statstutor

community project

encouraging academics to share statistics support resources



# Statistical Methods 9. Nonparametric Testing

Based on materials provided by Coventry University and Loughborough University under a National HE STEM Programme Practice Transfer Adopters grant

National HE STEM Programme



www.statstutor.ac.uk

Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Overview of workshop

- What are nonparametric methods?
- When should you use them?
- Overview of nonparametric methods
- Comparing two groups:
  - Mann-Whitney U test
- Cross-tabulating two variables:
  - Chi-squared test of association
  - Fisher’s exact test



Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Parametric v. nonparametric methods

- Theoretical statistical distributions are constructed using parameters:
  - These parameters determine the location and shape of the frequency distribution
  - For example, the mean and standard deviation of the normal distribution – see Workshop 6
- Methods that assume observations come from a certain distribution are called **parametric methods**
- Methods that make no distributional assumptions about observations are called **nonparametric methods**

# When should non-parametric tests be used?

- When scale-based data does not satisfy the assumptions of the appropriate test or the conditions of its robust use
- For testing hypotheses with categorical (nominal/ordinal) data
- The advantage of nonparametric tests is that we do not have to test any assumptions or robustness conditions


Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Limitations

- Nonparametric methods are **less powerful** than parametric methods
- For example, for normally distributed data, a two sample t-test will detect a smaller real difference than the corresponding non-parametric test
- Application of nonparametric methods is difficult or impossible for complex data structures
- Nonparametric methods mainly involve hypothesis testing: less descriptive statistics can be calculated, for example, only maximum, minimum, median and quartiles for scale or ordinal data

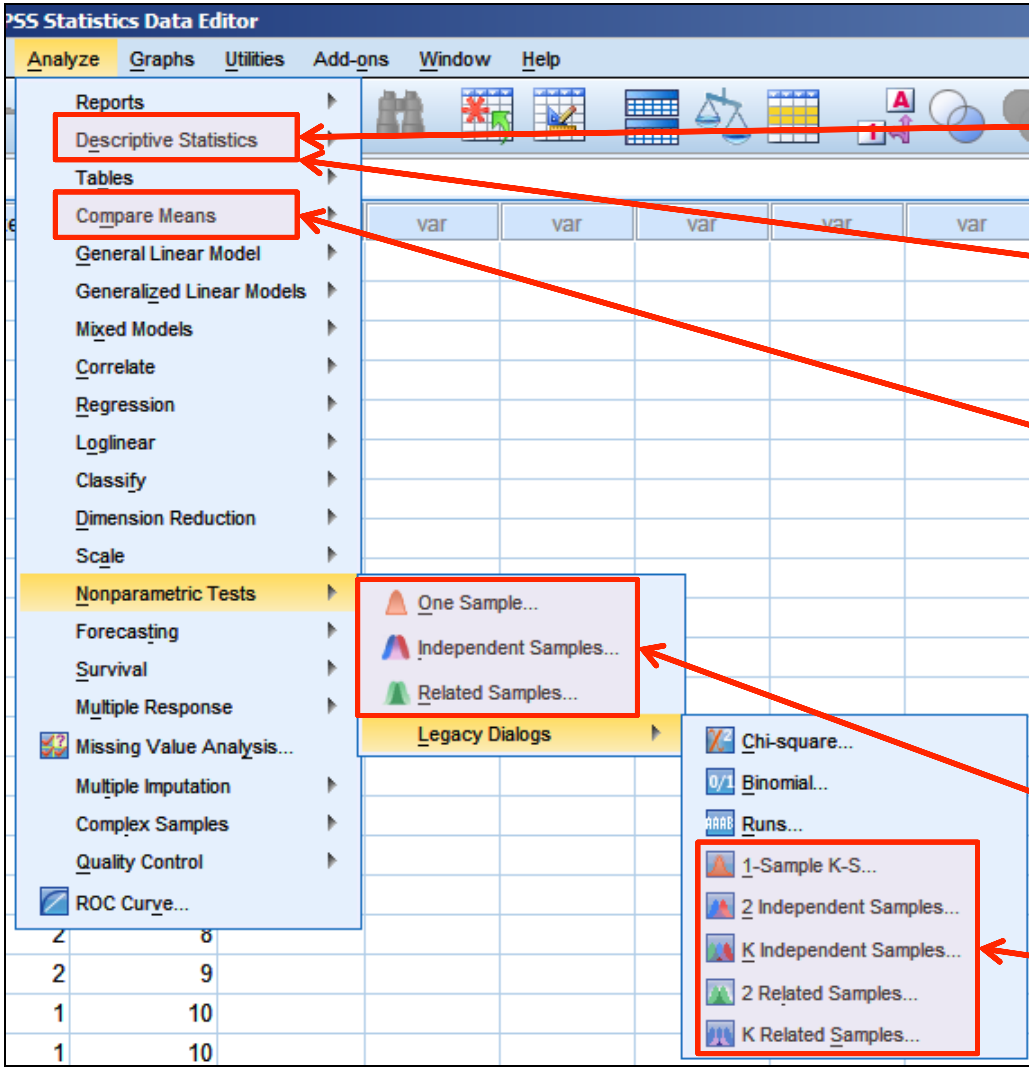
# Ranks

- Many nonparametric methods replace scale/ ordinal observations with ranks:

| Observation | 170 | 112 | 29 | 125 | 224 | 78 |
|---|---|---|---|---|---|---|
| Rank | 5 | 3 | 1 | 4 | 6 | 2 |

- Nonparametric test statistics are then based on ordering the data and working with the ranks
- SPSS takes care of all the details

# Nonparametric statistics in SPSS




Chi-squared and Spearman's correlation under Crosstabs

Normality tests under Explore

Alternative K Independent samples tests under One-Way ANOVA

Two choices with other nonparametric tests – new or legacy:

- New algorithms are better and output is more 'helpful'
- Legacy versions are less strict with data types



Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Nonparametric tests in SPSS

- **One sample:** Tests whether a sample of data follows a particular distribution
- **2 independent samples:** Compares two groups of cases (like an independent samples t-test – see Workshop 8)
- **K independent samples:** Compares two or more groups of cases (like a one-way ANOVA – see Workshop 10)
- **2 related samples:** Compares two paired groups of cases (like a paired-samples t-test – see Workshop 8)
- **K related samples:** Compares two or more related groups of cases (like repeated measures ANOVA)
- **2 ordinal/scale variables**: Spearman correlation and chi-squared test for association, and Fisher's exact test

# Example 1: Female stroke patients

- 100 female stroke patients were randomly assigned to one of two groups:
  - A standard physical therapy group (*Control*)
  - A group with the standard therapy plus emotional therapy (*Treatment*)
- Three months later the patients were evaluated on their ability to perform common Activities of Daily Life (ADL):

| Code | Travel | Cooking | Housekeeping |
|---|---|---|---|
| 0 | Same as before illness | Plans and prepares meals | As before |
| 1 | Gets out if someone else drives | Some cooking but less than before | Does at least half usual |
| 2 | Gets out in wheelchair | Gets food out if prepared by others | Occasional dusting of small jobs |
| 3 | Home or hospital bound | Does nothing for meals | No longer keeps house |
| 4 | Bed-ridden | Never did any | Never did any |

# Research question

- Does the additional therapy have an affect on any of these three measures activities for daily life?
- Null hypotheses:
  - The distribution of values for each measure does not depend on the group
- Alternative hypotheses:
  - The distribution of values for each measure does depend on the group

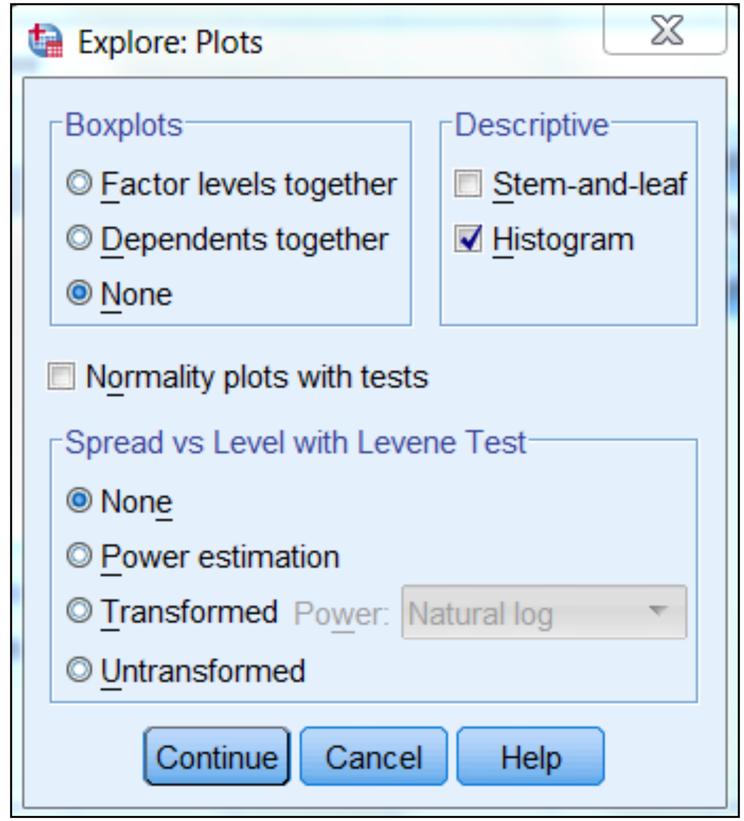


Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Step 1: descriptive statistics

- Open the file ADL.sav associated with this presentation
- Select Analyze – Descriptive Statistics – Explore…
- Select the three measures in the Dependent List and *Group* as the Factor List
- Select Plots… and choose None under boxplots and Histogram under descriptive

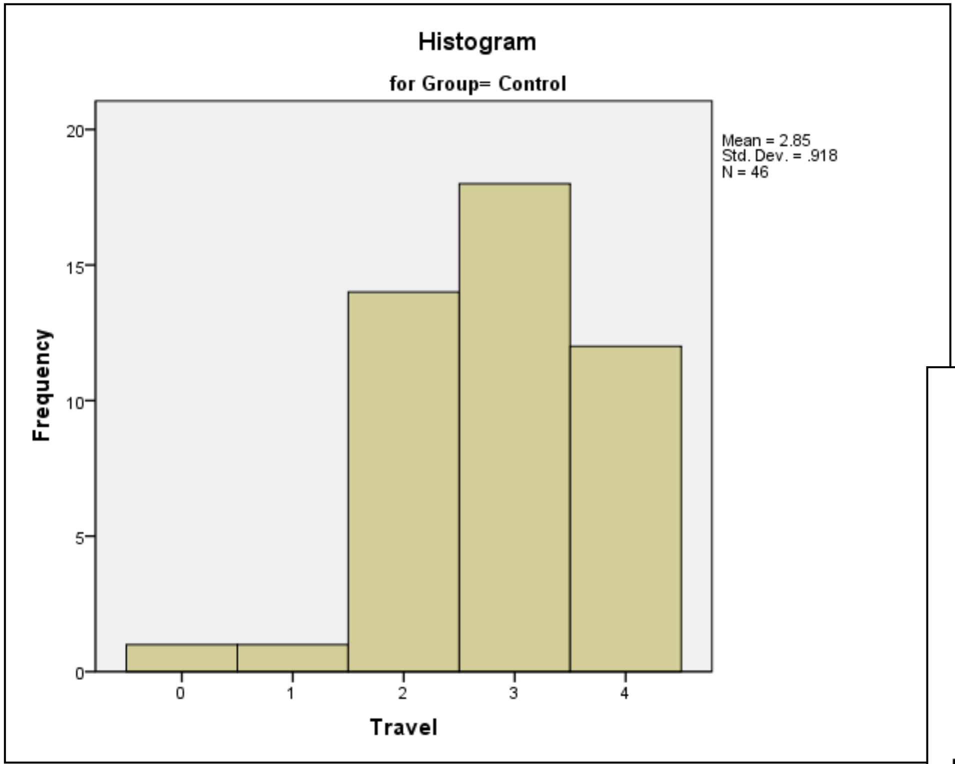




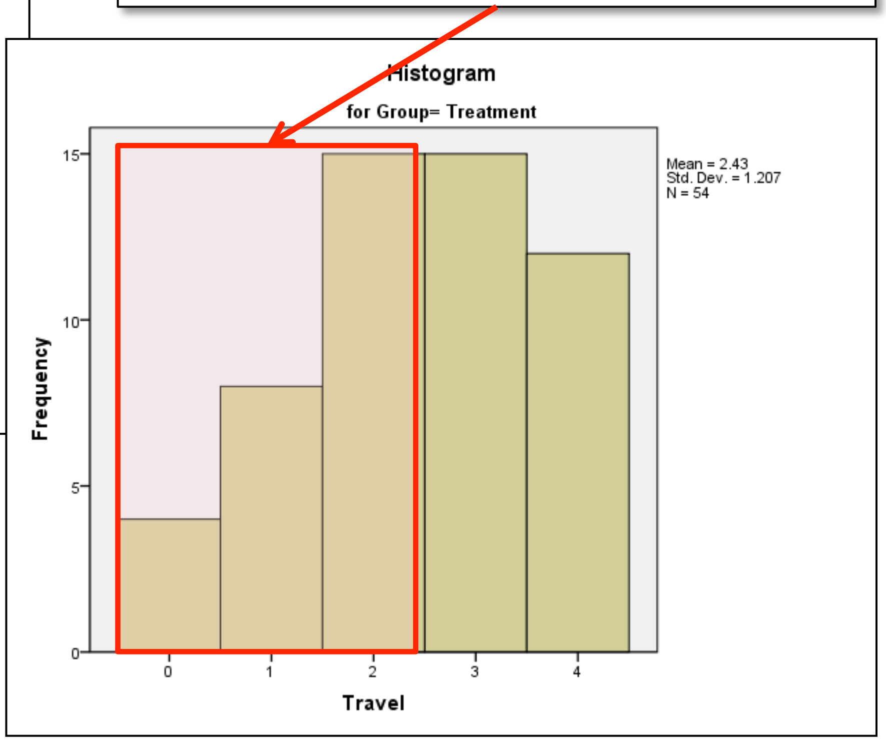

Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Distribution of *Travel* by *Group*




Percentage frequency of Treatment group is relatively higher for lower values of Travel




Thus the Treatment appears to be having a positive effect

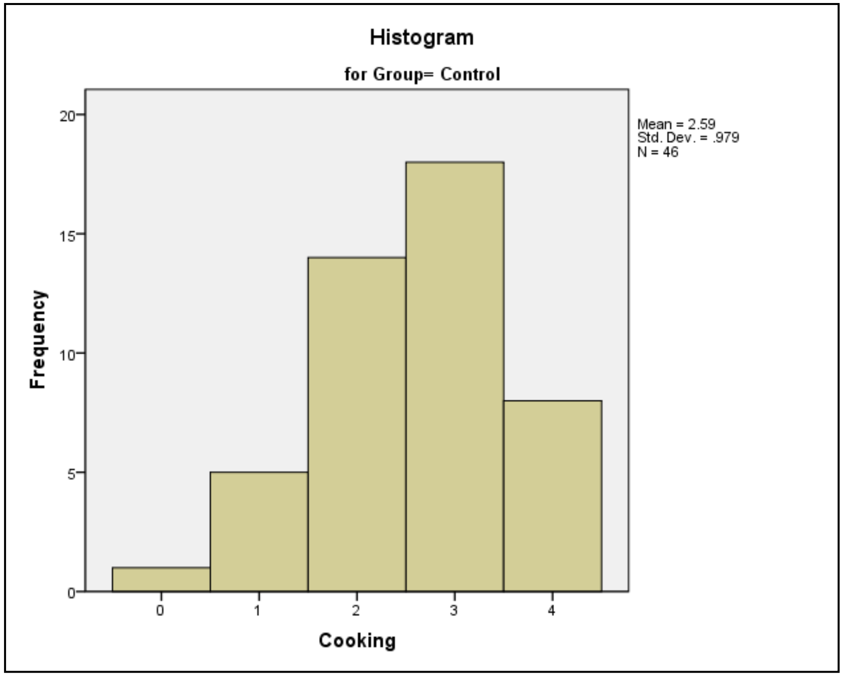


www.statstutor.ac.uk

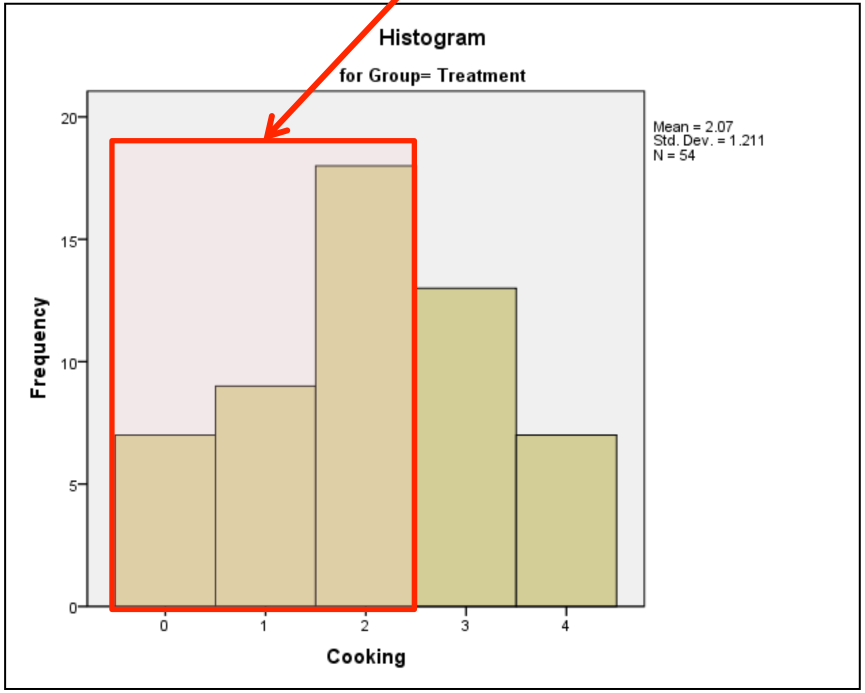
Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Distribution of *Cooking* by *Group*

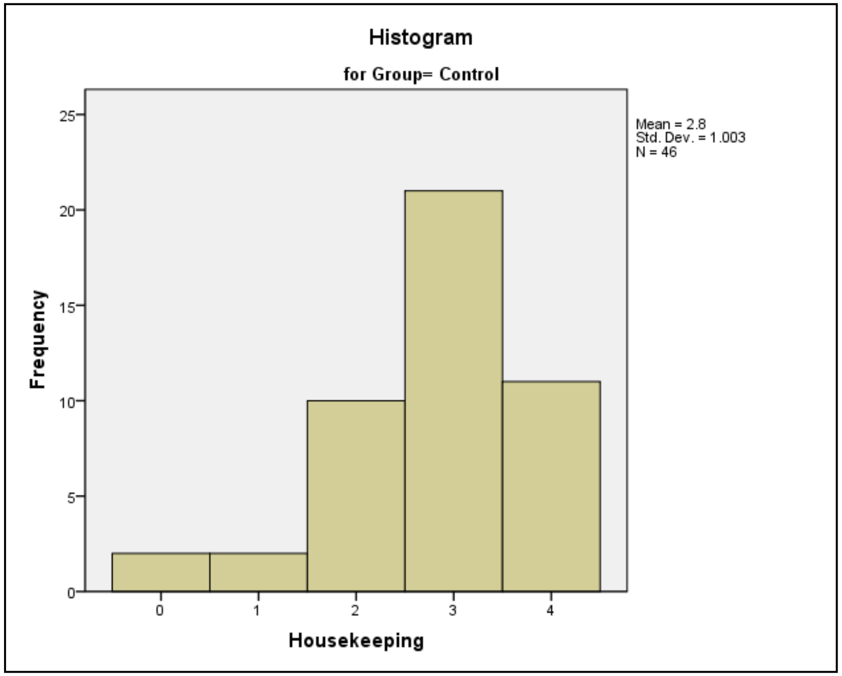



Difference in percentage frequencies bigger than for Travel




Thus the Treatment appears to be having a bigger positive effect

# Distribution of *Housekeeping* by *Group*




Difference in percentage frequencies even bigger than for Cooking


Thus the Treatment appears to be having an even bigger positive effect





Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Step 2: which test?

- The data is ordinal so we need to use a nonparametric test
- There are 2 independent groups
- The descriptive statistics suggest we should compare higher/lower ordinal values for the groups rather than general differences in shape
- The appropriate test is therefore the Mann-Whitney U test
- **Note**: If there had been other kinds of shape difference we should have used the chi-squared test
- **Note**: The category 4 data should be removed from each variable before it is tested because it corresponds to ‘not applicable’





Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Mann-Whitney U test

- A **non-parametric** test of two independent samples of ordinal or scale-based data
- Generally needs at least 5 data categories for the ordinal variable (here we only have 4 when the last case is removed, which is a bit of a problem)
- Alternative to an independent samples t-test for scale-based data if the test assumptions or robustness assumptions are not met
- Samples can be different sizes
- **Null hypothesis**: Values of *Travel* for the *Control* group are equally likely to be higher or lower than the values of *Travel* for the *Treatment* group

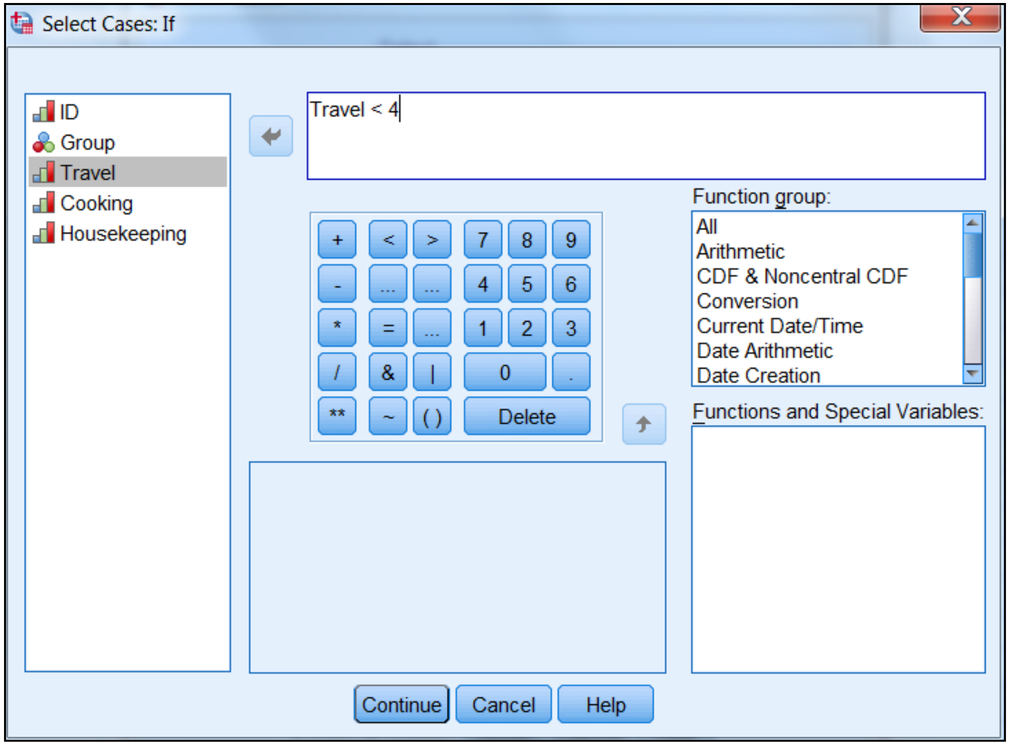


Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Step 3: remove the cases with *Travel* = 4

- ❑ Select Data – Select Cases
- ❑ Click on the If condition is satisfied radio button
- ❑ Select the If… button
- ❑ Select *Travel* then click on ‘<‘ and 4
- ❑ Select Continue etc.



# Step 4: run the test

- Select Analyze – Nonparametric Tests – Legacy Dialogs – 2 Independent Samples…
- **Note**: we cannot use the new version of this test with an ordinal data type for the test variable
- Select Travel for the Test Variable List and Group for the Grouping Variable
- Click on Define Groups and select 0 for Group 1 and 1 for Group 2
- The Mann-Whitney U test is selected by default



Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Step 5: interpret the output

- Significance value of test is 0.032
- We reject the null hypothesis with 95% confidence, i.e. there is evidence that treatment is having an effect. The effect is also positive (negative Z value).
- Repeat the same analysis for *Cooking* and *Housekeeping* (first select the appropriate cases)
- Same result obtained for *Cooking*, p-value is slightly smaller (so evidence is slightly stronger)
- Even stronger result obtained for *Housekeeping* – null hypothesis rejected with 99% confidence – strong evidence that treatment is having an effect

**Test Statistics**[a]

| | Travel |
|---|---|
| Mann-Whitney U | 524.000 |
| Wilcoxon W | 1427.000 |
| Z | -2.139 |
| Asymp. Sig. (2-tailed) | .032 |

a. Grouping Variable: Group

**Test Statistics**[a]

| | Cooking |
|---|---|
| Mann-Whitney U | 646.000 |
| Wilcoxon W | 1774.000 |
| Z | -2.310 |
| Asymp. Sig. (2-tailed) | .021 |

a. Grouping Variable: Group

**Test Statistics**[a]

| | Housekeeping |
|---|---|
| Mann-Whitney U | 553.500 |
| Wilcoxon W | 1681.500 |
| Z | -2.700 |
| Asymp. Sig. (2-tailed) | .007 |

a. Grouping Variable: Group

# Example 2 : Smartphone purchasing survey

**92 people were asked:**

**Q1**: What is your gender?

- Male
- Female

**Q2**: What is you age?

*This is recorded in the following categories:*

*17-24, 25-29, 30-39 and 40+*

**Q3**: On a scale of 0 to 10 how important do you consider brand when purchasing a smartphone?

(where 0 = extremely unimportant and 10 = extremely important)

# Research questions

1. Is the importance of brand when purchasing a smartphone gender related?
2. Is the importance of brand when purchasing a smartphone age related?

**Null hypotheses:**

1. Brand importance is not gender related
2. Brand importance is not age related

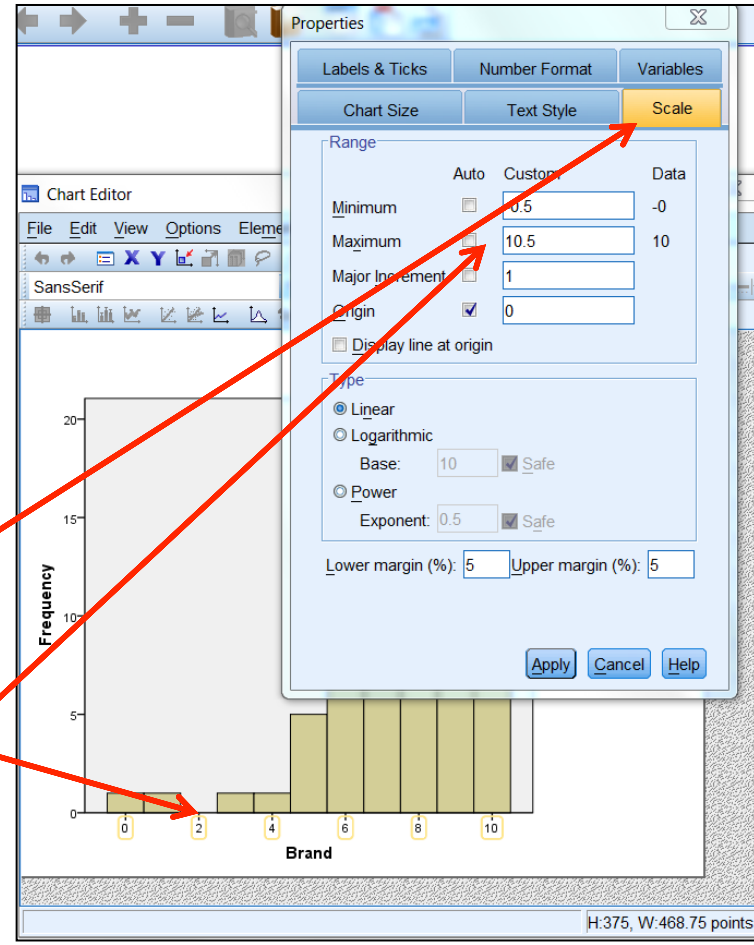


Peter Samuels
Birmingham City University

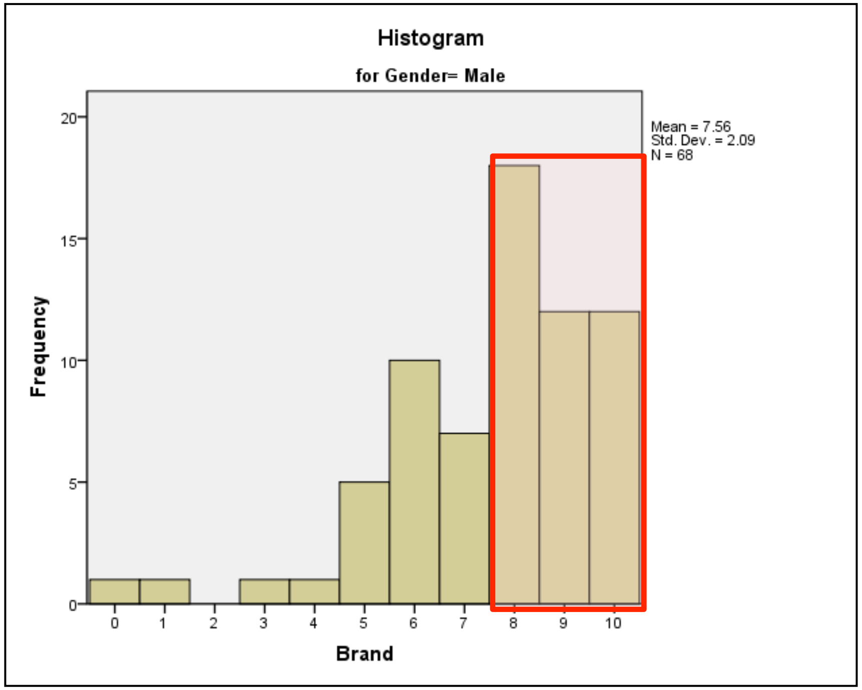
Reviewer: Ellen Marshall
University of Sheffield

# Step 1: descriptive statistics

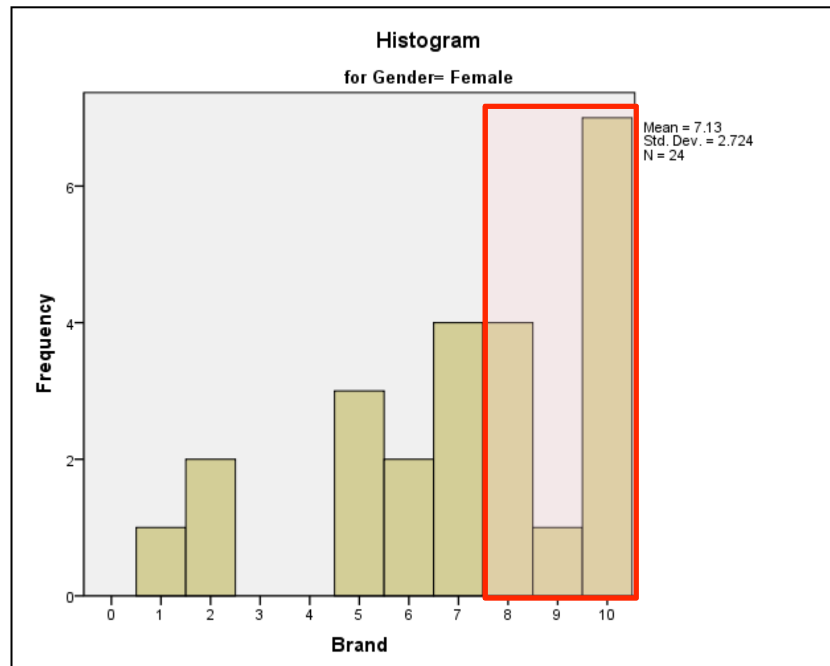
- ❑ Upload the file Smartphone.sav associated with this presentation
- ❑ Select Analyze – Descriptive Statistics – Explore…
- ❑ Select *Brand* in the Dependent List and *Gender* in the Factor List
- ❑ Select Plots… and choose None under boxplots and Histogram under descriptive
- ❑ Double click on each graph
- ❑ Double click on the values on the horizontal axis in the chart editor
- ❑ Select the Scale tab
- ❑ Change the Minimum to 0, the Maximum to 10 and the Major increment to 1



# Distribution of *Brand* by *Gender*




Both distributions clearly not normal with different sample sizes ⇒ non-parametric test needed




Shape of data different for Brand = 8, 9 and 10, but not in the same direction



www.statstutor.ac.uk

Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# A nonparametric test of association

- The data consist of counts of subjects with particular profiles
  - Profiles are formed by scale and categorical variables
  - Often referred to as a contingency table
- The test is whether there is an association between the categorical variables
- Equivalent to dropping pebbles at random in a grid where the row and column totals are already known

# Chi-squared ($\chi^2$) test

- Works with:
  - 2-way tables with known row and column totals, or
  - Measuring a sequence of observed values against expected values (not covered here)
- Based on calculating expected values for the table frequencies and comparing these with the observed values
- Only valid when most (≥80%) of the expected values are sufficiently large (≥5) and none has expected value <1
- Null hypothesis: the observed values are randomly distributed based on the expected values (i.e. there is no association between the two variables)

# Chi-squared test in SPSS

- ❑ Select *Analyze – Descriptive Statistics – Crosstabs…*
- ❑ Select *Gender* for the rows and *AgeCategory* for the columns
- ❑ Select *Statistics…* then *Chi-square* then *Continue*
- ❑ Select *Cells…* then *Expected* then *Continue*

Analysis is invalid because 15 out of 22 cells (way more than 20%) have an expected frequency less than 5

We need to combine some of the columns together and retest

**Gender * Brand Crosstabulation**

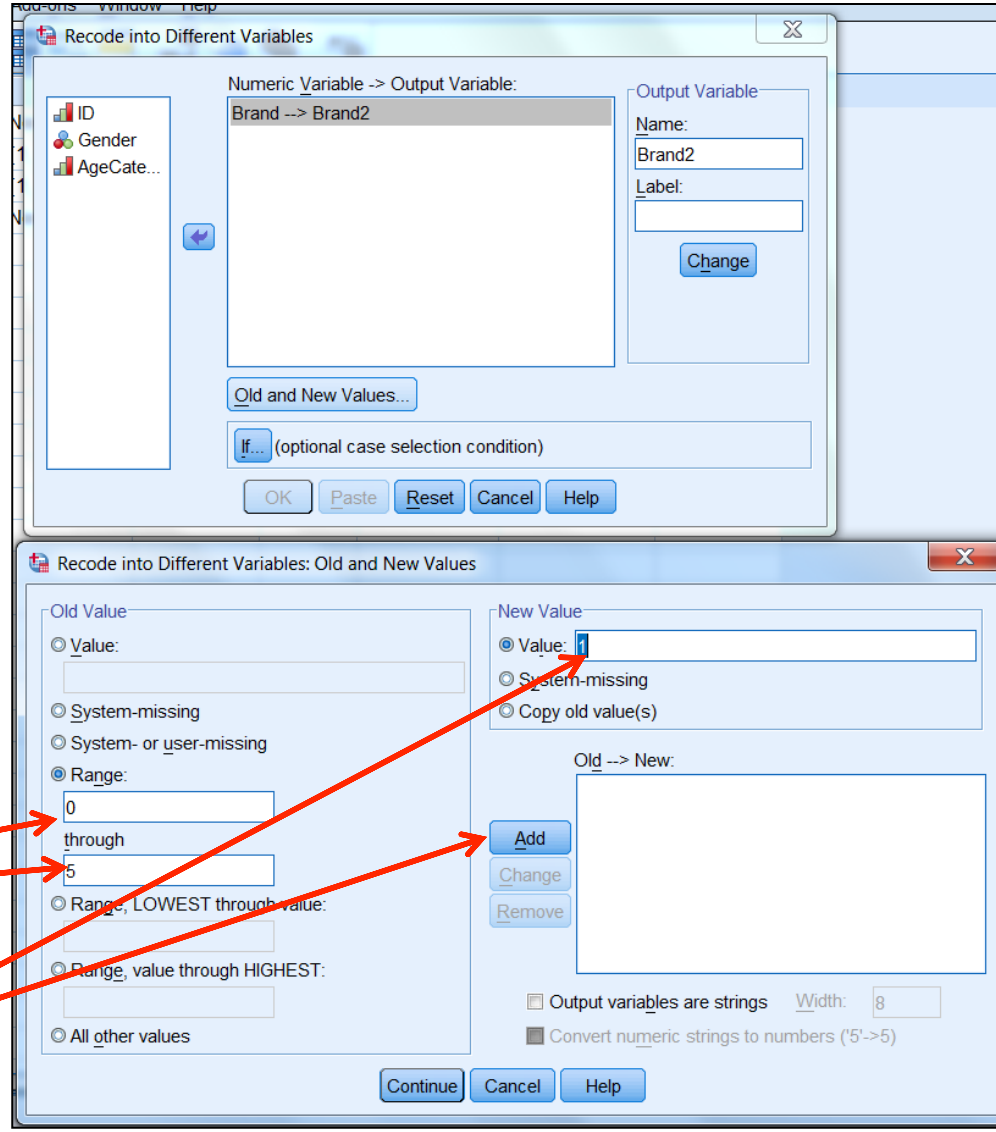
| | | | Brand | | | | | | | | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| Gender | Male | Count | 1 | 1 | 0 | 1 | 1 | 5 | 10 | 7 | 18 | 12 | 12 | 68 |
| | | Expected Count | .7 | 1.5 | 1.5 | .7 | .7 | 5.9 | 8.9 | 8.1 | 16.3 | 9.6 | 14.0 | 68.0 |
| | Female | Count | 0 | 1 | 2 | 0 | 0 | 3 | 2 | 4 | 4 | 1 | 7 | 24 |
| | | Expected Count | .3 | .5 | .5 | .3 | .3 | 2.1 | 3.1 | 2.9 | 5.7 | 3.4 | 5.0 | 24.0 |
| Total | | Count | 1 | 2 | 2 | 1 | 1 | 8 | 12 | 11 | 22 | 13 | 19 | 92 |
| | | Expected Count | 1.0 | 2.0 | 2.0 | 1.0 | 1.0 | 8.0 | 12.0 | 11.0 | 22.0 | 13.0 | 19.0 | 92.0 |


Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Recode *Brand* into a new variable

- Select Transform – Recode into Different Variables…
- Select *Brand* from the list
- Enter *Brand2* as the output variable name and press Change
- Select Old and New Values…
- Under Old Value, select Range and enter the values 0 and 5
- Under New Value, enter 1 and select Add








Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

- Repeat with the range 6 through 7 going to 2
- Recode the value 8 as 3, 9 as 4, and 10 as 5
- On the Variable View, change the number of decimal places of *Brand2* to 0 and its data type to Ordinal
- Add values to *Brand2* to explain these settings
- Re-run the chi-squared test by changing the variable from *Brand* to *Brand2* and keeping all the other options the same

Analysis still not valid as 3 out of 10 cells (>20%) still have expected frequency <5

**Gender * Brand2 Crosstabulation**

| | | | Brand2 | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | | 0-5 | 6-7 | 8 | 9 | 10 | |
| Gender | Male | Count | 9 | 17 | 18 | 12 | 12 | 68 |
| | | Expected Count | 11.1 | 17.0 | 16.3 | 9.6 | 14.0 | 68.0 |
| | Female | Count | 6 | 6 | 4 | 1 | 7 | 24 |
| | | Expected Count | 3.9 | 6.0 | 5.7 | 3.4 | 5.0 | 24.0 |
| Total | | Count | 15 | 23 | 22 | 13 | 19 | 92 |
| | | Expected Count | 15.0 | 23.0 | 22.0 | 13.0 | 19.0 | 92.0 |

Result not significant as P-value for chi-squared >0.05

**Chi-Square Tests**

| | Value | df | Asymp. Sig. (2-sided) |
|---|---|---|---|
| Pearson Chi-Square | 5.640[a] | 4 | .228 |
| Likelihood Ratio | 6.096 | 4 | .192 |
| Linear-by-Linear Association | .183 | 1 | .669 |
| N of Valid Cases | 92 | | |

a. 3 cells (30.0%) have expected count less than 5. The minimum expected count is 3.39.

Recoding again and re-running would probably not improve the result

# Fisher's exact test

- Applies to 2x2 contingency tables
- Works with smaller samples than chi-squared:
  - Sample size > 40 ⇒ chi-squared can be used
  - Sample size between 20 and 40 and the smallest expected frequency ≥ 5, ⇒ chi-squared can be used
  - Otherwise Fisher's exact test **must** be used
- A one-sided test with a 2-sided normal approximation
- Provided automatically by SPSS when you cross-tabulate and select chi-squared


Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

- Go back to the first data set
- Select Transform - Recode into Different Variable to recode *Housekeeping* into *Housekeeping2* with:
  - 0, 1 and 2 recoded as 1
  - 3 recoded as 2
  - (We are leaving out 4)
- Change *Housekeeping2* so that it has zero decimal places and the values are as above
- Select Analyze – Descriptive Statistics – Crosstabs and choose Group and Housekeeping2 and the chi-squared test

There are 82 valid cases so chi-squared would be valid here

Chi-squared P-value is 0.011 – slightly weaker than the Mann-Whitney U test result (0.007)

Fisher's exact P-value is 0.010 for one sided ($H_0$: Housekeeping is the same or worse) and 0.014 for 2-sided

**Group * Housekeeping2 Crosstabulation**

| | | | Housekeeping2 | | |
|---|---|---|---|---|---|
| | | | Still keeps house | No longer keeps house | Total |
| Group | Control | Count | 14 | 21 | 35 |
| | | Expected Count | 19.6 | 15.4 | 35.0 |
| | Treatment | Count | 32 | 15 | 47 |
| | | Expected Count | 26.4 | 20.6 | 47.0 |
| Total | | Count | 46 | 36 | 82 |
| | | Expected Count | 46.0 | 36.0 | 82.0 |

**Chi-Square Tests**

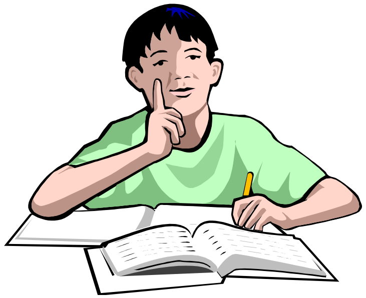
| | Value | df | Asymp. Sig. (2-sided) | Exact Sig. (2-sided) | Exact Sig. (1-sided) |
|---|---|---|---|---|---|
| Pearson Chi-Square | 6.425[a] | 1 | .011 | | |
| Continuity Correction[b] | 5.335 | 1 | .021 | | |
| Likelihood Ratio | 6.477 | 1 | .011 | | |
| Fisher's Exact Test | | | | .014 | .010 |
| Linear-by-Linear Association | 6.347 | 1 | .012 | | |
| N of Valid Cases | 82 | | | | |

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 15.37.
b. Computed only for a 2x2 table



Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield



# Activities

1. Run a chi-squared test with the three measures of ADL against the treatment *Group* with the first data set, recoding each measure into fewer categories if necessary. Are the results more or less significant? Explain.
2. Run a Mann-Whitney U test of *Brand* against *Gender* with the second data set. Are the results more or less significant? Explain.
3. Describe *Brand* against *AgeCategory* (e.g. a boxplot or multiple histograms), decide on the best way to test this association, carry out the test, ensuring it is valid, interpret your findings, repeating if necessary.


Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Recap

We have covered:

- What are nonparametric methods?
- When should you use them?
- Overview of nonparametric methods
- Comparing two groups:
  - Mann-Whitney U test
- Cross-tabulating two variables:
  - Chi-squared test of association
  - Fisher's exact test



Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield

# Bibliography

Field, A. (2013) *Discovering Statistics using SPSS: (And sex and drugs and rock 'n' roll),* 4th ed., London: SAGE, Chapter 6 (for Mann-Whitney U test) and Chapter 18 (for Chi-squared and Fisher's exact tests).

Pallant, J. (2010) *SPSS Survival Manual: A step by step guide to data analysis using SPSS*, 5th ed., Maidenhead: Open University Press, Chapter 16.

statstutor (n.d.) *Calculating Expected Frequencies in Two Way Tables resources*. Available at: http://www.statstutor.ac.uk/topics/chi-squared-tests-of-association/calculating-expected-frequencies-in-two-way/ [Accessed 8/01/14].

statstutor (n.d.) *Chi-Squared Tests for Two-Way (Contingency) Tables resources*. Available at: http://www.statstutor.ac.uk/topics/chi-squared-tests-of-association/chi-squared-tests-for-two-way-tables/ [Accessed 8/01/14].

statstutor (n.d.) *Wilcoxon Mann-Whitney Test resources*. Available at: http://www.statstutor.ac.uk/topics/nonparametric-methods/wilcoxon-mann-whitney-test/ [Accessed 8/01/14].



Peter Samuels
Birmingham City University

Reviewer: Ellen Marshall
University of Sheffield