www.epilepsy.va.gov/Statistics

Statistics in Evidence Based Medicine

Lecture 7: Research Questions for Two Groups

Rizwana Rehman, PhD
Regional Statistician
Southeast Epilepsy Center of Excellence
Durham VA Medical Center, Durham NC
Rizwana.Rehman@va.gov
(919)286-0411 ext: 5024

Epilepsy Center of Excellence
SOUTHEAST

Summary of One Group Methods

Nature of problem	Test
Interval & normal	One –sample t test
Ordinal or interval, non normal	Sign test
Dependent groups; interval & normal	Paired t test
Dependent groups; ordinal or interval	Wilcoxon signed rank test, Sign test
Single proportion	z approximation, exact binomial
Dependent groups; proportion	McNemar

Overview

- Difference of means in two normally distributed groups
- Nonparametric test for two independent means
- Statistical difference between computed proportions
- Fisher's Exact test

Normal Data with Same Standard Deviations

- Are two studied groups from the same population?
- t test is used for quantitative data under the following assumptions.
  - Both samples are normally distributed.
  - Both come from distributions with same value of standard deviation.
  - Observations are independent of each other.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} \quad \text{with } n_1 + n_2 - 2 \text{ d.f, } s_p \text{ is pooled SD.}$$

Why Assumptions Matter?

- When assumptions are not met, α increases.
  - If t test is used on highly skewed data with less than 30 observations, then confidence intervals are erroneously narrow and *p* value is smaller than it should be.
  - Same happens when observations are not independent due to less effective sample size.
  - Similarly pooled SD is not accurate if standard deviations are markedly different increasing type I error rate.

How To Check for Assumptions?

- Normality
  - For large data sets normality does not matter
  - For small data
    - Eyeball data sets
    - A normality test can be run
- Independence
  - One's decision
- Equal standard deviations
  - For equal sample sizes equal SD is not needed.
  - If ratio of larger to smaller SD is less than 2, OK
  - Run a test such as F test for equal variance

What if Standard Deviations are Different?

- If standard deviations are noticeably different, then use another approximate test by Sattherwaite also known as Welch test.
- Openstat gives results for both equal and unequal standard deviations.

Openstat for Two Independent Means

Assume that FEV_1 levels for following data are normally distributed. Are FEV_1 levels different in two groups?

$$\text{Asthmatics' } FEV_1: \quad n_1 = 5, \bar{x}_1 = 1.86, s_1 = 0.378$$

$$\text{Controls' } FEV_1: \quad n_2 = 6, \bar{x}_2 = 2.51, s_2 = 0.210$$

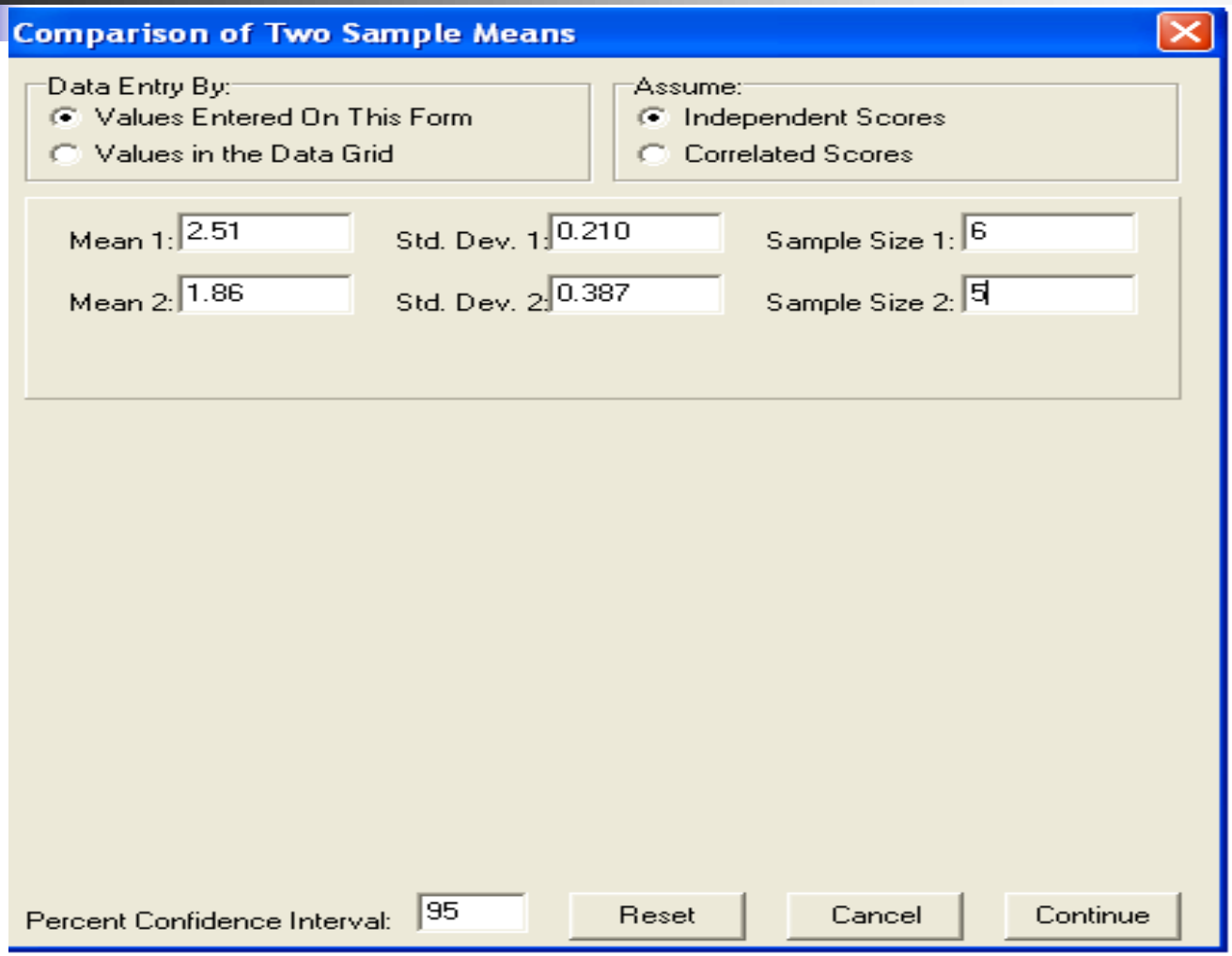
Medical Statistics A Common Sense Approach A5

Two Independent Groups t Test in Openstat

H_0: μ1 = μ2

H_A: μ1 ≠ μ2

Entering Data in Openstat



Results from Openstat

Results Window

Return

COMPARISON OF TWO MEANS

Variable	Mean	Variance	Std.Dev.	S.E.Mean	N
Group 1	2.51	0.04	0.21	0.09	6
Group 2	1.86	0.15	0.39	0.17	5

Assuming equal variances, t = 3.557 with probability = 0.0061 and 9 degrees of freedom
Difference = 0.65 and Standard Error of difference = 0.18
Confidence interval = (0.24, 1.06)
Assuming unequal variances, t = 3.365 with probability = 0.0154 and 5.92 degrees of freedom
Difference = 0.65 and Standard Error of difference = 0.19
Confidence interval = (0.18, 1.12)
F test for equal variances = 3.396, Probability = 0.1062

NOTE: t-tests are two-tailed tests.

Reject H_0, means in two groups are statistically significantly different

Mean Difference for Non Normal Data

- Use Mann-Whitney using the following assumptions.
  - Data are ordinal or continuous.
  - Two groups are independent.
  - Two groups are symmetrically distributed.
- Can be used for median or mean.
- As powerful as t test.

Mann- Whitney U Test for $n_1, n_2 \geq 10$

- Combine both data sets and rank them in increasing order. Break for ties.
- Separate the ranks in each group and add ranks for each group.
- Call the smaller rank T and smaller group size n_1
- Use *z* approximation

$$z = \frac{|\, T - n_1(n_1 + n_2 + 1)/2 \,|}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}}$$

- For smaller samples use tables

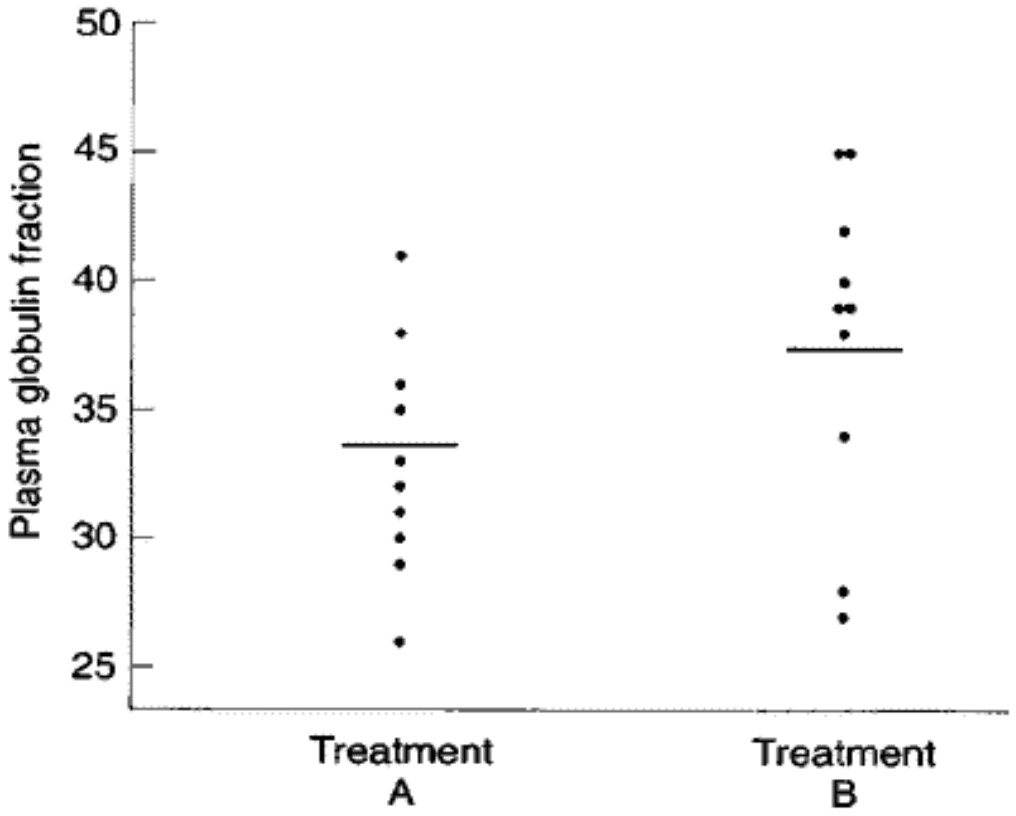
Example of Non Normal data

H_0: New treatment B does not change the plasma globulin.

H_A: New treatment B changes the plasma globulin.

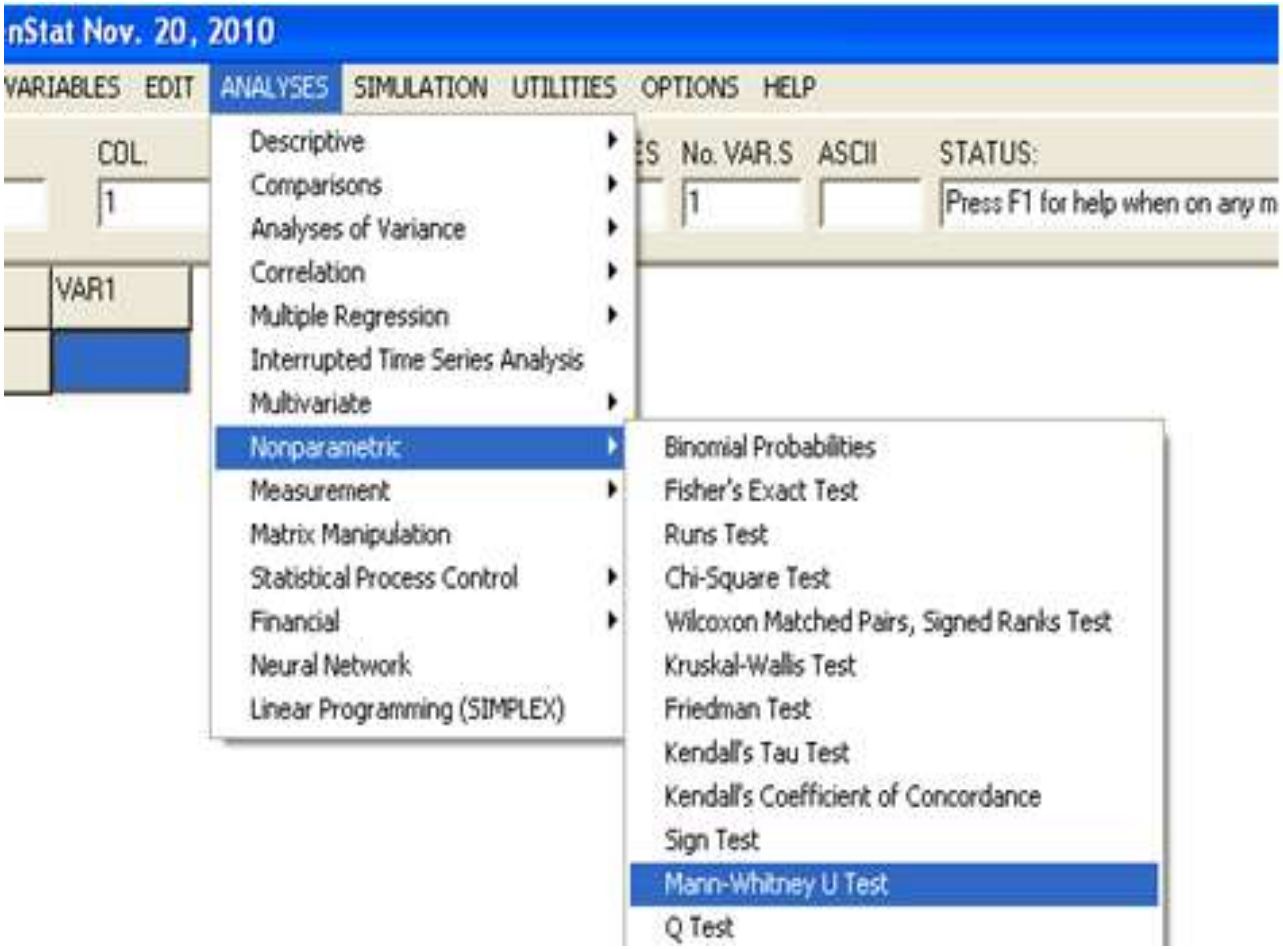
Plasma globulin fraction after randomization to treatments A or B										
Treatment A	38	26	29	41	36	31	32	30	35	33
Treatment B	45	28	27	38	40	42	39	39	40	45

Symmetry

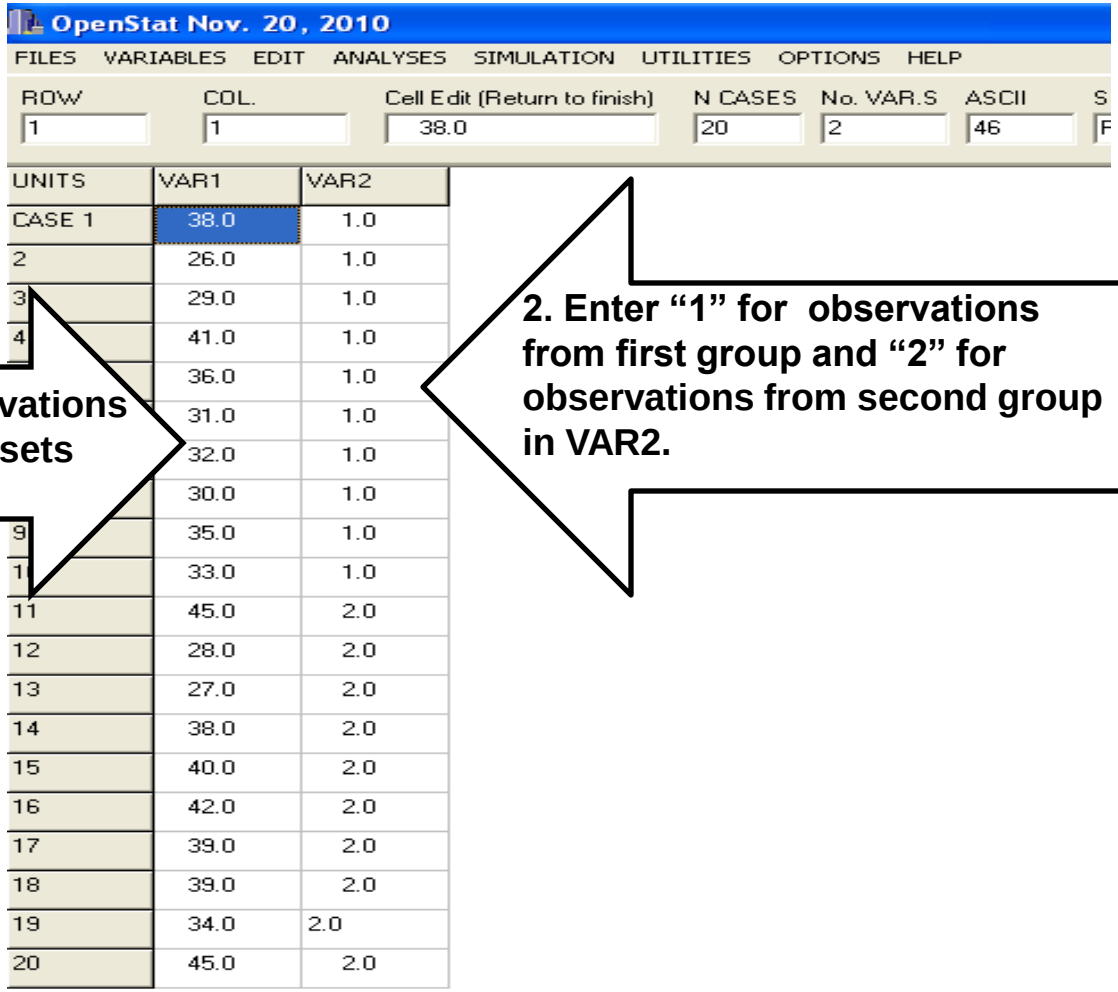


Plasma globulin fraction after treatments A or B with mean values.

Mann –Whitney in Openstat

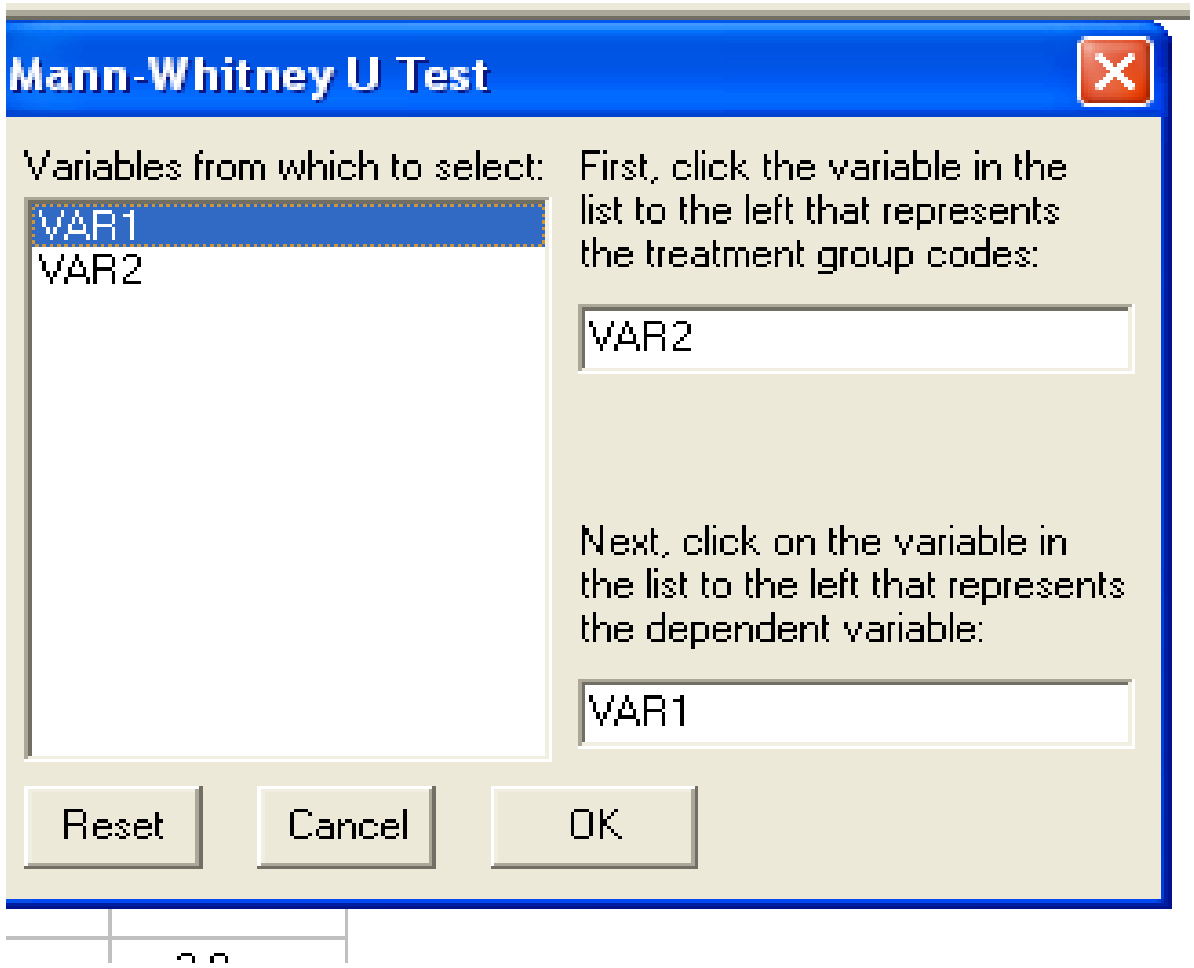


Entering Data for Analysis

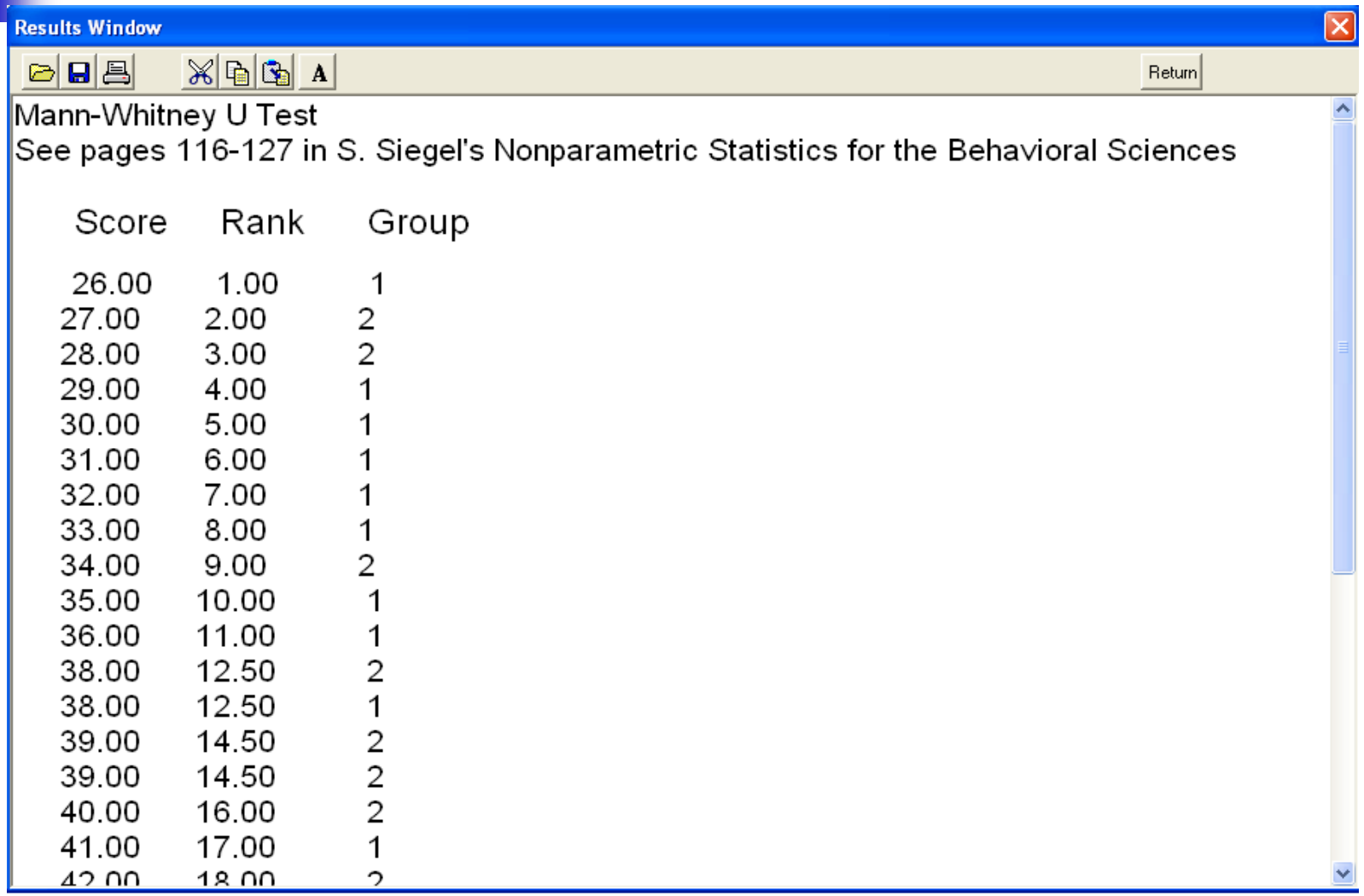


UNITS	VAR1	VAR2
CASE 1	38.0	1.0
2	26.0	1.0
3	29.0	1.0
4	41.0	1.0
	36.0	1.0
	31.0	1.0
	32.0	1.0
	30.0	1.0
9	35.0	1.0
1	33.0	1.0
11	45.0	2.0
12	28.0	2.0
13	27.0	2.0
14	38.0	2.0
15	40.0	2.0
16	42.0	2.0
17	39.0	2.0
18	39.0	2.0
19	34.0	2.0
20	45.0	2.0

Getting Results



Results


Results Window
Return
Mann-Whitney U Test
See pages 116-127 in S. Siegel's Nonparametric Statistics for the Behavioral Sciences
Score Rank Group
26.00 1.00 1
27.00 2.00 2
28.00 3.00 2
29.00 4.00 1
30.00 5.00 1
31.00 6.00 1
32.00 7.00 1
33.00 8.00 1
34.00 9.00 2
35.00 10.00 1
36.00 11.00 1
38.00 12.50 2
38.00 12.50 1
39.00 14.50 2
39.00 14.50 2
40.00 16.00 2
41.00 17.00 1
42.00 18.00 2

Results (cont'd)

Results Window

Return

35.00	10.00	1
36.00	11.00	1
38.00	12.50	2
38.00	12.50	1
39.00	14.50	2
39.00	14.50	2
40.00	16.00	2
41.00	17.00	1
42.00	18.00	2
45.00	19.50	2
45.00	19.50	2

Sum of Ranks in each Group

Group	Sum	No. in Group
1	81.50	10
2	128.50	10

No. of tied rank groups = 3
Statistic U = 26.5000
z Statistic (corrected for ties) = 1.7764, Prob. > z = 0.0378
z test is approximate. Use tables of exact probabilities in Siegel.
(Table J or K, pages 271-277)

For two tailed test

$p \approx 0.076$, can't reject H_0

Questions About Independent Proportions

- Test the hypothesis of equal proportions using the approximate z test.
- Test the hypothesis of expected frequencies using a chi-square test.

z Test for Independent Proportions

- We use the standard normal or z distribution as an approximation to binomial distribution for difference of proportions
- $$z = \frac{p_1 - p_2}{\sqrt{p(1-p)(1/n_1 + 1/n_2)}}$$

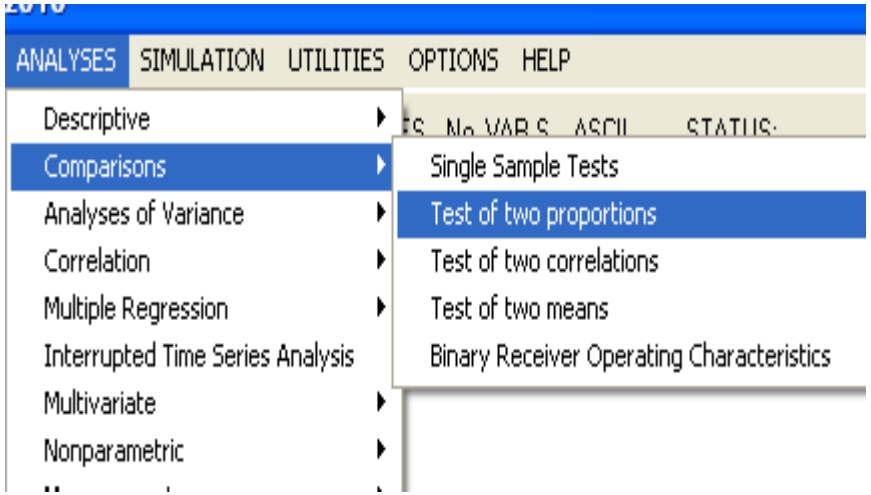
p is pooled or average proportion.

Example of Independent Proportions

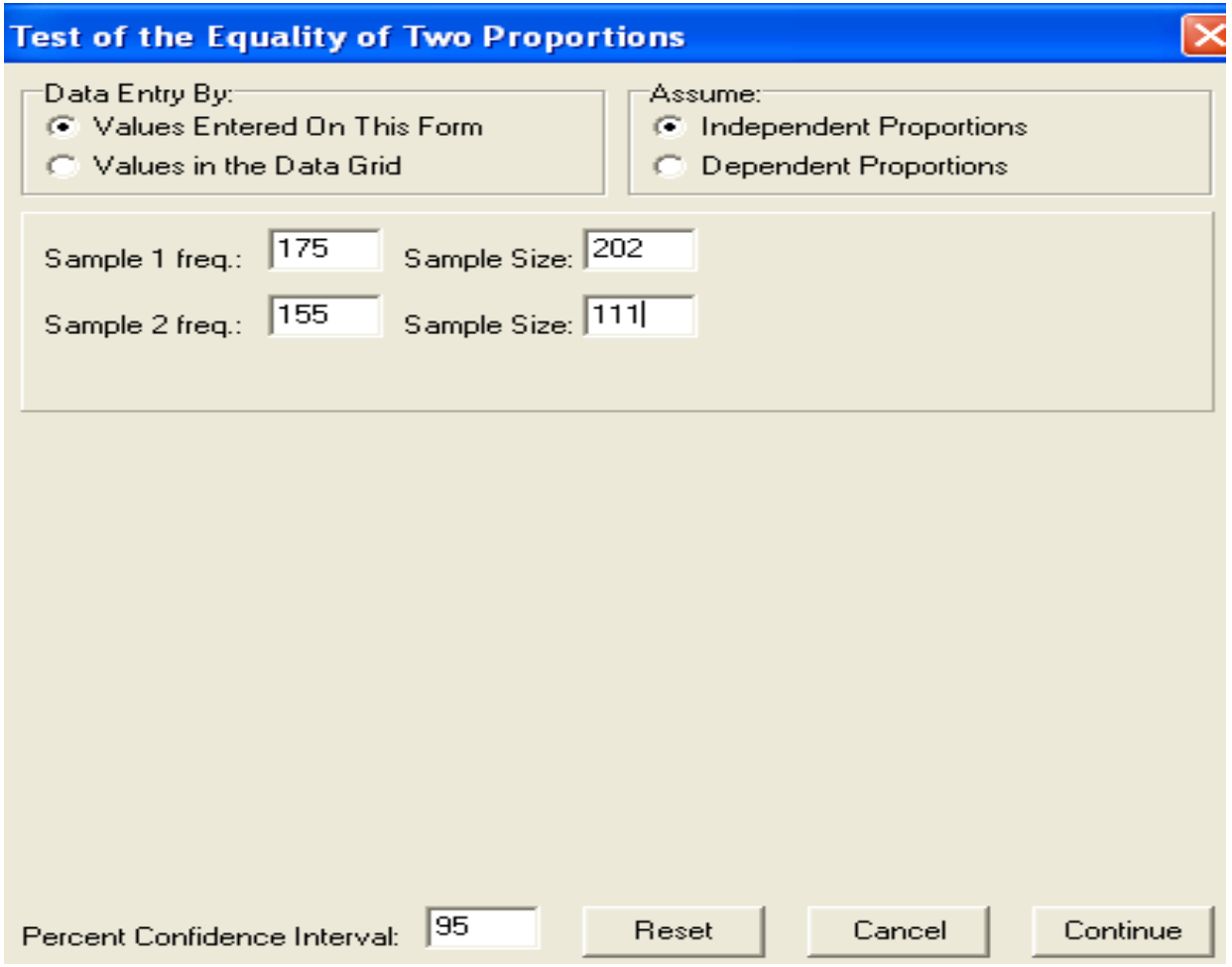
		Screen		
		Yes	No	Total
Previuos domestic violence (DV) training	Yes	175	27	202
	No	155	111	266

Basic & Clinical Biostatistics: Chapter 3

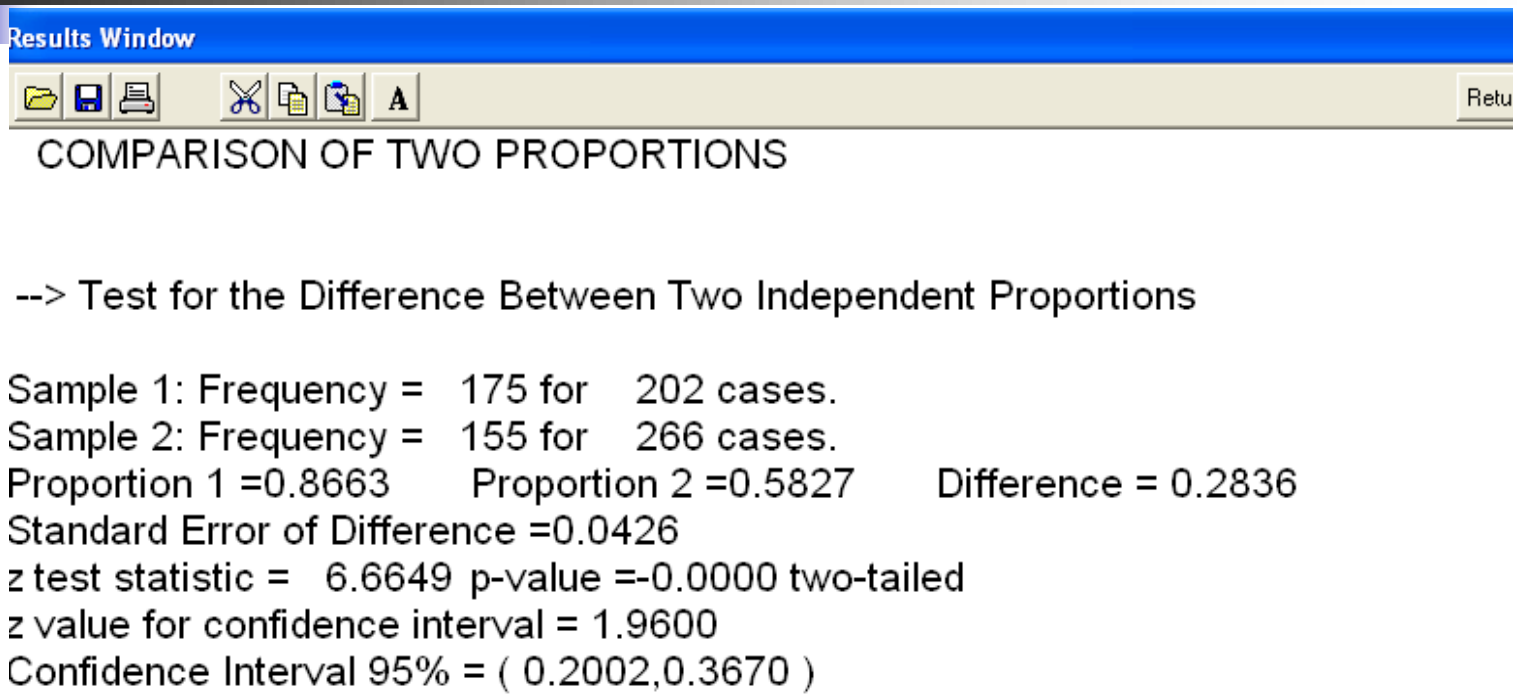
Openstat for z Approximation



Entering Data



Getting Results



Reject H_0, there is a statistically significant difference in proportions of two groups of physicians.

Chi-Square for DV Example

We can answer two questions with one chi-square test using a two by two table.

1) Is there a difference in proportions of physicians who screen and who don't screen for DV?

2) Is there a relationship (dependency) between a physician's prior DV training & whether the physicians screen for DV?

Logic behind Chi-Square

- The frequencies in two by two table are considered fixed known an marginal frequencies.
- Assuming rows and columns are independent, we calculate the expected frequencies that can occur by chance alone.
- The Chi-Square test compares the observed frequencies with expected frequencies. For no relationship, expected and observed should be close.
- For a two by two table a chi-square statistic has one degree of freedom.

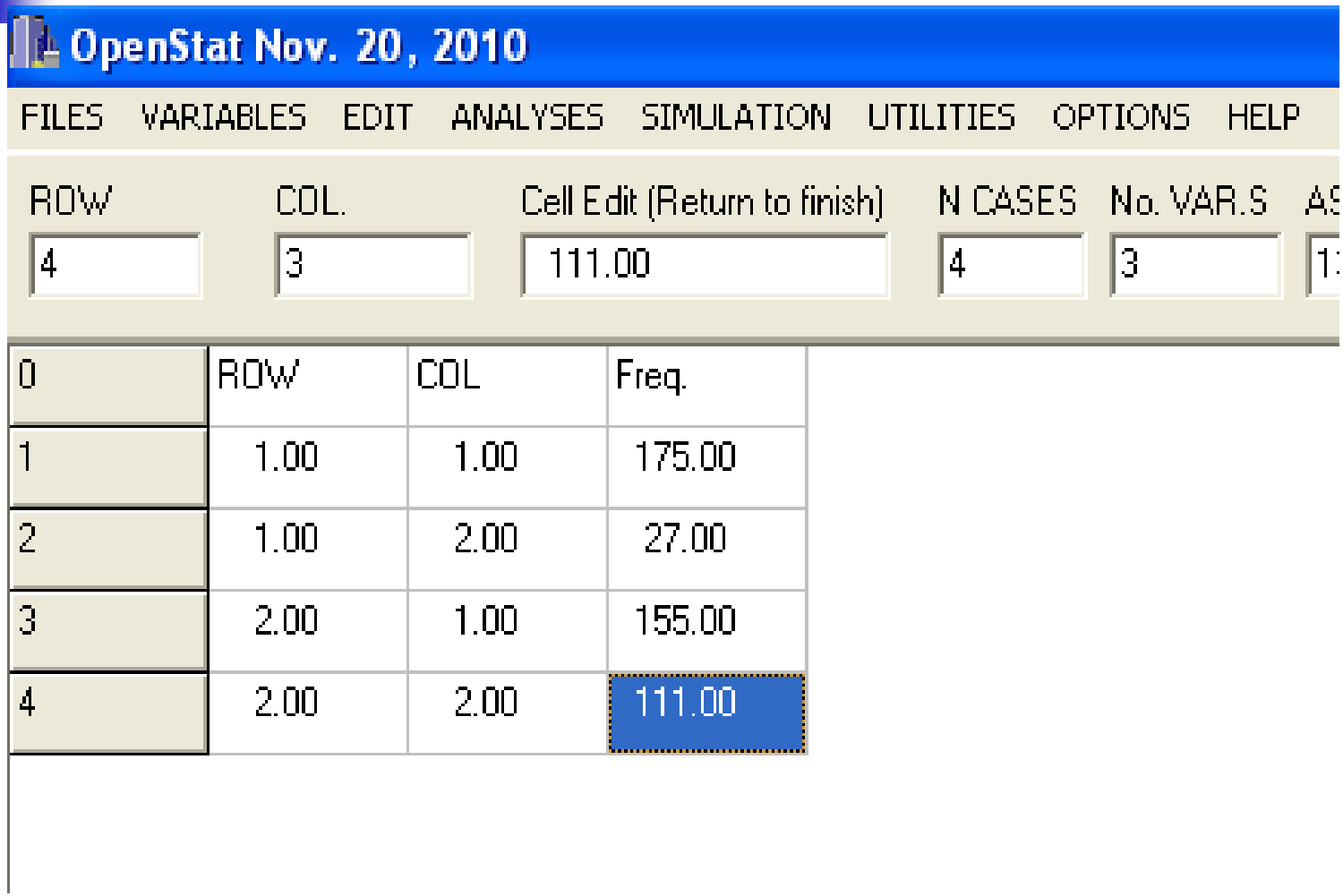
Assumptions of χ^2

- Sample size for two by two table is sufficiently large, total greater than 20.
- Expected value (not the original frequency) in any cell of two by two table is greater than 5.
- The data are independent.

What Happens if Assumptions are Violated?

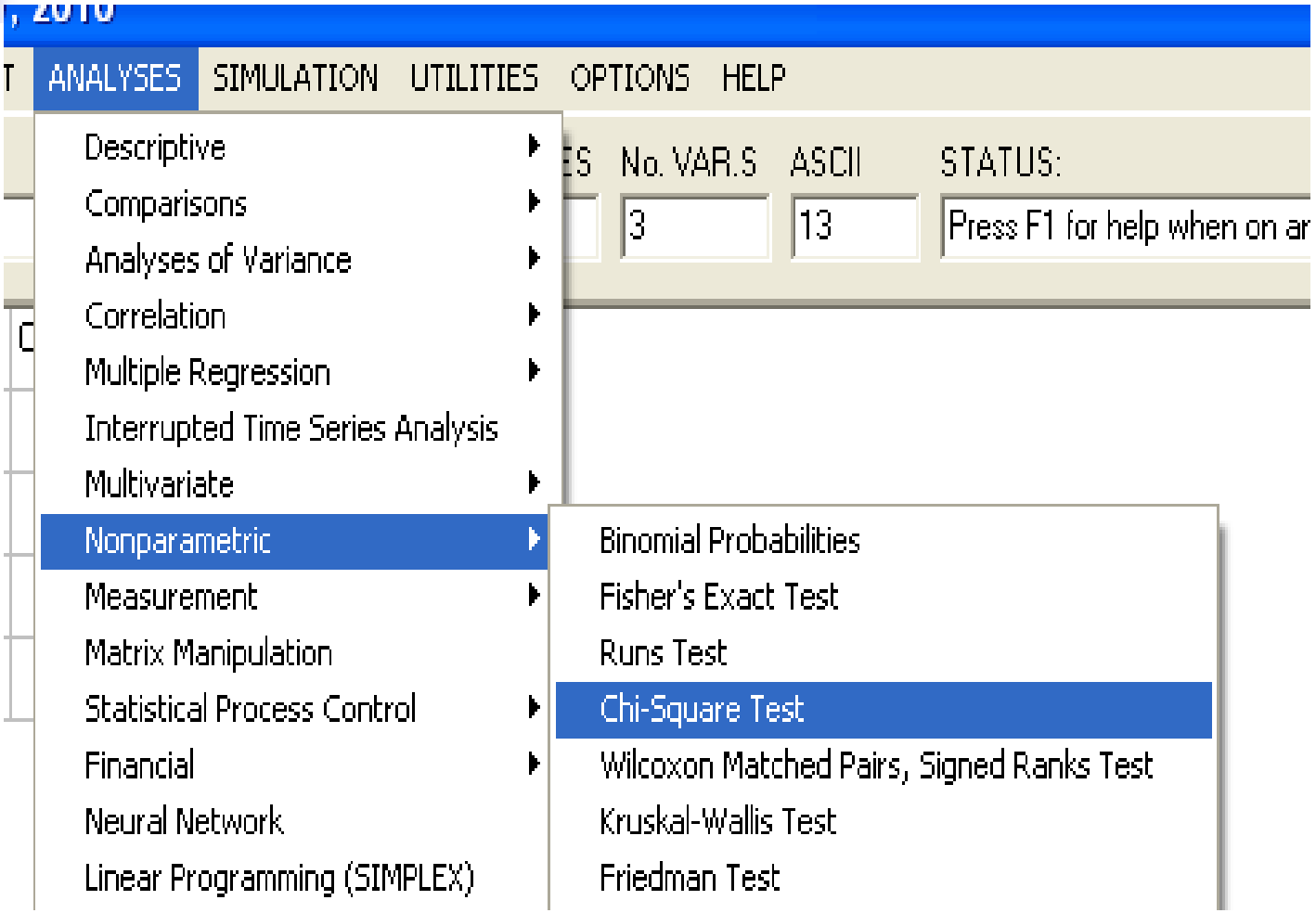
- The type II error rate β increases if chi-square is conducted on a smaller sample size.
- Can use Yates' continuity correction for small samples.
- Can use Fisher's exact test for small samples.

DV Example in Openstat

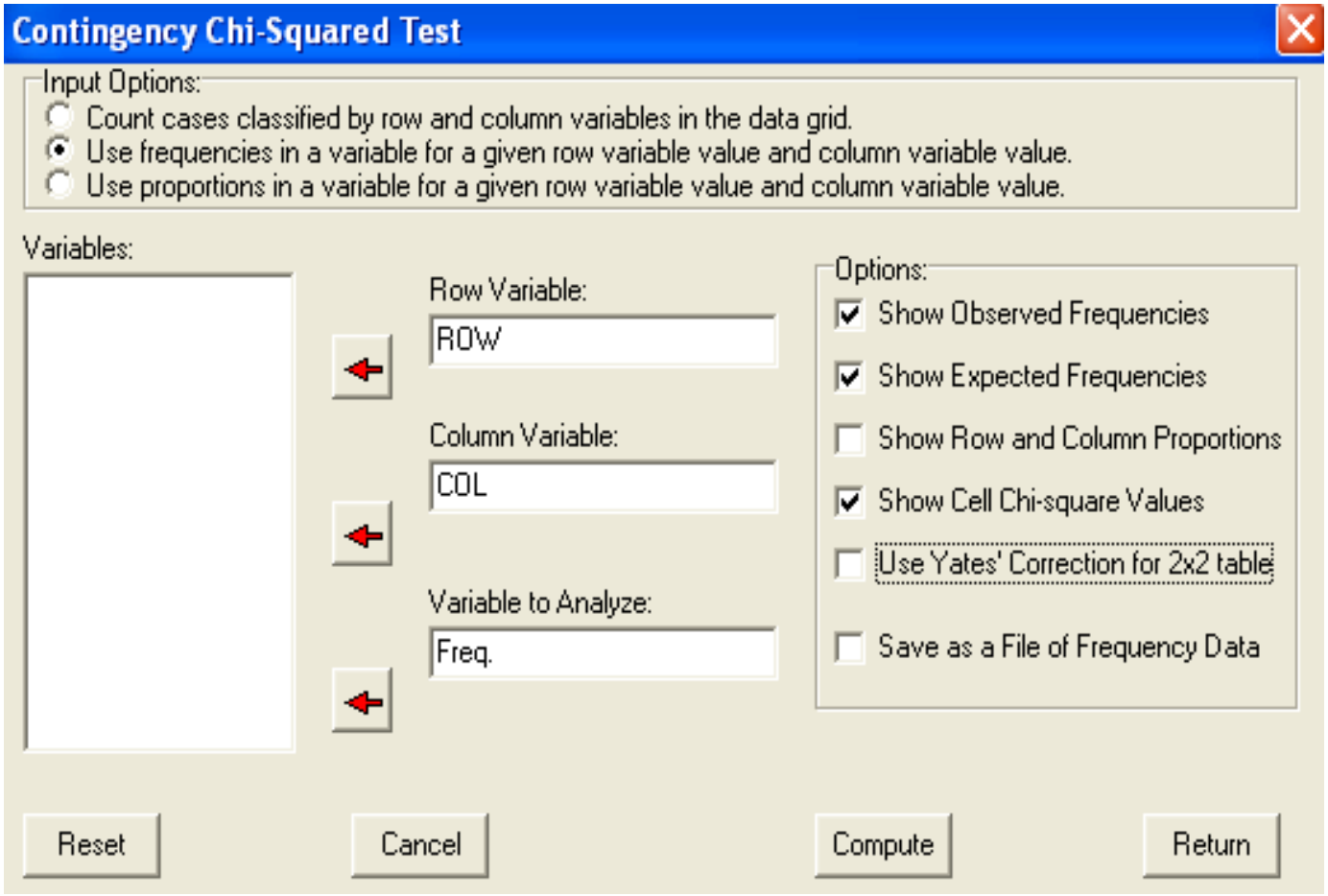


0	ROW	COL	Freq.
1	1.00	1.00	175.00
2	1.00	2.00	27.00
3	2.00	1.00	155.00
4	2.00	2.00	111.00

Chi-Square in Openstat



Getting Results



Chi-Square Results

Results Window

	COL. 1	COL. 2	Total
Row 1	175	27	202
Row 2	155	111	266
Total	330	138	468

EXPECTED FREQUENCIES

Expected Values

	COL. 1	COL. 2
Row 1	142.436	59.564
Row 2	187.564	78.436

CHI-SQUARED VALUE FOR CELLS

Chi-square Values

	COL. 1	COL. 2
Row 1	7.445	17.803
Row 2	5.654	13.520

Chi-square = 44.421 with D.F. = 1. Prob. > value = 0.000

Reject H_0, there is a difference in proportions of two groups of physicians.

Versatility of Chi-Square

- Distribution of a categorical variable in one sample can be compared with distribution of a categorical variable in another sample.
- Chi-Square can be used to compare an observed data with a theoretical distribution.
- Can be used to test difference between two counts.
- Can be extended to look at the effect of more than one input variable.

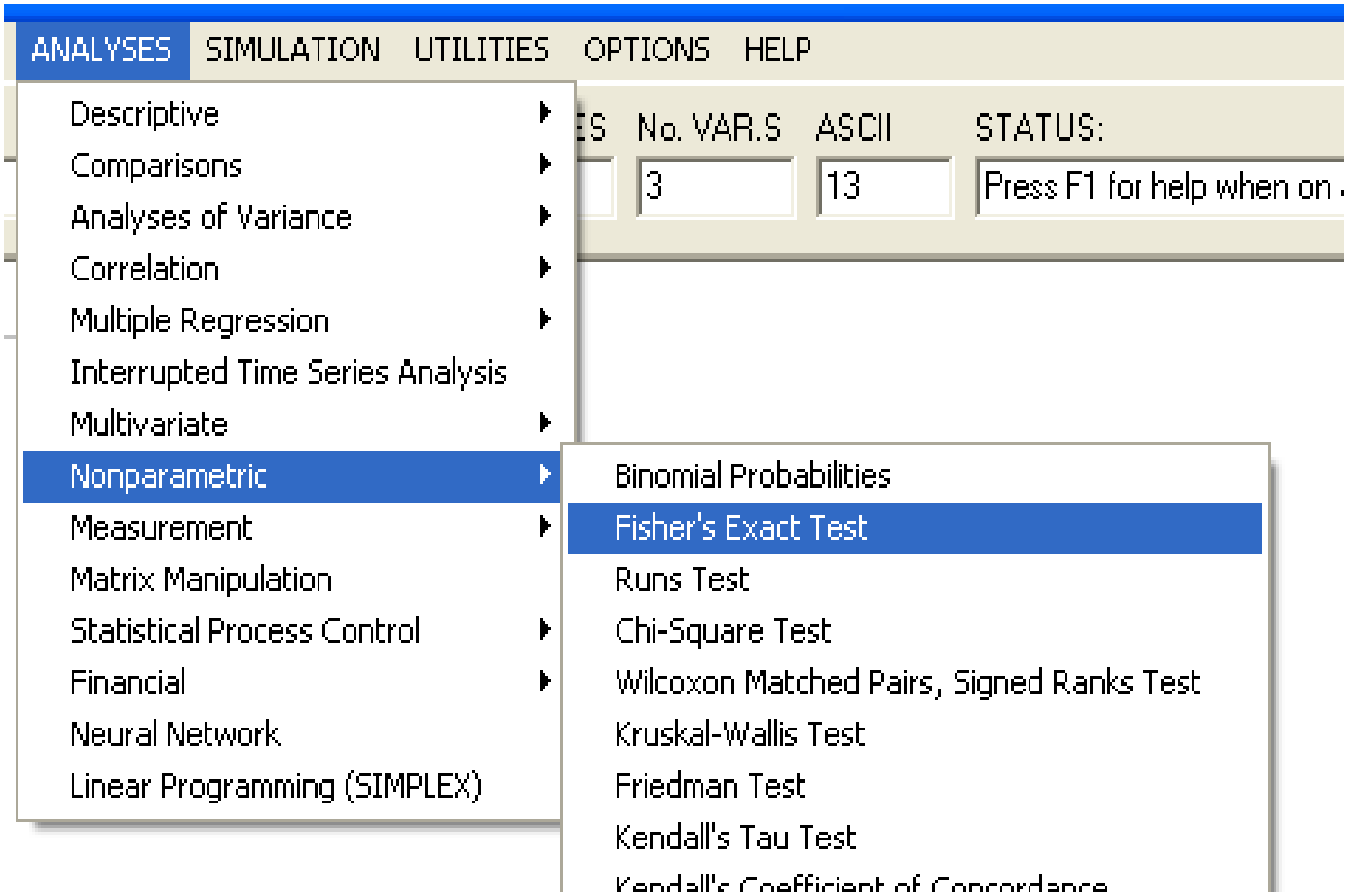
Fisher's Exact Test for Small Data Sets

Deaths in six months after fractured neck of femur in a specialised orthopedic ward (A) and a general ward (B)					
		Ward			
		A	**B**	**Total**	
Deaths	**Yes**	2	6	8	
	No	18	14	32	
Total		20	20	40	

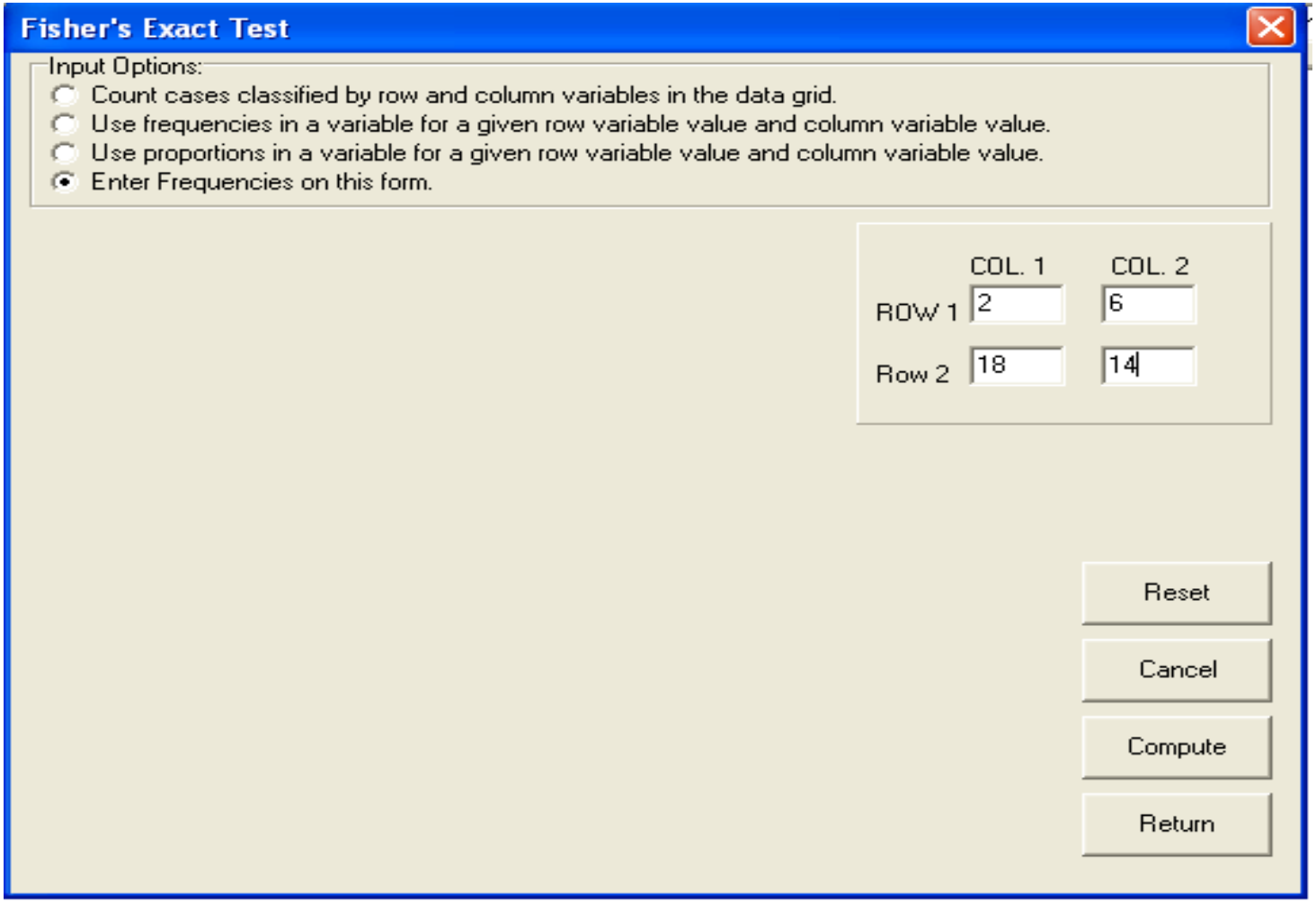
Medical Statistics A Commonsense Approach A8

H_0: There is no difference in deaths for two wards.

Fisher's Exact in Openstat



Entering Data



Results

Fisher Exact Probability Test

Accumulating Values of the Hypergeometric Distribution

Contingency Table for Fisher Exact Test

	Column	
Row	1	2
1	2	6
2	18	14

Probability = 0.0958

Cumulative Probability = 0.0958

Contingency Table for Fisher Exact Test

	Column	
Row	1	2
1	1	7
2	19	13

Probability = 0.0202

Results Cont'd

Probability = 0.0202

Cumulative Probability = 0.1159

Contingency Table for Fisher Exact Test

Row	Column 1	2
1	0	8
2	20	12

Probability = 0.0016

Cumulative Probability = 0.1176

Tocher ratio computed: 0.295
A random value of 0.549 selected was greater than the Tocher value.
Conclusion: Accept the null Hypothesis

Chi-squared = 2.500 with 1 d.f. and prob. > chi-square = 0.1138

Two tailed $p \approx 0.23$
can't reject H_0

Summary of Two Groups Methods

Nature of problem	Test
Interval & normal	Two independent sample t
Ordinal or interval, non normal	Mann-Whitney U
Independent Proportions	Z approximation
Independent proportions large sample size	Chi-Square
Independent Proportions small sample size	Fisher's Exact

www.epilepsy.va.gov/Statistics

Questions/Comments

Rizwana.Rehman@va.gov

(919) 286-0411 ext: 5024

Next lecture's highlights

Study Designs, Odds Ratio, Relative Risk