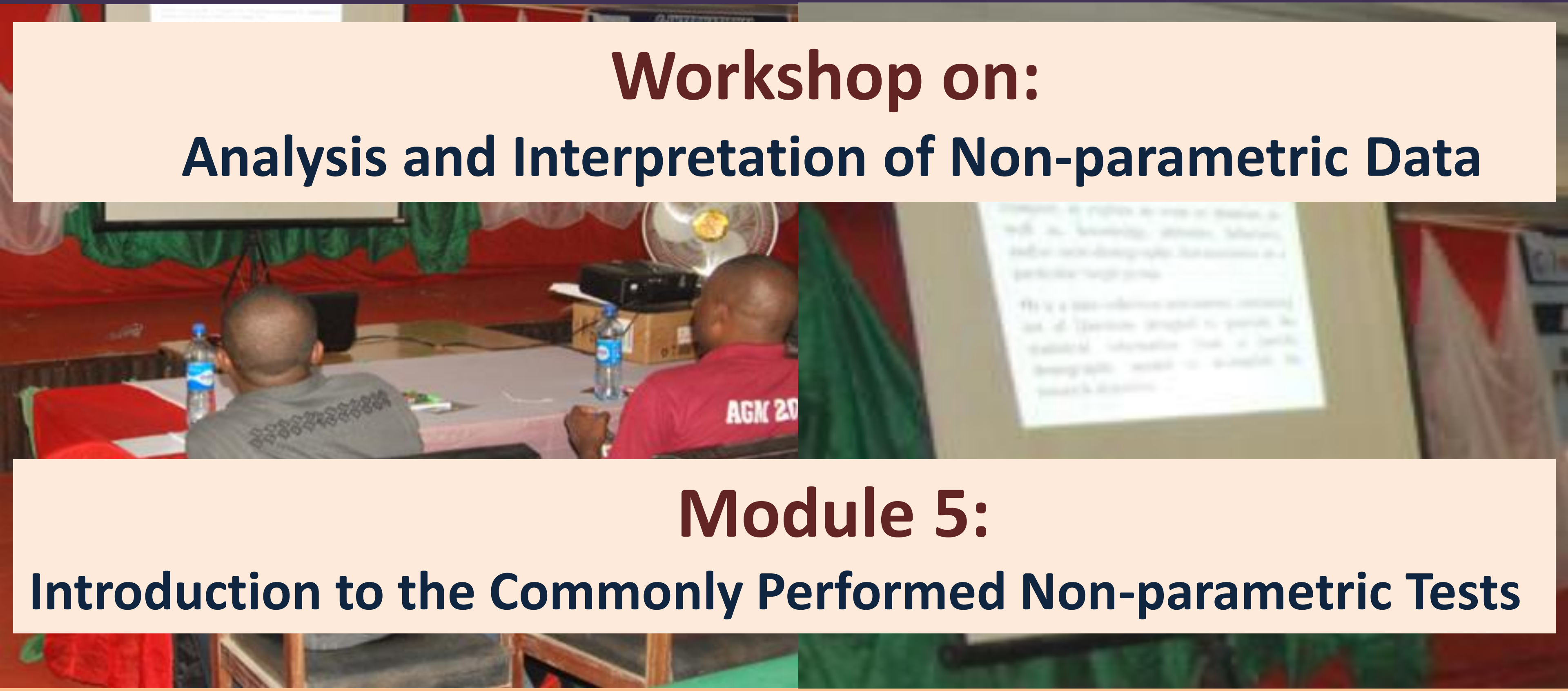
DOI: 10.33515/cint.sld/2023/003

C-International Tutors

Committed to bringing knowledge and skills to your doorsteps



Workshop on:

Analysis and Interpretation of Non-parametric Data

Module 5:

Introduction to the Commonly Performed Non-parametric Tests

https://cintarch.com/tutorials-videos/

https://www.facebook.com/cintarch.tutors/

cintarch.tutors@gmail.com

https://www.instagram.com/cintarchtutors/?hl=en

https://twitter.com/CInternational6

https://www.youtube.com/channel/UCPEhxtU4B3Tu0PtPyMGBU8Q?view_as=subscriber

Introduction to the Commonly Performed Non-parametric Tests

By

Prof. Awosan K.J.

Learning outcomes

After this session, you will know:

- The commonly performed non-parametric tests
- The purpose and types of the commonly performed non-parametric tests
- How to interpret the results of analysis of correlation and regression

Correlation and Regression Analysis

- Kendall's tau correlation
- Spearman rho correlation

Correlation and Regression

- Correlation and linear regression are used for investigating the relationship between two quantitative variables
- Correlation quantifies the strength of the linear relationship between a pair of variables, whereas regression expresses the relationship in the form of an equation
- The two common non-parametric tests of correlation are **Kendall's tau** and **Spearman rho tests**

Correlation and Regression contd.

- The first step in investigating a relationship between two quantitative variables is to show the data values graphically on a scatter diagram, but the strength of the relationship is measured by the means of an index known as **correlation coefficient** (represented by **r**)

- In examining the relationship between 2 quantitative variables, it is necessary to know which one influences the other

Correlation and Regression contd.

- For example in determining the association between height and weight, it is height that influences weight and not the other way round
- It can thus be said that weight is dependent on height, but height does not depend on weight
- Height is therefore the **independent variable** (and it is plotted on the **x axis**), while weight is the **dependent variable** (and it is plotted on the **y axis**) as shown in the scatter diagram (Figure 9.1)

Correlation and Regression contd.

- The correlation coefficient (r) ranges from -1 to 0, to +1

Interpretation

r = +ve means: **Positive correlation** (or **direct relationship**). For example, there is a positive correlation between height and weight. This means that as the height increases, the weight also increases (Figure 9.1).

r = -ve means: **Negative correlation** (or **inverse relationship**). For example, there is a negative correlation between the immunity status and the risk of disease. This means that as the immunity status increases, the risk of disease decreases (Figure 9.2).

r = 0 means: **No correlation** (or **no relationship**). For example, there is no correlation between the students' admission numbers and their examination scores (Figure 9.3)

Correlation and Regression contd.

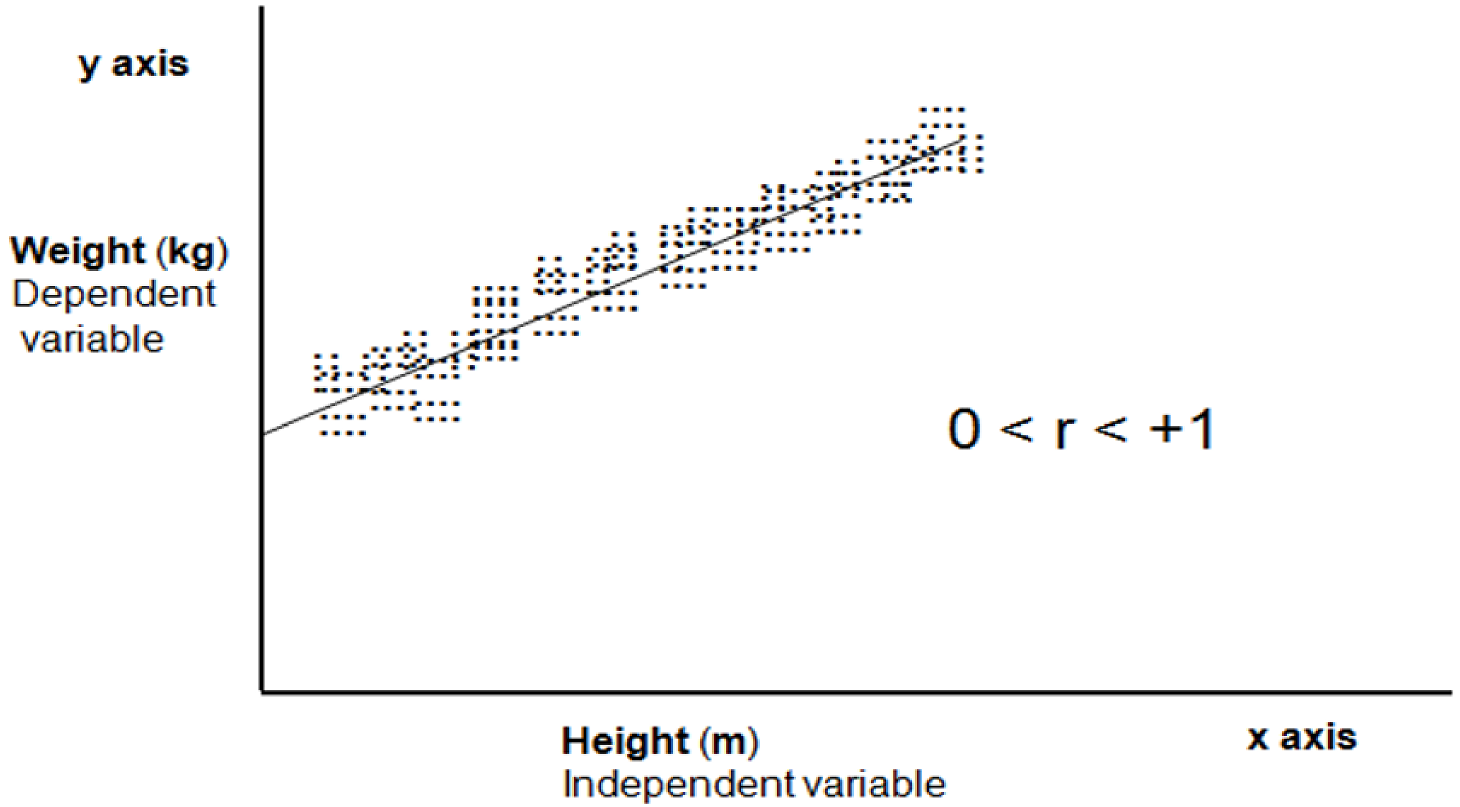


Figure 9.1: Positive correlation between weight and height

Source: Awosan (2020)

Correlation and Regression contd.

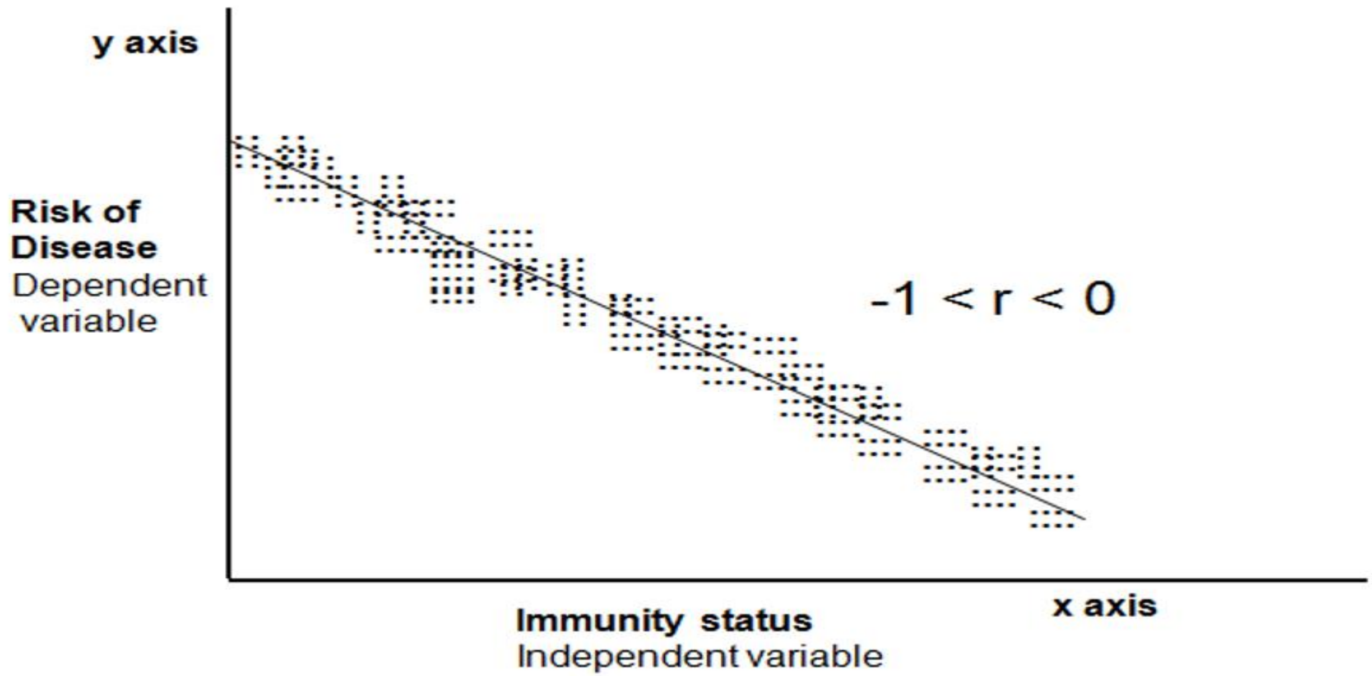


Figure 9.2: Negative correlation between immunity status and risk of disease

Source: Awosan (2020)

Correlation and Regression contd.

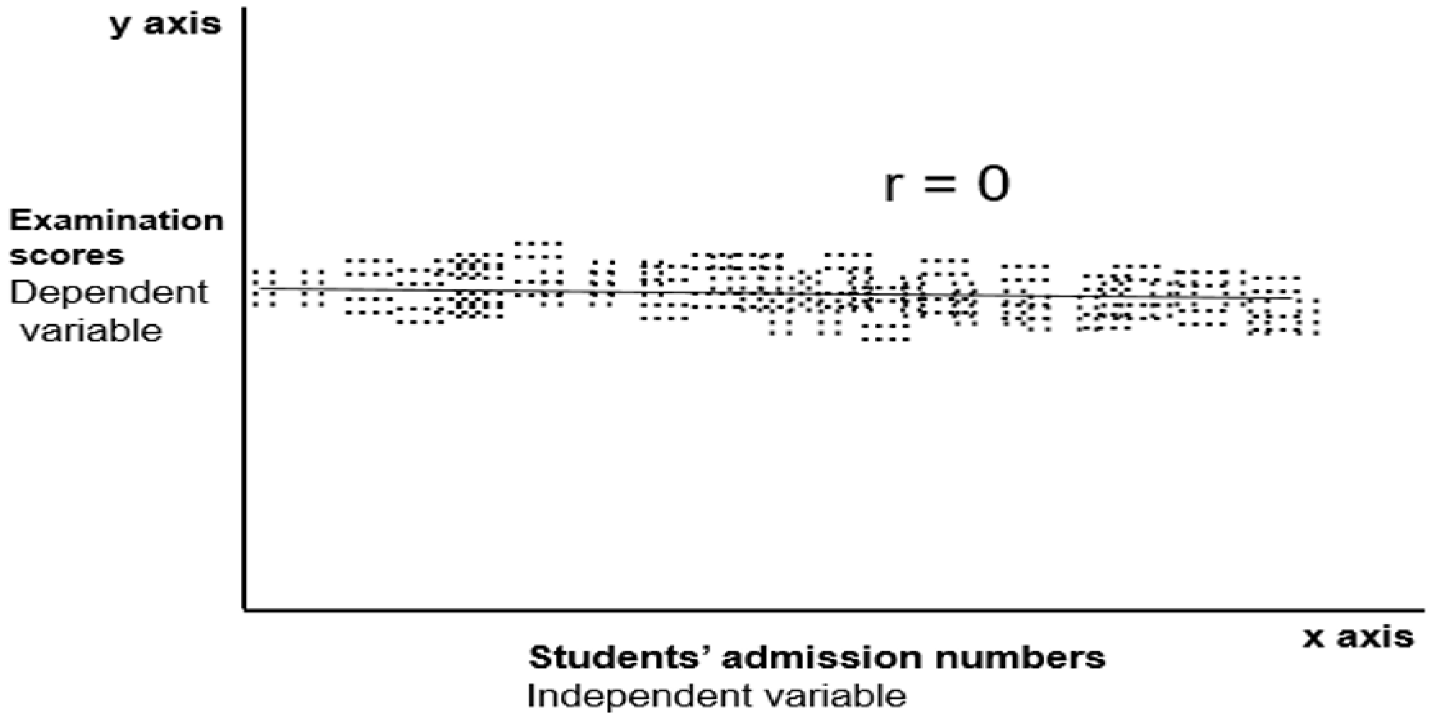


Figure 9.3: No correlation between students' admission numbers and examination scores

Source: Awosan (2020)

Correlation and Regression contd.

❖ Measurement of the strength of correlations

- The value of the correlation coefficient (**r**) shows the strength of the correlation between the variables concerned as shown in Table 9.1.

Table 9.1: Interpretation of the strength of correlations based on the r value

r value	Strength of correlation
+.70 or higher	Very strong positive
+.40 to +.69	Strong positive
+.30 to +.39	Moderate positive
+.20 to +.29	Weak positive
+.01 to +.19	Negligible
0	Nil
-.01 to -.19	Negligible
-.20 to -.29	Weak negative
-.30 to -.39	Moderate negative
-.40 to -.69	Strong negative
-.70 or higher	Very strong negative

Source: Awosan (2020)

Correlation and Regression contd.

❖ Measurement of effect size

- In research the value of the correlation coefficient is also used to measure the effect size, and this can be interpreted as indicated in Table 9.1, or interpreted in 3 levels as indicated in Table 9.2

Table 9.2: Interpretation of the effect size based on the r value

r value	**Effect size**
±.10 to .29	Small effect
±.30 to .49	Medium effect
±.50 or higher	Large effect

Source: Awosan (2020)

- It should be noted that strong correlation cannot prove causality because there may be other measured or unmeasured variables affecting the result. This is known as the **third variable problem** or **tertium quid**. Also, correlation coefficient does not specify which variable causes the other to change

Correlation and Regression contd.

❖ Coefficient of Determination

- In addition to the correlation coefficient, the coefficient of determination (**R^2**) is also used to measure the effect size. It is computed as the square of the correlation coefficient (i.e., **R^2** = **r** x **r**)
- It is a measure of the variability in one variable that is shared by the other variable

  For example:

  If **r** = .4410

  R^2 = $(.4410)^2$ = **0.194**

  And if converted to percentage gives **19.4%**

- It means **variable A** can account for 19.4% of the variation in **variable B**. It should be noted that variable A can only account for this magnitude of variation in variable B, but it does not necessarily cause the variation

Correlation and Regression contd.

❖ Types of correlation analysis

- The types of correlation analysis (which are usually performed using computer statistical packages) include:
- **Bivariate correlation:** This is a correlation between two variables
- **Partial correlation:** This looks at the relationship between two variables while controlling the effect of the third or other variables on the two variables being examined.
- The number of variables controlled for determines the order of the correlation as shown below:

Number of variables	Order
1 variable	First – order partial correlation
2 variables	Second - order partial correlation
3 variables	Third – order partial correlation

Correlation and Regression contd.

❖ Types of correlation analysis contd.

- **Partial correlation contd.:** If the effect of the third variable is controlled for in only one of the two variables being examined, it is called semi-partial (or part) correlation

- **Point – biserial correlation: P**oint – biserial correlation is used to determine the relationship between a continuous variable and a variable that is a discrete dichotomy (e.g. male or female)

- The appropriate test of correlation to use is determined by the nature of the data as shown in Table 9.3.

Table 9.3: Determination of the appropriate test of correlation to use

Type of data	Correlation test
Parametric (i.e., normally distributed data, usually on the interval scale)	Pearson (r)
Non-parametric data (i.e., data that are not normally distributed, or data on nominal and ordinal scales)	Kendall's tau, or Spearman rho

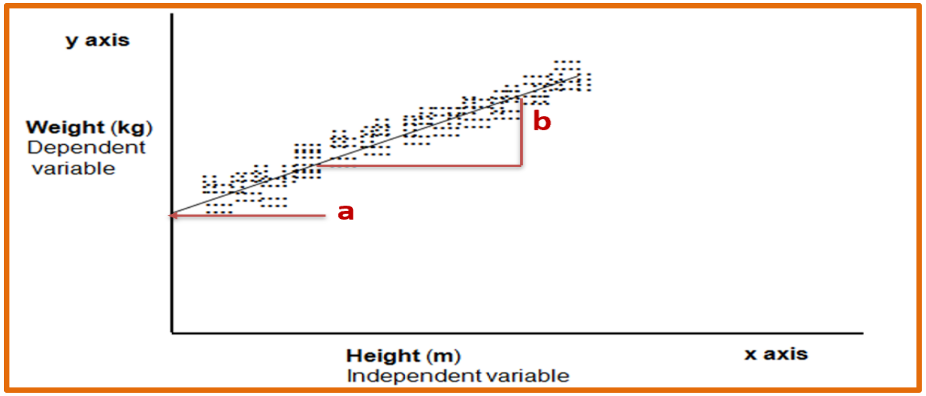
Source: Awosan (2020)

Correlation and Regression contd.

❖ **Predicting the value of the dependent variable for a particular value of the independent variable**

- If there is a strong correlation between the dependent variable (y) and the independent variable (x), the value of the dependent variable can be predicted for a particular value of the independent variable
- Given that the **dependent variable** is **y**, and the **independent variable** is **x**, the straight line relationship between them is defined by: **$y = a + bx$**

Where: **a** is the intercept on the y axis (i.e., the value of y when x = 0), and **b** is the regression coefficient (i.e., the slope)



Other commonly performed non-parametric tests

- One sample Wilcoxon signed-rank test
- Wilcoxon matched pair signed-rank test
- Mann-Whitney U test
- Kruskal-Wallis rank sum H test
- Friedman test

Common Non-parametric Tests

❖ Non-parametric tests and their parametric analogues

- The commonly performed non-parametric tests and their parametric analogues are shown in Table 4.3

Table 4.3 Parametric and non-parametric statistical tests

Nature of groups	Type of variables	Parametric test (Purpose of test)	Non-parametric test (Purpose of test)
One group	Quantitative	Pearson's correlation (Test for relationship)	Kendall's tau, or Spearman rho correlation (Test for relationship)
One group compared with a population	Quantitative	1 sample t-test, or Z test (Compare means)	1 Sample Wilcoxon signed-rank test, or Sign test (Compare medians)
Two independent groups	Quantitative	Independent (or Unpaired) t-test (Compare means)	Mann-Whitney U test, or Wilcoxon rank sum test (Compare medians)
Two related Groups	Quantitative	Paired t-test (Compare means)	Wilcoxon matched pair signed-rank test (Compare medians)
Three or more Independent groups	Quantitative	ANOVA (Compare means)	Kruskal-Wallis rank sum H test (Compare medians)
Three or more repeated measures in one group	Quantitative	Repeated measures ANOVA (Compare means)	Friedman test (Compare medians)

Source: Awosan (2020)

Common Non-parametric Tests contd.

❖ One sample Wilcoxon signed-rank test

- This is a non-parametric analogue to the one sample t-test. It is used when the data is not normally distributed
- The one sample Wilcoxon signed rank test is used to compare the median of a sample with that of a hypothetical population

Common Non-parametric Tests contd.

❖ Wilcoxon matched pair signed-rank test

- This is a non-parametric analogue to the paired sample t-test
- When one cannot assume that the data are from a normal distribution, then the Wilcoxon matched pair signed-rank test should be applied

Common Non-parametric Tests contd.

❖ Mann-Whitney U test

- This is a non–parametric analogue to the two samples t-test
- Just like in the other non-parametric procedures, the ranks of the measurements are used instead of the actual measurements

Common Non-parametric Tests contd.

❖ **Kruskal-Wallis Rank Sum H test**

- This is a non–parametric analogue to the One-way ANOVA test (i.e., ANOVA 1).

Common Non-parametric Tests contd.

❖ Friedman test

- This is a non–parametric analogue to the one-way ANOVA with repeated measures. For example when a group of patients is placed on a particular treatment and measurements are taken at intervals to assess the effect of the treatment but the data obtained are not normally distributed (in which case the one-way ANOVA with repeated measures cannot be used)

- It differs from the Kruskal-Wallis H test in the sense that whereas, the Kruskal-Wallis H test is used to test for **differences in measures obtained in 3 or more independent groups**, the Friedman test is used to test for **differences in 3 or more measures in a single group**

Non-parametric analysis in SPSS

- The commonly performed non-parametric analysis in SPSS shall be covered in Module 6
- The data sets can be accessed through the links below:

Cintarch Dataset_Non-parametric tests in SPSS 1_Correlation and Regression Analysis
https://drive.google.com/file/d/1rcTpXsZojD8rbRRh5ypCKCctmXYA8TCj/view?usp=sharing

Cintarch Dataset_Non-parametric tests in SPSS 2_Partial and Point Biserial Correlation_One Sample Wilcoxon_and Wilcoxon Matched Pair Signed Rank Test
https://drive.google.com/file/d/1bvUszJkWcresp9NtOMWwKl3ztMjXVBmL/view?usp=sharing

To access the videos, please visit: https://cintarch.com/tutorials-videos/

Non-parametric analysis in SPSS contd.

- The remaining data sets can be accessed through the links below:

Cintarch Dataset_Non-parametric tests in SPSS 3_Kruskal-Wallis H Test

https://drive.google.com/file/d/12KxXhYX2uFdvPbUrJM2GHYAzfmNAVk4K/view?usp=sharing

Cintarch Dataset_Non-parametric tests in SPSS 4_Friedman's Test

https://drive.google.com/file/d/11gv3UTxRKTVry4QDt9fsWV0rVqim7Czu/view?usp=sharing

To access the videos, please visit: https://cintarch.com/tutorials-videos/

Further Reading

Awosan KJ (2020). Student Friendly Statistics for Health, Life and Social Sciences. Ikeja, Lagos: Somerest Ventures

Now On Sale

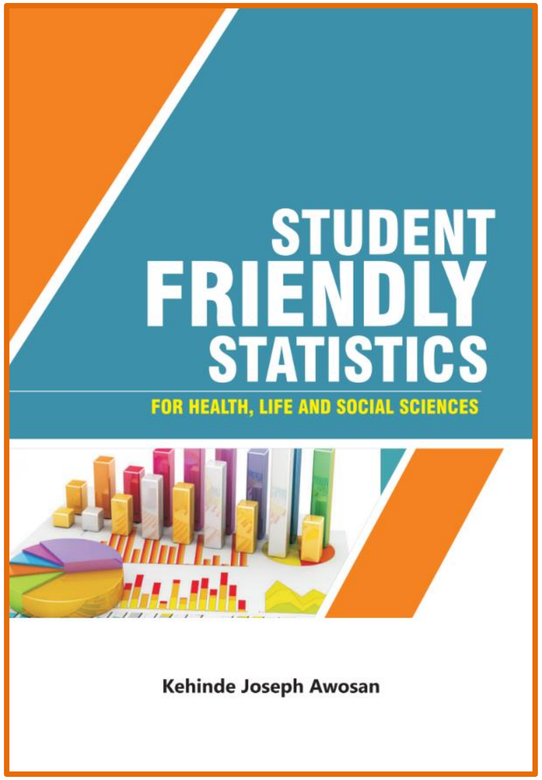
Student Friendly Statistics for Health, Life and Social Sciences

Please click on the link below to access the excerpts from the book:

https://cintarch.org/wp-content/uploads/erf_uploads/2022/02/CB_Excerpts-from-Student-Friendly-Statistics-for-Health-Life-and-Social-Sciences.pdf

To buy a copy or copies of the book online, please click on the link below to access the "ORDER FORM":

https://forms.gle/3J8dWn6Ng6oMKd2e8



C-International Archives' Journals

Please find the List of our Journals at:

https://cintarch.com/journals/

Please find our Recently Published Articles at:

https://cintarch.com/about-us/

Please Submit your Manuscripts at:

https://cintarch.com/submit-manuscripts/