Basic Statistical Analysis: Step by Step using SPSS

Author: Md. Fauzi Ahmad@Mohamad
Email: mohdfauzi@uthm.edu.my

Abstract: The main objective of this book is to provide step by step guidance in statistical analysis. It starts from explanation on type of statistical analysis, steps, results, interpretation and reporting. This book will improve statistical analysis understanding amongst students and researchers. This book contains twelve chapters. In each chapter, examples of the statistical analysis are provided for better understanding using SPSS. Furthermore, the interpretations of statistical result are explained in details for each of the examples. Hopefully this book will provide useful reference towards a better understanding in statistical analysis which shall lead to implementing successful research for students and researchers.

Keywords: SPSS, test, statistics, correlation, regression

A Simple Guide and Reference

Basic Statistical Analysis:

Step by Step using SPSS

"Research is easy if you can do it in a systematic way"

MD. FAUZI AHMAD@MOHAMAD

Penerbit UTHM

A Simple Guide and Reference

Basic Statistical Analysis:

Step by Step using SPSS

"Research is easy if you can do it in a systematic way"

MD. FAUZI AHMAD@MOHAMAD

Penerbit UTHM
2016

Perpustakaan Negara Malaysia Cataloguing—in—Publication Data

Md. Fauzi Ahmad @ Mohamad
Basic Statistical Analysis: Step by Step using SPSS: " Research is easy if you can do it in a systematic way" / MD. FAUZI AHMAD@MOHAMAD.
Includes index
Bibliography: page 189
ISBN: 978-967-0764-51-1
1. Social science--Statistical methods--Computer programs.
2. Statistical-Data processing.
Title.
519.50285536

Published by:
Penerbit UTHM
Universiti Tun Hussein Onn Malaysia
86400 Parit Raja,
Batu Pahat, Johor
No. Tel: 07-453 7051 / 7454
No. Faks: 07-453 6145

Website: http://penerbit.uthm.edu.my
E-mail: pt@uthm.edu.my
http://e-bookstore.uthm.edu.my

Penerbit UTHM is a member of
Majlis Penerbitan Ilmiah Malaysia
(MAPIM)

Printed by:
AWIJAYA ENTERPRISE
No. 15, Jalan Budi Utara
Taman Wawasan Perindustrian
83000 Batu Pahat, Johor

Table of Contents

Acknowledgements

In preparing this book, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. My sincere appreciation extends to Associate Professor Dr. Nor Hazana and Ms. Fatan Adibah Hashim for editing this book. My sincere appreciation also extends to dean, department head, all my colleagues and others who have provided assistance at various occasions. Their views and tips are useful indeed. Unfortunately, it is not possible to list all of them in this limited space. I am grateful to all my family members. Without their understanding and support, I can not complete my book with success. I always pray that Allah can give me the direction for success in life.

This book is one of my small contributions to country and nation based on my experience, observation and research. Hopefully this book will be useful as a guidance to students and researchers for improving their organisation and assist Malaysia to achieve vision 2020. At last, thanks to UTHM for giving me the opportunity to publish this book.

"Research is easy if you can do it in a systematic way"

Preface

At last, I manage to complete writing this book. Thanks to God for giving me, patience and inspiration. From my observation, many students and researchers still do not understand statistical analysis completely and this has been identified as a major problem in doing research or project. The main objective of this book is to provide step by step guidance in statistical analysis. It starts from explanation on type of statistical analysis, steps, results, interpretation and reporting. I believe that this book will improve statistical analysis understanding amongst students and researchers. This book contains twelve chapters. In each chapter, examples of the statistical analysis are provided for better understanding using SPSS. Furthermore, the interpretations of statistical result are explained in details for each of the examples.

Hopefully this book will provide useful reference towards a better understanding in statistical analysis which shall lead to implementing successful research for students and researchers. At last, thanks to UTHM for giving me the opportunity to publish this book , particularly to share my experience and knowledge by putting it in such a wonderful manuscript.

"Research is easy if you can do it in a systematic way"

Dr. Md.Fauzi Bin Ahmad @ Mohamad
Production and Operation Department,
Faculty of Technology Management and Business,
Universiti Tun Hussein Onn, Malaysia.

List Of Abbreviations

ABBREVIATION	
SPSS	Statistical Package for the Social Sciences
KMO	Kaiser-Meyer-Olkin
EFA	Exploratory Factor Analysis
ANOVA	Analysis of Variance

CHAPTER 1

Statistical Analysis

1.1 Statistics

Statistical Package for the Social Sciences (SPSS) has been used to tabulate and analyse responses from survey in the forms of frequency, percentages, cumulative frequency, and cumulative percentages to describe the characteristics of the sample with respect to the demographic (descriptive analysis). In order to test hypotheses, inferential analysis such as independent sample T-test, one-way ANOVA and correlation could be performed using the SPSS software. This book will cover some topics of inferential statistics and types of sampling. The details of inferential statistics procedure using SPSS are explained step by step with examples.

Statistics is a study of the collection, organization, analysis, interpretation and presentation of data from the samples. Basically, **descriptive and inferential statistics** are used to analyse data and draw conclusions to the population. Descriptive statistics (mean, standard deviation) that are applied to populations are called **parameters** as they represent the whole population. **Population** is group of data that includes all the data in the population. However, many researchers normally do not have access to the whole population because of limited resources such as time and financial. Descriptive statistics of samples, such as the

CHAPTER 2
Preparation of Data Files

2.1 Create and Open a File

This section explains how to create a new file in SPSS.

Example:
Twenty performance data have been collected based on 7-likert scale from Non-Japanese and Japanese companies. You need to create new file in SPSS as shown in Table 2.1:

Table 2.1: Performance

Respondent	Ownership	Performance
1	Non-Japanese	5.00
2	Non-Japanese	5.00
3	Non-Japanese	5.00
4	Non-Japanese	5.00
5	Non-Japanese	6.00
6	Non-Japanese	5.00
7	Non-Japanese	6.00
8	Non-Japanese	6.00
9	Non-Japanese	5.00

CHAPTER 3
Normality Test

3.1 Concept of Normality

Normality test is used to determine whether parametric tests could be used or not. Normal data refer to data that are drawn from a **normally distributed population** (Sekaran & Bougie, 2010). This distribution is inarguably the most important and the most frequently used distribution in both the theory and application of statistics. Normal data has the familiar bell-shaped form as shown in Figure 3.1. **Kolmogorov-Smirnov Test** and **Shapiro-Wilk** have been used for testing the normality of data. If the sample size is larger than 50, Kolmogorov-Smirnov test should be used. If the sample size were 50 or less, Shapiro-Wilk would yield better result. If p value is more than 0.05, it can be concluded that the data is normal and parametric test could be used.

CHAPTER 4
Reliability and Validity Test

4.1 Reliability Test

Reliability is a test of how consistently a measuring instrument measures whatever concept it is measuring, whereas validity is a test of how well an instrument that is developed measures the particular concept it is intended to measure. as shown in Figure 4.1 (Sekaran & Bougie, 2010; Hair, 2010). The **Cronbach's alpha coefficient** is used to assess the inter-item consistency for measurement items in a construct, in which the alpha values should be above 0.6, as suggested by Nunnally & Bernstein (1994). Specifically, 0.6 is satisfying for a relatively new measurement instrument while **0.7 is sufficient** (Nunally, 1978). The alpha value of above 0.70 indicates that the scales are internally consistent (Chi et al., 2011; Hair, 2010). A composite reliability (CR) of 0.70 or greater is acceptable (Fornell and Larcker, 1981), which shows that the measurements are reliable.

CHAPTER 5
Descriptive Statistics

5.1 Concept of Descriptive Statistics

The basic analysis in a research is descriptive statistics. **Descriptive statistics are used to describe the sample characteristics in forms of percentage, frequency, averages, and standard deviation**. Descriptive statistics allow simpler interpretation for describing the data. Basically, there are **three types** of statistic that are used to describe data:

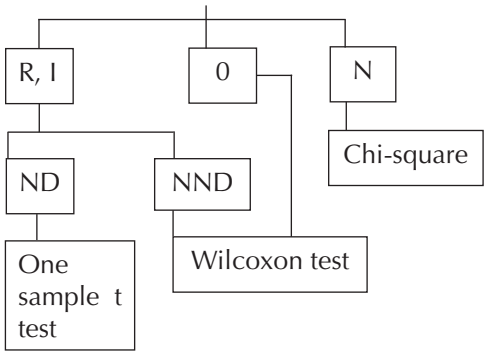
- **Measures of distribution**: It is describing a summary of respondents' demographic. For instance, researcher describes the distribution of type of ownership, listing the number or percentage of ownership. Or, researcher describes gender by listing the number or percentage of males and females.

- **Measures of central tendency**: It is describing the central position of a group of data using mode, median, and mean. Mode and median are used to describe categorical data while mean is used to describe continous data.

- **Measures of spread**: It is describing how spread out the data using spread using standard deviation, variance and range. Standard deviation and variance are used to for coninous data while range is used for categorical data.

CHAPTER 6

Comparing a Group to Hypothesised Mean



R,I = Ratio and Interval data O= Ordinal data N= Nominal data
N= Normal distribution NND= Non normal distribution

(Source: Jaykaran (2010))

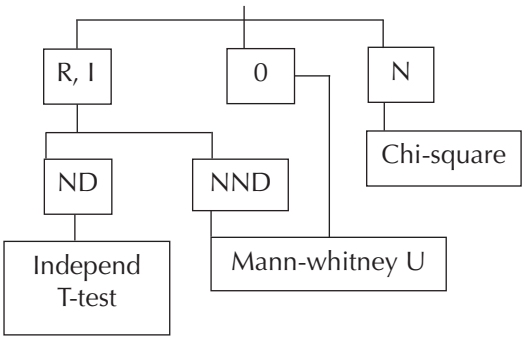
6.1 One-Sample T-test (Normal Data)

One-Sample T-test is a parametric test. The objective of this section is to compare the sample means of a data set to hypothesised mean as shown in Figure 6.1 .

One group: Interval or ratio data (Normal Data)

CHAPTER 7

Comparing Two Groups (Independent Group)



R,I = Ratio and Interval data O= Ordinal data N= Nominal data
N= Normal distribution NND= Non normal distribution

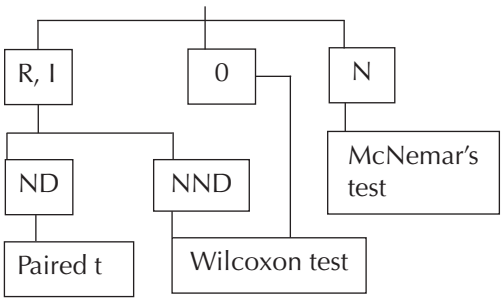
(Source: Jaykaran (2010))

7.1 T-test (Normal Data)

T-test is a parametric test. The main objective of this section is to compare the means between **two scores** under different group of **the independent variable** when you have data from **two groups of respondents as show Figure 7.1 and Figure 7.2**.

CHAPTER 8

Comparing Two Groups (Related Samples)



R,I = Ratio and Interval data O= Ordinal data N= Nominal data
N= Normal distribution NND= Non normal distribution

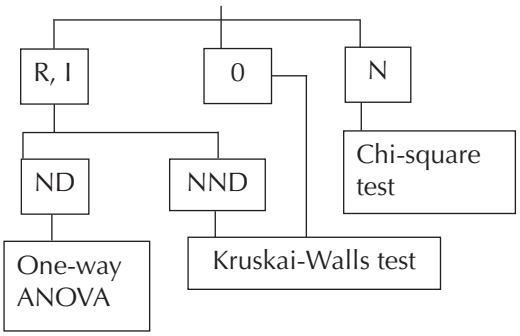
(Source: Jaykaran (2010))

8.1 Paired T-test (Normal Data)

Paired T-test is a parametric test. The main objective is to compare the means between **two score** under different levels of the independent variable when you have data from only **one group of respondents**.

CHAPTER 9

Comparing Tree Groups (Independent Group)



R,I = Ratio and Interval data O= Ordinal data N= Nominal data
N= Normal distribution NND= Non normal distribution

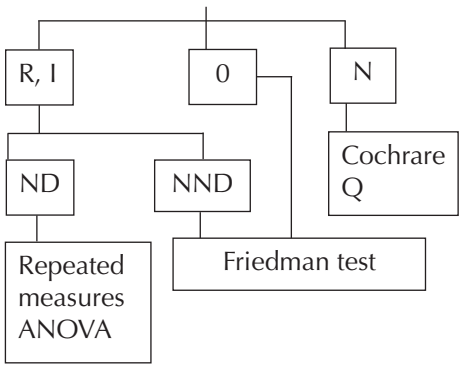
(Source: Jaykaran (2010))

9.1 One-way Anova (Normal Data)

One-way Anova is a parametric test. The main objective is to compare the means amongst more than two (**three and above) scores** under different groups of **independent variable** when you have data from **three groups and above of respondents** as shown in Figure 9.1, 9.2 and 9.3. ANOVA stands for analysis of variance.

CHAPTER 10

Comparing Tree Groups (Related Samples)



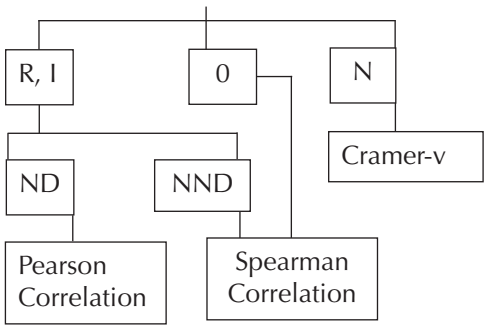
R,I = Ratio and Interval data O= Ordinal data N= Nominal data
N= Normal distribution NND= Non normal distribution

(Source: Jaykaran (2010))

10.1 Repeat Anova (Normal Data)

Repeat Anova is a parametric test. The main objective is to compare the means between **three** and above scores under different levels of the independent variable when you have data from **only one group of respondents** as shown in Figure 10.1.

CHAPTER 11
Relationship: Correlation



R,I = Ratio and Interval data O= Ordinal data N= Nominal data
N= Normal distribution NND= Non normal distribution

(Source: Jaykaran (2010))

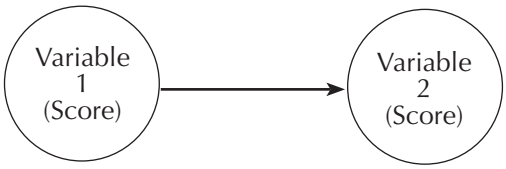
11.1 Pearson correlation (Normal Data)

Pearson correlation is a parametric test. The main objective is to describe the relationship between **two variables** when you have data from the **same respondents**.

CHAPTER 12
Causal and Prediction: Regression

12.1 Linear Regression (Normal Data)

Linear regression is a parametric test. The main objective is to examine **the impact or effect** of independent variables on dependent variable. Regression analysis is used to **infer causal relationships** between the independent and dependent variables.



Independent variable and dependent variables are **interval or ratio data**.

Assumption of Testing

Each statistical test has their assumptions that must be fulfilled before performing the analysis.

1. There should be **no significant outliers.**
2. Independent variable and dependent variable should be **interval or ratio data.**

Bibliography

Awang, Z. (2012). *A Handbook on SEM* (5th ed.). Kota Bharu: UITM.

Brown, R. L. (1994). Efficacy of the indirect approach for estimating structural equation models with missing data: A comparison of five methods. *Structural Equation Modelling*, *1*(4), 287–316.

Byrne, B. . (2010). *Structural Equation Modeling with AMOS*. *Structural Equation Modeling* (2nd ed.). New York: Routledge Francis & Taylor.

Chi, A., Bahjat, A., & Matsui, Y. (2011). Quality management practices and competitive performance : Empirical evidence from Japanese manufacturing companies. *Intern. Journal of Production Economics*, *133*(2), 518–529. doi:10.1016/j.ijpe.2011.01.024

Chinna, K. (2009). *Structural Equation Modeling Using AMOS* (p. 20). Kuala Lumpur: Malaysia: Lecture Note for SPSS User' Group.

Demirbag, M., Tatoglu, E., Tekinkus, M., & Zaim, S. (2006). An analysis of the relationship between TQM implementation and organizational performance: Evidence from Turkish SMEs. *Journal of Manufacturing Technology Management*, *17*(6), 829–847. doi:10.1108/17410380610678828

Field, A. (2009). *Discovering Statistics Using SPSS*. *Statistics* (2nd ed.). London: Sage Publication.

Galbreath, J. (2005). Which resources matter the most to firm success? An exploratory study of resource-based theory. *Technovation*, *25*(9), 979–987. doi:10.1016/j.technovation.2004.02.008

Graham, J. W., Hofer, S. M., Donaldson, S. L., MacKinnon, D. P., & Schafer, J. L. (1997). *The Science of Prevention: Methodological advances from alcohol and substance abuse research* (pp. 325–366). Washington, DC: American Psychological Association.

Gunday, G., Ulusoy, G., Kilic, K., & Alpkan, L. (2011). Effects of innovation types on firm performance. *International Journal of Production Economics*, *133*(2), 662–676. doi:10.1016/j.ijpe.2011.05.014

Hair, J. F. (2010). Multivariate Data Analysis. New York: Pearson Prentice Hall.

Jaykaran (2010). How to select appropriate statistical test?. Journal of pharmaceutical Negative Results, Oct (2), 61-63.

Kline, R. . (2011). *Principles and practices of structural equation modeling* (3rd ed.). New York: The Guilford Press.

Krejcie, R. V., & Morgan, D. W. (1970). *Determining sample size for research activities*. Educ Psychol Meas.

Nunally, J. C. (1978). *Psychometric Theory* (Second ed.). New York: McGRAW-HILL.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd editio.). New York: McGRAW-HILL.

Rohani, J. M., Yusof, S. M., & Mohamad, I. (2009). The relationship between statistical process control critical success factors and performance: A structural equation modeling approach. *2009 IEEE International Conference on Industrial Engineering and Engineering Management*, 1352–1356. doi:10.1109/IEEM.2009.5373033

Sekaran, U. (2003). *Research methods for business: a skill building approach*. Hoboken, NJ: John Wiley & Sons.

Sekaran, U., & Bougie, R. (2010). *Research methods for business: a skill building approach*. UK: Wiley.