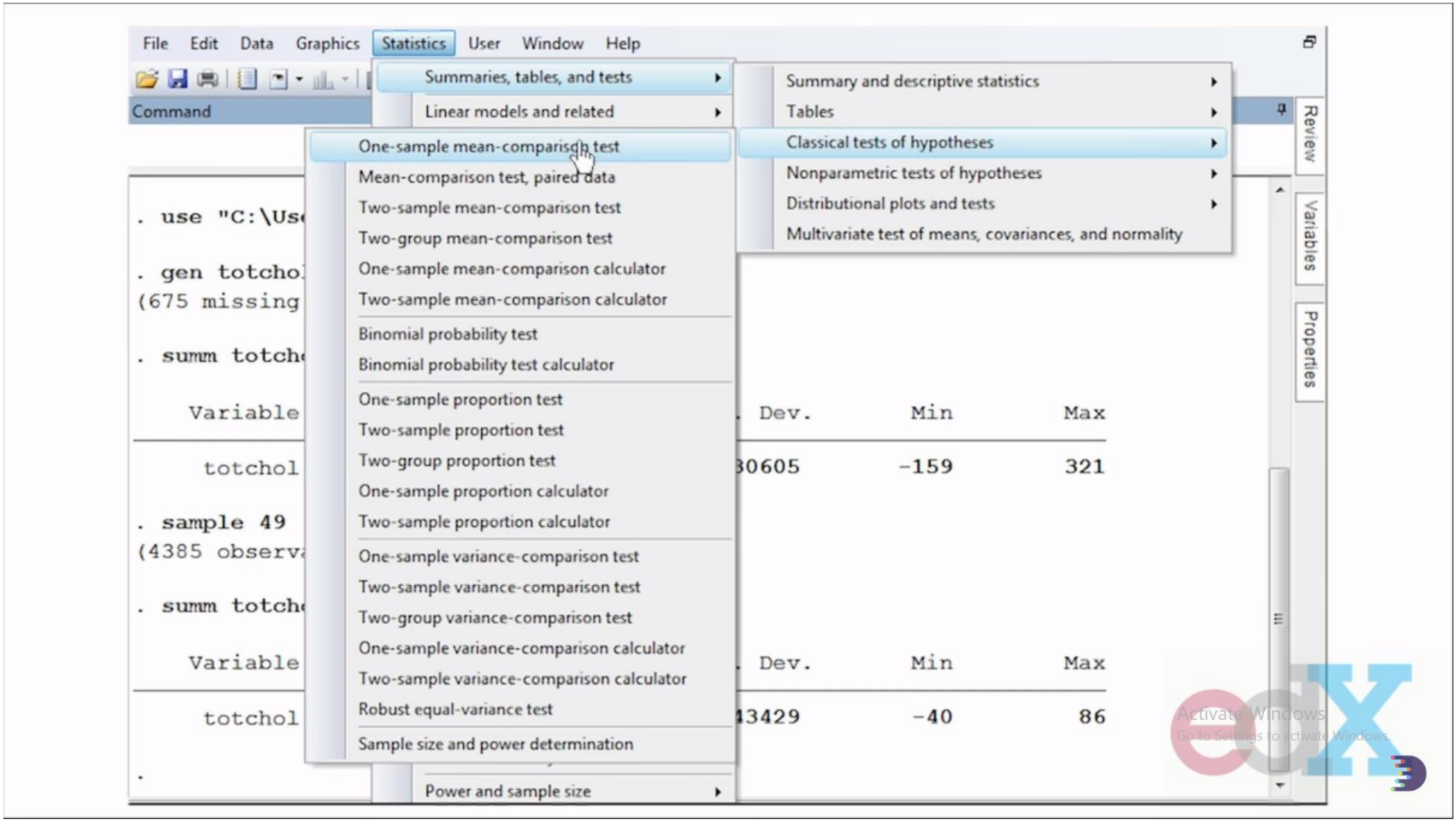

File Edit Data Graphics Statistics User Window Help
Command
gen totchol = totchol2 - totchol1
(R)
12.1
Copyright 1985-2011 StataCorp LP
Statistics/Data Analysis
StataCorp
4905 Lakeway Drive
MP - Parallel Edition
College Station, Texas 77845 USA
800-STATA-PC
http://www.stat
> a.com
979-696-4600
stata@stata.com
979-696-4601 (fax)
Notes:
1. (/v# option or -set maxvar-) 5000 maximum variables
. use "C:\Users\Marcello\Desktop\fhs.dta", clear
.
Review
Variables
Properties

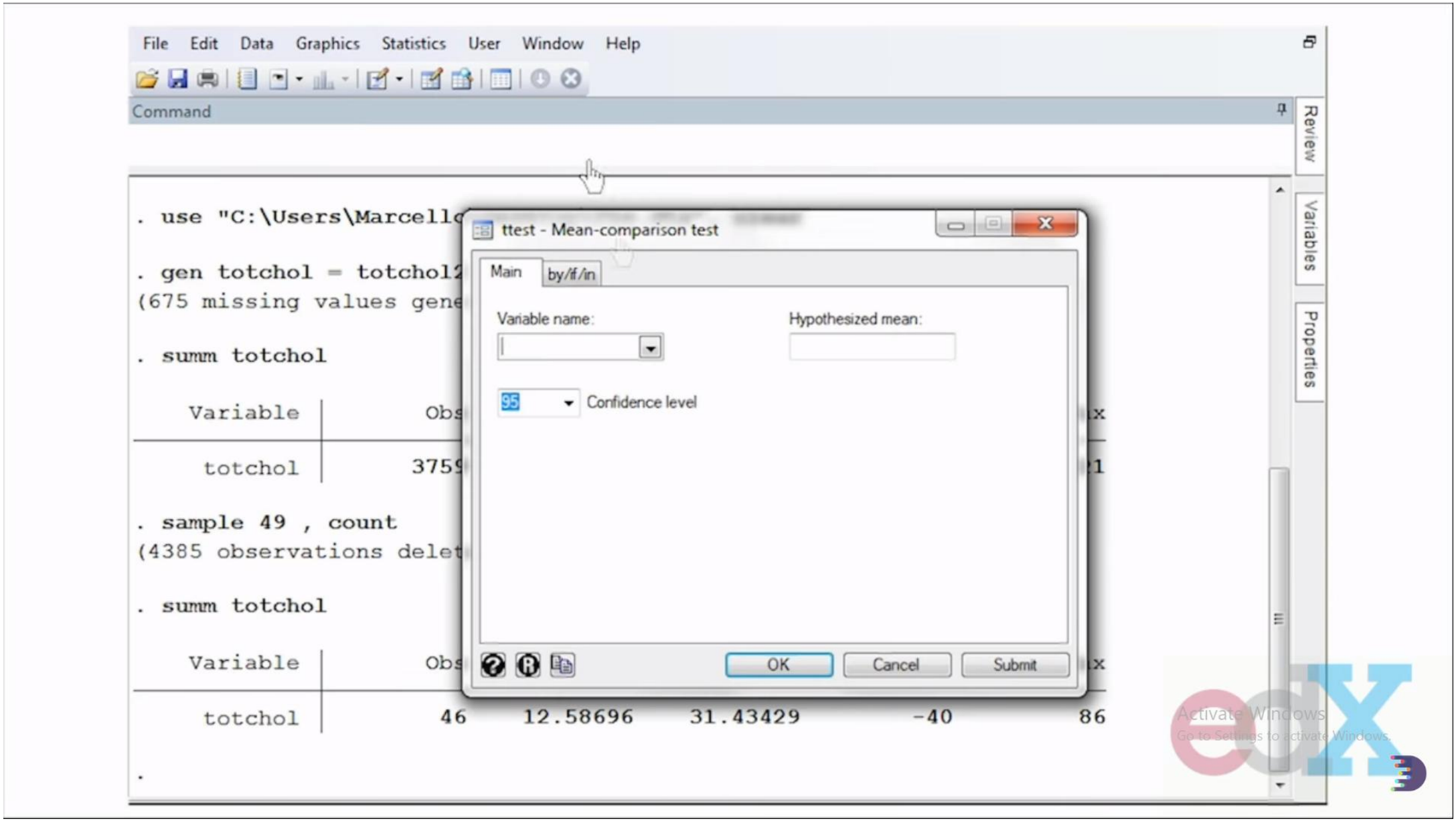

File Edit Data Graphics Statistics User Window Help
Command
Review
Variables
Properties
Statistics/Data Analysis
StataCorp
4905 Lakeway Drive
MP - Parallel Edition
College Station, Texas 77845 USA
800-STATA-PC
http://www.stat
> a.com
979-696-4600
stata@stata.com
979-696-4601 (fax)
Notes:
1. (/v# option or -set maxvar-) 5000 maximum variables
. use "C:\Users\Marcello\Desktop\fhs.dta", clear
. gen totchol = totchol2 - totchol1
(675 missing values generated)
.
Activate Windows
Go to Settings to activate Windows.

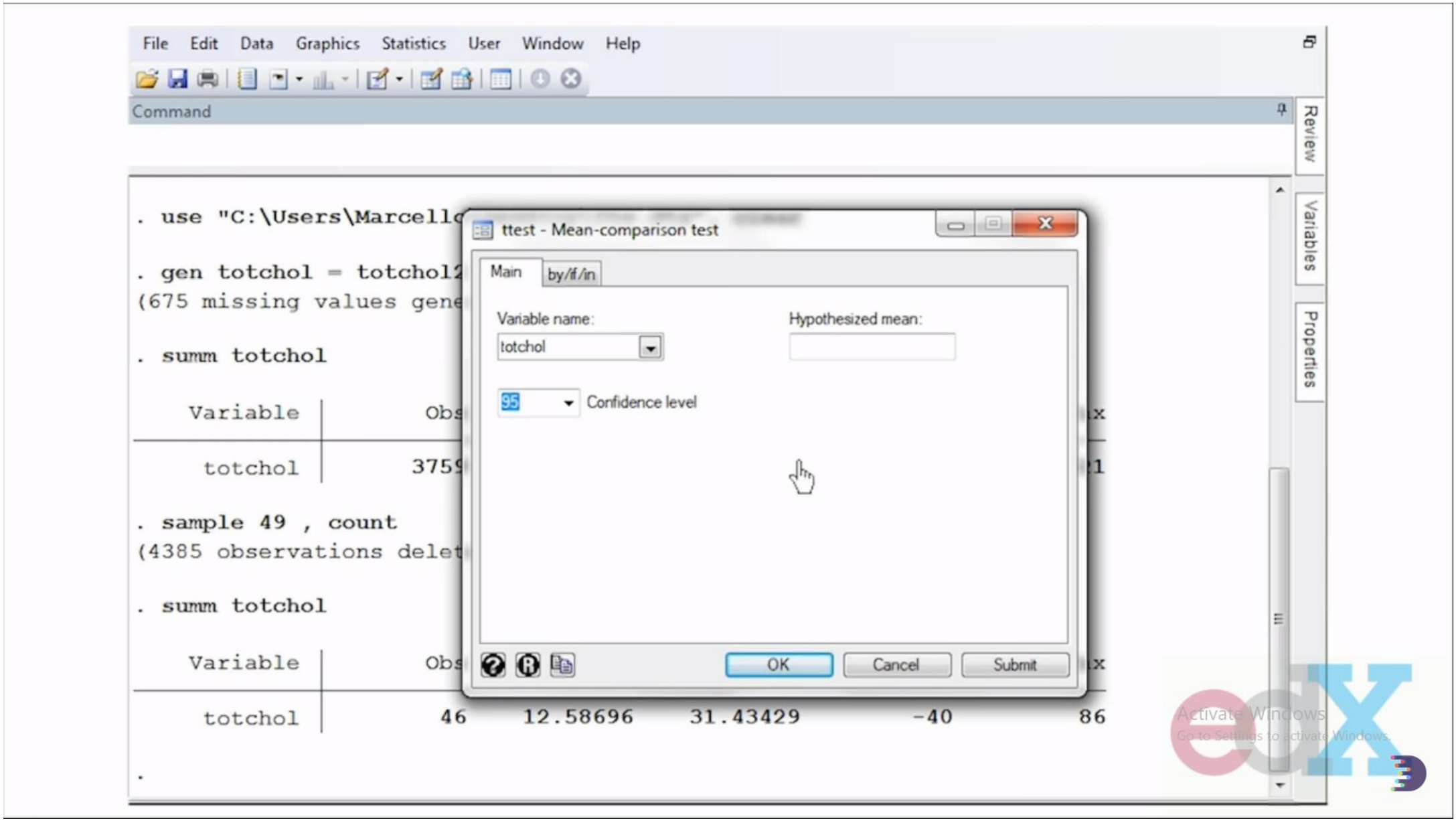

File Edit Data Graphics Statistics User Window Help
Command
Review
Variables
Properties
979-696-4601 (fax)
Notes:
1. (/v# option or -set maxvar-) 5000 maximum variables
. use "C:\Users\Marcello\Desktop\fhs.dta", clear
. gen totchol = totchol2 - totchol1
(675 missing values generated)
. summ totchol
Variable | Obs Mean Std. Dev. Min Max
totchol | 3759 13.16574 33.30605 -159 321
.
Activate Windows
Go to Settings to activate Windows.

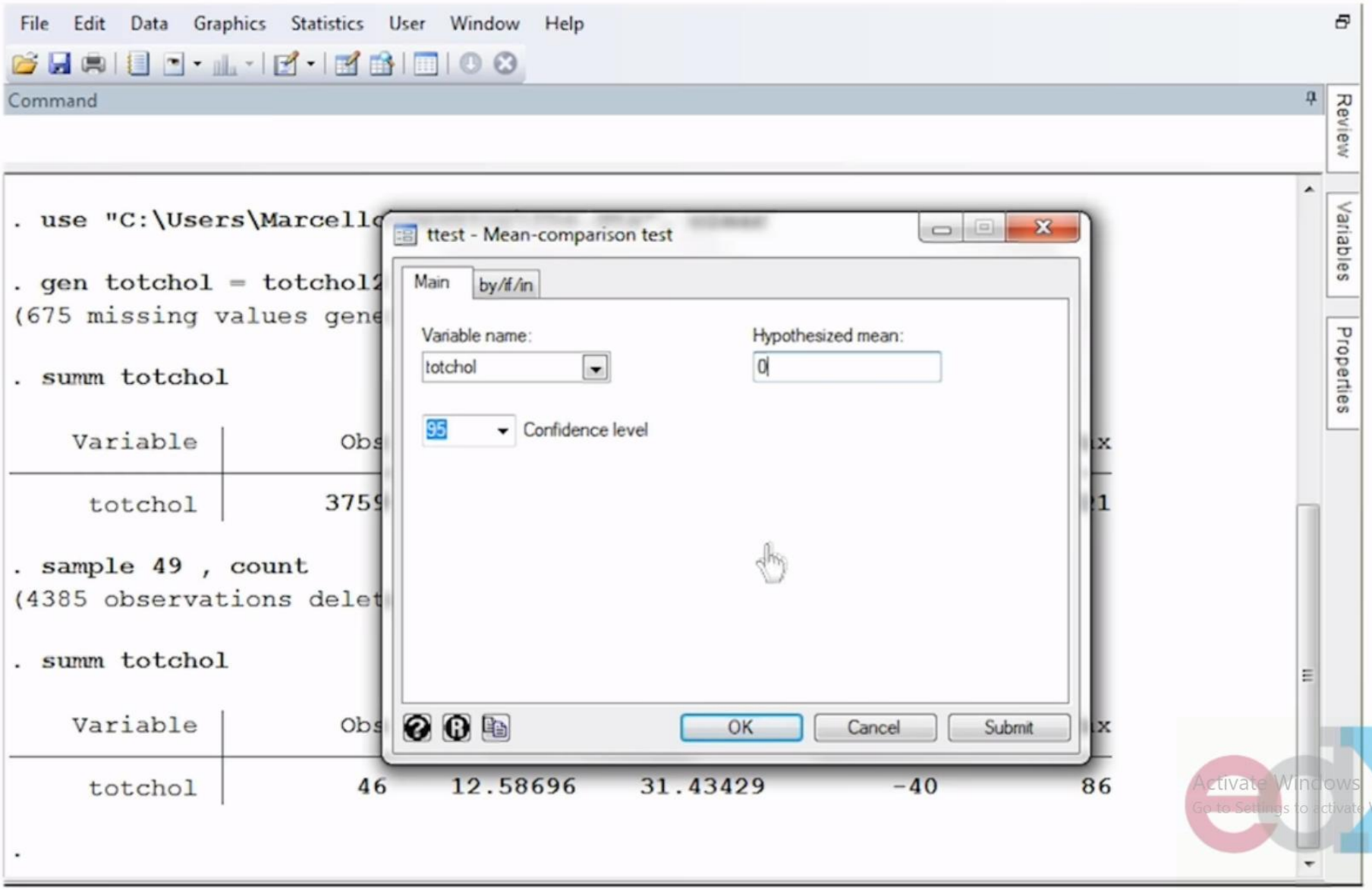

File Edit Data Graphics Statistics User Window Help
Command
Review
Variables
Properties
Notes:
1. (/v# option or -set maxvar-) 5000 maximum variables
. use "C:\Users\Marcello\Desktop\fhs.dta", clear
. gen totchol = totchol2 - totchol1
(675 missing values generated)
. summ totchol
Variable | Obs Mean Std. Dev. Min Max
totchol | 3759 13.16574 33.30605 -159 321
. sample 49 , count
(4385 observations deleted)
.

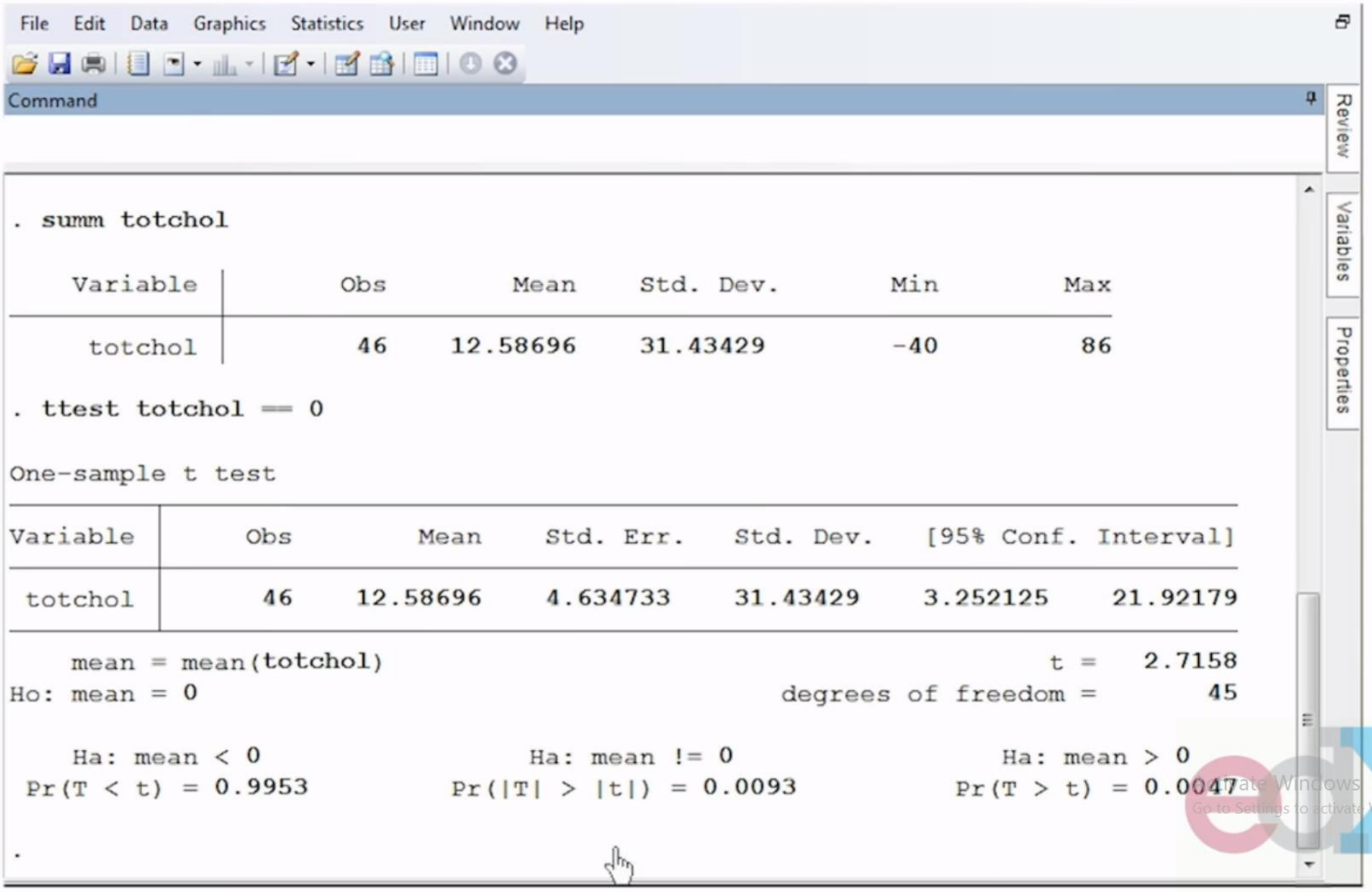

File Edit Data Graphics Statistics User Window Help
Command
Review
Variables
Properties
. use "C:\Users\Marcello\Desktop\fhs.dta", clear
. gen totchol = totchol2 - totchol1
(675 missing values generated)
. summ totchol
Variable | Obs Mean Std. Dev. Min Max
totchol | 3759 13.16574 33.30605 -159 321
. sample 49 , count
(4385 observations deleted)
. summ totchol
Variable | Obs Mean Std. Dev. Min Max
totchol | 46 12.58696 31.43429 -40 86
.
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
Command
Summaries, tables, and tests
Linear models and related
Summary and descriptive statistics
Tables
Classical tests of hypotheses
Nonparametric tests of hypotheses
Distributional plots and tests
Multivariate test of means, covariances, and normality
One-sample mean-comparison test
Mean-comparison test, paired data
Two-sample mean-comparison test
Two-group mean-comparison test
One-sample mean-comparison calculator
Two-sample mean-comparison calculator
Binomial probability test
Binomial probability test calculator
One-sample proportion test
Two-sample proportion test
Two-group proportion test
One-sample proportion calculator
Two-sample proportion calculator
One-sample variance-comparison test
Two-sample variance-comparison test
Two-group variance-comparison test
One-sample variance-comparison calculator
Two-sample variance-comparison calculator
Robust equal-variance test
Sample size and power determination
Power and sample size
. use "C:\Us
. gen totcho
(675 missing
. summ totch
Variable
Dev.
Min
Max
totchol
30605
-159
321
. sample 49
(4385 observa
. summ totch
Variable
Dev.
Min
Max
totchol
43429
-40
86
.
Review
Variables
Properties
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
Command
Review
Variables
Properties
. use "C:\Users\Marcell
. gen totchol = totchol2
(675 missing values gene
. summ totchol
Variable | Obs
totchol | 3759
. sample 49 , count
(4385 observations delet
. summ totchol
Variable | Obs
totchol | 46 12.58696 31.43429 -40 86
.
ttest - Mean-comparison test
Main
by/if/in
Variable name:
Hypothesized mean:
95
Confidence level
OK
Cancel
Submit
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
Command
Review
Variables
Properties
. use "C:\Users\Marcell
. gen totchol = totchol2
(675 missing values gene
. summ totchol
Variable | Obs
totchol | 3759
. sample 49 , count
(4385 observations delet
. summ totchol
Variable | Obs
totchol | 46 12.58696 31.43429 -40 86
.
ttest - Mean-comparison test
Main
by/if/in
Variable name:
totchol
Hypothesized mean:
95
Confidence level
OK
Cancel
Submit
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
Command
Review
Variables
Properties
. use "C:\Users\Marcell
. gen totchol = totchol2
(675 missing values gene
. summ totchol
Variable | Obs
totchol | 3759
. sample 49 , count
(4385 observations delet
. summ totchol
Variable | Obs
totchol | 46 12.58696 31.43429 -40 86
.
ttest - Mean-comparison test
Main
by/if/in
Variable name:
totchol
Hypothesized mean:
0
95
Confidence level
OK
Cancel
Submit


File Edit Data Graphics Statistics User Window Help
Command
Review
Variables
Properties
. summ totchol
Variable | Obs Mean Std. Dev. Min Max
totchol | 46 12.58696 31.43429 -40 86
. ttest totchol == 0
One-sample t test
Variable | Obs Mean Std. Err. Std. Dev. [95% Conf. Interval]
totchol | 46 12.58696 4.634733 31.43429 3.252125 21.92179
mean = mean(totchol) t = 2.7158
Ho: mean = 0 degrees of freedom = 45
Ha: mean < 0 Ha: mean != 0 Ha: mean > 0
Pr(T < t) = 0.9953 Pr(|T| > |t|) = 0.0093 Pr(T > t) = 0.0047
.

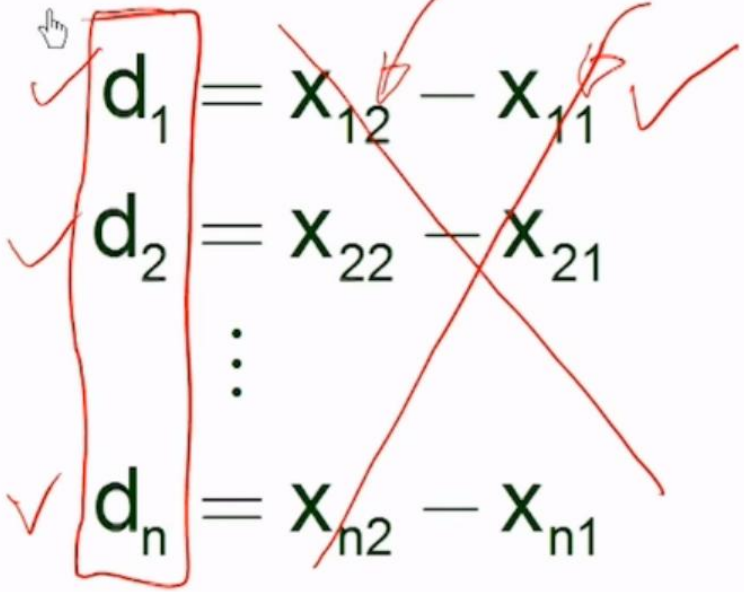
1 Dependent

$$d_1 = x_{12} - x_{11}$$
$$d_2 = x_{22} - x_{21} \quad \rangle \quad \Delta = \mu_2 - \mu_1$$
$$\vdots$$
$$d_n = x_{n2} - x_{n1}$$

So a test about the difference in the means is a test about the mean of the differences.

$$H_0 : \Delta = ?$$

1 Dependent



$$d_1 = x_{12} - x_{11}$$
$$d_2 = x_{22} - x_{21}$$
$$\vdots$$
$$d_n = x_{n2} - x_{n1}$$

$$\Rightarrow \Delta = \mu_2 - \mu_1$$

So a test about the difference in the means is a test about the mean of the differences.

$$H_0 : \Delta = ?$$

So treat the d's as the data and perform a one-sample t-test:

$$\bar{d} = \frac{1}{n}\sum_{j=1}^{n} d_j$$

$$s^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(d_j - \bar{d}\right)^2$$

$$t = \frac{\bar{d} - \mu_D}{s / \sqrt{n}} \qquad (n-1) \quad d.f.$$

So treat the d's as the data and perform a one-sample t-test:

$$\bar{d} = \frac{1}{n}\sum_{j=1}^{n} d_j$$

$$s^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(d_j - \bar{d}\right)^2$$

$$t = \frac{\bar{d} - \mu_D}{s / \sqrt{n}} \qquad (n-1) \quad d.f.$$

2 Independent

$$H_0 : \mu_1 - \mu_2 = \Delta \quad ?$$

$$t = \frac{\bar{X}_1 - \bar{X}_2 - \Delta}{\text{standard deviation}}$$

2 Independent

2 Independent

Population (Normal)

Pop. 1	Pop. 2
μ_1	μ_2
σ_1	σ_2

Sample

n_1	n_2
$\bar{x}_1$	$\bar{x}_2$
s_1	s_2

$H_0 : \mu_1 - \mu_2 = \Delta$

2 Independent

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

2 Independent

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

In order to decide the s^2s and the degrees of freedom we need to know whether, or not, $\sigma_1 = \sigma_2$

2 Independent

(a) Homoscedastic Case

If $\sigma_1 = \sigma_2$ (which can be tested) we can use a common value:

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

$$\text{d.f.} = n_1 + n_2 - 2$$

2 Independent

(a) Homoscedastic Case

If $\sigma_1 = \sigma_2$ (which can be tested) we can use a common value:

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

$$\text{d.f.} = n_1 + n_2 - 2$$

d.f. dependent $n-1$

2 Independent

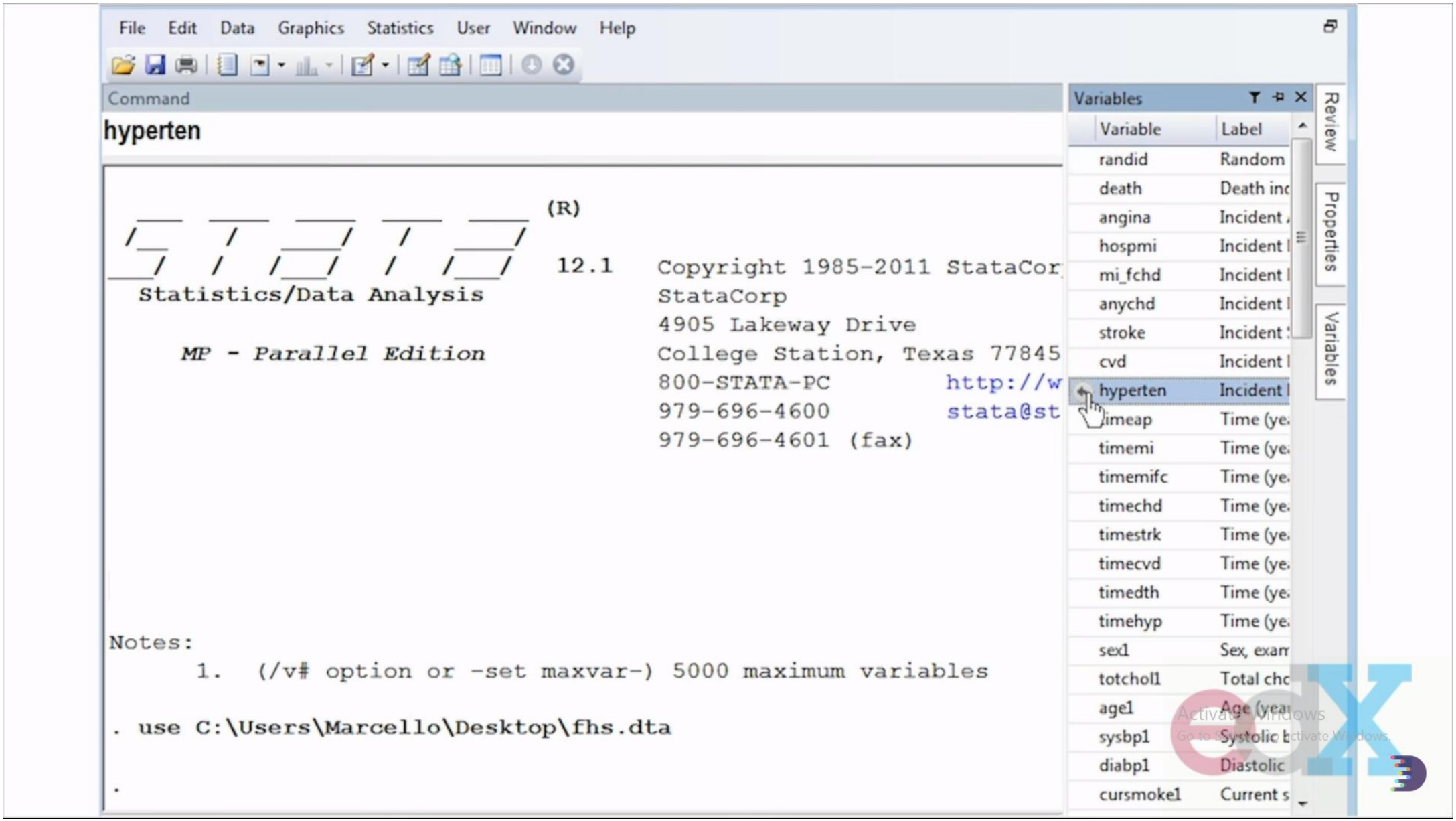
(b) Heteroscedastic Case

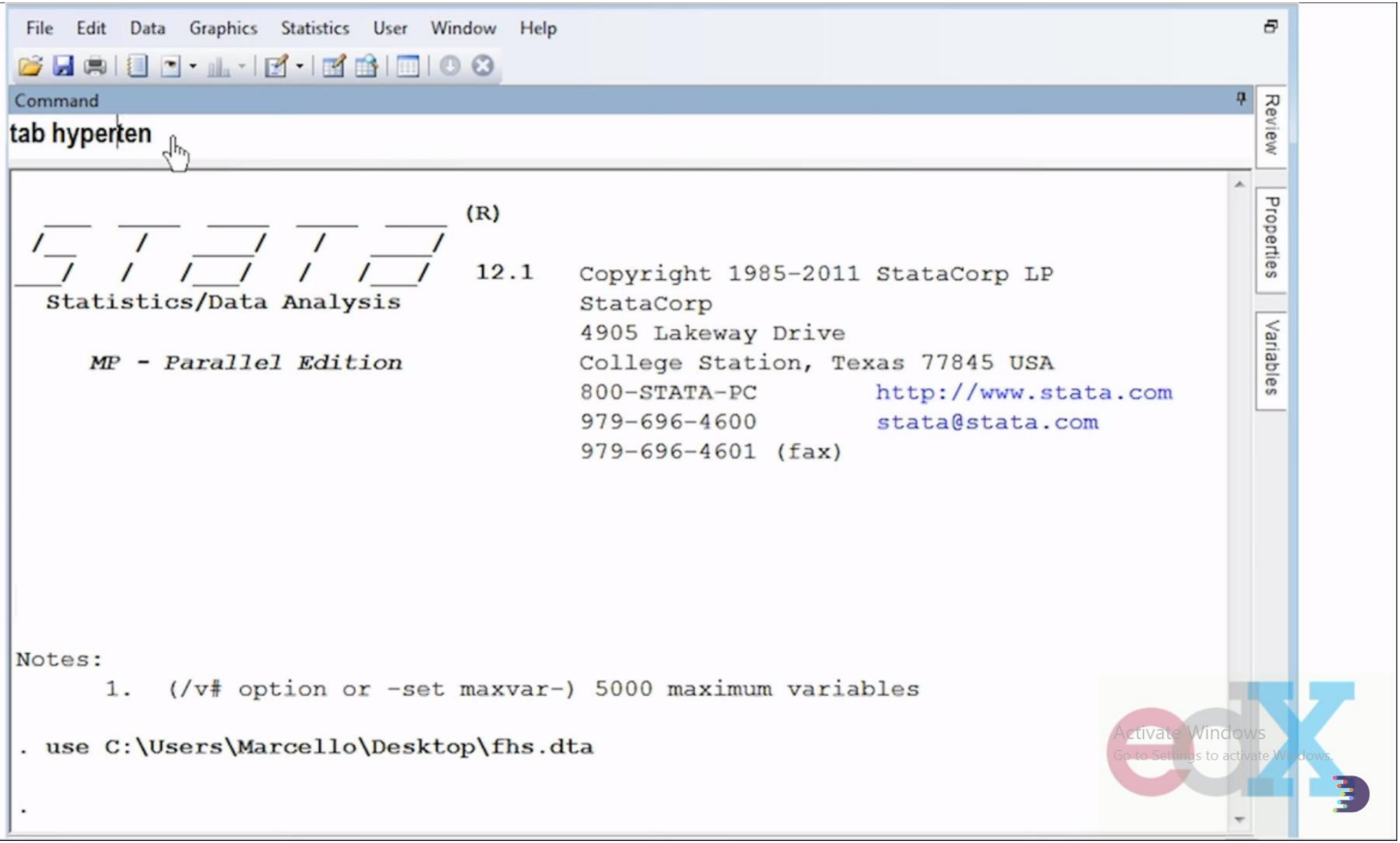
If $\sigma_1 \neq \sigma_2$ (recommended)

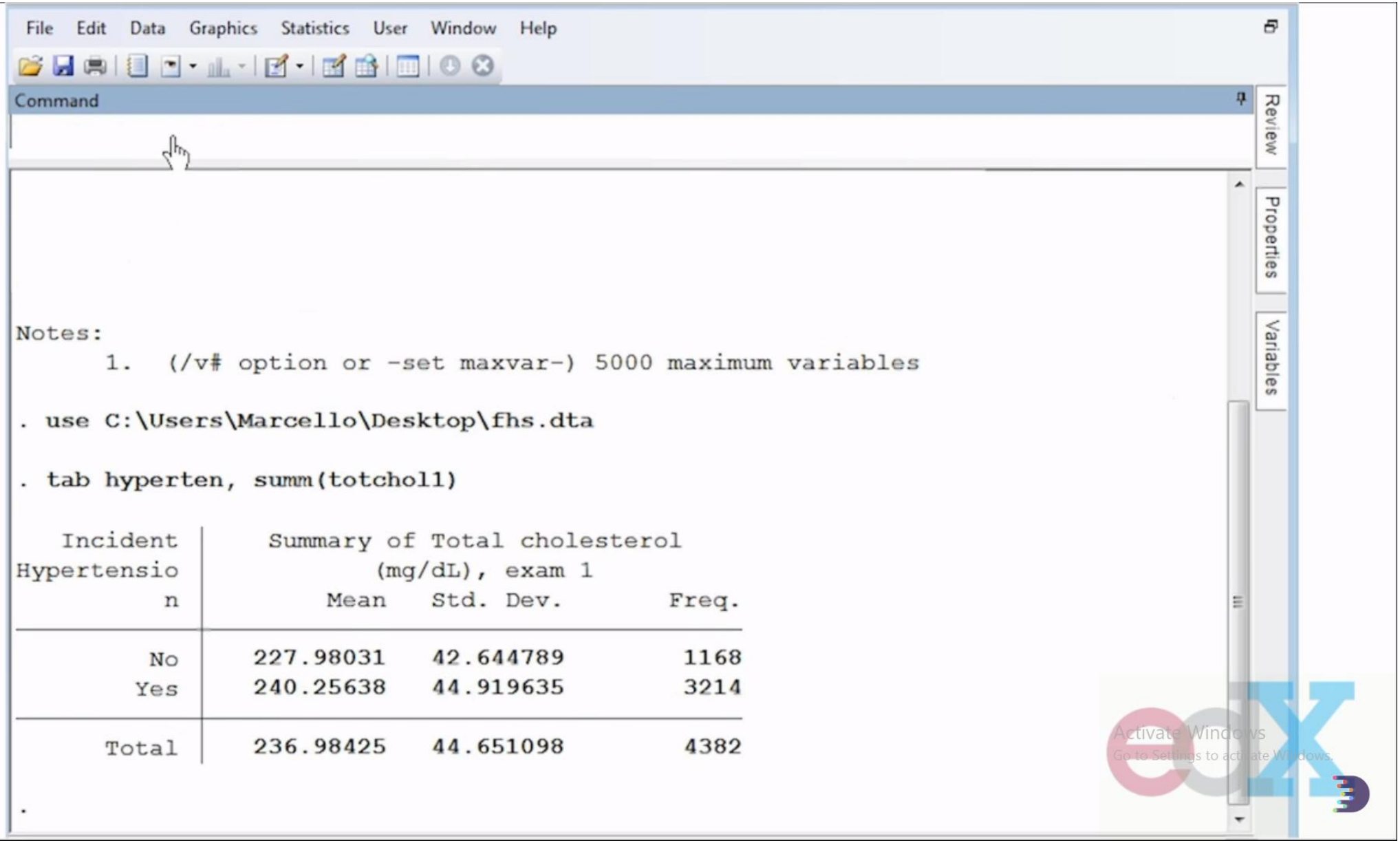
Use individual sample standard deviations and degrees of freedom, ν :

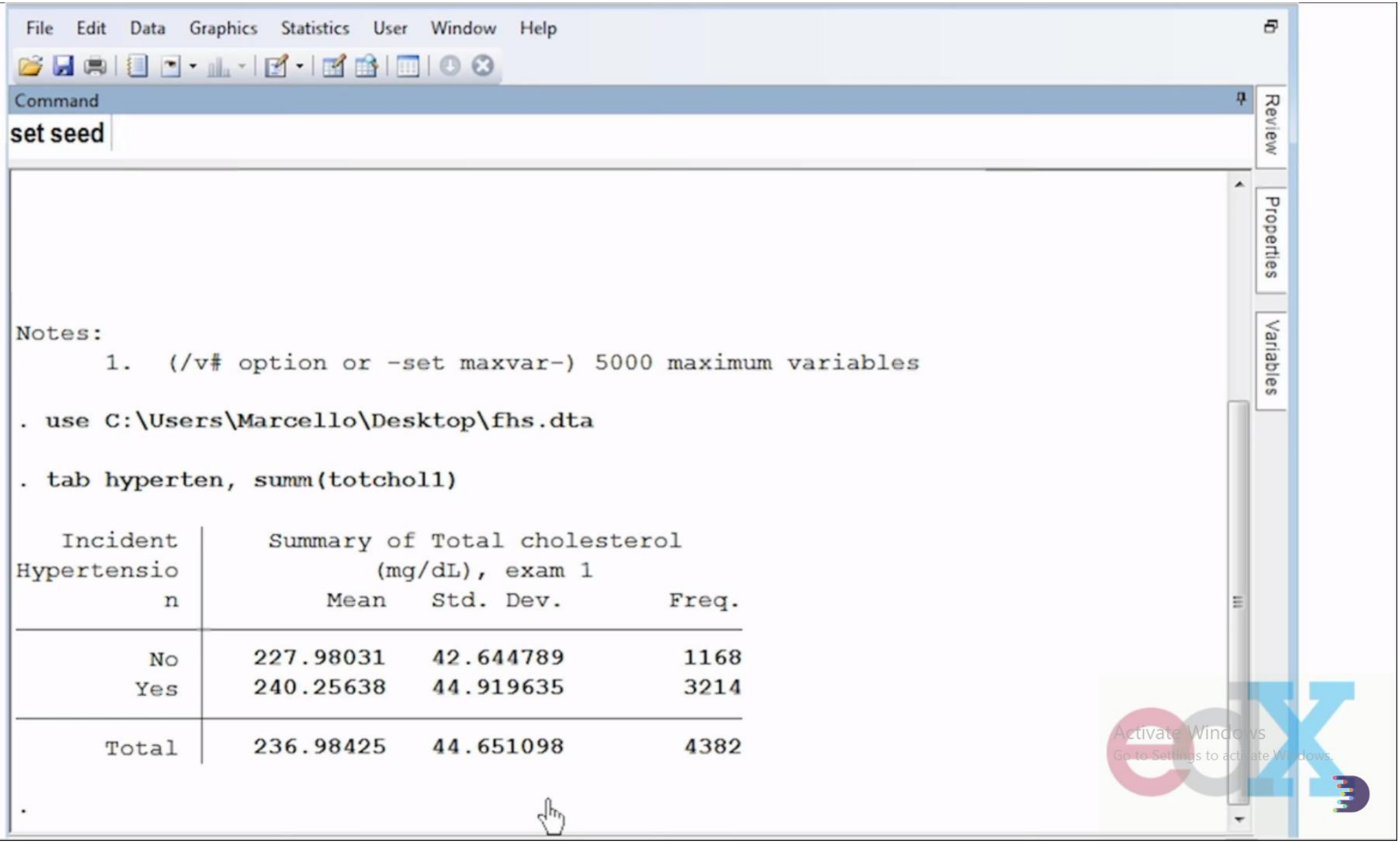
$$a = \frac{s_1^2}{n_1} \quad \text{and} \quad b = \frac{s_2^2}{n_2}$$

$$\nu = \frac{(a+b)^2}{\frac{a^2}{(n_2-1)} + \frac{b^2}{(n_2-1)}}$$


File Edit Data Graphics Statistics User Window Help
Command
hyperten
(R)
12.1
Copyright 1985-2011 StataCor
StataCorp
4905 Lakeway Drive
Statistics/Data Analysis
MP - Parallel Edition
College Station, Texas 77845
800-STATA-PC
http://w
979-696-4600
stata@st
979-696-4601 (fax)
Notes:
1. (/v# option or -set maxvar-) 5000 maximum variables
. use C:\Users\Marcello\Desktop\fhs.dta
.
Variables
Variable
Label
randid
Random
death
Death inc
angina
Incident
hospmi
Incident
mi_fchd
Incident
anychd
Incident
stroke
Incident
cvd
Incident
hyperten
Incident
timeap
Time (ye
timemi
Time (ye
timemifc
Time (ye
timechd
Time (ye
timestrk
Time (ye
timecvd
Time (ye
timedth
Time (ye
timehyp
Time (ye
sex1
Sex, exam
totchol1
Total cho
age1
Age (yea
sysbp1
Systolic b
diabp1
Diastolic
cursmoke1
Current s
Review
Properties
Variables


File Edit Data Graphics Statistics User Window Help
Command
tab hyperten
(R)
12.1
Copyright 1985-2011 StataCorp LP
Statistics/Data Analysis
StataCorp
4905 Lakeway Drive
MP - Parallel Edition
College Station, Texas 77845 USA
800-STATA-PC
http://www.stata.com
979-696-4600
stata@stata.com
979-696-4601 (fax)
Notes:
1. (/v# option or -set maxvar-) 5000 maximum variables
. use C:\Users\Marcello\Desktop\fhs.dta
.
Review
Properties
Variables
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
Command
Review
Properties
Variables
Notes:
1. (/v# option or -set maxvar-) 5000 maximum variables
. use C:\Users\Marcello\Desktop\fhs.dta
. tab hyperten, summ(totchol1)
Incident Hypertensio n
Summary of Total cholesterol (mg/dL), exam 1
Mean Std. Dev. Freq.
No 227.98031 42.644789 1168
Yes 240.25638 44.919635 3214
Total 236.98425 44.651098 4382
.
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
Command
set seed
Review
Properties
Variables
Notes:
1. (/v# option or -set maxvar-) 5000 maximum variables
. use C:\Users\Marcello\Desktop\fhs.dta
. tab hyperten, summ(totchol1)
Incident Hypertensio n
Summary of Total cholesterol (mg/dL), exam 1
Mean Std. Dev. Freq.
No 227.98031 42.644789 1168
Yes 240.25638 44.919635 3214
Total 236.98425 44.651098 4382
.
Activate Windows
Go to Settings to activate Windows.

File Edit Data Graphics Statistics User Window Help

Command

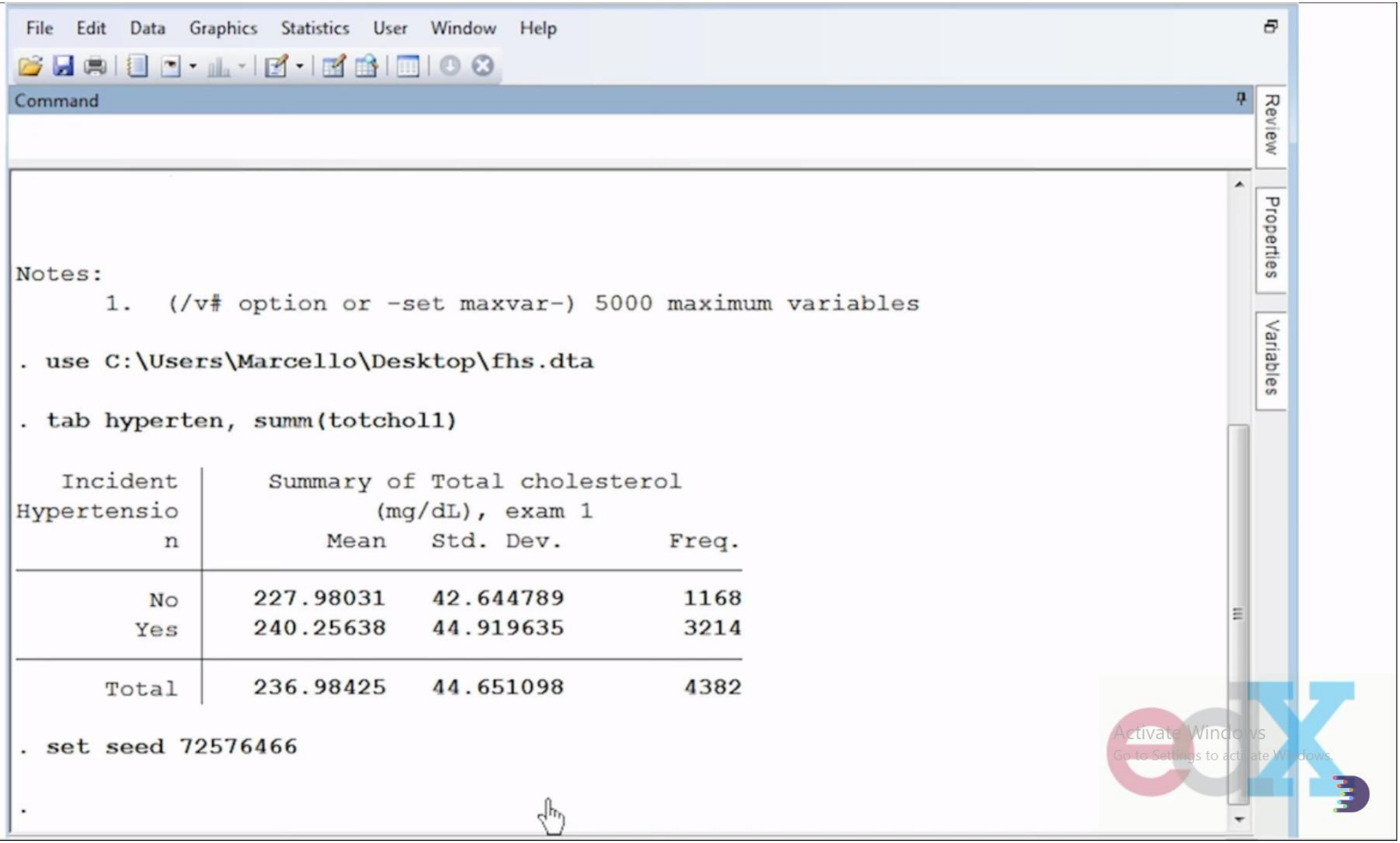
```
set seed 725
```

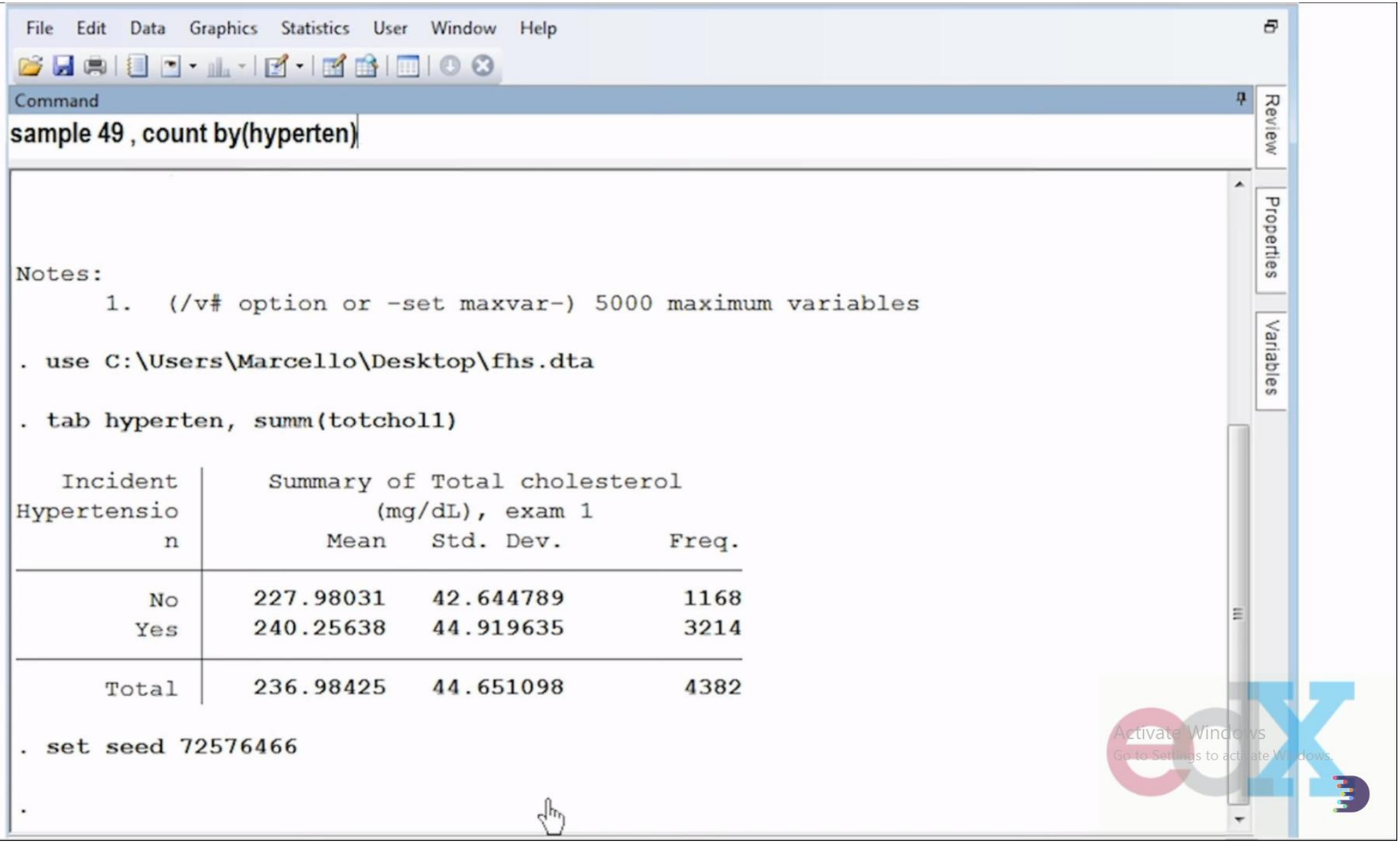
```
Notes:
      1.   (/v# option or -set maxvar-) 5000 maximum variables

. use C:\Users\Marcello\Desktop\fhs.dta

. tab hyperten, summ(totchol1)
```

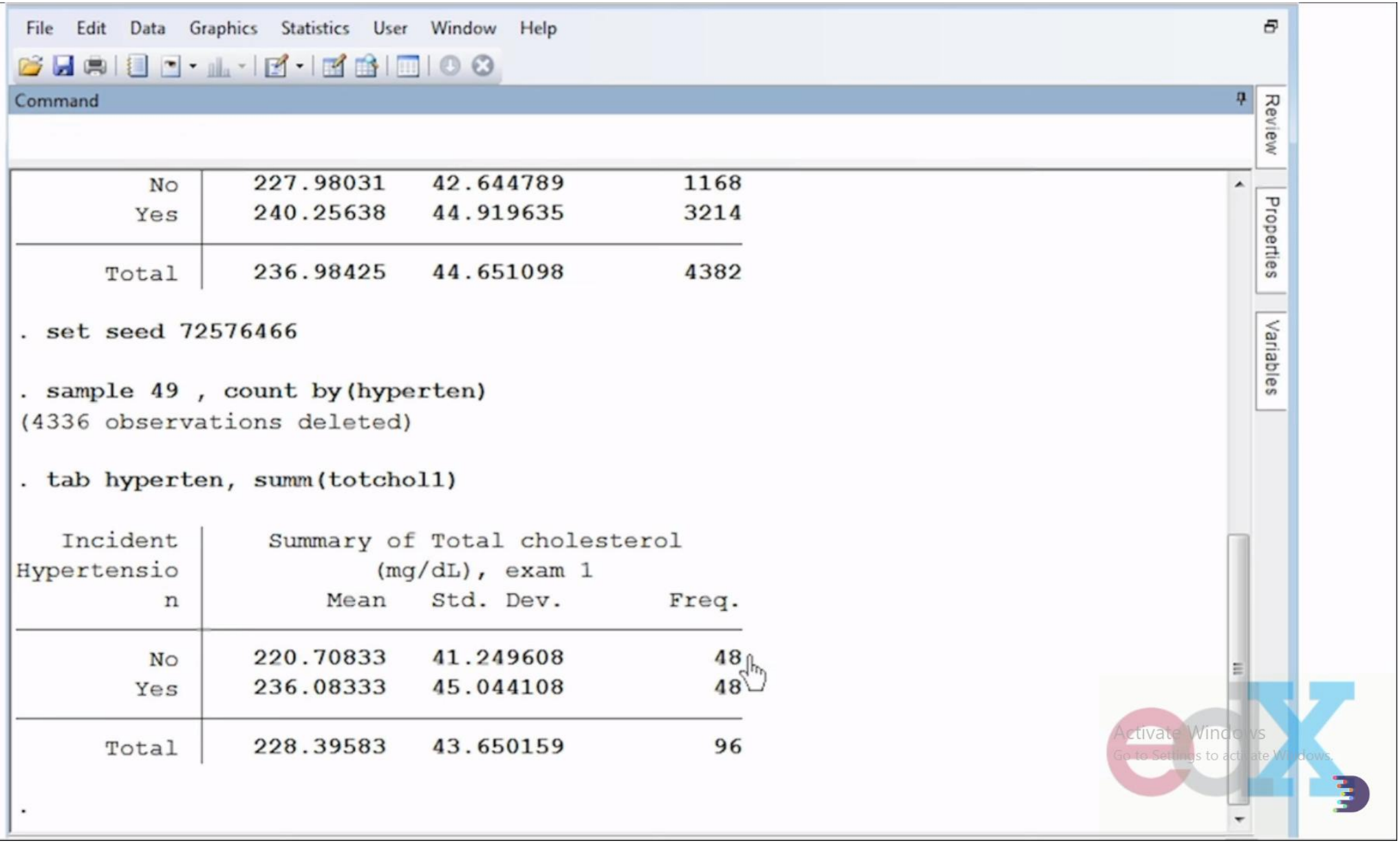
Incident Hypertension	Summary of Total cholesterol (mg/dL), exam 1 Mean	Std. Dev.	Freq.
No	227.98031	42.644789	1168
Yes	240.25638	44.919635	3214
Total	236.98425	44.651098	4382

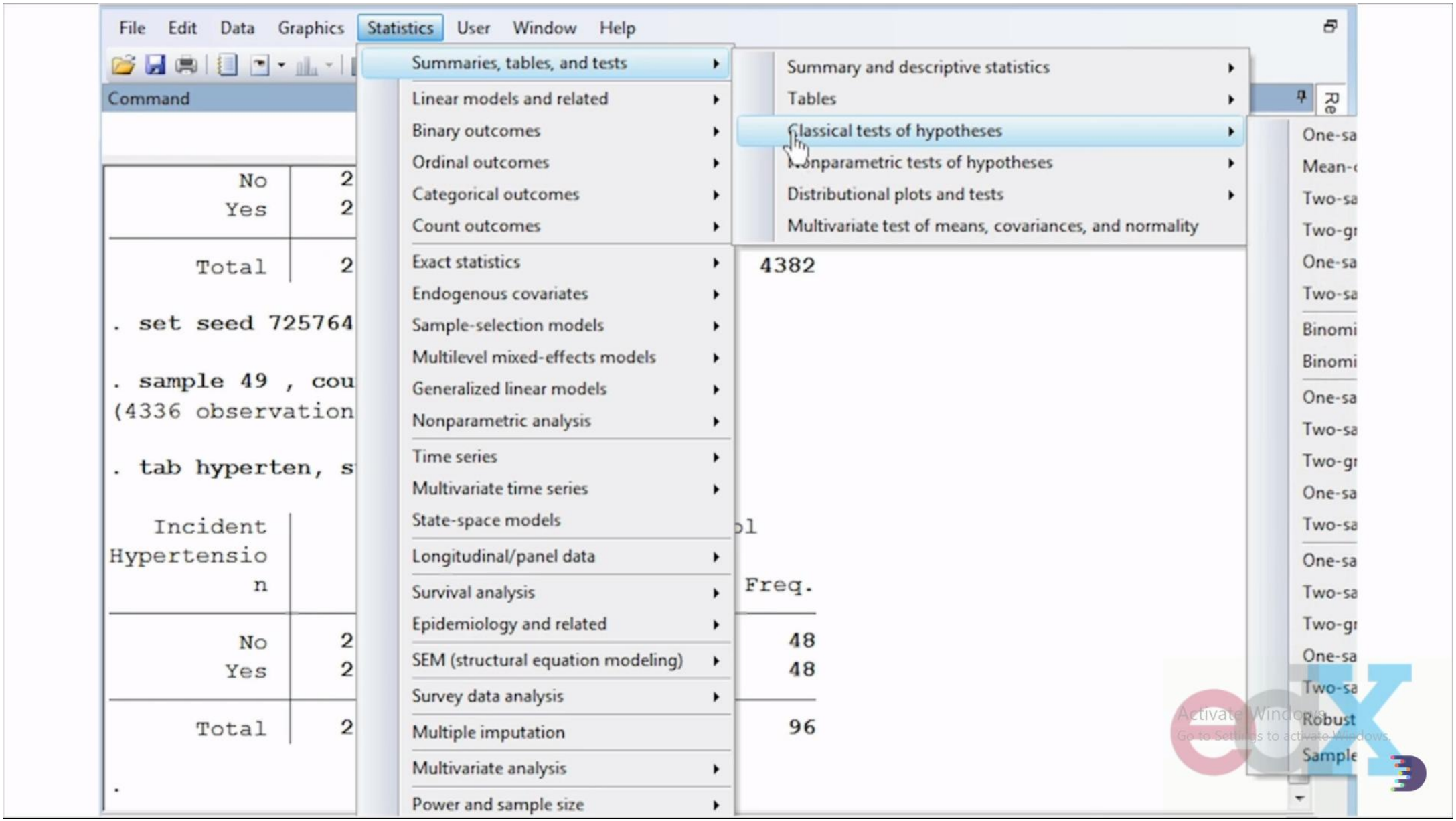
```
.
```

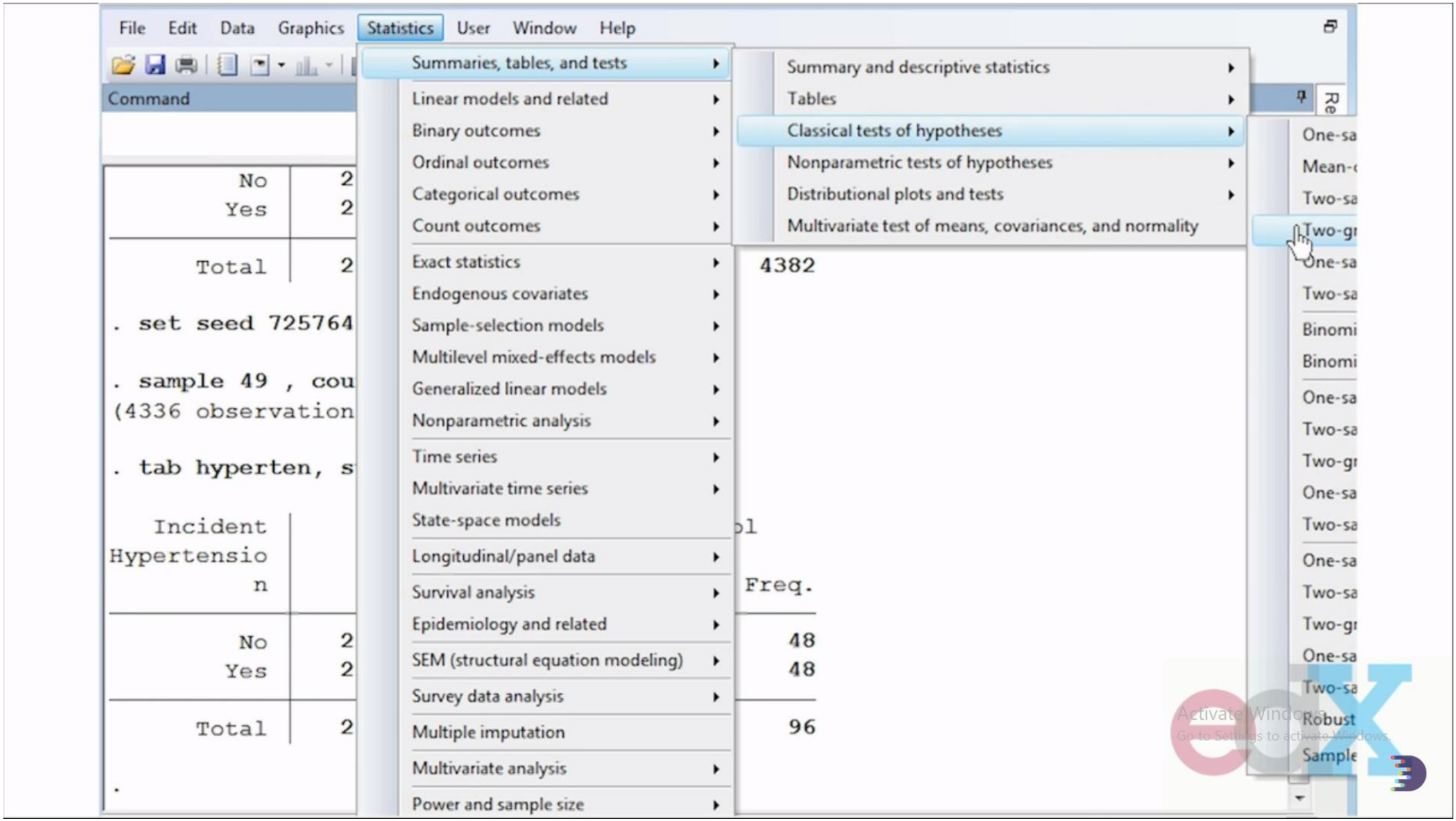
Review Properties Variables

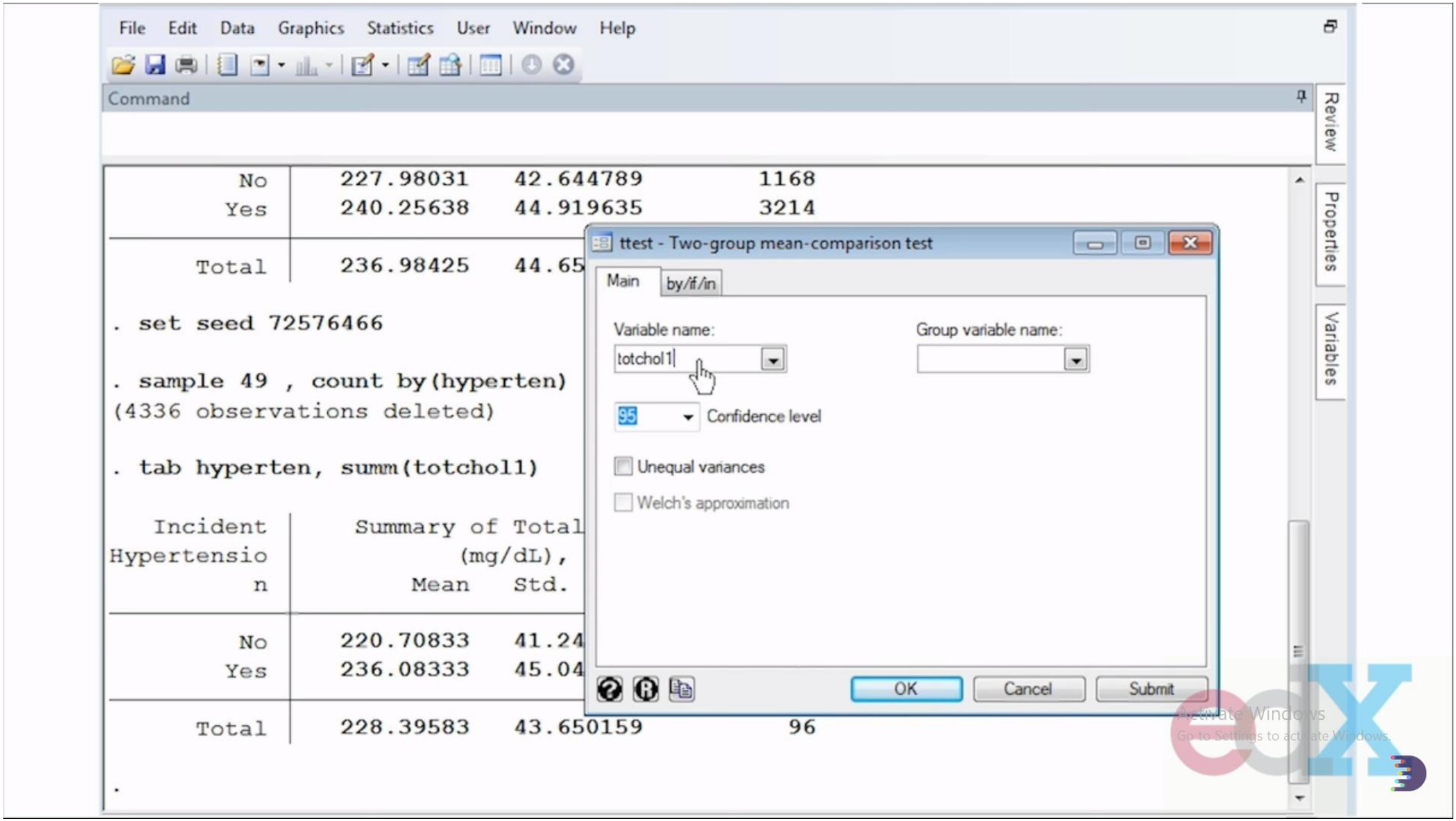

File Edit Data Graphics Statistics User Window Help
Command
Review
Properties
Variables
Notes:
1. (/v# option or -set maxvar-) 5000 maximum variables
. use C:\Users\Marcello\Desktop\fhs.dta
. tab hyperten, summ(totchol1)
Incident Hypertensio n
Summary of Total cholesterol (mg/dL), exam 1
Mean Std. Dev. Freq.
No 227.98031 42.644789 1168
Yes 240.25638 44.919635 3214
Total 236.98425 44.651098 4382
. set seed 72576466
.
Activate Windows
Go to Settings to activate Windows.

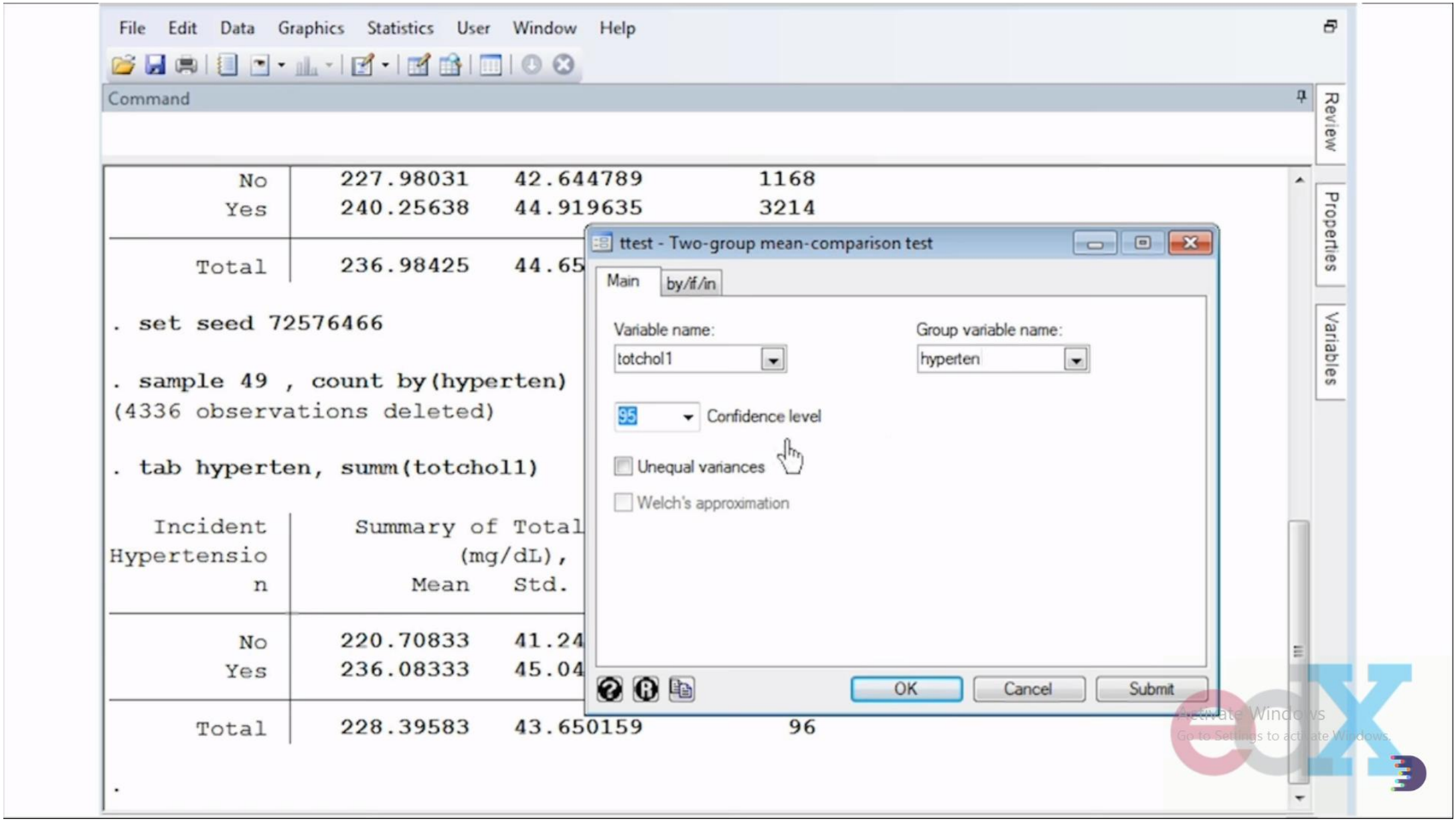

File Edit Data Graphics Statistics User Window Help
Command
sample 49 , count by(hyperten)
Review
Properties
Variables
Notes:
1. (/v# option or -set maxvar-) 5000 maximum variables
. use C:\Users\Marcello\Desktop\fhs.dta
. tab hyperten, summ(totchol1)
Incident Hypertensio n
Summary of Total cholesterol (mg/dL), exam 1
Mean Std. Dev. Freq.
No 227.98031 42.644789 1168
Yes 240.25638 44.919635 3214
Total 236.98425 44.651098 4382
. set seed 72576466
.
Activate Windows
Go to Settings to activate Windows.

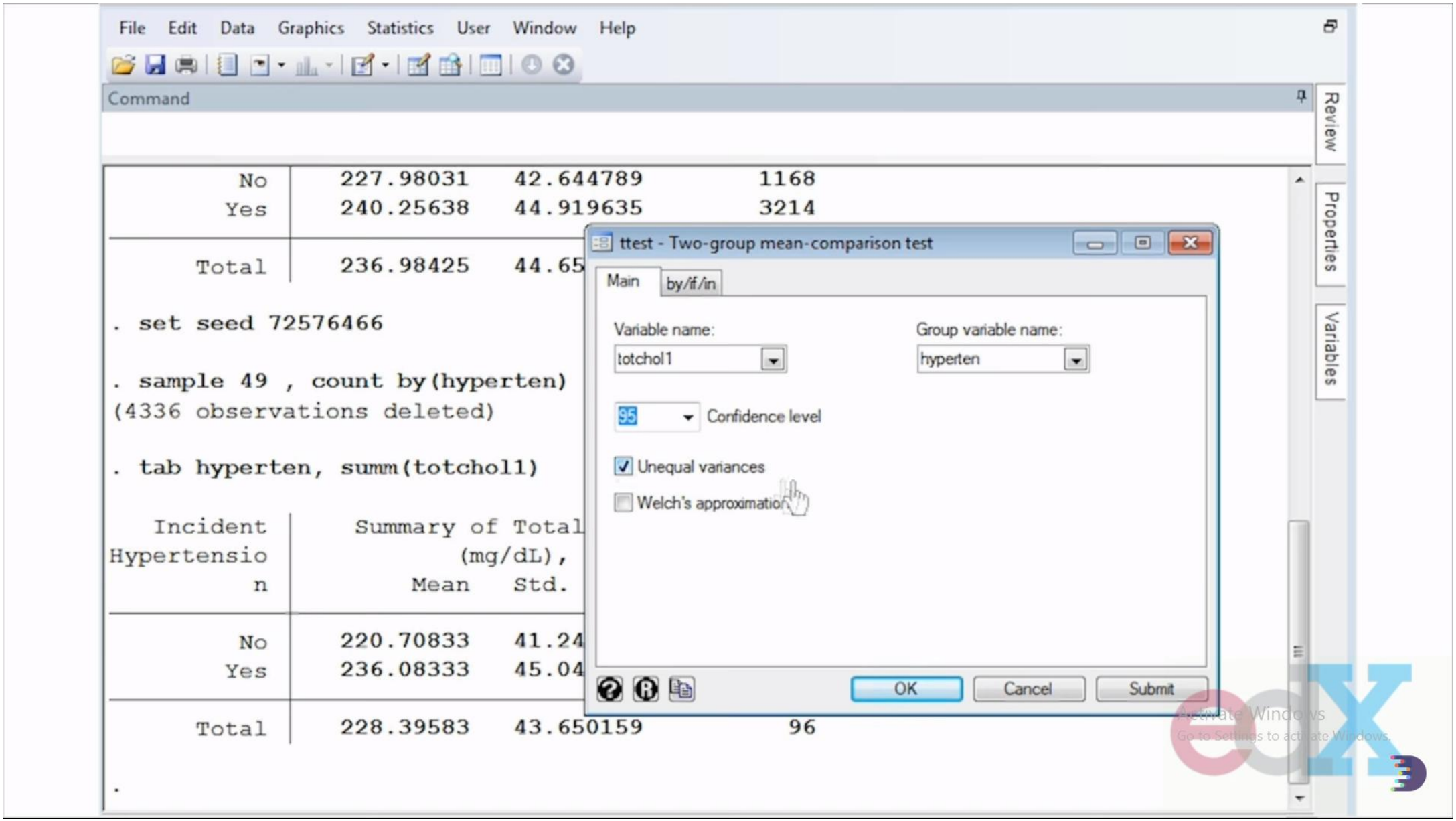

File Edit Data Graphics Statistics User Window Help
Command
Review
Properties
Variables
No 227.98031 42.644789 1168
Yes 240.25638 44.919635 3214
Total 236.98425 44.651098 4382
. set seed 72576466
. sample 49 , count by(hyperten)
(4336 observations deleted)
. tab hyperten, summ(totchol1)
Incident Hypertensio n
Summary of Total cholesterol (mg/dL), exam 1
Mean Std. Dev. Freq.
No 220.70833 41.249608 48
Yes 236.08333 45.044108 48
Total 228.39583 43.650159 96
.
Activate Windows
Go to Settings to activate Windows.

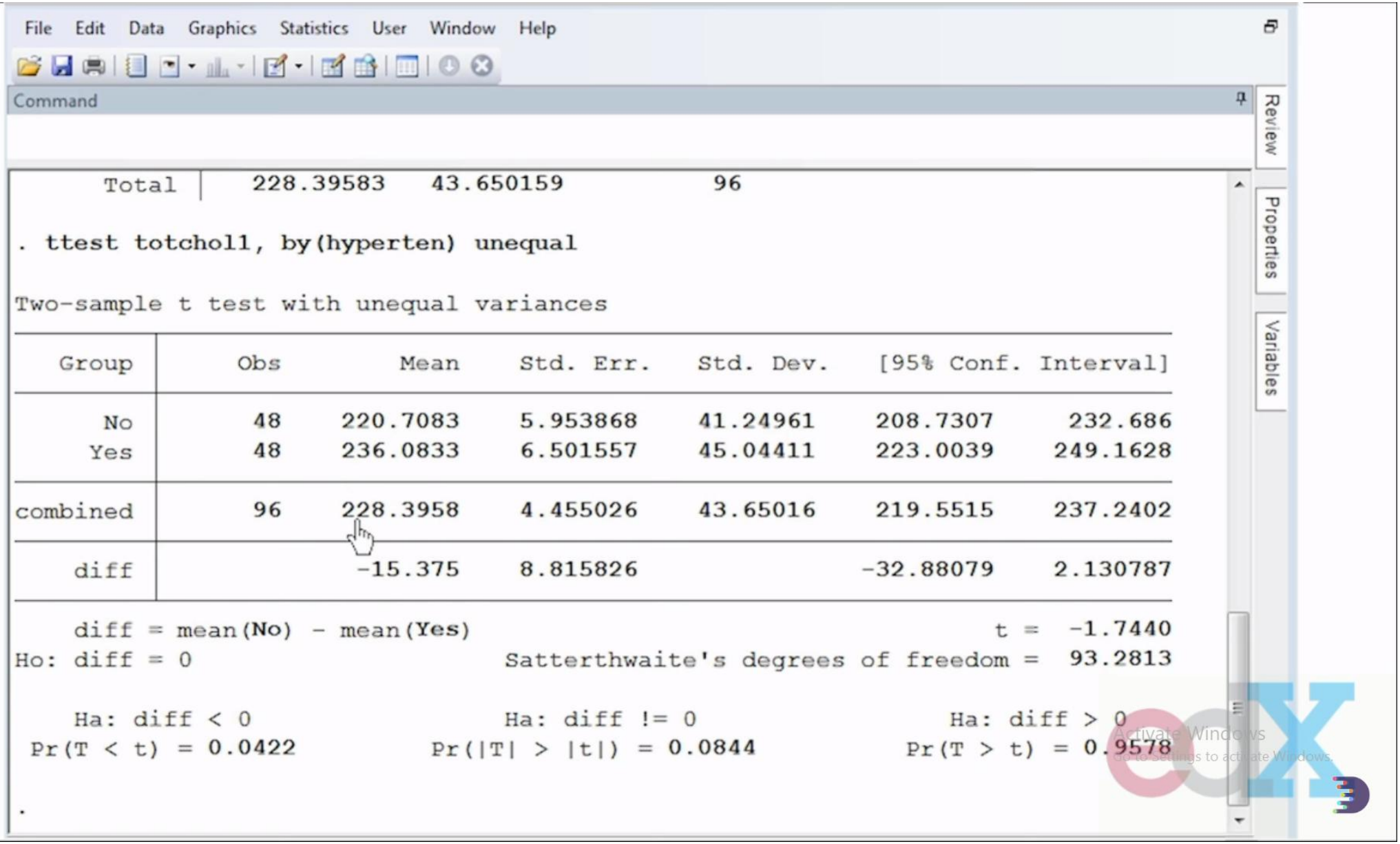

File Edit Data Graphics Statistics User Window Help
Command
Summaries, tables, and tests
Linear models and related
Binary outcomes
Ordinal outcomes
Categorical outcomes
Count outcomes
Exact statistics
Endogenous covariates
Sample-selection models
Multilevel mixed-effects models
Generalized linear models
Nonparametric analysis
Time series
Multivariate time series
State-space models
Longitudinal/panel data
Survival analysis
Epidemiology and related
SEM (structural equation modeling)
Survey data analysis
Multiple imputation
Multivariate analysis
Power and sample size
Summary and descriptive statistics
Tables
Classical tests of hypotheses
Nonparametric tests of hypotheses
Distributional plots and tests
Multivariate test of means, covariances, and normality
No 2
Yes 2
Total 2
. set seed 725764
. sample 49 , cou
(4336 observation
. tab hyperten, s
Incident Hypertensio n
No 2
Yes 2
Total 2
4382
Freq.
48
48
96
One-sa
Mean-
Two-sa
Two-gr
One-sa
Two-sa
Binomi
Binomi
One-sa
Two-sa
Two-gr
One-sa
Two-sa
One-sa
Two-sa
Two-gr
One-sa
Two-sa
Robust
Sample
Activate Windows
Go to Settings to activate Windows.

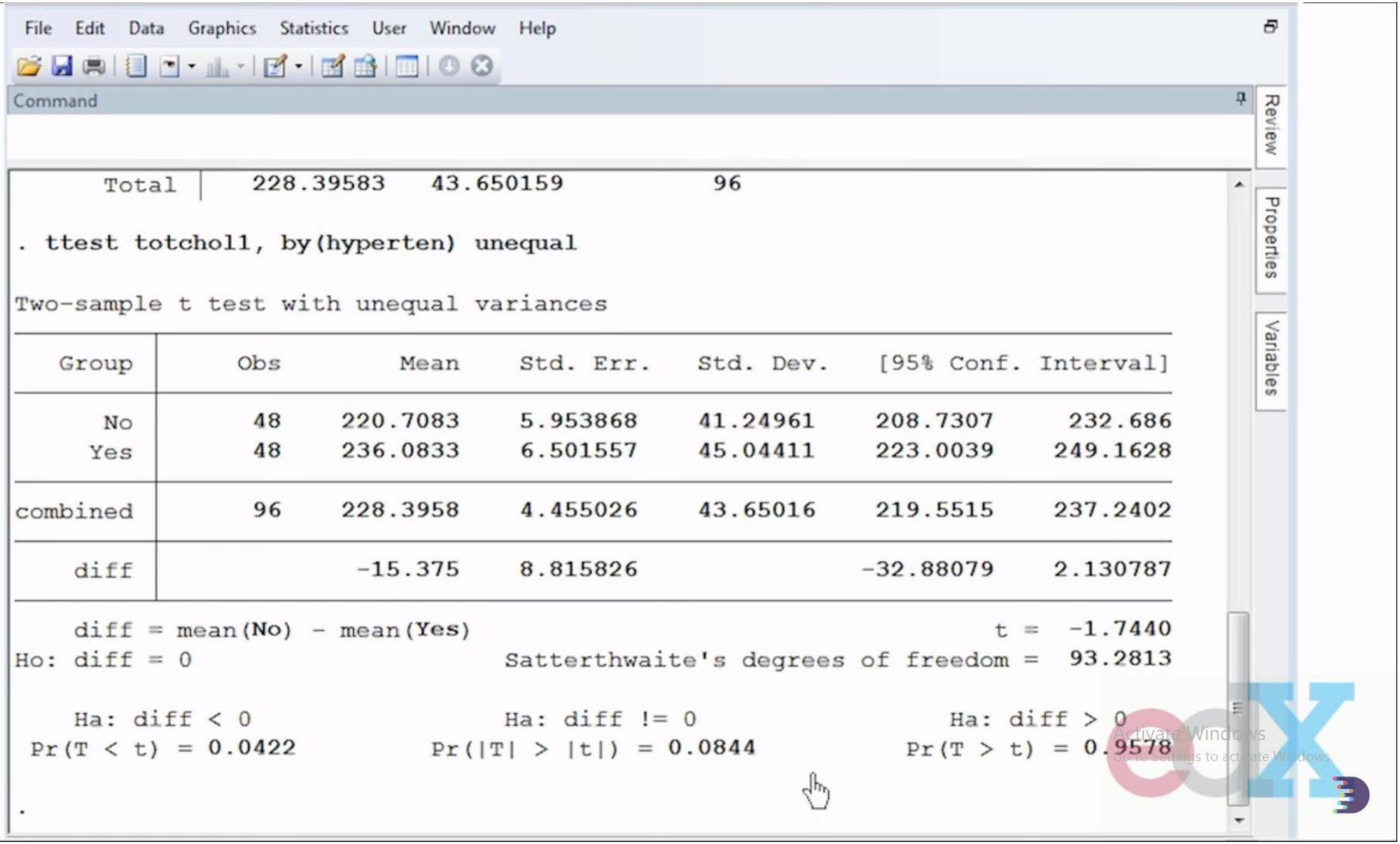

File Edit Data Graphics Statistics User Window Help
Command
Summaries, tables, and tests
Linear models and related
Binary outcomes
Ordinal outcomes
Categorical outcomes
Count outcomes
Exact statistics
Endogenous covariates
Sample-selection models
Multilevel mixed-effects models
Generalized linear models
Nonparametric analysis
Time series
Multivariate time series
State-space models
Longitudinal/panel data
Survival analysis
Epidemiology and related
SEM (structural equation modeling)
Survey data analysis
Multiple imputation
Multivariate analysis
Power and sample size
Summary and descriptive statistics
Tables
Classical tests of hypotheses
Nonparametric tests of hypotheses
Distributional plots and tests
Multivariate test of means, covariances, and normality
No 2
Yes 2
Total 2
. set seed 725764
. sample 49 , cou
(4336 observation
. tab hyperten, s
Incident Hypertensio n
No 2
Yes 2
Total 2
.
4382
Freq.
48
48
96
One-sa
Mean-
Two-sa
Two-gr
One-sa
Two-sa
Binomi
Binomi
One-sa
Two-sa
Two-gr
One-sa
Two-sa
One-sa
Two-sa
Two-gr
One-sa
Two-sa
Robust
Sample
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
Command
Review
Properties
Variables
No 227.98031 42.644789 1168
Yes 240.25638 44.919635 3214
Total 236.98425 44.65
. set seed 72576466
. sample 49 , count by(hyperten)
(4336 observations deleted)
. tab hyperten, summ(totchol1)
Incident Hypertensio n
Summary of Total (mg/dL), Mean Std.
No 220.70833 41.24
Yes 236.08333 45.04
Total 228.39583 43.650159 96
.
ttest - Two-group mean-comparison test
Main
by/if/in
Variable name:
totchol1
Group variable name:
95
Confidence level
Unequal variances
Welch's approximation
OK
Cancel
Submit


File Edit Data Graphics Statistics User Window Help
Command
No 227.98031 42.644789 1168
Yes 240.25638 44.919635 3214
Total 236.98425 44.65
. set seed 72576466
. sample 49 , count by(hyperten)
(4336 observations deleted)
. tab hyperten, summ(totchol1)
Incident Hypertensio n
Summary of Total (mg/dL), Mean Std.
No 220.70833 41.24
Yes 236.08333 45.04
Total 228.39583 43.650159 96
.
ttest - Two-group mean-comparison test
Main
by/if/in
Variable name:
totchol1
Group variable name:
hyperten
95
Confidence level
Unequal variances
Welch's approximation
OK
Cancel
Submit
Review
Properties
Variables


File Edit Data Graphics Statistics User Window Help
Command
Review
Properties
Variables
No 227.98031 42.644789 1168
Yes 240.25638 44.919635 3214
Total 236.98425 44.65
. set seed 72576466
. sample 49 , count by(hyperten)
(4336 observations deleted)
. tab hyperten, summ(totchol1)
Incident Hypertensio n
Summary of Total (mg/dL), Mean Std.
No 220.70833 41.24
Yes 236.08333 45.04
Total 228.39583 43.650159 96
.
ttest - Two-group mean-comparison test
Main
by/if/in
Variable name:
totchol1
Group variable name:
hyperten
95
Confidence level
Unequal variances
Welch's approximatio
OK
Cancel
Submit
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
Command
Review
Properties
Variables
Total | 228.39583 43.650159 96
. ttest totchol1, by(hyperten) unequal
Two-sample t test with unequal variances
Group | Obs Mean Std. Err. Std. Dev. [95% Conf. Interval]
No 48 220.7083 5.953868 41.24961 208.7307 232.686
Yes 48 236.0833 6.501557 45.04411 223.0039 249.1628
combined 96 228.3958 4.455026 43.65016 219.5515 237.2402
diff -15.375 8.815826 -32.88079 2.130787
diff = mean(No) - mean(Yes) t = -1.7440
Ho: diff = 0 Satterthwaite's degrees of freedom = 93.2813
Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 0.0422 Pr(|T| > |t|) = 0.0844 Pr(T > t) = 0.9578
.


File Edit Data Graphics Statistics User Window Help
Command
Review
Properties
Variables
Total | 228.39583 43.650159 96
. ttest totchol1, by(hyperten) unequal
Two-sample t test with unequal variances
Group | Obs Mean Std. Err. Std. Dev. [95% Conf. Interval]
No 48 220.7083 5.953868 41.24961 208.7307 232.686
Yes 48 236.0833 6.501557 45.04411 223.0039 249.1628
combined 96 228.3958 4.455026 43.65016 219.5515 237.2402
diff -15.375 8.815826 -32.88079 2.130787
diff = mean(No) - mean(Yes) t = -1.7440
Ho: diff = 0 Satterthwaite's degrees of freedom = 93.2813
Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 0.0422 Pr(|T| > |t|) = 0.0844 Pr(T > t) = 0.9578
.

One sample hypothesis testing about the mean

1. Set up null hypothesis

$$(H_0 : \mu = \mu_0)$$

2. Set up the alternative

$$(H_A : \mu \neq \mu_0)$$

3. Choose α-level

$$(\alpha = 0.05)$$

4. Take a sample and calculate

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \quad \text{or} \quad t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

$\Pr(\text{rejecting } H_0 \mid H_0 \text{ is correct}) = \alpha$

$\Pr(\text{not rejecting } H_0 \mid H_A \text{ is correct}) = \beta$

$\Pr(\text{rejecting } H_0 \mid H_A \text{ is correct}) = \text{Power}$

Remember: $\text{Power} = 1 - \beta$

- Omniflox, broad spectrum antibiotic, recalled < 6 mos, severe reactions; 3 deaths
- Versed, a sedative, < 18 mos, 86 adverse reactions including 46 deaths.
- Fenoterol, many years, relieve asthma attacks, increased the risk of death.
- Oraflex, antiinflammatory (arthritis) < 3 mos, 72 deaths (USA & UK.)

2000: Lotronex -- Risks: Intestinal damage from reduced blood flow

2000: Propulsid -- Risks: Fatal heart rhythm abnormalities

2000: Rezulin -- Risks: Severe liver toxicity

1999: Hismanal -- Risks: With other drugs or high dose can cause fatal heart rhythm

1999: Raxar -- Risks: Fatal heart rhythm abnormalities

1998: Posicor -- Risks: Dangerous interaction with other drugs

1998: Duract -- Risks: Severe liver damage

1998: Seldane -- Risks: Fatal heart rhythm abnormalities

1997: Pondimin -- Risks: Heart valve abnormalities

1997: Redux -- Risks: Heart valve abnormalities

Bendectin Story

- Hyperemisis gravidarum—morning sickness
- 1956 Bendectin introduced (FDA approved) (known as Debendox in the UK and Diclectin in Canada) is a mixture of pyridoxine (Vitamin B-6), and doxylamine.
- 1979 *National Enquirer* attributes "hideous birth defects" to bendectin. 'Experts' compare it to thalidomide.
- 1983 After millions of dollars in litigation costs for alleged birth defects (including *Daubert v. Merrell Dow Pharmaceuticals* (1993)) Bendectin removed from market.

Aftermath

- 30 epidemiological studies, WHO, FDA & March of dimes, concur that Bendectin is safe
- Not one court case lost
- Since 1983, CDC
  - no significant decrease in incidence of birth defect
  - hospitalizations for hyperemesis gravidarum has doubled

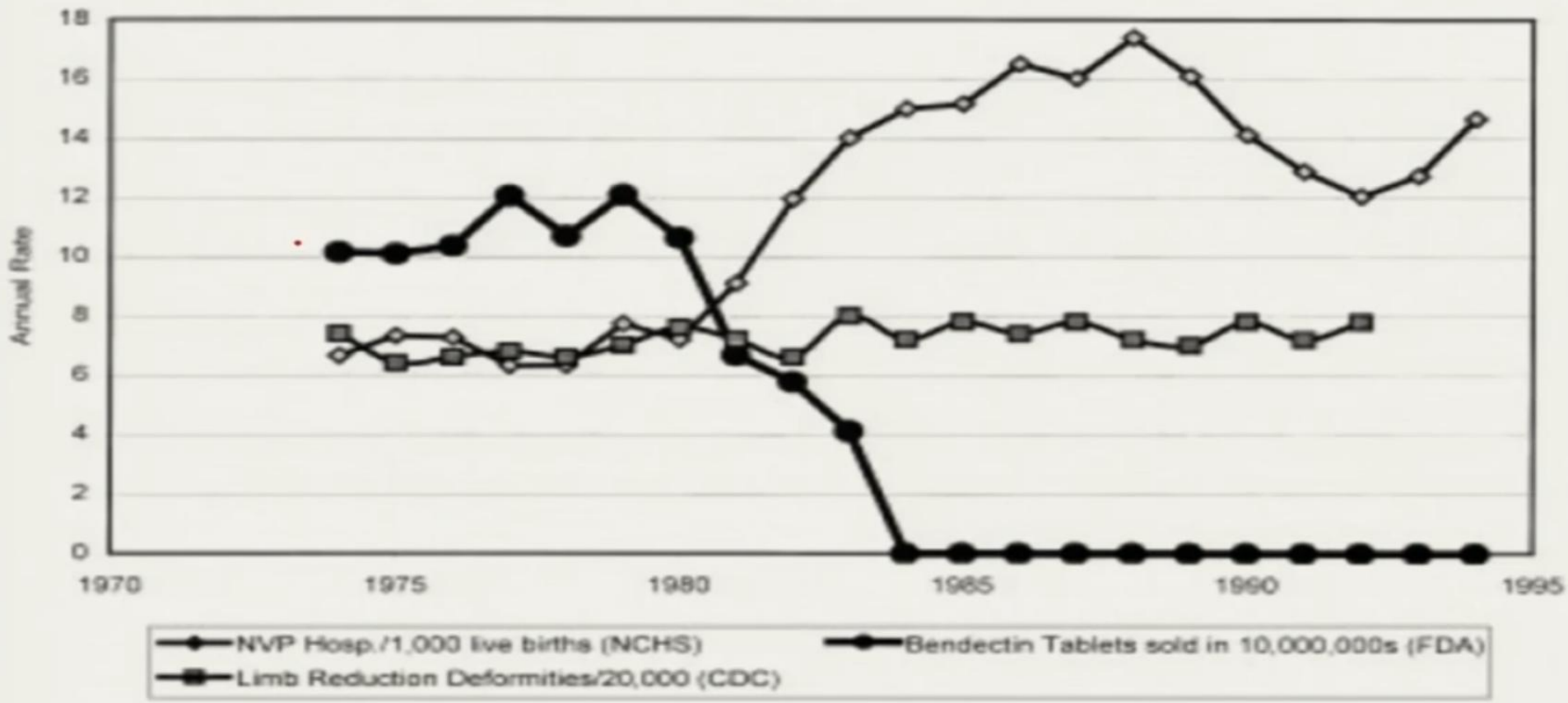


Figure 12. Public health data related to Bendectin therapy.

JS Kutcher, A Engle, J Firth, SH. Lamm ,
Birth Defects Research (Part A) 67:88–97 (2003)

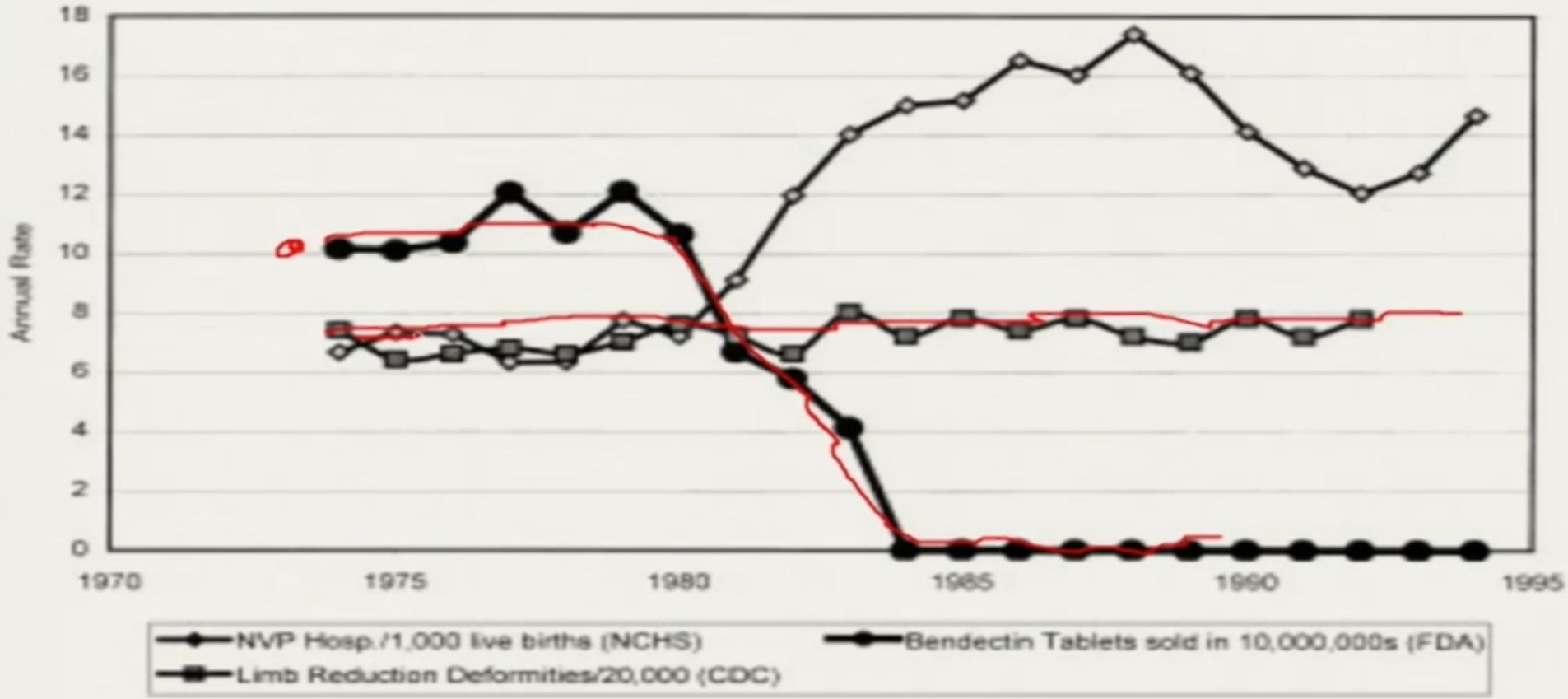


Figure 12. Public health data related to Bendectin therapy.

JS Kutcher, A Engle, J Firth, SH. Lamm ,
Birth Defects Research (Part A) 67:88–97 (2003)

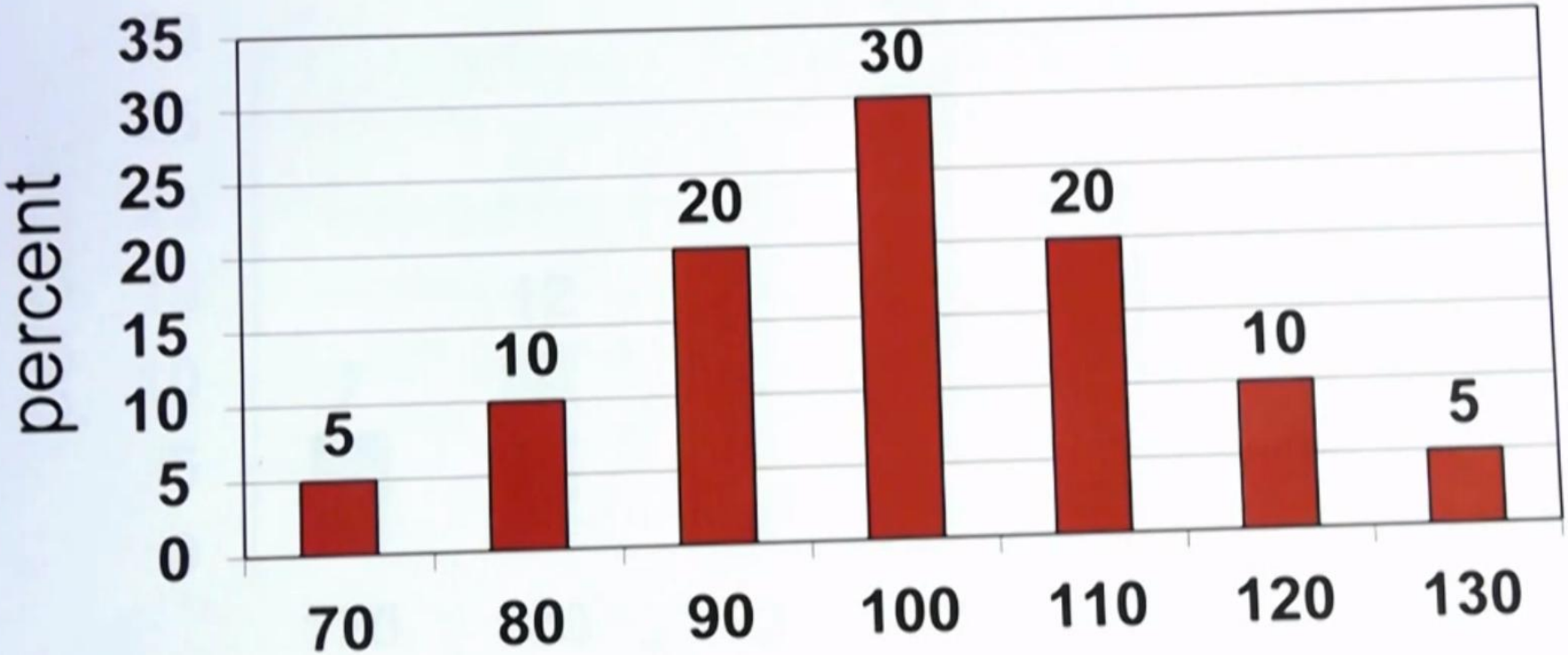
Melvin Belli
1907-1996

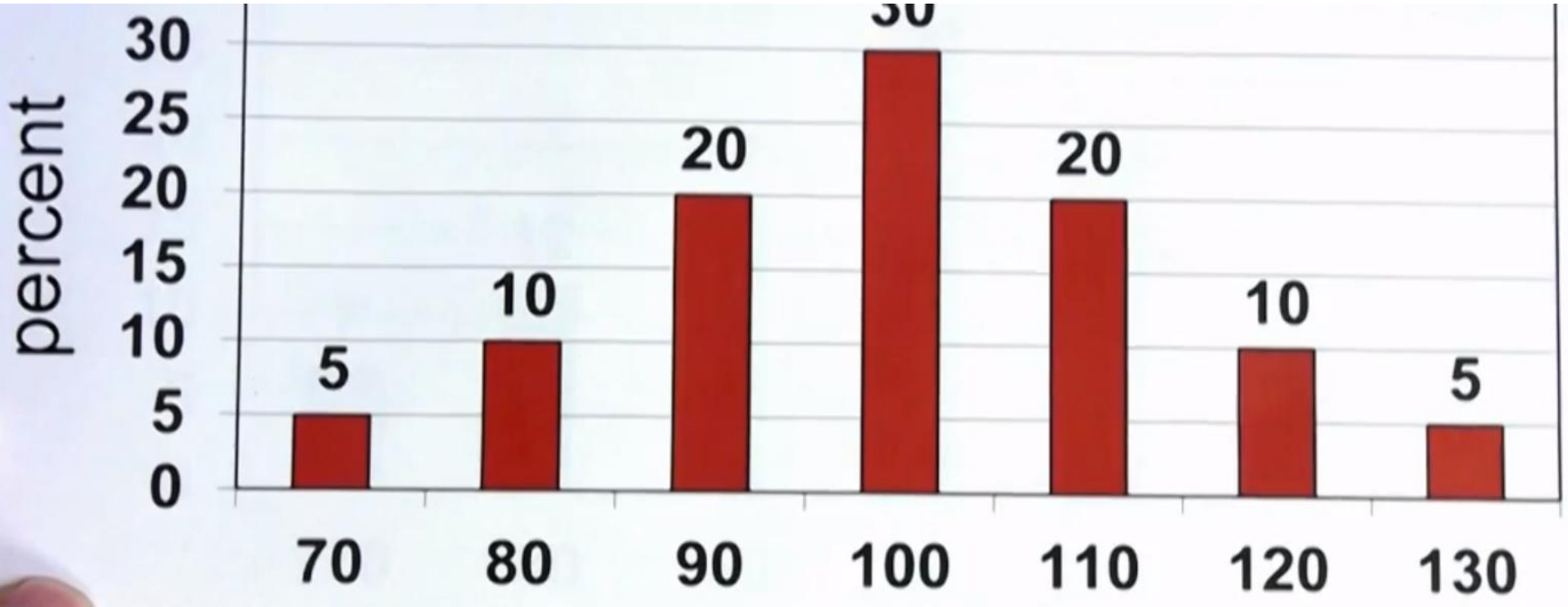
Bendectin and Birth Defects
The Challenges of Mass Toxic Substances Litigation
(1997) University of Pennsylvania Press p. 106, 124

Michael D. Green


SHEIN
COMPRA ORA

Flabrats

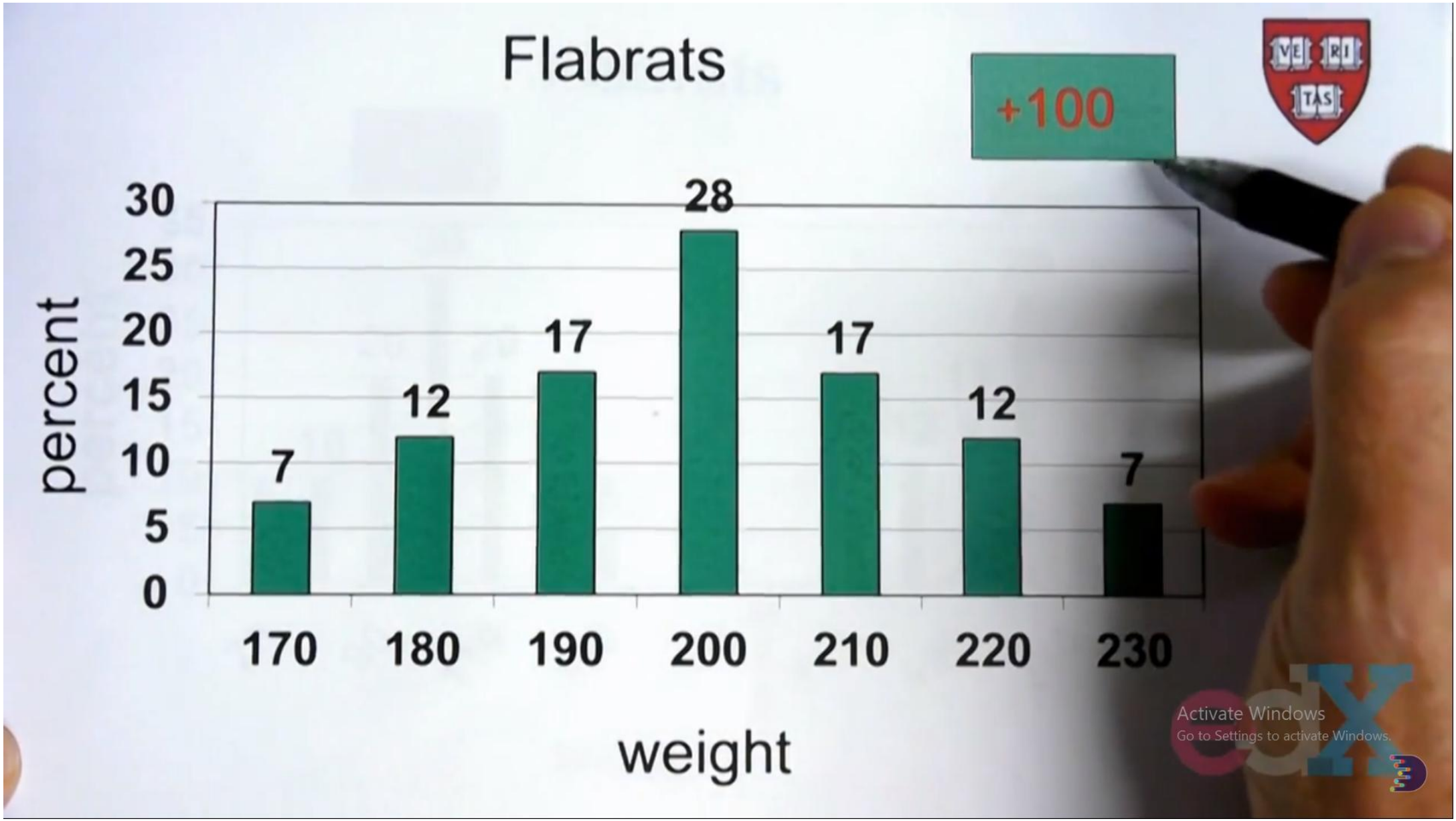


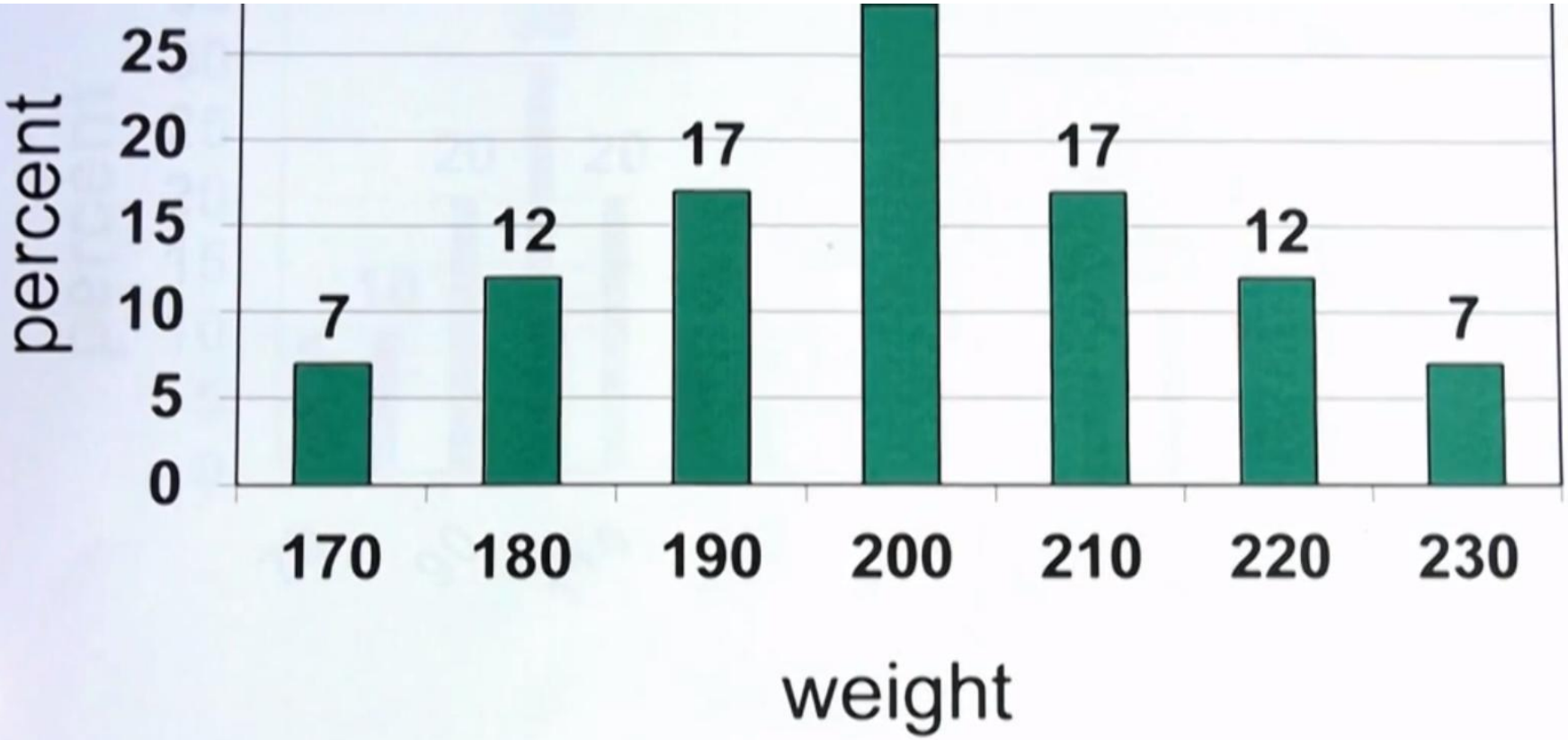


Mean = 100

Std. dev. = 14.5

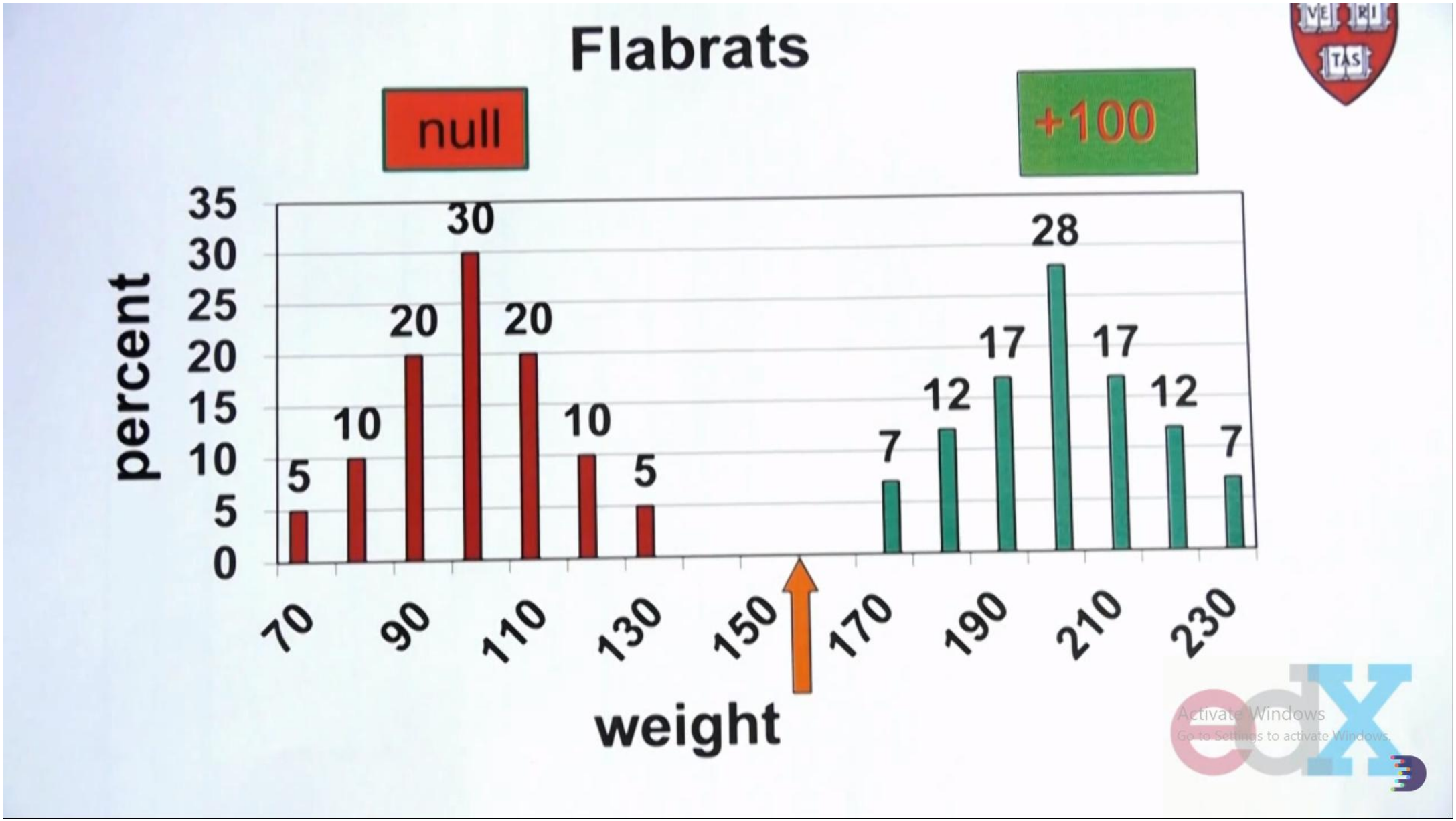
(null)

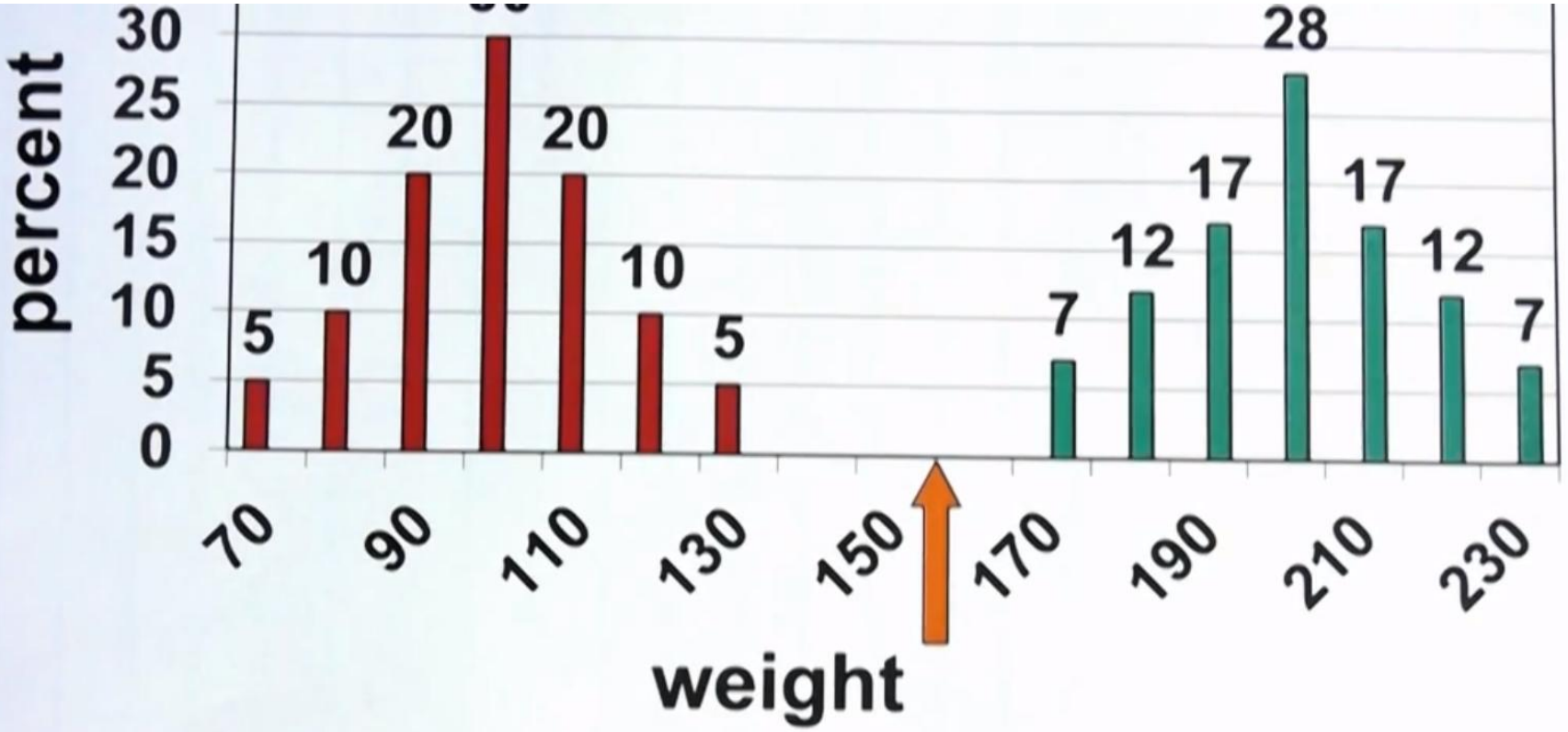

Flabrats
+100
percent
30
25
20
15
10
5
0
7
12
17
28
17
12
7
170
180
190
200
210
220
230
weight
Activate Windows
Go to Settings to activate Windows.



Mean = 200

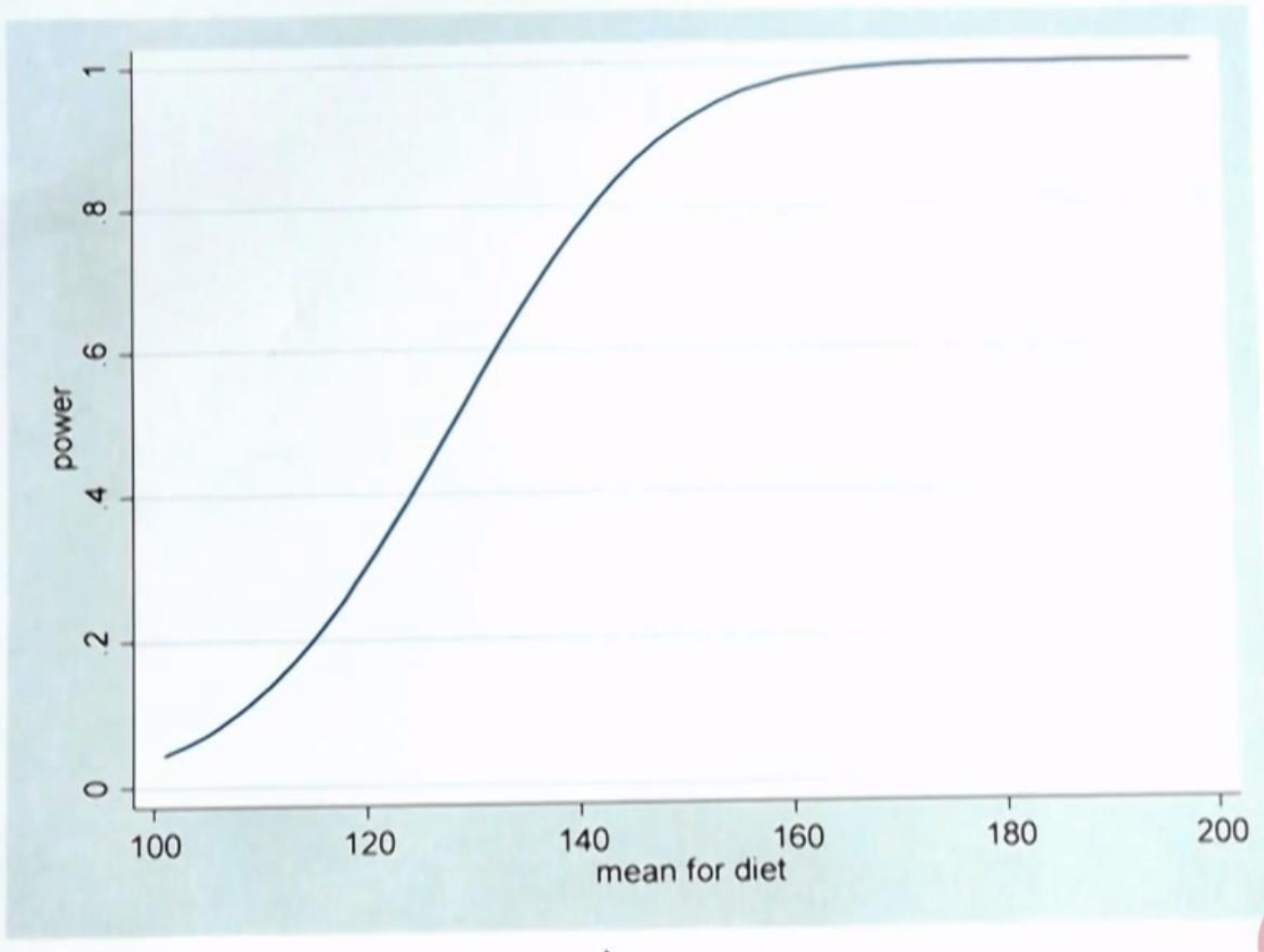
Std. dev. = 16

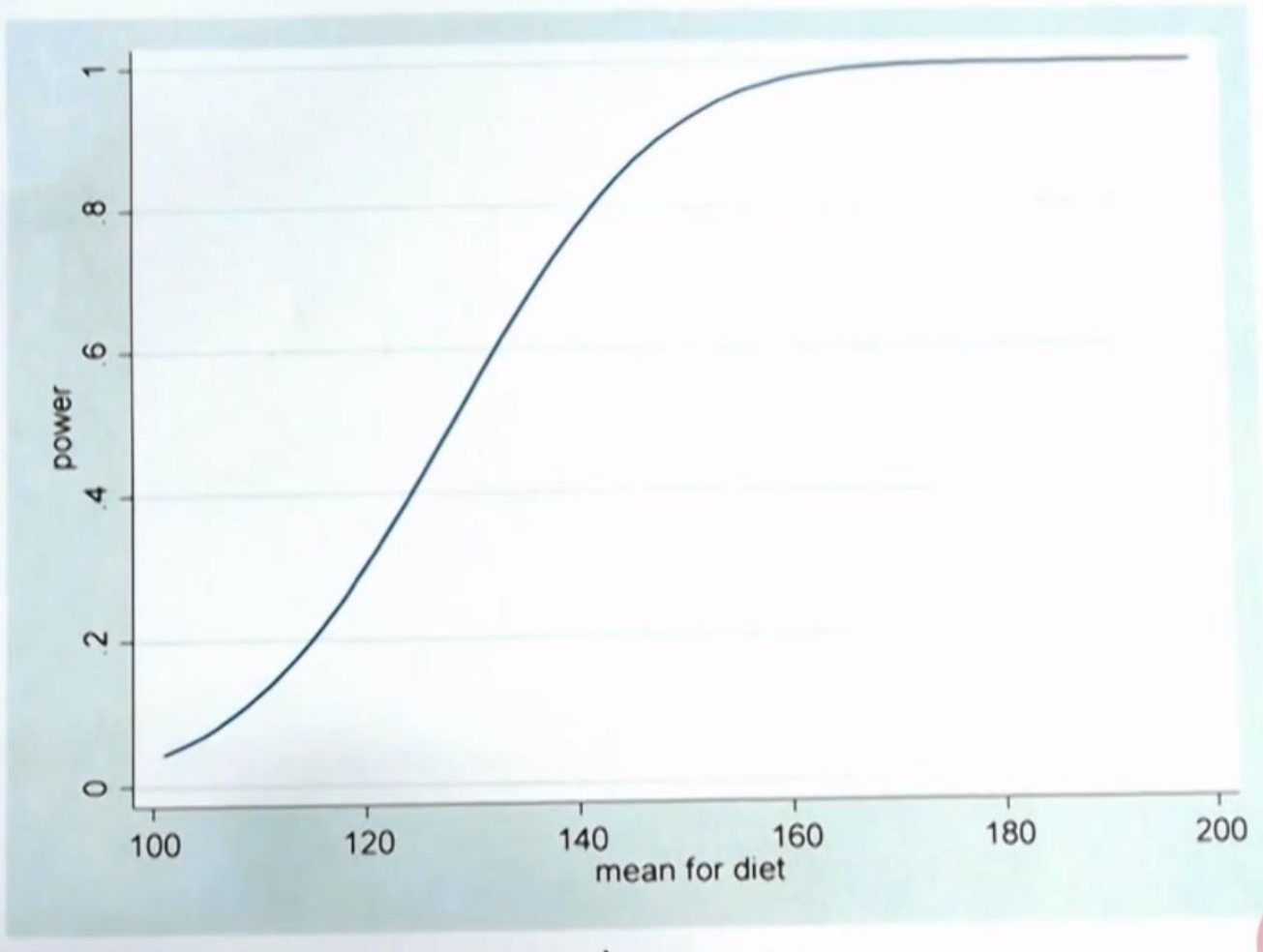

Flabrats
null
+100
percent
35
30
25
20
15
10
5
0
5
10
20
30
20
10
5
7
12
17
28
17
12
7
70
90
110
130
150
170
190
210
230
weight

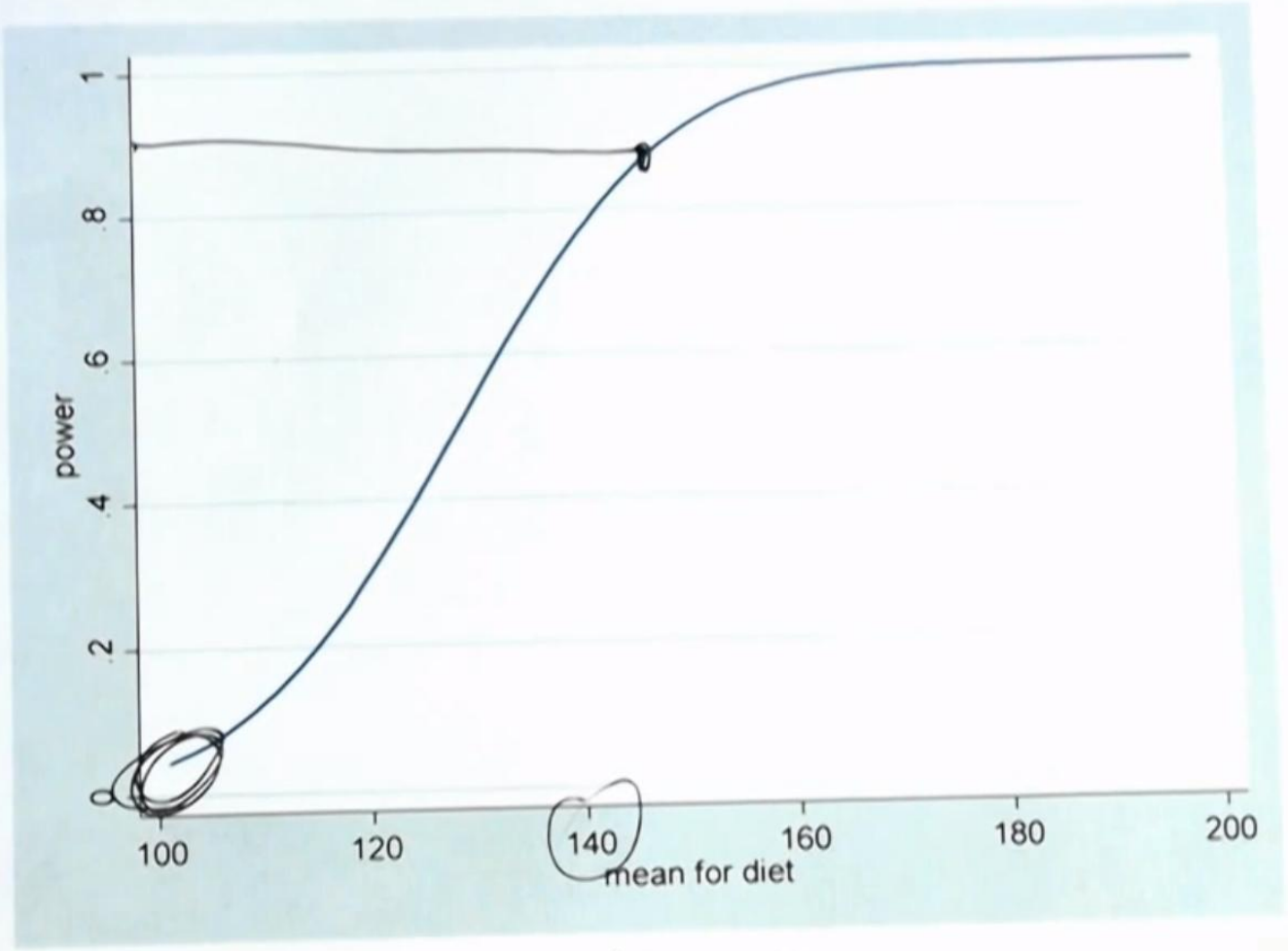


If new mean = 200, then α = 0 and β = 0 (power=1) ;

i.e. no errors if classifying a single fla rat.



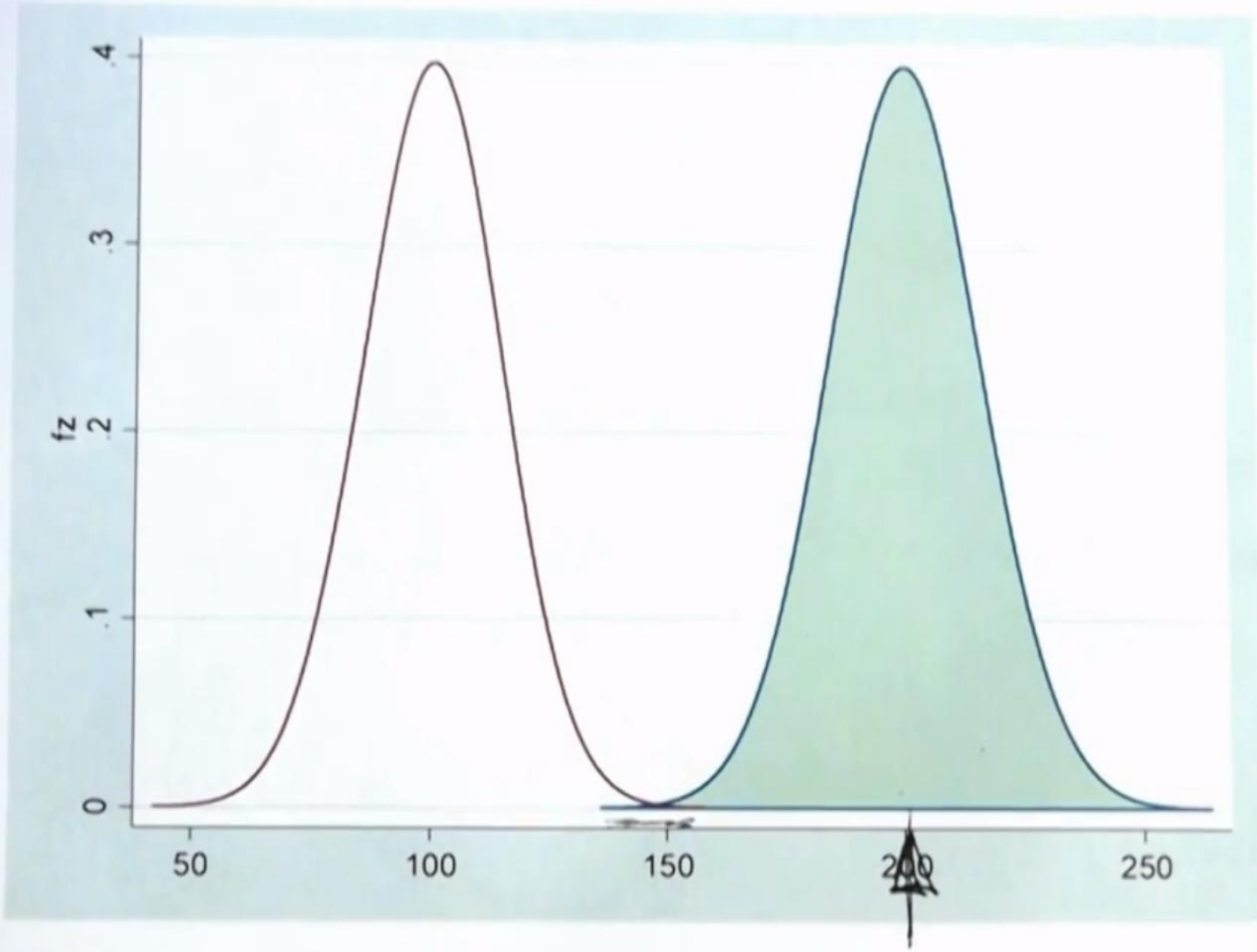


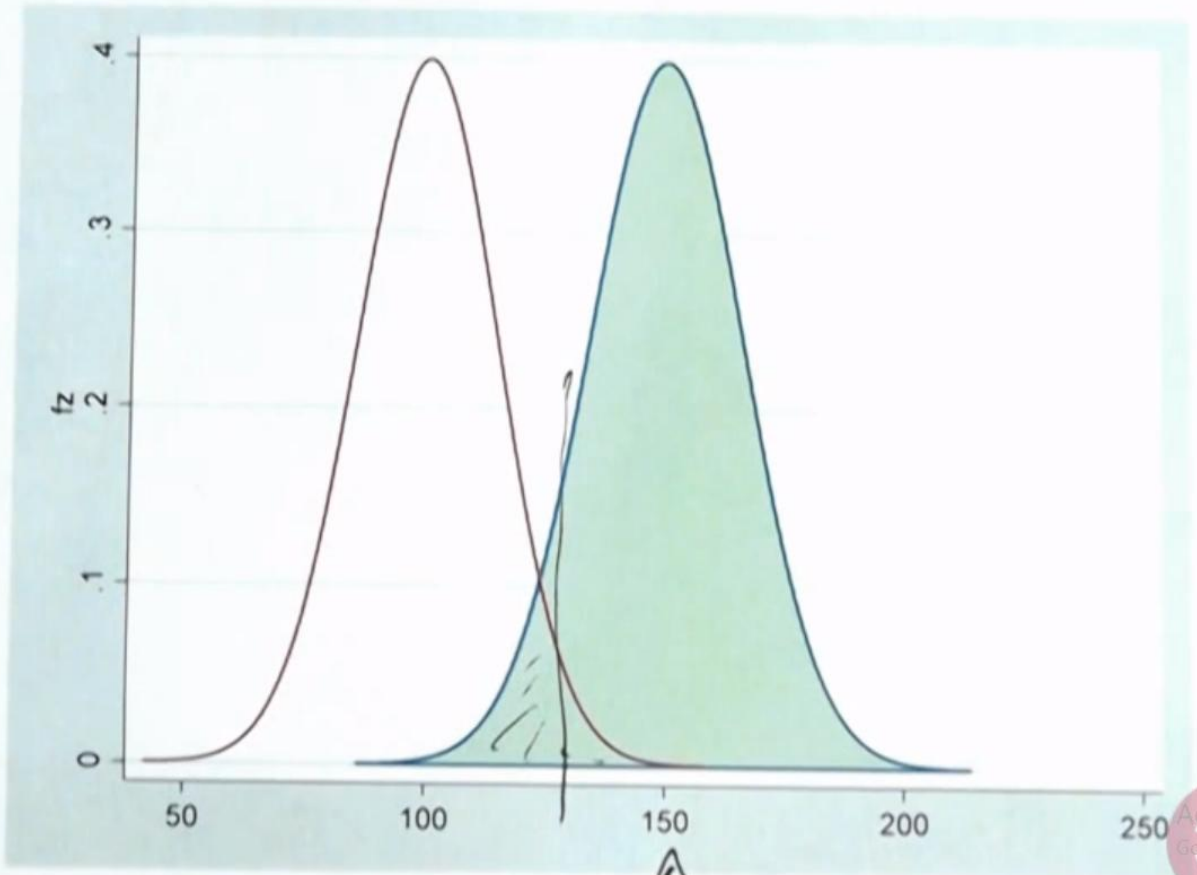


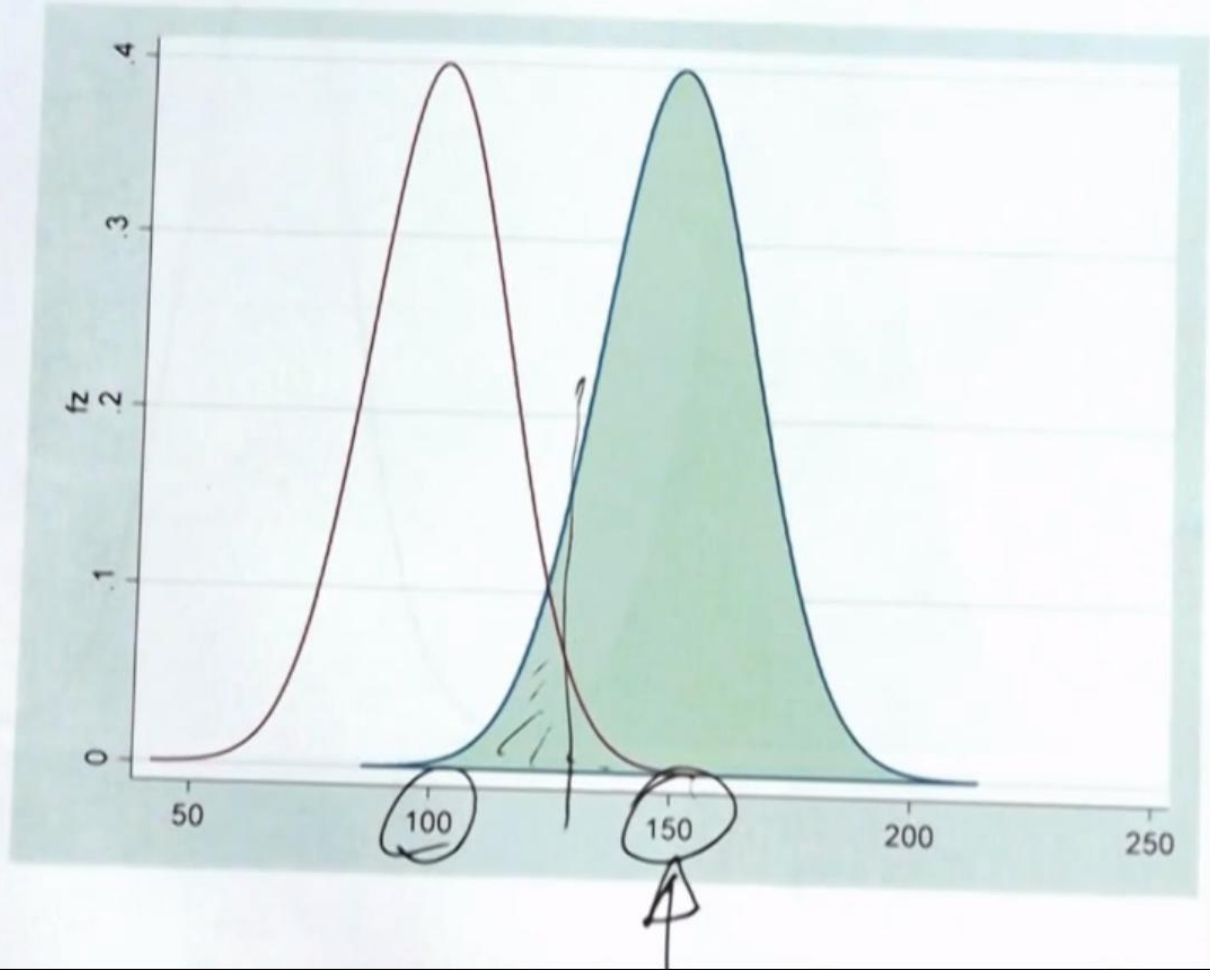
Power curve

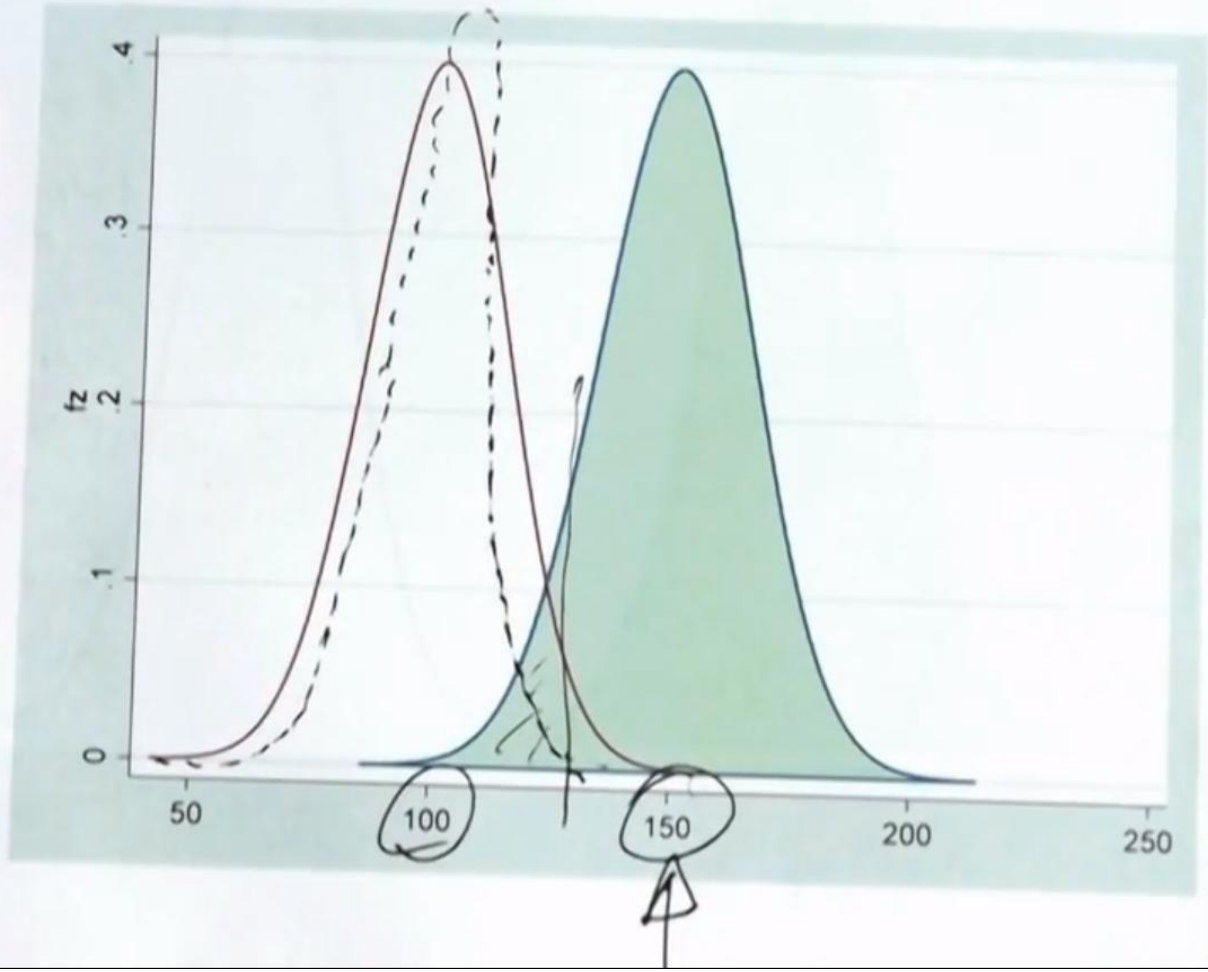
The power increases as the two population means become further apart – (Δ increases)

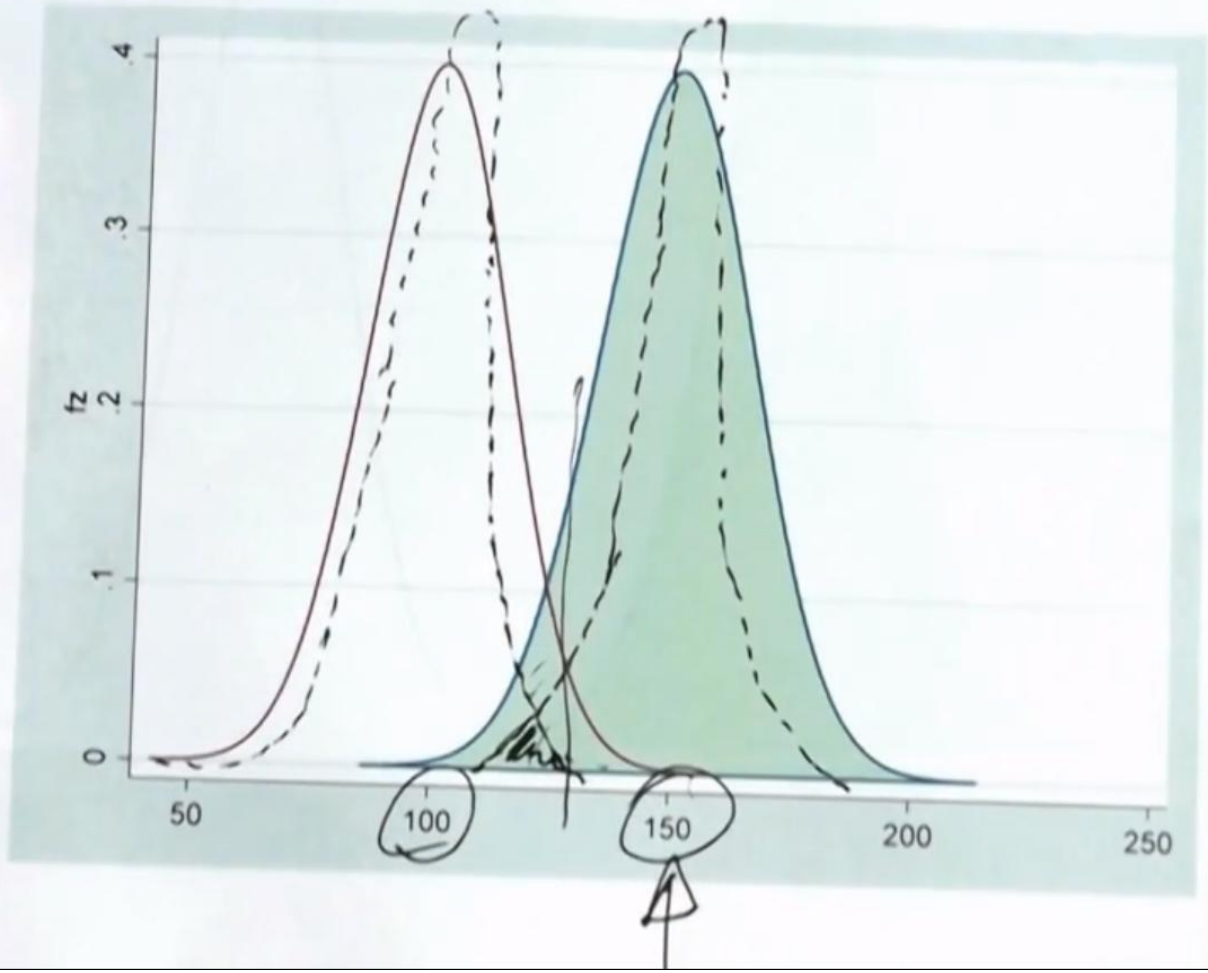
The power increases as the two population means become further apart – (Δ increases) holding α constant

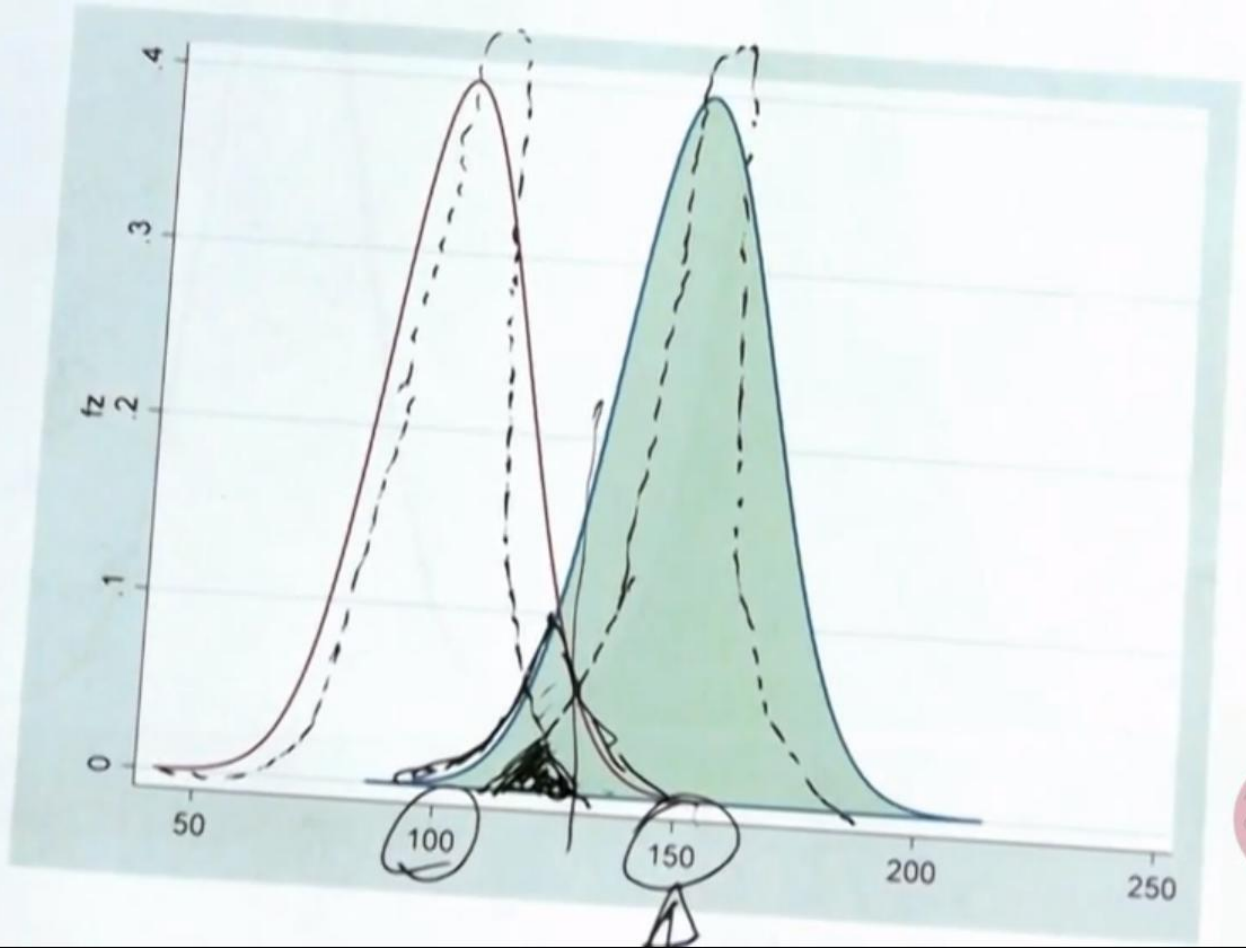




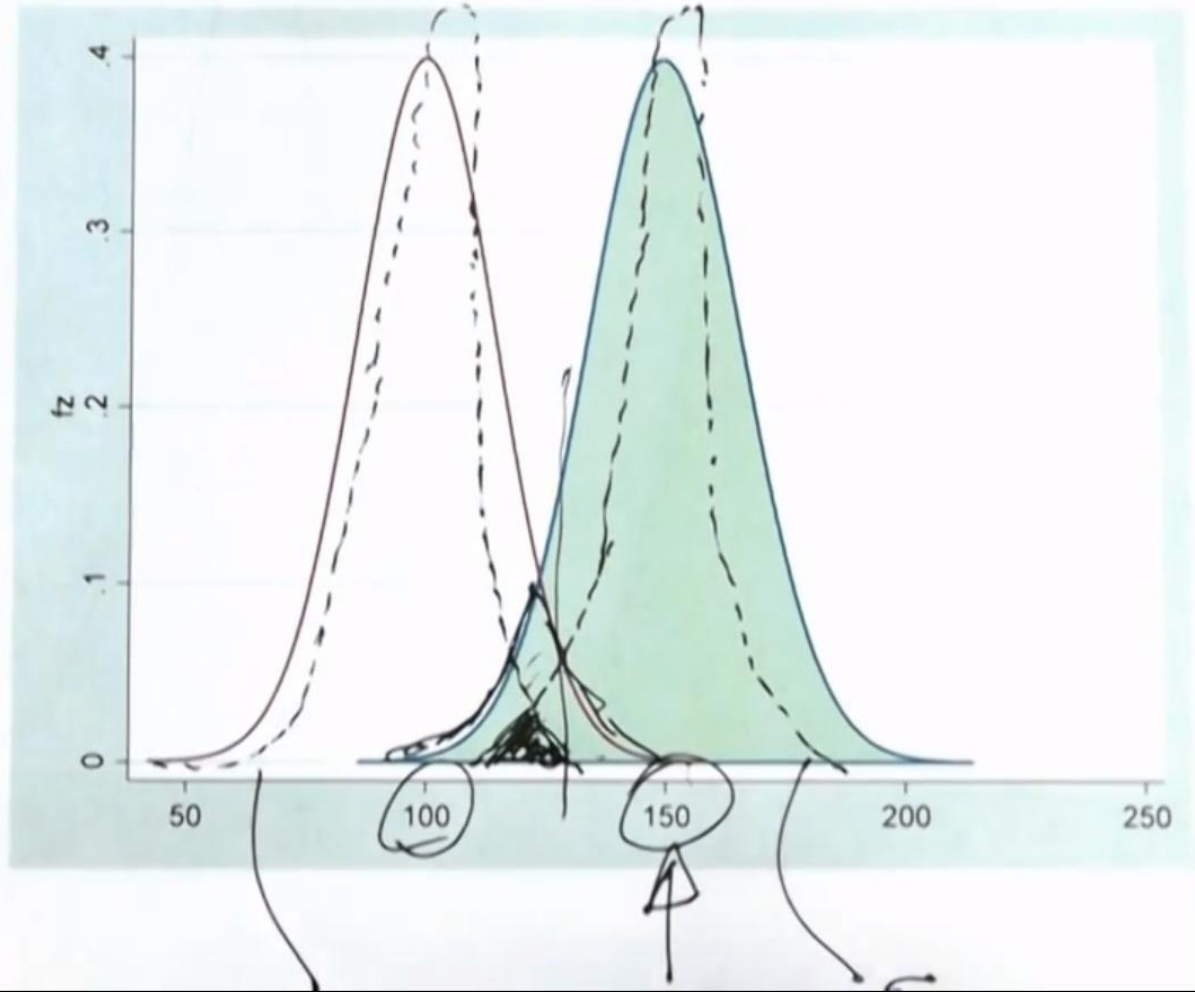

.4
.3
.2
.1
0
fz
50
100
150
200
250

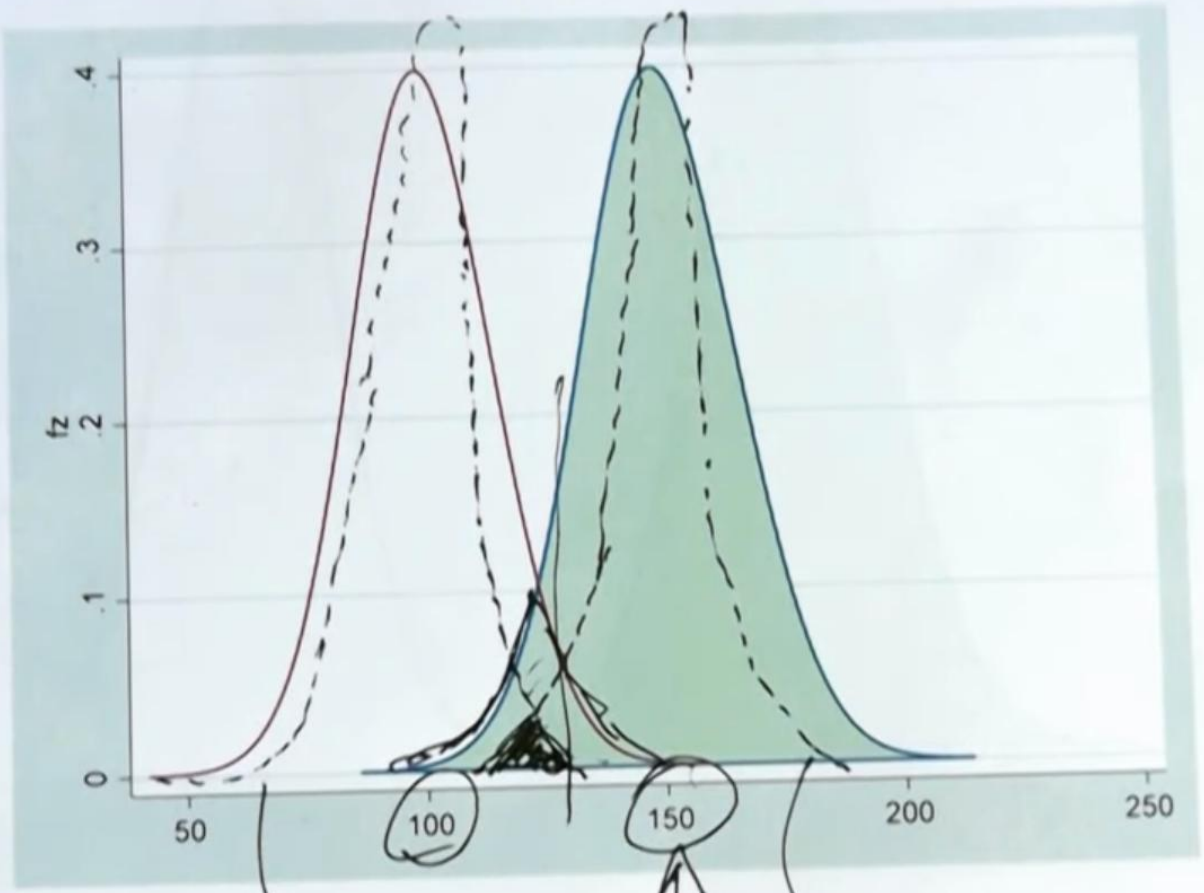

.4
.3
.2
.1
0
fz
50
100
150
200
250

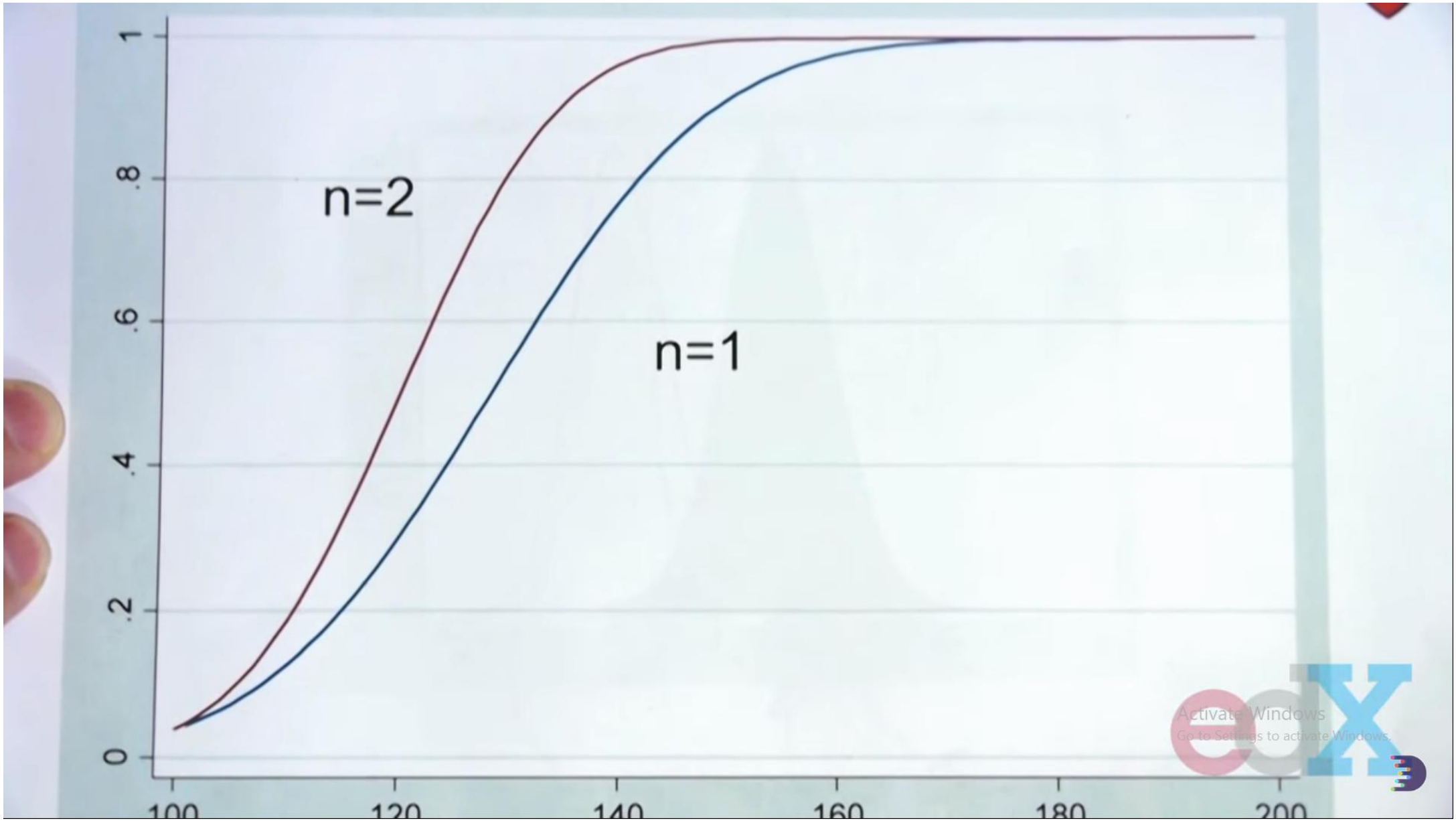

.4
.3
.2
.1
0
fz
50
100
150
200
250

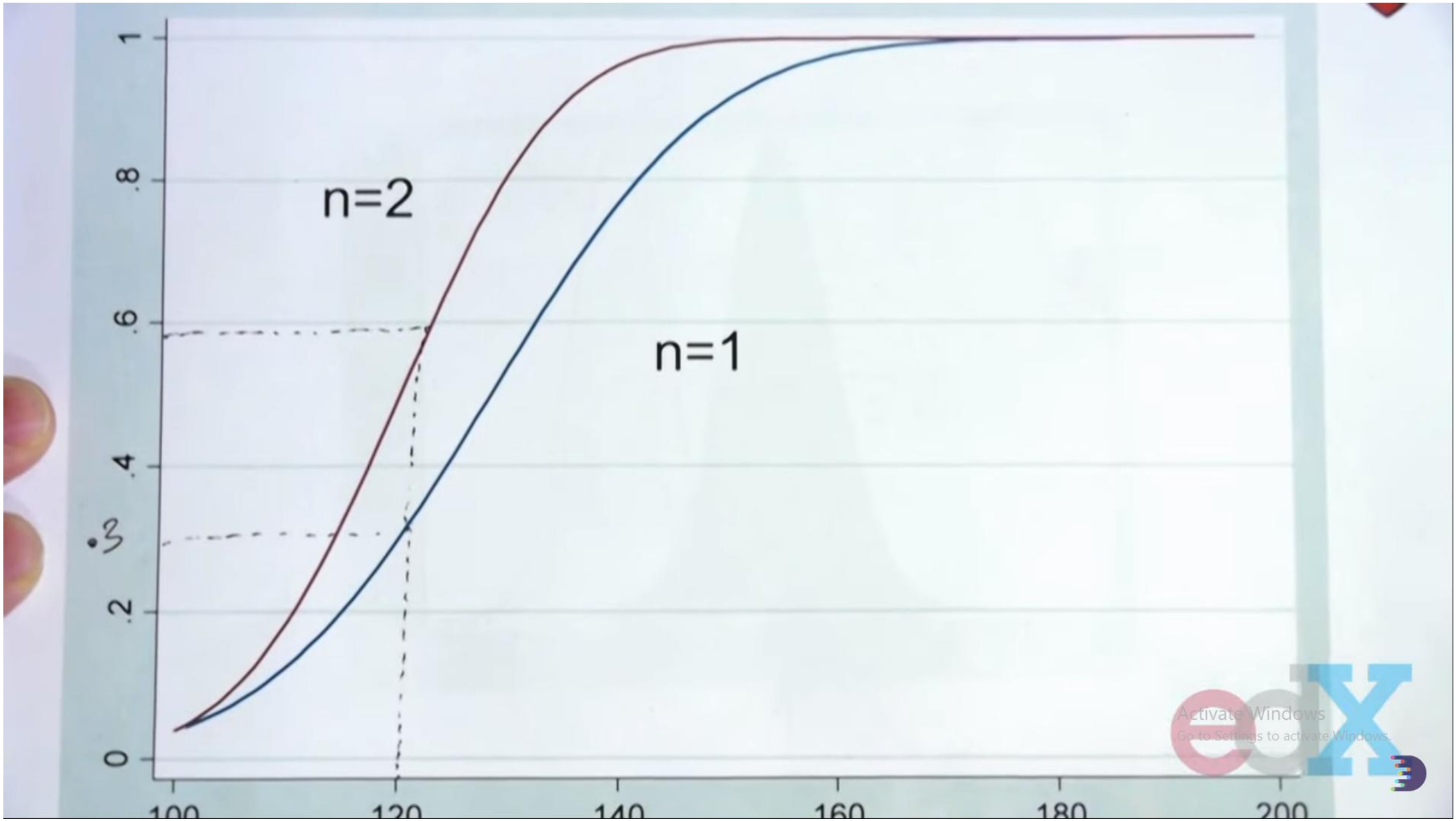

.4
.3
.2
.1
0
fz
50
100
150
200
250

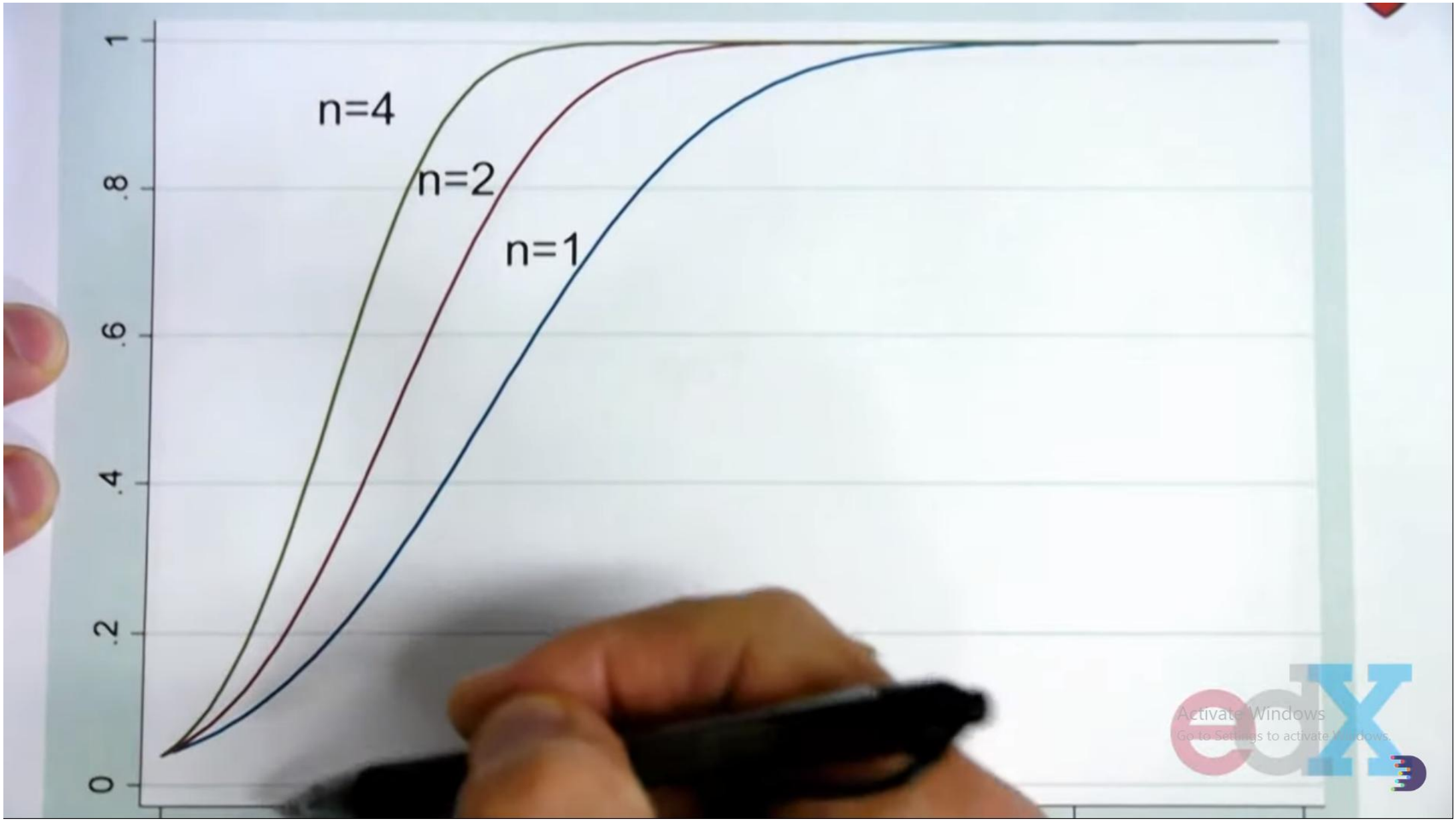
$$\frac{\sigma}{\sqrt{2}}$$

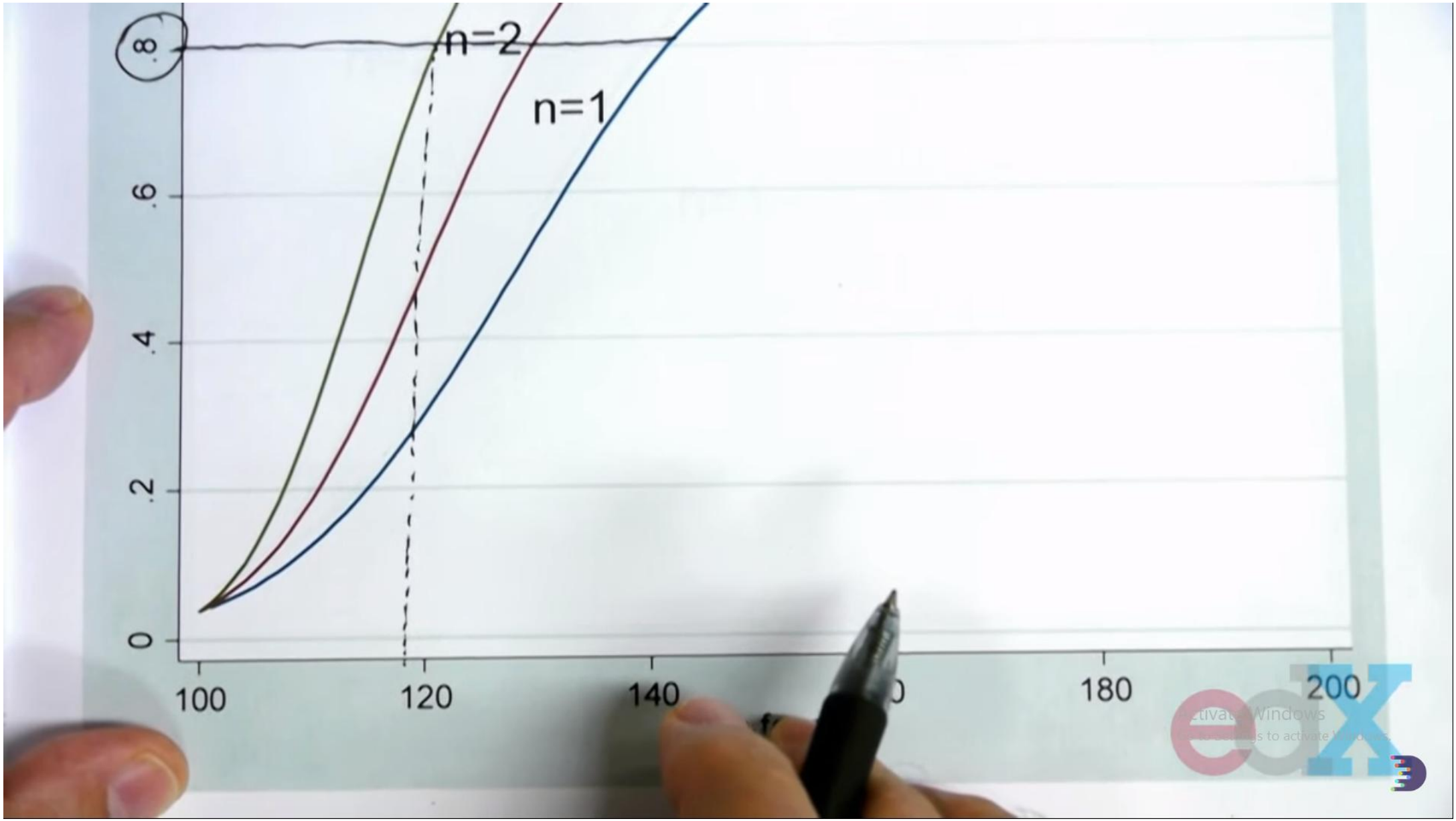


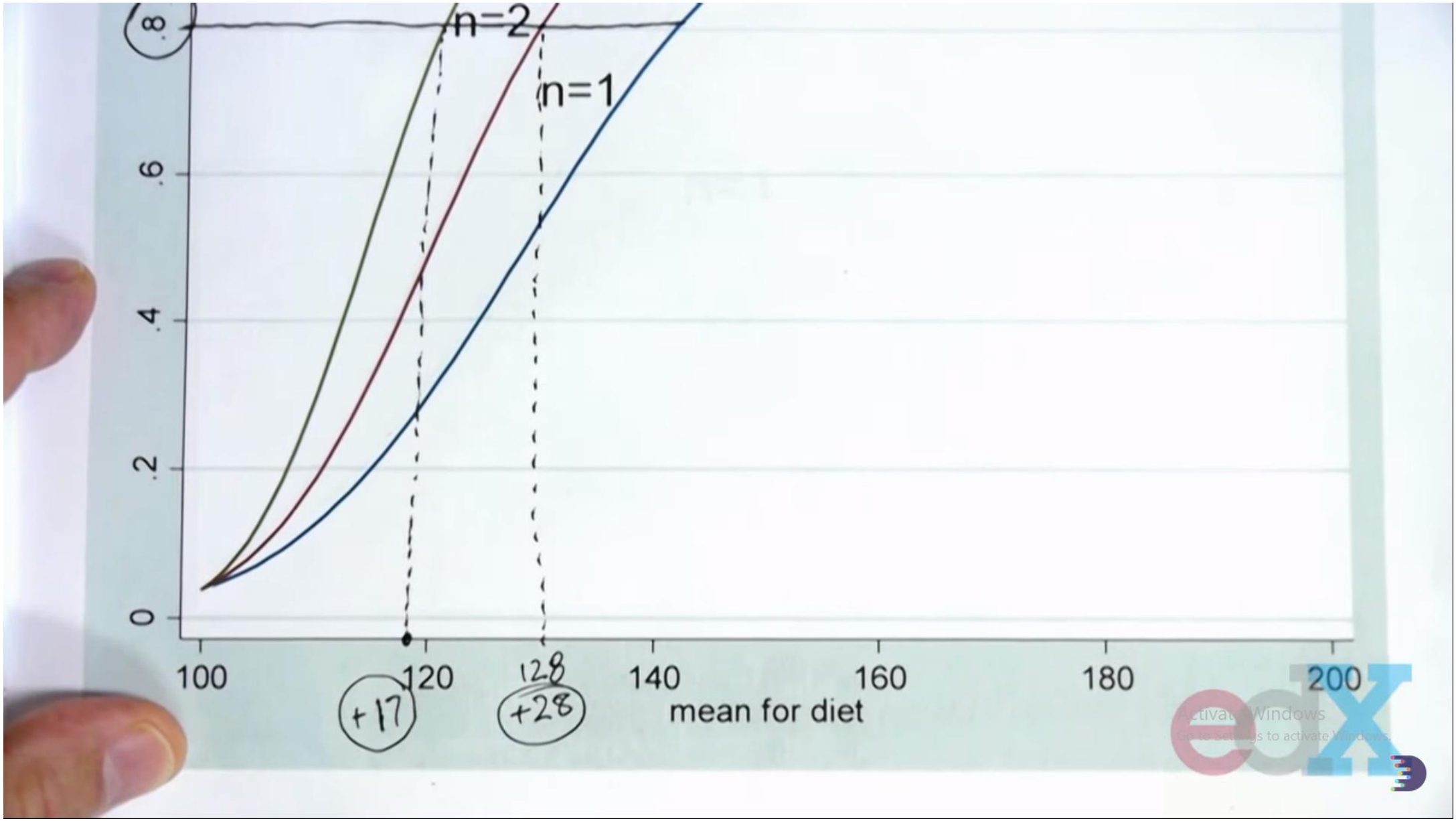


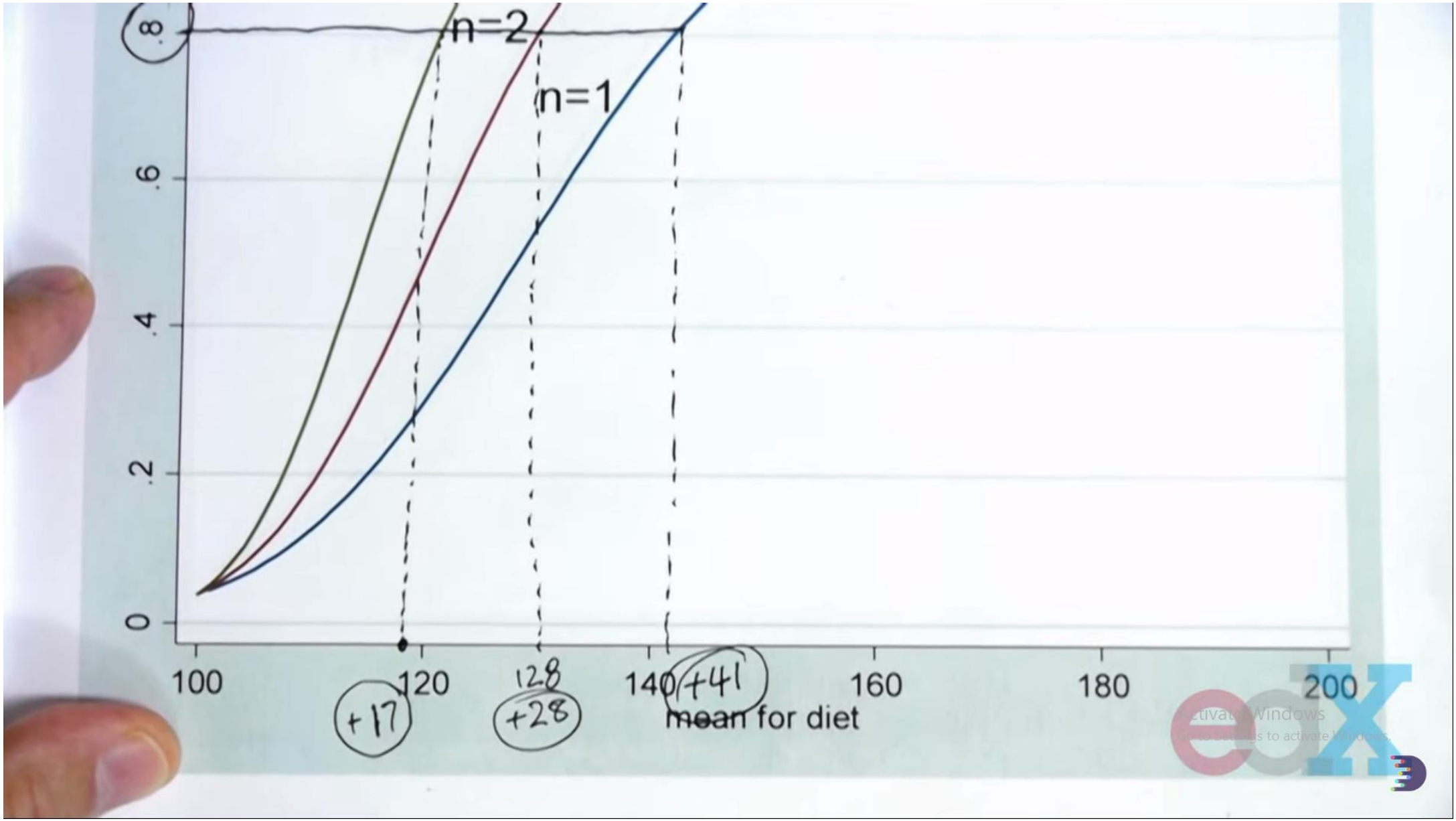

n=2
n=1
0
.2
.4
.6
.8
1
100
120
140
160
180
200


n=2
n=1
1
.8
.6
.4
.2
0
100
120
140
160
180
200


n=4
n=2
n=1
1
.8
.6
.4
.2
0


n=2
n=1
.8
.6
.4
.2
0
100
120
140
180
200


.8
.6
.4
.2
0
n=2
n=1
100
120
128
140
160
180
200
mean for diet
+17
+28


.8
.6
.4
.2
0
n=2
n=1
100
120
128
140
160
180
200
+17
+28
+41
mean for diet

One population:

$$H_0 : \mu = \mu_0$$

One population:

$$H_0 : \mu = \mu_0$$

Two populations:

$$H_0 : \mu_1 - \mu_2 = \Delta$$

Multiple populations:

$$H_0 : \mu_1 = \ldots = \mu_k \quad (k \geq 2)$$

Data, k independent, random samples:

Data, k independent, random samples:

Population:	1	2	⋯	k
	$x_{1,1}$	$x_{2,1}$	⋯	$x_{k,1}$
	$x_{1,2}$	$x_{2,2}$	⋯	$x_{k,2}$
	⋮	⋮	⋮	⋮
Sample size	n_1	n_2	⋯	n_k
Sample mean	$\bar{x}_1$	$\bar{x}_2$	⋯	$\bar{x}_k$
Sample sd's	s_1	s_2	⋯	s_k

Populations:

means: $\mu_1, \ldots, \mu_k$

s.d.'s : $\sigma_1, \ldots, \sigma_k$

Populations:

means: $\mu_1, \ldots, \mu_k$

s.d.'s : $\sigma_1, \ldots, \sigma_k$

Null hypothesis:

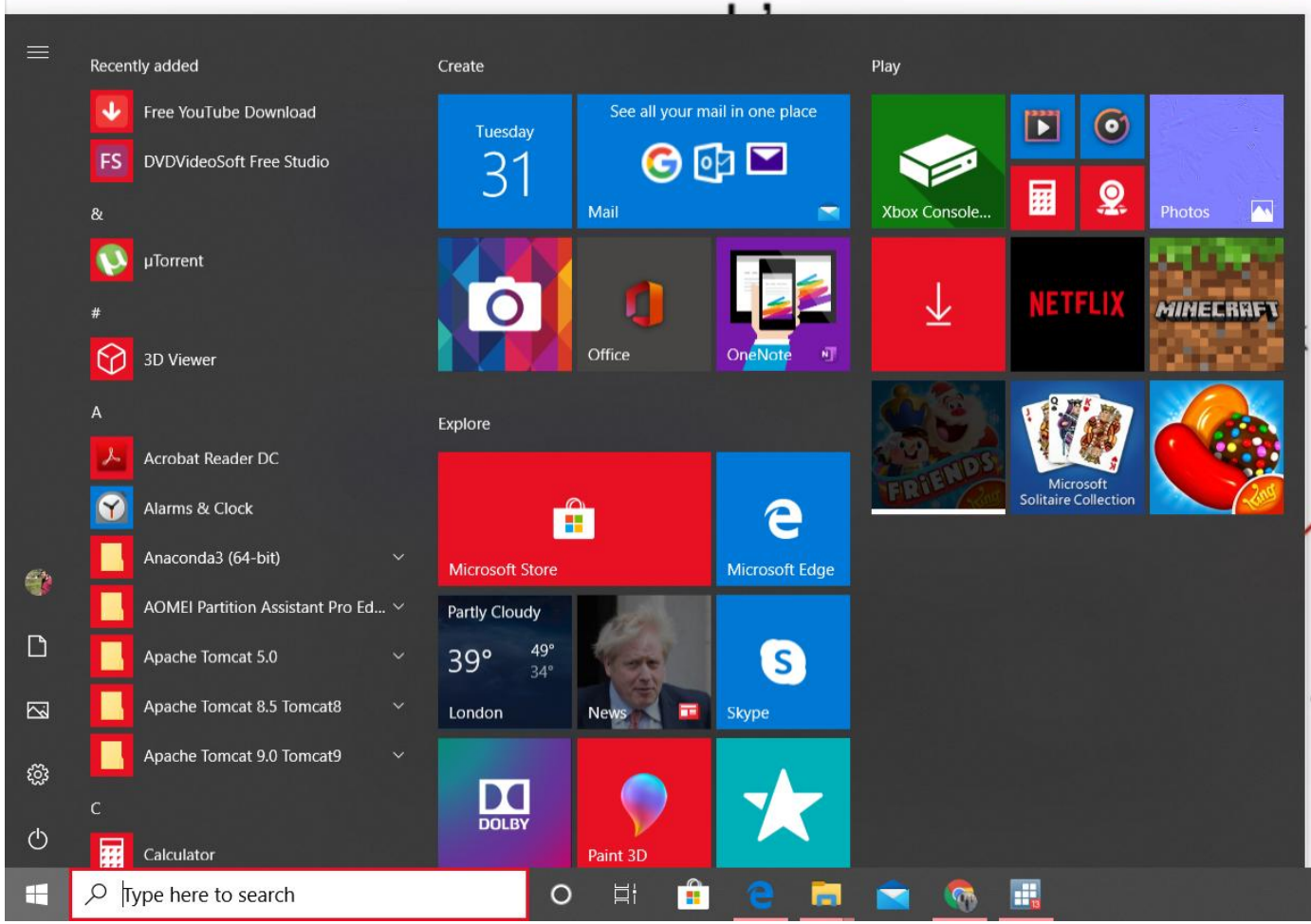
$$H_0 : \mu_1 = \ldots = \mu_k$$

Assumptions:

Populations:

means: $\mu_1, \ldots, \mu_k$

$\cdot = \sigma_k$



Populations:

means: $\mu_1, \ldots, \mu_k$

s.d.'s : $\sigma_1, \ldots, \sigma_k$

Null hypothesis:

$$H_0 : \mu_1 = \ldots = \mu_k$$

Assumptions:

1. Populations are normal.
2. Homoscedasticity: $\sigma_1 = \ldots = \sigma_k$
3. Independent samples (k)

We could test the null hypothesis by testing for each pair i,j the hypotheses:

$$H_0 : \mu_i = \mu_j \qquad i,j = 1,2,\ldots,k$$

We could test the null hypothesis by testing for each pair i,j the hypotheses:

$$H_0 : \mu_i = \mu_j \qquad i,j = 1,2,\ldots,k$$

$\mu_1 - \mu_2 = 0$

$\mu_1 - \mu_3 = 0$

$\mu_2 - \mu_3 = 0$

But if we did proceed in this fashion consider the type I error rate: e.g. $k = 5$ so there are 10 hypotheses we need to test.

We could test the null hypothesis by testing for each pair i,j the hypotheses:

$$H_0 : \mu_i = \mu_j \qquad i,j = 1,2,\ldots,k$$

$\binom{5}{2} = 10$

But if we did proceed in this fashion consider the type I error rate: e.g. $k = 5$ so there are 10 hypotheses we need to test.

We could test the null hypothesis by testing for each pair i,j the hypotheses:

$$H_0 : \mu_i = \mu_j \qquad i,j = 1,2,\ldots,k$$

$\binom{5}{2} = 10$

But if we did proceed in this fashion consider the type I error rate: e.g. $k = 5$ so there are 10 hypotheses we need to test.

$1 - \alpha = 95\%$

We could test the null hypothesis by testing for each pair i,j the hypotheses:

$$H_0 : \mu_i = \mu_j \qquad i,j = 1,2,\ldots,k$$

$\binom{5}{2} = 10$

But if we did proceed in this fashion consider the type I error rate: e.g. $k = 5$ so there are 10 hypotheses we need to test.

$1 - \alpha = 0.95$

Then ask ourselves, what is the probability that we get it right 10 times, even if the null hypothesis is true?

Hint : $(0.95)^{10} = 0.6$

Alternatively,.....

Homoscedasticity assumption and within variance

From the assumptions we have that s_1 , s_2 , …, s_k all estimate σ the common value of the standard deviation in each of the groups.

So, combine to get a better estimate:

$$s_W^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2 \cdots + (n_k - 1)\, s_k^2}{n_1 + n_2 + \ldots + n_k - k}$$

This is the “within” variance estimator.

Homoscedasticity assumption and within variance

From the assumptions we have that s_1 , s_2 , …, s_k all estimate σ the common value of the standard deviation in each of the groups.

So, combine to get a better estimate:

$$s_W^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2 \cdots + (n_k - 1)\, s_k^2}{n_1 + n_2 + \ldots + n_k - k}$$

This is the "within" variance estimator.

Homoscedasticity assumption and between variance

IF the null hypothesis is true (all means are equal).

Looking at the k groups, it's as if we were sampling k times from the same population.

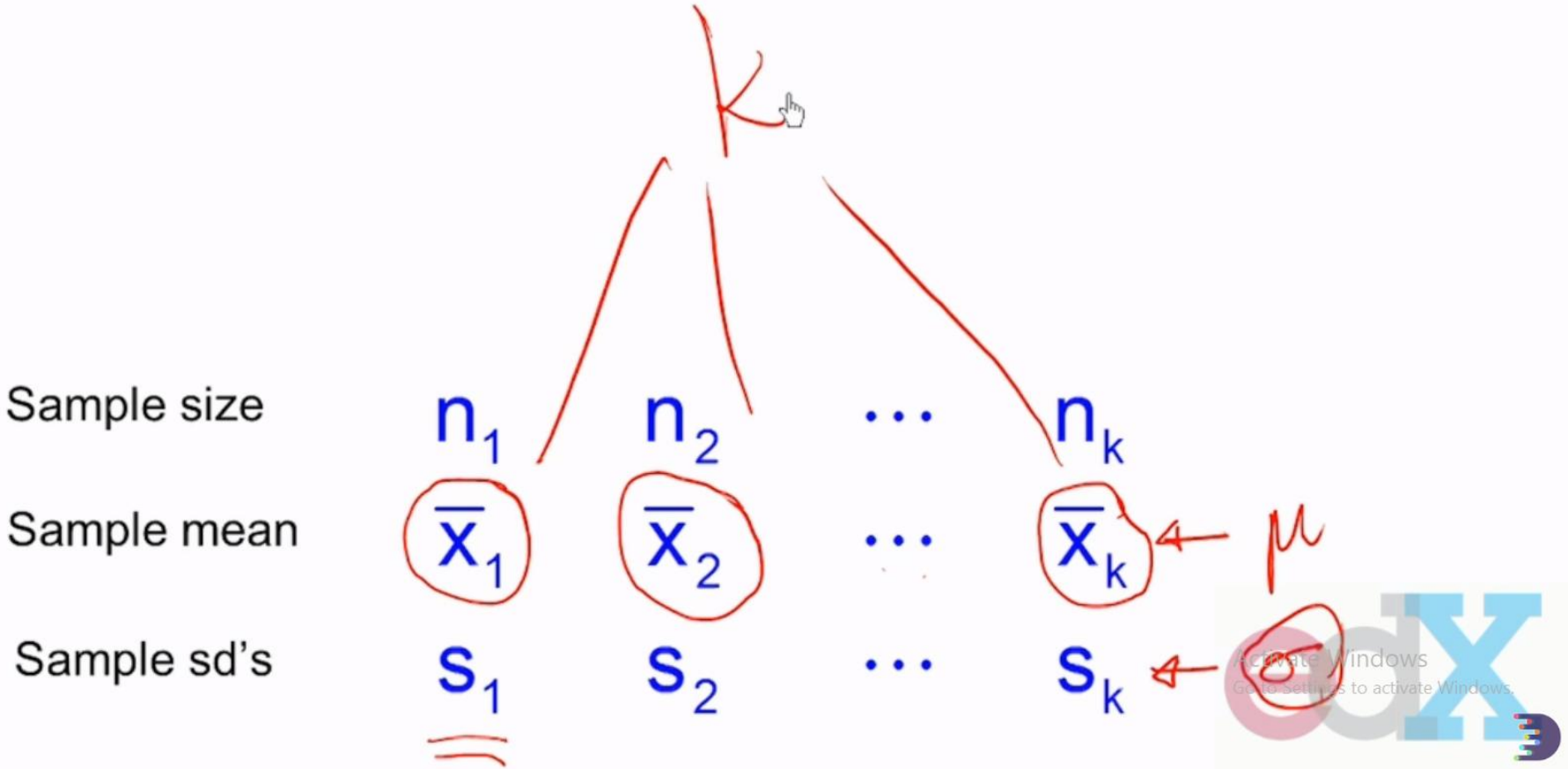
Data, k independent, random samples:

Population:	1	2	⋯	k
	$x_{1,1}$	$x_{2,1}$	⋯	$x_{k,1}$
	$x_{1,2}$	$x_{2,2}$	⋯	$x_{k,2}$
	⋮	⋮	⋮	⋮
Sample size	n_1	n_2	⋯	n_k
Sample mean	$\bar{x}_1$	$\bar{x}_2$	⋯	$\bar{x}_k$
Sample sd's	s_1	s_2	⋯	s_k

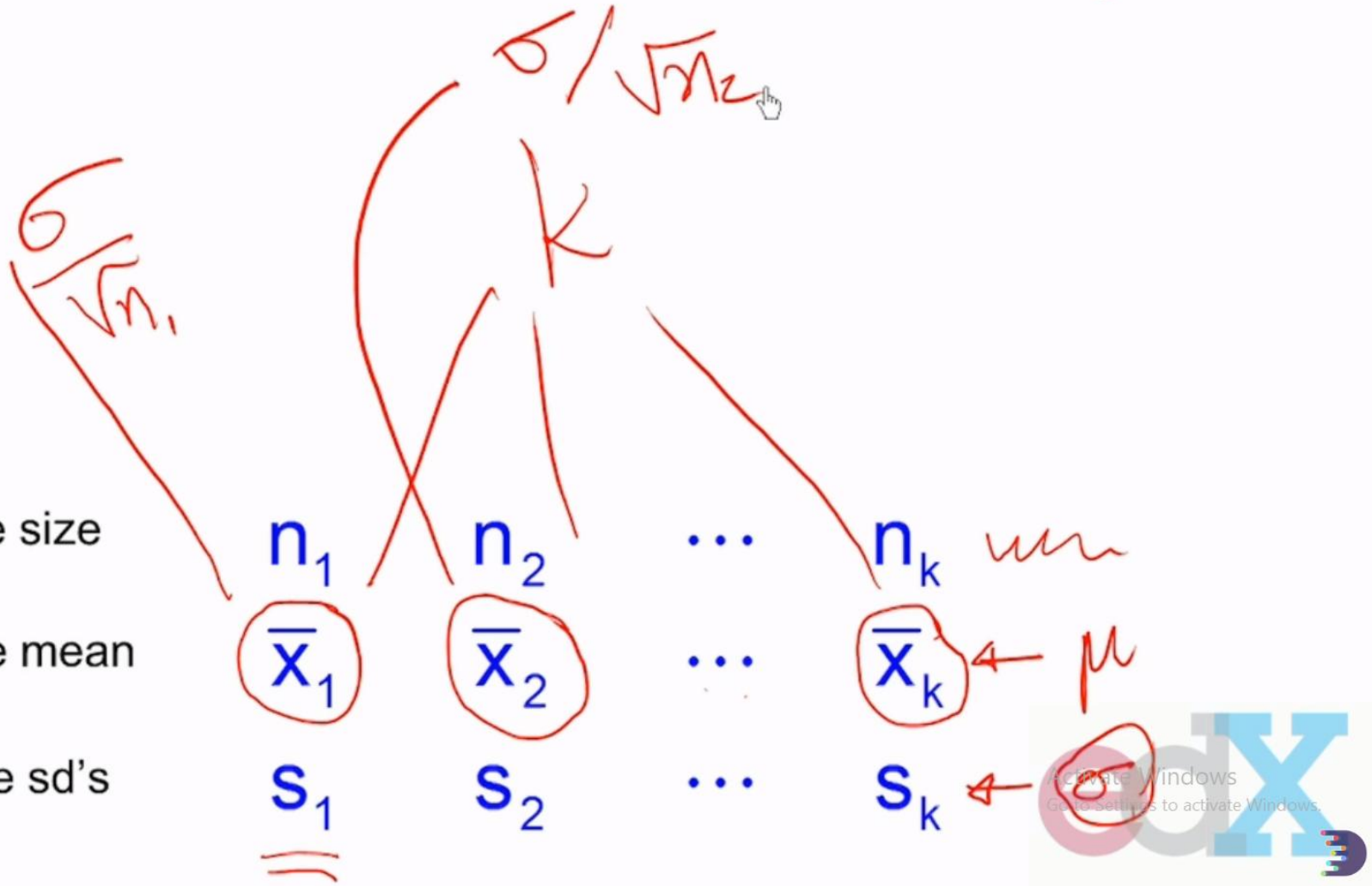
Data, k independent, random samples:

Sample size	n_1	n_2	$\cdots$	n_k
Sample mean	$\overline{x}_1$	$\overline{x}_2$	$\cdots$	$\overline{x}_k$
Sample sd's	s_1	s_2	$\cdots$	s_k

Data, k independent, random samples:



Data, k independent, random samples:



So we can get a better estimate of μ by combining all k estimators:

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + \ldots + n_k\bar{x}_k}{n_1 + n_2 + \ldots + n_k}$$

$$= \frac{\sum_{\text{all}} x}{n} \qquad \text{where } n = n_1 + n_2 \ldots + n_k$$

And another estimator of σ^2 , the "between" estimator

$$s_B^2 = \frac{n_1(\bar{x}_1 - \bar{x})^2 + \ldots + n_k(\bar{x}_2 - \bar{x})^2}{k-1}$$

So we can get a better estimate of μ by combining all k estimators:

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + \ldots + n_k\bar{x}_k}{n_1 + n_2 + \ldots + n_k}$$

$$= \frac{\sum_{\text{all}} x}{n}$$

where $n = n_1 + n_2 \ldots + n_k$

And another estimator of σ^2 , the "between" estimator

$$s_B^2 = \frac{n_1(\bar{x}_1 - \bar{x})^2 + \ldots + n_k(\bar{x}_k - \bar{x})^2}{k-1}$$

Ronald Fisher
1890--1962

If the null hypothesis is true, these two estimates of σ , namely s_B and s_W, should be about the same. So as a measure of the null hypothesis, compare them:

$$F = \frac{s_B^2}{s_W^2}$$

is Snedecor's F with (k-1) and (n-k) degrees of freedom.

Ronald Fisher
1890--1962

If the null hypothesis is true, these two estimates of σ , namely s_B and s_W, should be about the same. So as a measure of the null hypothesis, compare them:

$$F = \frac{s_B^2}{s_W^2}$$

is Snedecor's F with (k-1) and (n-k) degrees of freedom.

File Edit Data Graphics Statistics User Window Help

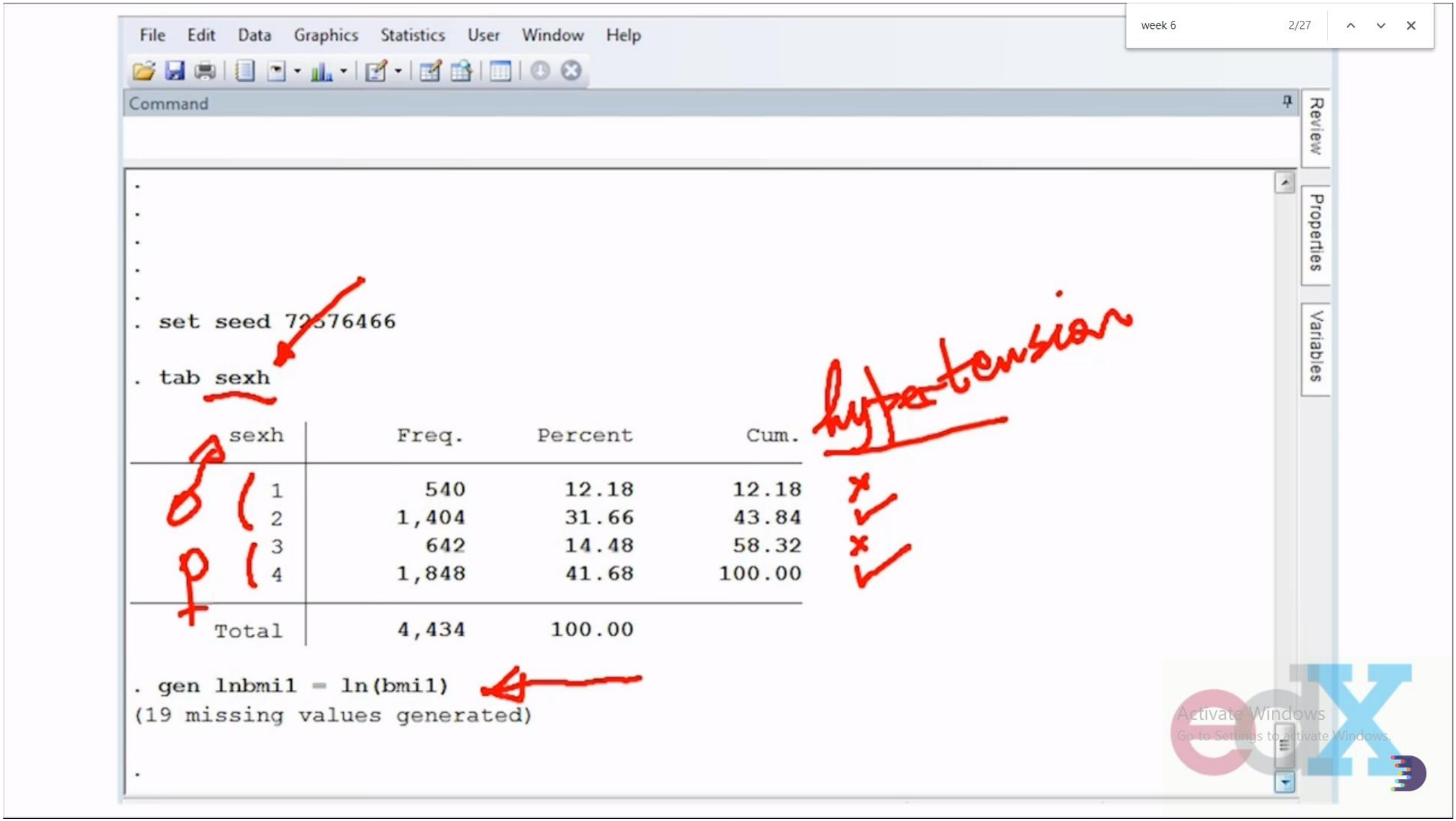
Command

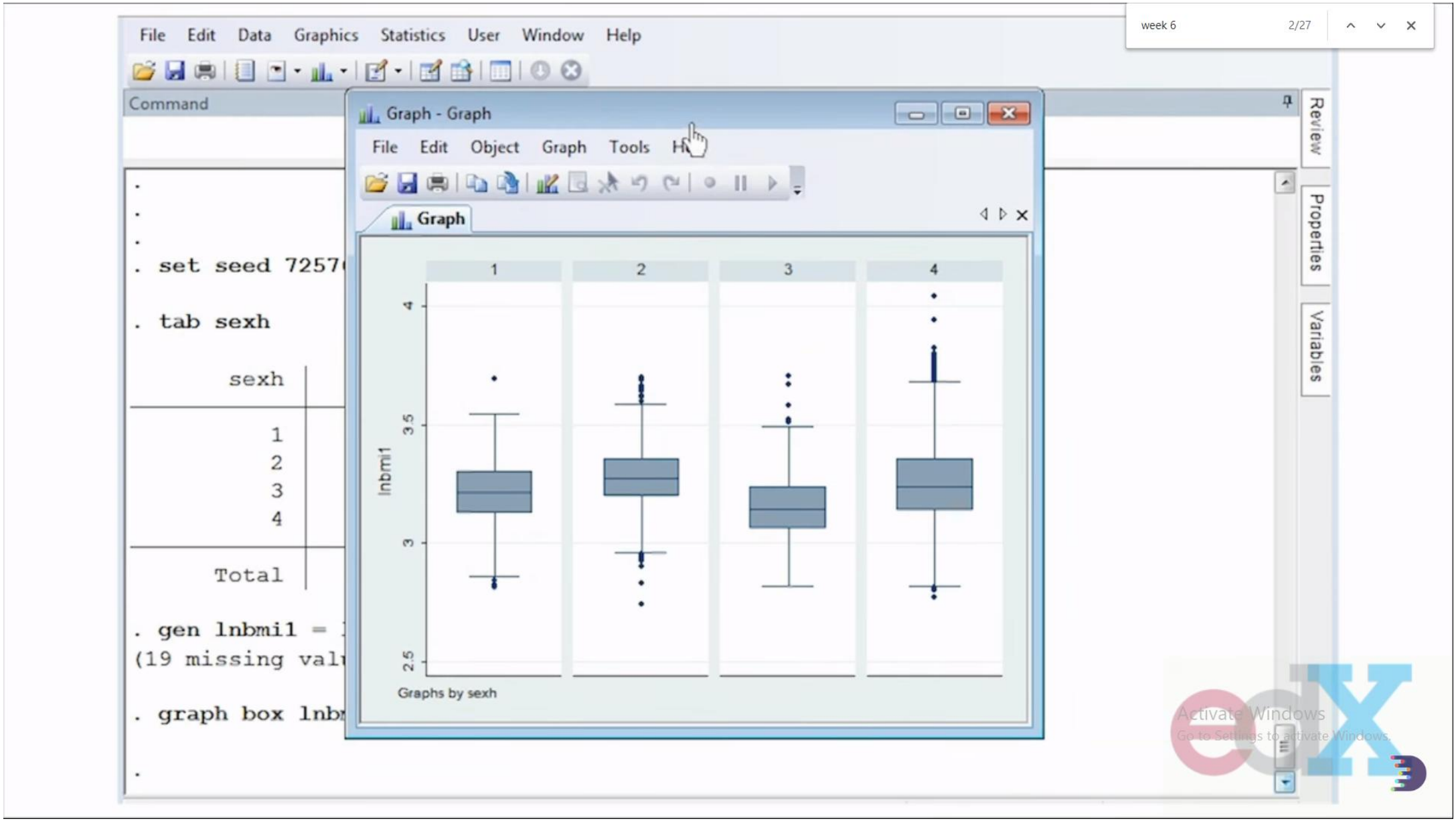
```
.
.
.
.
.
. set seed 72576466

. tab sexh
```

sexh	Freq.	Percent	Cum.
1	540	12.18	12.18
2	1,404	31.66	43.84
3	642	14.48	58.32
4	1,848	41.68	100.00
Total	4,434	100.00	

```
. gen lnbmi1 = ln(bmi1)
(19 missing values generated)

.
```

Review Properties Variables


File Edit Data Graphics Statistics User Window Help
Command
Review
Properties
Variables
. set seed 72576466
. tab sexh
sexh | Freq. Percent Cum.
1 | 540 12.18 12.18
2 | 1,404 31.66 43.84
3 | 642 14.48 58.32
4 | 1,848 41.68 100.00
Total | 4,434 100.00
. gen lnbmi1 = ln(bmi1)
(19 missing values generated)
hypertension
week 6 2/27
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
week 6 2/27
Command
. set seed 7257
. tab sexh
sexh
1
2
3
4
Total
. gen lnbmi1 =
(19 missing valu
. graph box lnbr
Graph - Graph
File Edit Object Graph Tools H
Graph
1 2 3 4
lnbmi1
4 3.5 3 2.5
Graphs by sexh
Review
Properties
Variables
Activate Windows
Go to Settings to activate Windows.

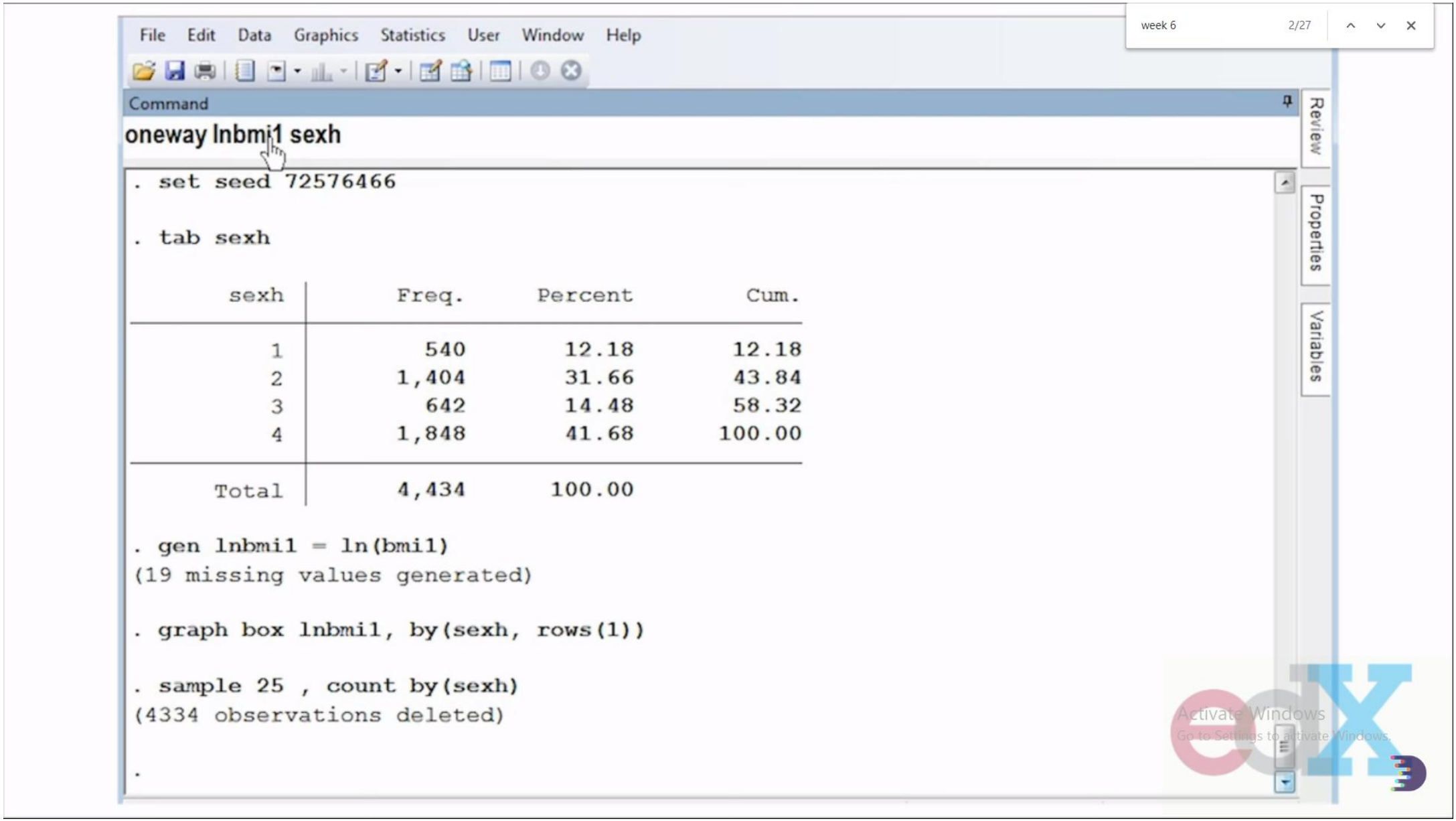
File Edit Data Graphics Statistics User Window Help

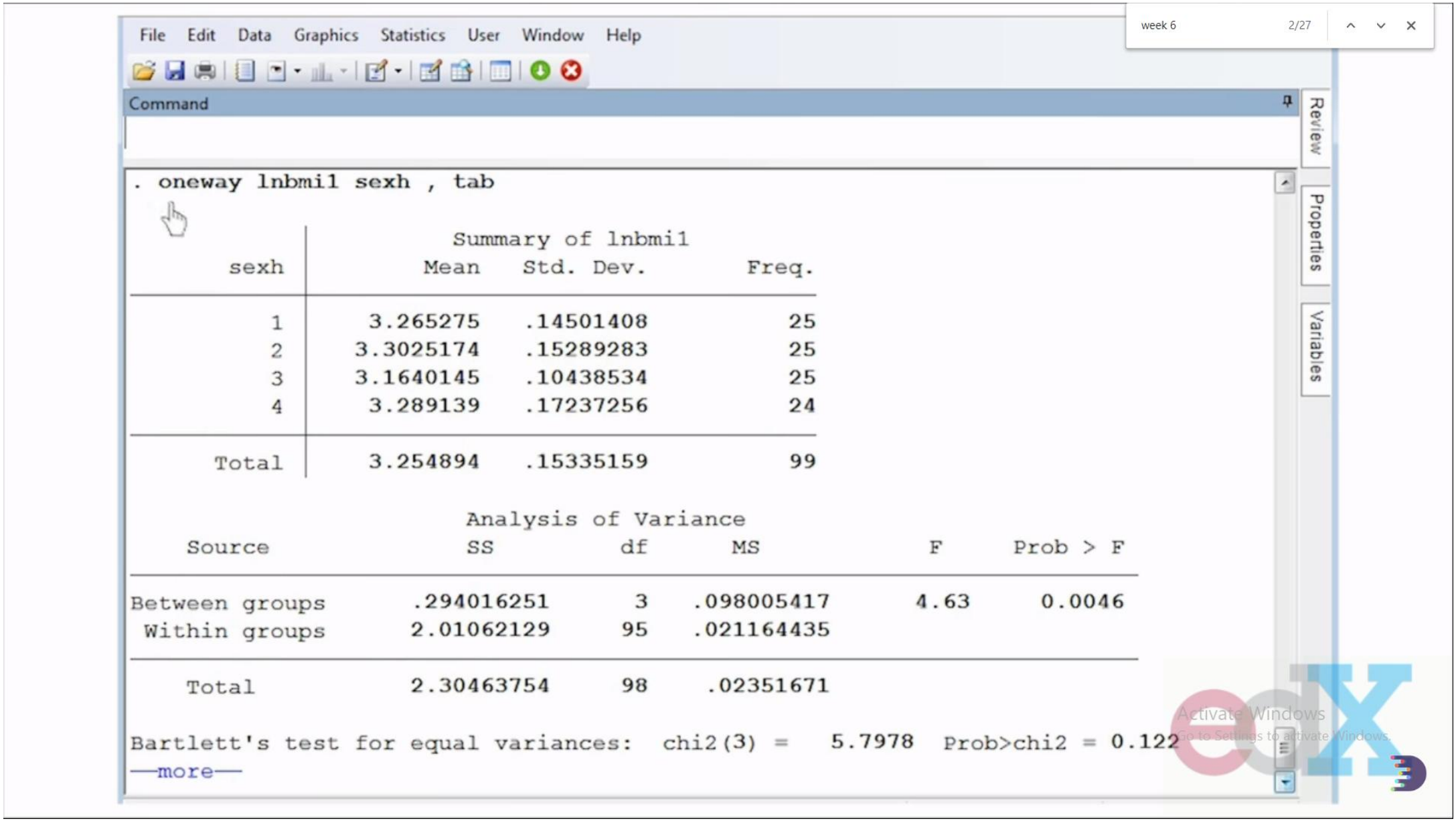
Command

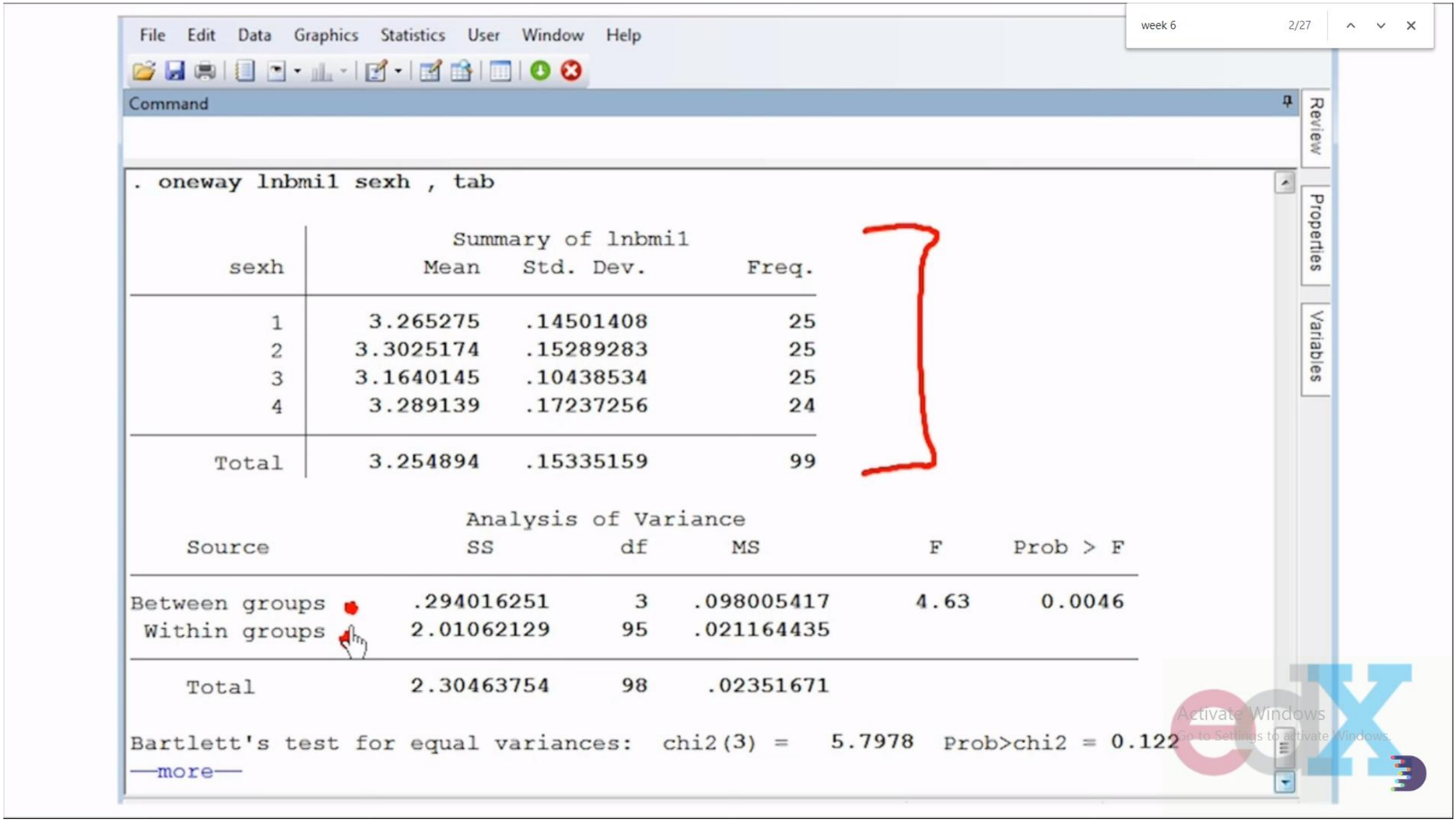
```
sample 25 , count by(sexh)
```

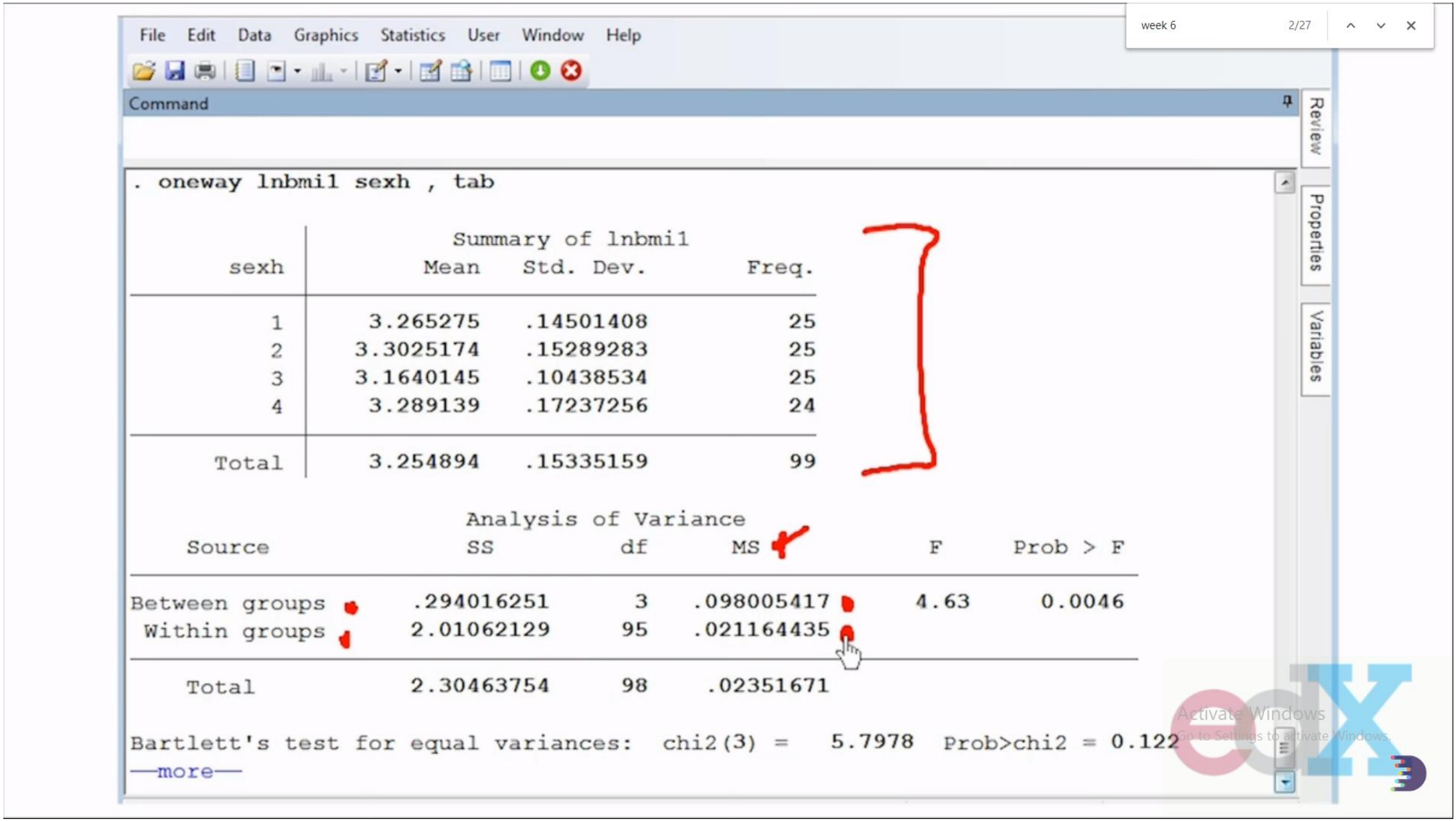
```
.
.
.
. set seed 72576466

. tab sexh

       sexh |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        540       12.18       12.18
          2 |      1,404       31.66       43.84
          3 |        642       14.48       58.32
          4 |      1,848       41.68      100.00
------------+-----------------------------------
      Total |      4,434      100.00

. gen lnbmi1 = ln(bmi1)
(19 missing values generated)

. graph box lnbmi1, by(sexh, rows(1))

.
```

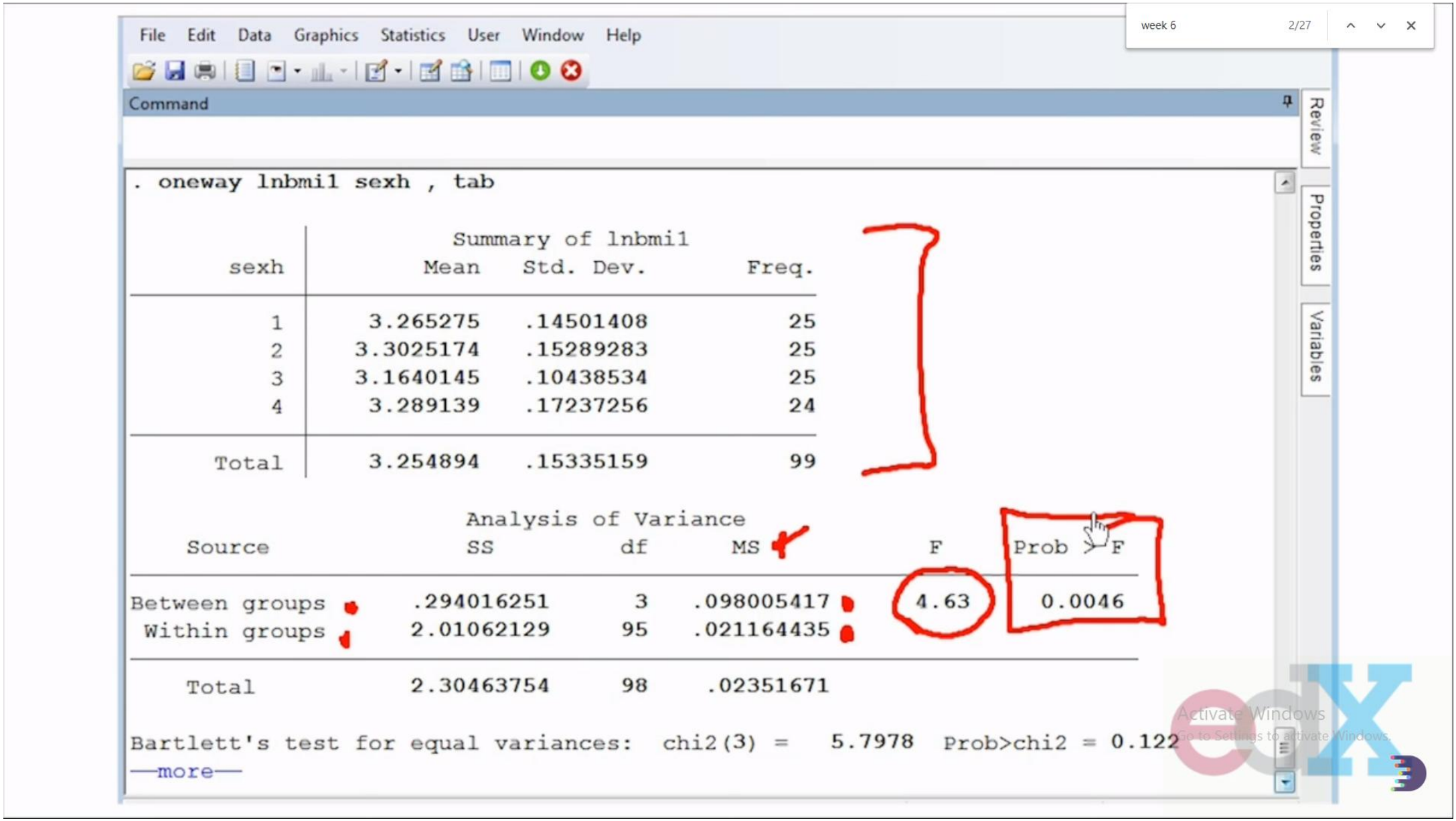
Review Properties Variables


week 6 2/27
File Edit Data Graphics Statistics User Window Help
Command
oneway lnbmi1 sexh
. set seed 72576466
. tab sexh
sexh | Freq. Percent Cum.
1 | 540 12.18 12.18
2 | 1,404 31.66 43.84
3 | 642 14.48 58.32
4 | 1,848 41.68 100.00
Total | 4,434 100.00
. gen lnbmi1 = ln(bmi1)
(19 missing values generated)
. graph box lnbmi1, by(sexh, rows(1))
. sample 25 , count by(sexh)
(4334 observations deleted)
.
Review
Properties
Variables
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
week 6 2/27
Command
Review
Properties
Variables
. oneway lnbmi1 sexh , tab
Summary of lnbmi1
sexh | Mean Std. Dev. Freq.
1 | 3.265275 .14501408 25
2 | 3.3025174 .15289283 25
3 | 3.1640145 .10438534 25
4 | 3.289139 .17237256 24
Total | 3.254894 .15335159 99
Analysis of Variance
Source SS df MS F Prob > F
Between groups .294016251 3 .098005417 4.63 0.0046
Within groups 2.01062129 95 .021164435
Total 2.30463754 98 .02351671
Bartlett's test for equal variances: chi2(3) = 5.7978 Prob>chi2 = 0.122
—more—
Activate Windows
Go to Settings to activate Windows.


week 6 2/27
File Edit Data Graphics Statistics User Window Help
Command
Review
Properties
Variables
. oneway lnbmi1 sexh , tab
Summary of lnbmi1
sexh Mean Std. Dev. Freq.
1 3.265275 .14501408 25
2 3.3025174 .15289283 25
3 3.1640145 .10438534 25
4 3.289139 .17237256 24
Total 3.254894 .15335159 99
Analysis of Variance
Source SS df MS F Prob > F
Between groups .294016251 3 .098005417 4.63 0.0046
Within groups 2.01062129 95 .021164435
Total 2.30463754 98 .02351671
Bartlett's test for equal variances: chi2(3) = 5.7978 Prob>chi2 = 0.122
--more--
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
Command
. oneway lnbmi1 sexh , tab
Summary of lnbmi1
sexh Mean Std. Dev. Freq.
1 3.265275 .14501408 25
2 3.3025174 .15289283 25
3 3.1640145 .10438534 25
4 3.289139 .17237256 24
Total 3.254894 .15335159 99
Analysis of Variance
Source SS df MS F Prob > F
Between groups .294016251 3 .098005417 4.63 0.0046
Within groups 2.01062129 95 .021164435
Total 2.30463754 98 .02351671
Bartlett's test for equal variances: chi2(3) = 5.7978 Prob>chi2 = 0.122
—more—
Review
Properties
Variables
week 6
2/27
Activate Windows
Go to Settings to activate Windows.



File Edit Data Graphics Statistics User Window Help

week 6 2/27

Command

Review Properties Variables

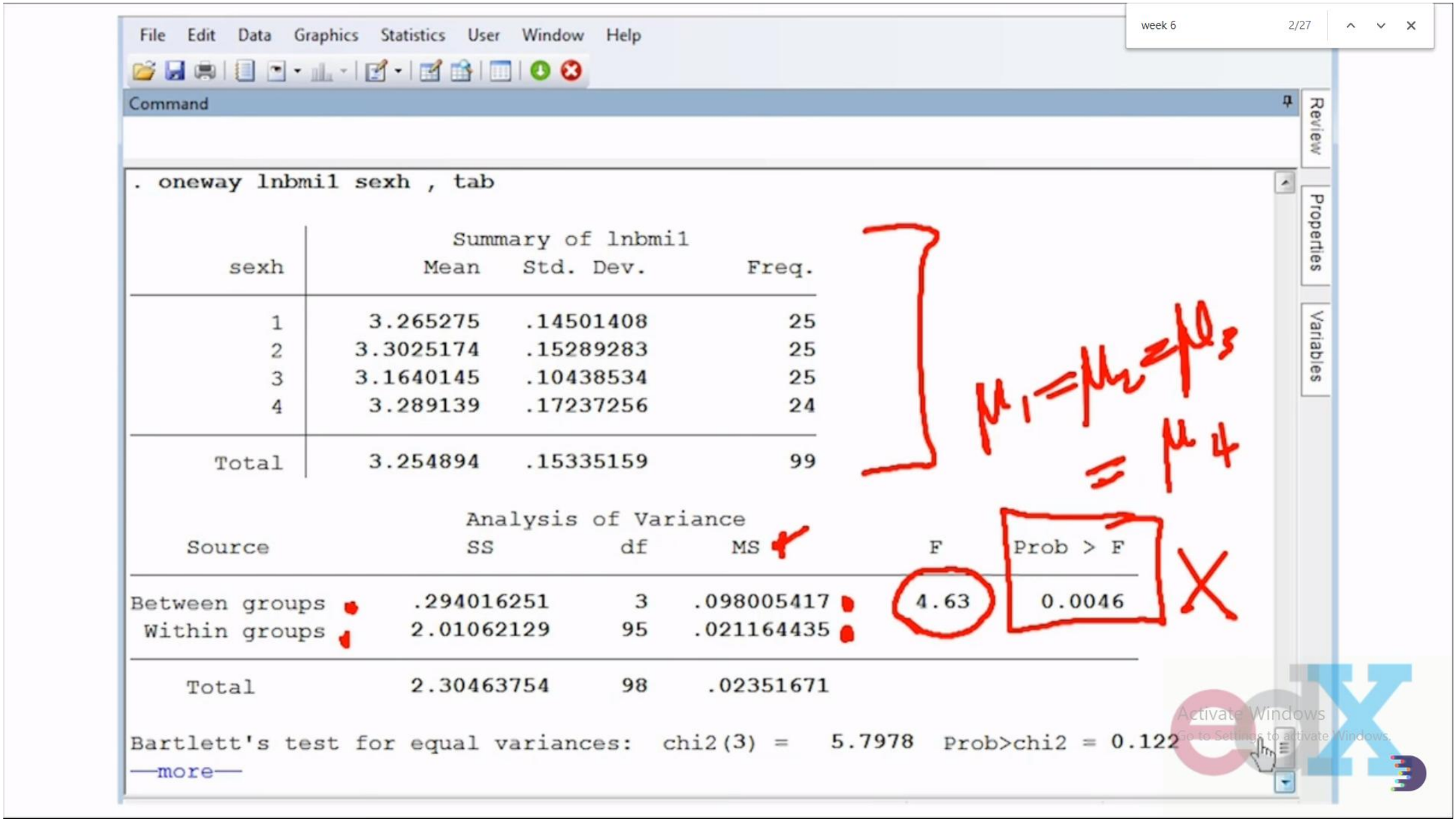
```
. oneway lnbmi1 sexh , tab
```

sexh	Summary of lnbmi1 Mean	Std. Dev.	Freq.
1	3.265275	.14501408	25
2	3.3025174	.15289283	25
3	3.1640145	.10438534	25
4	3.289139	.17237256	24
Total	3.254894	.15335159	99

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	.294016251	3	.098005417	4.63	0.0046
Within groups	2.01062129	95	.021164435		
Total	2.30463754	98	.02351671		

```
Bartlett's test for equal variances:  chi2(3) =   5.7978  Prob>chi2 = 0.122
--more--
```


week 6 2/27
File Edit Data Graphics Statistics User Window Help
Command
Review
Properties
Variables
. oneway lnbmi1 sexh , tab
Summary of lnbmi1
sexh | Mean Std. Dev. Freq.
1 | 3.265275 .14501408 25
2 | 3.3025174 .15289283 25
3 | 3.1640145 .10438534 25
4 | 3.289139 .17237256 24
Total | 3.254894 .15335159 99
Analysis of Variance
Source SS df MS F Prob > F
Between groups .294016251 3 .098005417 4.63 0.0046
Within groups 2.01062129 95 .021164435
Total 2.30463754 98 .02351671
Bartlett's test for equal variances: chi2(3) = 5.7978 Prob>chi2 = 0.122
—more—
μ1=μ2=μ3 = μ4
Activate Windows
Go to Settings to activate Windows.

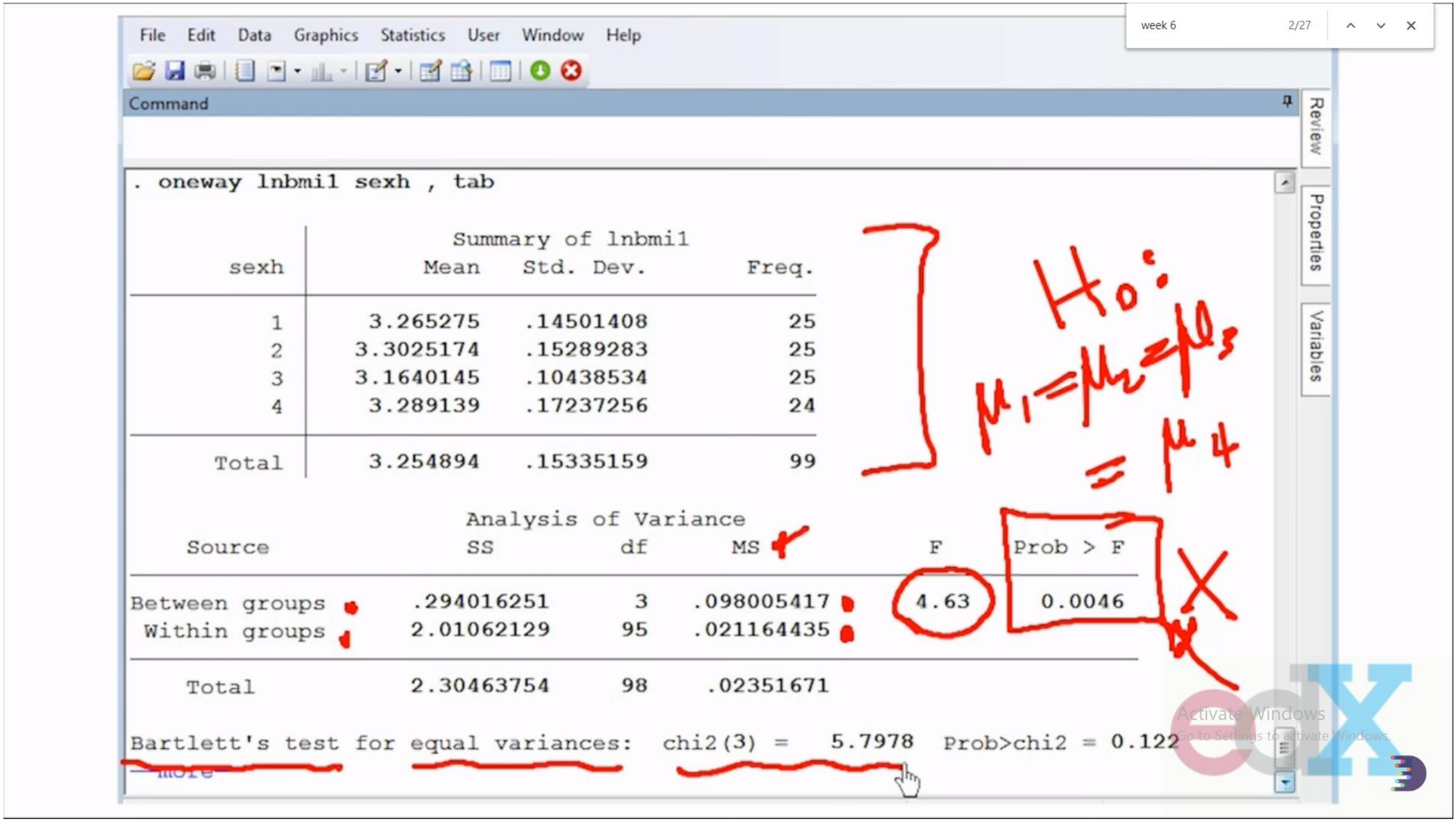
File Edit Data Graphics Statistics User Window Help

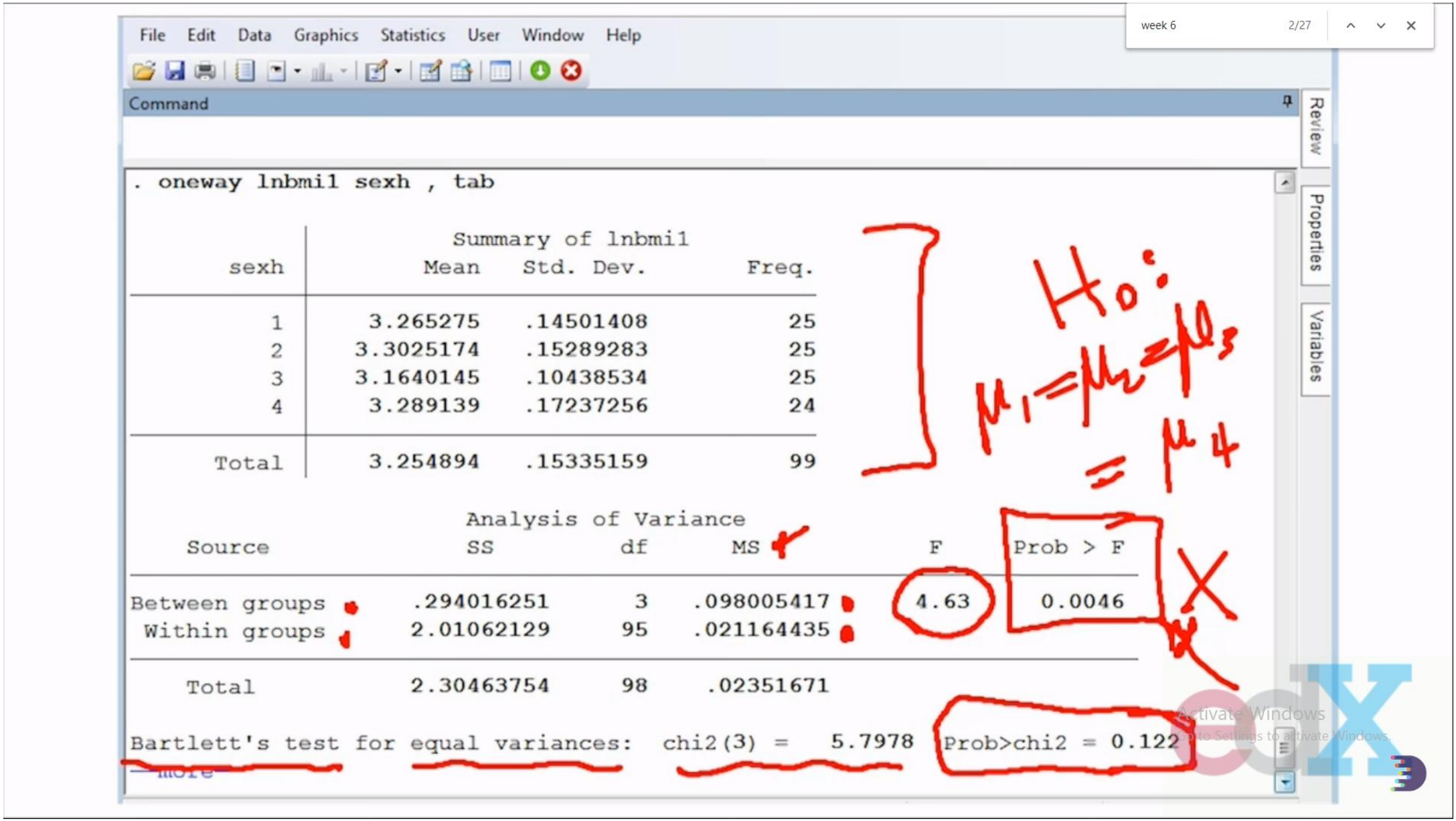
Command

```
. oneway lnbmi1 sexh , tab

                Summary of lnbmi1
      sexh |      Mean   Std. Dev.       Freq.
-----------+------------------------------------
         1 |  3.265275   .14501408          25
         2 | 3.3025174   .15289283          25
         3 | 3.1640145   .10438534          25
         4 |  3.289139   .17237256          24
-----------+------------------------------------
     Total |  3.254894   .15335159          99

                        Analysis of Variance
    Source              SS         df      MS            F     Prob > F
------------------------------------------------------------------------
Between groups      .294016251      3    .098005417      4.63     0.0046
 Within groups      2.01062129     95    .021164435
------------------------------------------------------------------------
    Total           2.30463754     98     .02351671

Bartlett's test for equal variances:  chi2(3) =   5.7978  Prob>chi2 = 0.122
--more--
```

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

X


week 6 2/27
File Edit Data Graphics Statistics User Window Help
Command
Review
Properties
Variables
. oneway lnbmi1 sexh , tab
Summary of lnbmi1
sexh | Mean Std. Dev. Freq.
1 | 3.265275 .14501408 25
2 | 3.3025174 .15289283 25
3 | 3.1640145 .10438534 25
4 | 3.289139 .17237256 24
Total | 3.254894 .15335159 99
Analysis of Variance
Source SS df MS F Prob > F
Between groups .294016251 3 .098005417 4.63 0.0046
Within groups 2.01062129 95 .021164435
Total 2.30463754 98 .02351671
Bartlett's test for equal variances: chi2(3) = 5.7978 Prob>chi2 = 0.122
H0: μ1=μ2=μ3=μ4
Activate Windows
Go to Settings to activate Windows.


File Edit Data Graphics Statistics User Window Help
Command
. oneway lnbmi1 sexh , tab
Summary of lnbmi1
sexh | Mean Std. Dev. Freq.
1 | 3.265275 .14501408 25
2 | 3.3025174 .15289283 25
3 | 3.1640145 .10438534 25
4 | 3.289139 .17237256 24
Total | 3.254894 .15335159 99
Analysis of Variance
Source SS df MS F Prob > F
Between groups .294016251 3 .098005417 4.63 0.0046
Within groups 2.01062129 95 .021164435
Total 2.30463754 98 .02351671
Bartlett's test for equal variances: chi2(3) = 5.7978 Prob>chi2 = 0.122
Review
Properties
Variables
week 6
2/27

```
. oneway lnbmi1 sexh, tab bon
```

sexh	Summary of lnbmi1 Mean	Std. Dev.	Freq.
1	3.265275	.14501408	25
2	3.3025174	.15289283	25
3	3.1640145	.10438534	25
4	3.289139	.17237256	24
Total	3.254894	.15335159	99

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	.294016251	3	.098005417	4.63	0.0046
Within groups	2.01062129	95	.021164435		
Total	2.30463754	98	.02351671		

Bartlett's test for equal variances: chi2(3) = 5.7938 Prob>chi2 = 0.122

$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

```
. oneway lnbmi1 sexh, tab bon
```

sexh	Summary of lnbmi1 Mean	Std. Dev.	Freq.
1	3.265275	.14501408	25
2	3.3025174	.15289283	25
3	3.1640145	.10438534	25
4	3.289139	.17237256	24
Total	3.254894	.15335159	99

μ_4

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	.294016251	3	.098005417	4.63	0.0046
Within groups	2.01062129	95	.021164435		
Total	2.30463754	98	.02351671		

Bartlett's test for equal variances: chi2(3) = 5.7978 Prob>chi2 = 0.122

```
. oneway lnbmi1 sexh, tab bon
```

Summary of lnbmi1

sexh	Mean	Std. Dev.	Freq.
1	3.265275	.14501408	25
2	3.3025174	.15289283	25
3	3.1640145	.10438534	25
4	3.289139	.17237256	24
Total	3.254894	.15335159	99

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	.294016251	3	.098005417	4.63	0.0046
Within groups	2.01062129	95	.021164435		
Total	2.30463754	98	.02351671		

Bartlett's test for equal variances: chi2(3) = 5.7978 Prob>chi2 = 0.122

Carlo Bonferroni
1892--1960

Bonferroni Correction:

If we wish to perform all possible pairs of comparisons, then there are $\binom{k}{2}$ such comparisons. So to have an overall level of α, one needs to perform each individual test at level

$$\alpha^* = \frac{\alpha}{\binom{k}{2}} \quad \text{or} \quad \alpha^* \frac{k!}{2!\,(k-2)!} = \alpha$$

Carlo Bonferroni
1892--1960

Bonferroni Correction:

1 vs 2
1 vs 3
1 vs 4
2 vs 3
2 vs 4
3 vs 4

If we wish to perform all possible pairs of comparisons, then there are $\binom{k}{2}$ such comparisons. So to have an overall level of α, one needs to perform each individual test at level

$$\alpha^* = \frac{\alpha}{\binom{k}{2}}$$

or

$$\alpha^* \frac{k!}{2!\,(k-2)!} = \alpha$$

Comparison of lnbmi1 by sexh
(Bonferroni)

Row Mean-Col Mean	1	2	3
2	.037242		
	1.000		
3	-.10126	-.138503	
	0.094	0.007	
4	.023864	-.013378	.125125
	1.000	1.000	0.020

Comparison of lnbmi1 by sexh
(Bonferroni)

Row Mean-Col Mean	1	2	3
2	.037242		
	1.000		
3	-.10126	-.138503	
	0.094	0.007	
4	.023864	-.013378	.125125
	1.000	1.000	0.020

Comparison of lnbmi1 by sexh
(Bonferroni)

Row Mean-Col Mean	1	2	3
2	.037242		
	1.000		
3	-.10126	-.138503	
	0.094	0.007	
4	.023864	-.013378	.125125
	1.000	1.000	0.020

Comparison of lnbmi1 by sexh
(Bonferroni)

Row Mean- Col Mean	1	2	3
2	.037242 1.000		
3	-.10126 0.094	-.138503 0.007	
4	.023864 1.000	-.013378 1.000	.125125 0.020

Comparison of lnbmi1 by sexh
(Bonferroni)

Row Mean- Col Mean	1	2	3
2	.037242 1.000		
3	-.10126 0.094	-.138503 0.007	
4	.023864 1.000	-.013378 1.000	.125125 0.020