COMPARISON OF SOME PARAMETRIC AND NON-PARAMETRIC STATISTICAL METHODS

BY

IBRAHIM FAROUK ABDUL

DEPARTMENT OF MATHEMATICS
AHMADU BELLO UNIVERSITY, ZARIA
NIGERIA

DECEMBER, 2008

COMPARISON OF SOME PARAMETRIC AND NON-PARAMETRIC STATISTICAL METHODS

BY

IBRAHIM FAROUK ABDUL
M.Sc/SCI/44404

A THESIS SUBMITTED TO THE POSTGRADUATE SCHOOL, AHMADU BELLO UNIVERSITY, ZARIA NIGERIA

IN PARTIAL FULFILLMENT FOR THE AWARD OF THE MASTERS OF SCIENCE IN STATISTICS

DEPARTMENT OF MATHEMATICS
AHMADU BELLO UNIVERSITY, ZARIA
NIGERIA

DECEMBER, 2008

DECLARATION

I declare that the work in this thesis entitled 'Comparison of some Parametric and Non-parametric' has been performed by me in the Department of Mathematics under the supervision of Dr. Usman Gulumbe and Prof. G.U. Garba.

The information derived from the literature has been duly acknowledged in the text and a list of references provided. No part of this thesis was previously presented for another Degree or Diploma at any University.

Ibrahim Farouk. Abdul.		December 1, 2008
Name of student	Signature	Date

CERTIFICATION

This thesis entitled **COMPARISON OF SOME PARAMETRIC AND NON-PARAMETRIC STATISTICAL METHODS** by IBRAHIM, Farouk. Abdul. Meets the regulations governing the award of Degree of Master of Science of Ahmadu Bello University, Zaria, and is approved for its contribution to knowledge and literary presentation.

_____ Date:_____
Dr. S.U Gulumbe
Major Supervisor

_____ Date:_____
Prof. G.U Garba
Minor Supervisor

_____ Date:_____
Prof. R. A. Ipinyomi
External Examiner

_____ Date:_____
Prof. G.U Garba
Head of Department

_____ Date:_____
Prof. S.A. Nkom
Dean of Postgraduate School

DEDICATION

This research work is dedicated to my beloved father, the late Mal. Abdul Enebe, my mother, Madam Aishat Abdul, my aunt, Madam Hadizat Apatewa, my brothers: yakubu Enebe the late Abubakar, Isah, Musa and my sister Mariyetu Abdul.

All the afore-mentioned people and others whose names for space's sake I could not mention, may ALLAH reward them abundantly (Amin).

ACKNOWLEDGEMENT

Praise be to ALLAH, the cherisher and sustainer of the universe. He alone has complete moral excellence for His guidance from the dawn to the dusk of this research work.

I humbly acknowledge the efforts of my noble wife, whose balm of spiritual, financial and moral support soothes my churning emotion, without her, I am a living corpse. It is better to be unborn than to be untaught. She stood firmly by me even when in difficulties, gave me much needed encouragement, throughout this programme. My dear, you have tried in your domains.

I owe my profound gratitude to Mrs. Mary Gabadi who contributed immensely to my academic progress. My sincere appreciation goes to my supervisor Dr. S.U Gulumbe for his guidance and encouragement throughout the period of this research exercise.

I do sincerely thank the Director of Universal Basic Education Board, Alh. Halilu H. Pai and all my intimate honourable friends: Alh. Yahaya Osada, Mr. Matthew Aminu, and Suleimen Anuhi.

I am also indebted to my beloved classmates notably Shettima, Bukar and Mrs. Adewuyi.

I also seize this opportunity to thank my well wishers: Dr. Abdulraham Usman, Mrs T. O. Mbonu, Madam Maryam Gomina, Mrs. Ebere Obina, Danladi Maji.

Finally, my heartfelt gratitude goes to the Head of Department, Prof. G. U. Garba and the rest of staff of the Department who assisted me in no small measures.

ABSTRACT

Multivariate parametric and nonparametric techniques are two broad statistical methods for testing independence among continuous random variables. In this thesis, multivariate parametric and nonparametric techniques were performed to tests for independence using multivariate data collected from Human Development Report 2005. Three parametric tests, Hotelling's T^2, Wilks Lambda and Canonical correlation analysis were performed. Wilcoxon, Friedman, Kendall's and Cochran are the nonparametric analysis performed. The results show that parametric and nonparametric tests of hypothesis yield the same results.

TABLE OF CONTENTS

Content		Page No.
TITLE PAGE		i
DECLARATION		iii
CERTIFICATION		iv
DEDICATION		v
ACKNOWLEDGEMENT		vi
ABSTRACT		viii
TABLE OF CONTENTS		ix
LIST OF TABLES		x
CHAPTER ONE	**INTRODUCTION**	
1.01	INTRODUCTION	1
1.02	Multivariate Parametric Test	2
1.03	Multivariate Nonparametric Tests	4
1.04	Advantages of Nonparametric Test.	6
1.05	Disadvantages of Nonparametric Tests.	6
1.06	Purpose of the Study.	7
1.07	Limitation of the Study	7
1.08	Aims and Objectives	**Error! Bookmark not defined.**
CHAPTER TWO	**LITERATURE REVIEW**	
2.01	INTRODUCTION	9
2.02	Multinormality Theory	9
2.03	Parametric versus non-parametric statistics in the analysis of randomized trials with non-normally distributed data.	10
2.04	Background of the study	**Error! Bookmark not defined.**
2.05	A review of method of comparison using Mann-Whitney and ANCOVA	15
2.06	Their Results	19
2.07	Their-Observations	19
2.08	Testing for Independence in High Dimension	20
2.09	Testing Independence of Multivariate data using Boot strap	22
CHAPTER THREE	**MATERIALS USED FOR THE STUDY**	
3.01	Data Set	24
3.02	Software Used.	26
3.03	Methodology	27

3.04 Wilks' lambda Test 27
3.05 Canonical Correlation Analysis 29
3.06 Wilcoxon signed-ranks test 29
3.07 Friedman two way analysis of variance test. 30
3.08 Kendall's W Test. 32
3.09 Cochran Q Test 33

CHAPTER FOUR ANALYSIS
4.01 Hoteling's-T^2-Test 35
4.02 Wilks' Lasmbda test 36
4.03 Canonical correlation analysis 37
4.04 The Wilcoxon signed-ranks sum test 39
4.05 The Friedman two way analysis of variance test 41
4.06 Kendall's W test 42
4.07 Cochran Q test 43

CHAPTER FIVE CONCLUSION AND RECOMMENDATION
5.01 SUM MARY AND CONCLUSION 44
5.02 Summary 44
5.03 Conclusion 44
5.04 Recommendation 45

References 46
APPENDIX-A 49
Appendix-B 50

LIST OF TABLES

TABLE TITLE PAGE

1: Hotelling's T^2 test of independence 37

2: Wilks Lambda test of independence 38

3: Canonical correlation analysis of independence 39

4: Wilcoxon signed-rank test of independence 41

5: Friedman test of independence 43

6: Kendall test of independence 44

7: Cochran test of independence 45

CHAPTER ONE

1.01 INTRODUCTION

Testing hypothesis is very important in any statistical analysis. However, many considerations enter into the choice of statistical inference. Statistical testing often refers to the concepts of parameter (mean, standard deviation, correlation) and reference population (multinomial distribution). However, the multinormal distribution and others play a key role in using parametric tests. If the data are not normal, however, it does not make sense to compute statistics such as the mean, or standard deviation. To make a good choice we must consider the manner in which the sample of scores were drown, the nature of the measurement or scaling which was employed in the definitions of the variables involved (i.e. in the scores). All the above assumptions are useful in determining which statistical test is most appropriate for analyzing one set of data.

In practice, we often encounter more complex situations when measures of associations for more variables are required. In that situations, the multivariate test procedures are very useful for us. However, in some cases measurements may represent attributes of psychological characteristics while others represent attributes of physical characteristics. We then require a test of independence between pairs of vectors, where the vectors have different measurement scales and dimensions. For instance if $x_i^T = (x_i\ x_j)$ i =

1, 2, ... n, j = 1, 2 . . . . . n for n denotes a sample size of vector pairs, where x_1^2 and x_1^2 were continuous vector of dimension p and q respectively. The test is Ho x_i^1 and x_j^2 are independent against the alternative. Hı x_i^1 and x_i^2 are not independent.

Testing for independence between continuous random variables is an old approach but important and has been the subject of much interest. Many thorough investigations on their efficiency have been given by many authors. Gieser and Randles (1997) Takinen, kanlainen and Oja (2003a) taskinen et al (2003b) Genest and Verret (2003) Chernoff Savage and Hodges-Lehrnann (1956) Hallin and Puri (1958) Hallin and Paindanceine (2002a, b and 2003) on multivariate inferences using both or either parametric or nonparametric methods.

All these reviewed seems to accord much interest on an adequate generalization of rank and signs. The reason for this interest is the confusion people make since ranks and sign are independent and intimately related and exchangeable. They (ranks and signs) were inherently confined to the analysis of independent observation.

1.02 Multivariate Parametric Test

A parametric statistical test specifies conditions about the parameters of the population from which the sample was drawn. This means that the

parametric model considers that the population distribution should have a particular form (e.g. a normal distribution) and also involves hypothesis about population parameters. The meaningfulness of the results of a parametric test depends on the validity of these assumptions.

However, the most powerful tests are those having most extensive assumptions. The parametric tests, for example the T^2 test, for it to be applied, the underlying conditions should be satisfied.

- The observations must be drawn from normally distributed populations.
- The observation must be independent
- The mean effects must be additive
- The populations must have the same variance-covariance matrices and
- The variables involved must have been measured in at least an interval scale.

If these conditions are met in the data under analysis, one can choose T^2 or Wilks lambda test. In addition to the above conditions, data measured in interval (data identify subjects by non-relative rank in a scale: that can be represented along a number line, there is no Absolute – Zero and the number are non-relative e.g. IQ achievement, Aptitude, etc), or ratio scales

(data identified subjects by quantification, number that can represent equal units, and the range from an absolute zero point e.g. time, weight, distance, temperature) may be analyzed by parametric methods.

The meaningfulness of the results of a parametric test depends on the validity of these assumptions. Therefore, when those assumptions are valid, these tests (parametric) are most likely of all tests to reject H_0 when it is false. Parametric tests are also generally more flexible, capable for example of dealing with a mixture of categorical and continuous variable, which is not possible with nonparametric tests. Bralley JV (1968)

1.03 Multivariate Nonparametric Tests

Nonparametric tests are distribution free, that is to say they do not assume that the samples were drawn from a population with a specific distribution and because of this, nonparametric statistics are useful not only when variables are categorical, but also when parametric data do not conform to the multinomal distribution.

Nonparametric method is defined as a statistical method which can be used on data that is:

a. ordinal scale of measurement

b. a nominal scale of measurement

c. an interval or ratio scale of measurement where the distribution function of the random variable producing the data are unspecified

(or specified except for an infinite number of unknown parametric). Conover (1999).

Thus, a nonparametric statistical tests (otherwise called distribution-free test) is a test whose model does not specify conditions about the parameters of the population from which the sample was drawn, hence, it is not meaningful to look for normalization or distribution on an interval scale. Nevertheless, certain assumptions are associated with most nonparametric statistical tests, that is, the observations are independent and the variables under consideration have underlying continuity.

These assumptions are not as strong as those associated with parametric tests. Moreover, nonparametric tests do not require strong measurement as required for the parametric tests. Most nonparametric tests apply to data in an ordinal scale (data identified subjects by position in a series range by height, order finished in a race, academic standing in class, measurement of average, mean) and some apply also in a nominal scale (data which identify subjects by category or name, gender, ethnicity, class, hair colour). Nonparametric data types are called nominal or ordinal. However, if one or more of the assumptions (of parametric) have been violated, then it is capital important to transform data into a format that is compatible with the appropriate nonparametric test. This is because the parametric tests generally provide a more powerful test of an alternative hypothesis than

nonparametric counterparts. If one or more of the underlying parametric test assumption is or (are) violated, the power advantage may be negated.

1.04 Advantages of Nonparametric Test.

The nonparametric tests have the following advantages among others. Dallal GE (1988).

a. Nonparametric tests are stable for treating sample made up of observations from several different populations.

b. Nonparametric tests are distribution free i.e. they do not require any assumption to be made about population following normal or any other distribution.

c. Generally they are simple to understand and easy to apply when the sample sizes are small.

d. Many nonparametric methods make it possible to work with very small samples. This is particularly helpful to the researcher collecting pilot study data or to medical researcher working with a rare disease.

1.05 Disadvantages of Nonparametric Tests.

Nonparametric tests have some disadvantages, and the disadvantages are as follows:

i. The major disadvantages of nonparametric techniques is contained in its name. Because the procedures are nonparametric, there are no parameters to describe and it becomes more difficult to make

quantitative statements about the actual difference between populations.

ii. Nonparametric procedures throw away information. The sign test for example, uses only the signs of the observations. Ranks preserve information about the order of the data but discard, the actual values. Because information is discarded, nonparametric procedures can never be as powerful as their parametric counterparts when parametric tests can be used.

1.06 Purpose of the Study.

i. To find out if there exist any relationship between indexes of human development report using both parametric and nonparametric tests.

ii. To equally test for independence using both multivariate parametric method and nonparametric method to see if it can produce the same result.

iii. To find out similarities/difference between the two statistical methods based on the result of the analysis used in this thesis.

iv. To analyze statistically, multivariate data by using parametric and nonparametric tests for independence

1.07 Limitation of the Study

The study contained in this thesis extends from these existing literatures providing both parametric and nonparametric multivariate tests of independence. Asymptotic tests based on person's spearman's Kendall's Coefficient, Wilks test, signed rank test, Wilcoxon rank sum test, kruskal-wallis test, serial rank tests, locally most powerful tests of independence for

specific copula model, testing independence for multivariate data using Bootstrap, testing for independence in high dimensions etc. Thus, while the emphasis is on test of hypothesis, it is useful to consider their asymptotic relative efficiency in order to demonstrate the excellent performance of these methods.

CHAPTER TWO

LITERATURE REVIEW

2.01 INTRODUCTION

Multivariate analysis deals with observation on more than one variable where there are some inherent interdependence between variables. There are wide variety of multivariate techniques. The choice of the most appropriate method depends on the type of data, the problem, and the sort of objectives which are envisaged for analysis. The review in this chapter extends from the existing literature by providing both multivariate parametric and nonparametric test for independence.

2.02 Multinormality Theory

Multivariate analysis lays too much interest on the assumption that all random vectors come from multivariate normal distribution. By definition, the probability density function of a normal variable with mean μ and variance σ^2 is given by

$$f(x) = (2\pi\sigma^2) \exp - \tfrac{1}{2} (x-\mu)(\sigma^2)^{-1}(x-\mu)$$

Then the extension to p-variate is

$$f(x) = (2\pi)^{-\frac{n}{2}} \left|\sum\right|^{-\frac{1}{2}} \exp - \frac{1}{2}(x-\mu)^{1} \sum^{-1} (x-\mu)$$

The reasons for its (normal distribution) preference in the multivariate case are among others. (Hollander M and Wolfe DA, 1973)

(a) The multivariate normal distribution is entirely defined by its first and second moments.

(b) The multivariate distribution is an easy generalization of its univariate counterpart, and the multivariate analysis runs almost parallel to the corresponding analysis based on univariate normality.

(c) Linear functions of a multinormal vector are themselves univariate normal.

(d) In the case of normal variables, zero correlation implies independence and pair wise independence implies total independence.

(e) Equiprobability contours of the multivariate distribution are simple ellipse, which by a suitable change of coordinates can be made into circle.

(f) When the original data is not multinormal, one can often appeal to central limit theorems which prove that certain functions such as the sample mean are normal for large samples.

2.03 Parametric versus non-parametric statistics in the analysis of randomized trials with non-normally distributed data.

The following ideas are the contributions and conclusions gotten from **Andrew J Vickers,** Wolfowitz, (1942) Siegel & Castellan, (1988) & Dr Matthew Ellis

(2002). It has generally been argued that parametric statistics should not be applied to data with non-normal distributions. Empirical research has demonstrated that Mann-Whitney (Wilcoxon) generally has greater power than the *t*-test unless data are sampled from the normal. In the case of randomized trials, we are typically interested in how an endpoint, such as blood pressure or pain, changes following treatment. Such trials should be analyzed using ANCOVA, rather than *t*-test. The objectives of this study were:

a) to compare the relative power of Mann-Whitney and ANCOVA;

b) to determine whether ANCOVA provides an unbiased estimate for the difference between groups;

c) to investigate the distribution of change scores between repeat assessments of a non-normally distributed variable.

Introductory statistics textbooks e.g A statistical approach to testing mutual independence of ICA Recovered sources, typically advise against the use of parametric methods, such as the *t*-test, for the analysis of randomized trials unless data is approximated to a normal distribution. Altman (1948) for example, states that "parametric methods require the observations within each group to have an approximately Normal distribution, if the raw data do not satisfy these conditions a non-parametric method should be used"

a. In some cases, central limit theorem is invoked such that parametric methods are said to be applicable if sample size is suitably large: "for

reasonably large samples (say, 30 or more observations in each sample) the *t*-test may be computed on almost any set of continuous data".

b. The rationale for recommending non-parametric over parametric methods, unless certain conditions are met, is rarely made explicit. But techniques for statistical inference from randomized trials can only fail in one of two ways: they can inappropriately reject the null hypothesis of no difference between groups (false positive or Type I error) or inappropriately fail to reject the null (false negative or Type II error). Hence any recommendation to favor one technique over another must be based on the relative rates of these two errors.

Empirical statistical research has clearly demonstrated that the *t*-test does not inflate Type I (false positive) error, Andrew JV Vickers (1942). In a typical study, Heeren et al examined the properties of the *t*-test to analyze small two-group trials where data are ordinal, such as from a five point scale, and thus non-normal.

c. They found that where there was truly no difference between groups, Siegel and Castellan (1988), the *t*-test would reject the null hypothesis close to 5% of the time. Thus, concern over the relative advantages of parametric and non-parametric methods has focussed on Type II error

d. Typically, researchers have created a large number of data sets, in which observations were created from a distribution incorporating a difference between groups. Each data set is then analyzed by both

parametric and non-parametric methods in order to calculate the proportion of times the null hypothesis is rejected (that is, the power). Mathew Ellis (2002).

Where data are sampled from a normal distribution, the *t*-test has very slightly higher power than Mann-Whitney, (Wilconxon) the non-parametric alternative. However, when data are sampled from any one of a variety of non-normal distributions, Mann-Whitney is superior, often by a large amount. Bridge and Sawilowsky, for example, concluded that" "the *t*-test was more powerful only under a distribution that was relatively symmetric, although the magnitude of the differences was trivial. In contrast, the [Mann-Whitney] held huge power advantages for data sets which presented skewness". Many workers have linked results showing the superiority of non-parametric methods for non-normal distributions to claims that data rarely follow a normal distribution.

It is arguable, however, that these prior investigations are flawed. The *t*-test and Mann-Whitney are used for continuous variables such as blood pressure, depression, weight or pain. Most commonly, we are interested in seeing how these variables *change* following an intervention. This reflects clinical practice where the patient presents with a problem and asks the doctor to help improve it. In a typical study, a patient with hypertension, obesity or chronic headache is randomized to drug or placebo to see whether the drug is effective for reducing blood pressure, weight or pain.

The researchers might report that, say, blood pressure fell by 5 mm in the placebo group but by 14 mm in the drug group. Indeed, trials in which we are interested only in post-treatment scores where change is not of interest, are rather rare, being primarily confined to iatrogenic symptoms such as post-operative pain or chemotherapy vomiting.

There are two implications for methodological research on the relative value of parametric and non-parametric techniques. First, we should worry about the distribution of change scores. It seems likely that change from baseline would approximate more closely to a normal distribution than the post-treatment score. This is because change scores are a linear combination and the Central Limit Theorem therefore applies. As a simple example, imagine that baseline and post-treatment score were represented by a single throw of a die. The post-treatment score has a flat (uniform) distribution, with each possible value having an equal probability. The change score has a more normal distribution: there is a peak in the middle at zero – the chance of a zero change score is the same as the chance of throwing the same number twice, that is 1 in 6 – with more rare events at the extremes – there is only a 1 in 18 chance of increasing or decreasing score by 5.

Moreover, where an endpoint is measured at baseline and again at follow-up, the *t*-test is not the recommended parametric method. Analysis of covariance (ANCOVA), where baseline score is added as a covariate in a

linear regression, has been shown to be more powerful than the *t*-test. It has several additional advantages:

(i) it adjusts for any chance baseline imbalances;

(ii) it can be extended to incorporate randomization strata as co-variates, which has been shown to increase power;

(iii) it can also be extended to incorporate time effects where measures are repeated.

In this paper, they used more rational comparison between parametric and non-parametric methods: ANCOVA and Mann-Whitney. Such a comparison does not appear to have been reported previously. He aimed to compare relative power of the two methods under a variety of distributions. As a secondary objective, He aimed to determine whether ANCOVA provided an unbiased estimate for the difference between groups where data did not follow a normal distribution. A third, overarching aim was to investigate the distribution of change scores between repeat assessments of a non-normally distributed variable.

2.04 A review of method of comparison using Mann-Whitney and ANCOVA

Polynomials were developed to simulate five archetypal non-normal distributions for baseline and post-treatment scores in a randomized trial. Simulation studies compared the power of Mann-Whitney and ANCOVA

for analyzing each distribution, varying sample size, correlation and type of treatment effect (ratio or shift) Andrew et al, (2002).

The starting point for this study was to obtain archetypal data sets for analysis. Andrew et al, (2002) follow Bridge in choosing empirical rather than theoretical distributions. They examined the distribution of a large number of empirical data sets and cross-referenced these with those described by Micceri, who systematically obtained 440 data sets from the psychological and educational domains. The most common distribution appeared one with moderate positive skew. As an exemplar, they used a headache severity index from a large (n = 401) randomized trial of headache prophylaxis. This distribution was also used with scores reversed, to create a distribution with moderate negative skew. A second pain data set, this time from a trial on athletes with shoulder pain, provides an example of a more uniform distribution. Data on Ki67, an antigen that is a marker for cell proliferation, were obtained from a randomized comparison of two normal treatments for breast cancer. The distribution for Ki67 is comparable to Micceri's “extreme asymmetry distribution”. For extreme negative skew, they used data from the physical functioning scale of the SF36, again taken from the headache trial. As a comparison group, data were also drawn from a normal distribution with a mean of 5 and a standard deviation of 1.

For each of the distributions, they created a polynomial that converted normal data to a distribution with an approximately similar shape. For example, the distribution with moderate positive skew in was simulated by sampling x from the normal and creating a new variable equal to $14.8+16.5x+7.5x^2-1.15x^3$, rounded, like the original scale, to the nearest 0.25. The simulation distributions were compared to the empirical distributions by visual inspection and comparison of the standard deviation, skewness and kurtosis.

To run the simulations, a bivariate normal (mean 0, standard deviation 1) with a specified correlation was created for a trial of a given sample size equally divided in two groups. The polynomial was applied and a treatment effect introduced. The treatment effect was one of two forms: a shift, for example, scores in the treatment group were reduced by two points; and a ratio, for example, treatment group scores were reduced by 20%. Results were then analyzed by Mann-Whitney and ANCOVA, with *p*-values obtained by asymptotic approximation for the Mann-Whitney test. In some simulations, *t*-tests and ANCOVA of log-transformed data were applied. The *t*-test and Mann-Whitney used the follow-up score if correlation was less than 0.5 and the change score otherwise. This maximizes the power of these tests and might be seen as favoring unadjusted tests on the grounds that the correlation between baseline and follow-up scores is not known when the protocol for statistical analysis is written. Note that the correlation

cited in the results is the correlation between baseline and follow-up in the control group. Some previous workers have used the overall correlation using both groups when investigating the properties of ANCOVA. The difference between these two values was small in the context of our simulations, for example, a correlation of 0.5 in the control group was equivalent to a correlation of 0.476 for both groups combined.

Simulations were repeated 1000 times for each combination of sample size (10, 20, 30, 40, 60, 100, 200, 400, 800) and correlation (0.1, 0.2, 0.3 ... 0.9) using Stata 8.2 (Stata Corp., College Station, Texas). The exception was extreme asymmetry data for the Ki67 biomarker. The baseline and post-treatment distributions had quite different shapes and different polynomials were used to model each. This constrained the range of possible correlations, hence only the empirical correlation observed in the original study was used, 0.4, with 5000 iterations.

Results were compared between different methods using the "relative efficiency" (RE) measure. This gives the relative number of patients required for a study analyzed using parametric methods so that power was equivalent to the non-parametric alternative. Hence an RE of 1.25 indicates that a particular trial analyzed by parametric statistics would have to accrue 25% more patients than if it were to be analyzed non-parametrically; an AE of 0.80 would indicate that the parametric method was superior by an equivalent amount. The RE is calculated from observed power of the tests,

that is, the proportion of simulations in which the *p*-value was less than the α of 5%. Where $(1-\beta_{np})$ and $(1-\beta_{p})$ are the observed powers from the simulations for the non-parametric and parametric test respectively.

Note that, although it is arguable that the null hypotheses for different tests, say the *t*-test and Mann-Whitney, are technically different, the conclusions drawn by investigators of a randomized trial given a particular *p*-value will be the same, regardless of the analytic method used. Hence direct comparison of the power of different tests is justified in this setting.

2.05 Their Results: Andrew J. Vickers et al (1942)

Change between skewed baseline and post-treatment data tended towards a normal distribution. ANCOVA was generally superior to Mann-Whitney in most situations, especially where log-transformed data were entered into the model. The estimate of the treatment effect from ANCOVA was not importantly biased.

2.06 Their Observations: Andrew J. Vickers et al (1942)

ANCOVA is the preferred method of analyzing randomized trials with baseline and post-treatment measures. In certain extreme cases, ANCOVA is less powerful than Mann-Whitney. Notably, in these cases, the estimate of treatment effect provided by ANCOVA is of questionable interpretability.

2.07 Testing for Independence in High Dimension

Since multivariate inference involves often a large number of variables, the proposed test is like a solution to this problem. Test for independence in high dimensions can be used either when the number of variables exceeds the sample size. In this section, testing for independence of the variable having a multivariate normal distribution with covariance matrix $\sum$ is considered. For example, suppose we have a random sample y_1 ……y_n which is used to compute the usual unbiased sample covariance matrix S and the sample correlation matrix R. The likelihood ratio test for complete independence, which tests for the diagonality of the covariance matrix, rejects this hypothesis of independence for small values of | R | Morrison (2005) so this procedure is not appropriate for high dimensional data since | R | = 0

Wherever m > n (m number of variable, n sample size). The general problem is if the vector y consists of k subvectors $y_{(i)}$ ……. $y_{(k)}$ with m_i variable in the ith subsector, and we wish to test the hypothesis that the k subsectors are independent. We are testing that the $\sum$ has the form

$$\sum = \begin{pmatrix} \sum_{11} & \sum_{12} & \sum_{1k} \\ \sum_{21} & \sum_{22} & \sum_{2k} \\ \sum_{k1} & \sum_{k2} & \sum_{kk} \end{pmatrix} = \begin{pmatrix} \sum_{11} & 0 & 0 \\ 0 & \sum_{22} & 0 \\ 0 & 0 & \sum_{kk} \end{pmatrix}$$

where $\sum_{ij}$ is M_i x M_j: the likelihood ratio test Morrison (2005) based on

$$W = \frac{|S|}{|S_{11}|..|S_{kk}|}$$

Where S has been portioned as was $\sum$ for complete independence.

The likelihood ratio test is not valid when m > n.

(i) Complete Independence

Suppose p_{ij} is the $(ij)^{th}$ element of the covariance matrix $\sum$. Then the null hypothesis of complete independence can be written as

Ho:$P_{ij} = 0$ for i > j

A statistic for testing this hypothesis is the sum of squared s_{ij}'s for i > j so, under Ho we have

$$E\left(\sum_{i=2}^{m}\sum_{j-1}^{i=1} S_{ij}^2\right) = \sum_{i=2}^{m}\sum_{j-1}^{i-1}\frac{P_{ij}P_{ij}}{n}$$

A statistics with mean zero under Ho is

$$t_{nm} = \sum_{i=2}^{m}\sum_{j-1}^{i-1}\left(S_{ij}^2 - n^{-1}S_{ii}S_{ij}\right)$$

Under Ho we have also

$$\rho_{mm}^2 = Var(t_{nm}) = \frac{2(n+2)(n-1)}{n^4}\sum_{i=2}^{m}\sum_{j-1}^{i-1} P_{ii}^2 P_{ij}^2$$

and so t_{nm}/P_{nm} will have mean 0 and variable 1 under H_0

(2) Independence of K sets of variables

our null hypothesis is H_0: $\sum_{ij} = 0$, i > j under H_0 we have

$$E\left(\sum_{i=2}^{m}\sum_{j-1}^{i=1} tr\left(S_{ij}S_{ij}\right)\right) = n^{-1}\sum_{i=2}^{k}\sum_{j-1}^{i-1} tr\left(\sum_{ii}\right)tr\left(\sum_{jj}\right)$$ Beerstecher, et al, (1950).

The statistics with 0 mean under H_0 is

$$U_{nm} = \left(\sum_{i=2}^{k} \sum_{j-1}^{i-1} tr\left(S_{ij} S_{ij}\right) \right) - n^{-1}\left(S_{ij}\right) tr\left(S_{jj}\right)$$

And also we have

$$PU_{nm}^{2} = \text{var}\left(U_{nm}\right) = \frac{2(n+2)(n-1)}{n^4} \sum_{i=2}^{k} \sum_{j-1}^{i-1} tr\left(\sum ii^2 \right) tr\left(\sum {}_{jj}{}^{2} \right)$$

This method can be used in testing the biochemical data Beerstecher, et al, (1950).

2.08 Testing Independence of Multivariate data using Boot strap

The test can be used to check independence of many factors in an experiment. Both parametric (likelihood test) and nonparametric tests can be used for the same data.

The null hypothesis is $H_0 : \sum = \text{diag}\ (\lambda_{11,}\ \lambda_{22,}\ \lambda_{35}\ \ldots\ldots\lambda_{nm})$

Versus

H1: not H_0

Where $\sum = (P_{ij})_{nxm}$ i= 1……n, j= 1…….n

Under the null hypothesis the likelihood ratio test statistic is based on a random sample from

$$N_n\left(U \sum\ \right) isT = \frac{|A|}{\prod_{aij}}$$

where A = (a_{ij}), for i=1, . . . n

j=1, . . . m

is the maximum likelihood estimator of $\sum$.

CHAPTER THREE

MATERIALS USED FOR THE STUDY

3.01 Data Set

The data set used in this work, listed as Appendix A, consists of eight (8) indicators variables selected from 38 African Countries (HDR 2005).

These indicators are

1. HDI Human Development Index
2. EI Education Index
3. GDP/Index Gross Domestic Product Index
4. LEI Life Expectancy Index
5. HPI-I Human Poverty Index values (%)
6. PNUIWS Population not Using Improved Water Sources
7. AIR Adult Illiteracy Rate
8. PBNS Probability at birth of not surviving to age 40 (%).

In the indicators table, countries and areas are ranked in descending order by their human development index, value or by their human poverty index. The indicators present both either or both similarities or differences. The group of human development Index includes indicators such as LEI, EI, GDP, HDI and the second group includes HPI, PNUIWS, AIR, PBNS.

(1) The human development index (HDI): This measures 'the average achievements in a country in three basic dimensions of human development, once the dimension indices have been calculated, then

the HDI is a simple average of the three dimensions indices. HDI = ⅓ (Life expectancy Index LEI + Education Index El' + Gross Domestic Product Index 'GDP')

(2) The Human Poverty Index (HPI): While the HDI measures average achievement, the HPI measures deprivations in the three basic dimensions of human development captured in the HDI.

The formula for calculating the HP1-I include

(i) Probability at birth of not surviving to age 40(times 100) (PBNS).

(ii) Adult Illiteracy rate (AIR)

(iii) Unweighted average of population not using improved water sources and unweighted children under age five.

(3) Education Index (El). This is one of the three indices on which the human development index is built. It is based on the adult literacy rate and the combined primary, secondary and tertiary gross enrolment rate. In addition the adult illiteracy rate (AIR) can also be calculated as 100 minus the Adult literacy rate.

(4) GDP Index (Gross domestic product). It is based on GDP per capital (PPP) (Purchasing power parity).

(5) Population not using improved water sources (PNUIWS) is a percentage calculated as 100 minus the percentage of the population using improved water source.

The main purpose of Human Development Report is to assess the state of human development across the globe by providing a critical analysis of a specific theme each year. It combines thematic policy analysis with detailed country data that focus on human well-being.

The report presents statistical information that has been built up through the collective effort of many Organization such as:

- Food and Agricultural Organization (FAO)
- United Nations Educational, Scientific and Cultural Organization (UNESCO)
- United Nations Statistics Division (UNSD)
- World Bank
- World Health Organization (WHO)
- Government Documents
- Journal articles and other scholarly publications.

These organizations collect data from national sources or through their own survey and make it internationally comparable as possible. These data allow classification of countries in four ways: In major world aggregate, by region, by human development level and income.

3.02 Software Used.

For the analysis of data collected SPSS 13 statistical programme was used for its appropriateness.

3.03 Methodology

3.04 Hotelling's T^2 Test

The Hotelling's T^2 test assesses the statistical significance of the difference on the means of two or more variables. Hotelling's T^2 test is the generalization to P variables of the univariate t-test. The test statistic is presented as: If X_1 …….X_n ~Np (U, ∑) with P = 1. The Statistic is

$T^2 = n\left(\overline{X} - U_0\right)^T S^{-1}\left(\overline{X} - U_0\right)$

where

$\overline{X} = \frac{1}{n}\sum_{i=1}^{m} X_1$ and $S = \frac{1}{n-1}\sum_{i=1}^{m}\left(X_i - \overline{X}\right)\left(X_i - \overline{X}\right)^T$

Then the significance of T^2 is determined by using the fact that, in the null hypothesis case of equal population mean, the transformed statistic is

$$T^2 \sim \left(\frac{n-1}{n-p}\right)^p F_{p,n-p}$$

The two sided Hotelling test rejects H_0 in favours of the

$H_o : \mu = \mu_o$ against $H_1 : \mu \neq \mu_0$

H1:U ≠ Uo at α level

if T^2 =n $(\overline{X} - U_0)^T S^{-1}\left(\overline{X} - U_0\right) > \left(\frac{n-1}{n-p}\right)^p F_{p,n-p,1-\partial}$

3.05 Wilks' lambda Test

Wilks' lambda test is test statistic that is used in multivariate hypothesis testing especially with regard to the likelihood-ratio test (MANOVA). It is

used to test whether there are differences between the means of identified groups of subjects on a combination of dependent variables. It is a direct measure of the proportion of variance in the combination of' dependent variable, that is unaccounted for by the dependent variable (the grouping variable or factor). If a large proportion of the variance is accounted for by the independent variable, then it suggests that there is an effect from the grouping variable and that the groups have different mean values. Wilks Lambda is an inverse criterion, so a lower lambda (α) value means that there is a greater treatment effect. This means that, when the dependent variable are not inter-correlated, less variance is explained by the within groups effects and more variance is explained by the between groups effect. Under the null hypothesis, the statistic is:

$$\wedge(P,m,n)=\frac{|A|}{|A+B|}\,then-\left(m-\frac{P-m+1}{2}\right)\log\wedge\approx X^{2}_{pm}$$

Wilk's lambda statistic can be transformed to a statistic which has approximately an F distribution. This makes the P-value calculated easier.

We shall use the distribution approximations of or X_2 to obtain significance levels. The null hypothesis is rejected at α level if

$$\frac{1-\wedge}{\wedge}-\left(\frac{m-p+1}{p}\right)\geq F_{p,m-p+1}$$

$$-\left(m-\frac{p-m+1}{2}\right)\log\wedge\geq X^{2}_{pn}$$

3.06 Canonical Correlation Analysis

Canonical correlation analysis is an additional procedure for assessing the relationship between variables. Specifically, this analysis allows us to investigate the relationship between two sets of variables. In this analysis, linear combinations of canonical variables are created for each data set such that the correlations between canonical variables of the two data sets are maximized. The hypothesis tested is $\sum_{12} = 0$, which means that the two sets of variables are uncorrelated with one another. The statistic is under the assumption of normality, and the likelihood ratio statistics for testing the null hypothesis, given by

$$\lambda^{\frac{2}{n}} = \left|1 - S_{22}^{-1} S_{21} S_{11}^{-1} S_{12}^{-1}\right| = \prod_{i=1}^{k} \left(1 - r_i^2\right) \approx X_{pq}^2$$

The null hypothesis is rejected for small $\lambda^{2/n}$ at α level.

3.07 Wilcoxon signed-ranks test

The test is used when it is assume that the difference is ordinal. The Wilcoxon signed ranks test, rank all the difference (d's) between pairs from the two groups and affix for each rank the sign of difference. Under the null hypothesis we expect that the sum of positive ranks will be equal to the sum of negative, otherwise the null hypothesis is rejected. The statistic is obtained using a linear function of the ranks of Euclidean norms of the transformed vectors.

Therefore, the test statistic for testing the null hypothesis of independence is obtained as follows for i =1, 2, …….., n for p and q dimensional standardized sign vectors based on the first and second components

$X_1^{(1)} and - X_1^{(2)}$

Let $\overset{\wedge(1)}{R}$ and $\overset{\wedge(2)}{R}$ denote the rank of $\overset{\wedge(1)}{r} = \left\| Z_i^{(1)} \right\|$ and $\overset{\wedge(2)}{r} = \left\| Z_i^{(2)} \right\|$ among $r_i^{(1)}$ ……$r_n^{(1)}$ and $r_1^{(2)}$ ……-$r_n^{(2)}$-respectively

The Wilcoxon test of independence with test statistic is then

$$T = \frac{9npq}{(n+1)^4} \left\| ave_i \left(\overset{\wedge(1)}{R}_1 \overset{\wedge(2)}{R}_1 \overset{\wedge(1)}{U}_1 \overset{\wedge(2)}{U}_1 \right) \right\|^2$$

For the value of T computed if the P associated is equal to or less than that α level, then the null hypothesis is rejected in favour to the alternative. To perform this test, the data of appendix A is transformed into two groups of variables (HDI and HPI) regarding to the sets of variables above, and we refer to it as appendix B. The eight variables become two broad variables.

3.08. Friedman two way analysis of variance test.

This method compares several related samples and can he used as **a** non-parametric alternative to the two ways ANOVA. The power of this method is with small samples but it is the best method for non-parametric two way analysis of variance with sample, sizes above five. This method allows a follow-up comparison between individual parameters, and it is therefore, recommended that this test be used for the analysis of ranks or scores. In this situation the data are cast in two way table having N(38) rows and K(8) columns. The data of the test are ranks. The scores in each row are ranked separately. The test determines whether it is likely that the different columns of ranks (samples) came from the same population. If the null

hypothesis (that the samples have all been drawn from the same population or populations with the same median) is true then the distribution of ranks in each column would be a matter of chance, and thus we will expect all the ranks from 1, 2,……..,N to appear in all columns with about equal frequency. Therefore if the null hypothesis is false, the rank totals would vary from one column to another.

The Friedman test determines whether the ranks totals for each parameter differ significantly from the values, which would be expected by chance and also tests whether the K related samples come from the same population with respect to ranks. When the number of rows and/or columns is large enough, the following statistic is approximately distributed as Chi-square (X_r^2) with K-l degree of freedom.

$$X_r^2 = \frac{12}{NK(k+1)} \sum_{i=1}^{k} (Rj)^2 - 3N(k+1)$$

and that X_r^2 is distributed approximately as Chi-Square with df= K-1 only when the number of rows/or columns is large.

The probability of occurrence under the null hypothesis associated with observed values of X_r^2 depends on the size of N and K and by reference to the X_r^2 distribution with df=K-l. Then if the probability is equal to or less than that a (table), the null hypothesis is rejected we will use this test to determine the independence of variables HDI, LEI, El, GDP, PNUIWs, AIR, PBNS, HPI.

3.09 Kendall's W Test.

Kendall's coefficient of concordance W measures the relationship among several rank-ordered variables for N objects or individuals. The analysis is conducted on a table which contain, in each column the ranks of the N objects on one of K variables. As a solution to the problem of measuring relation between several K-sets of rankings. Friedman (1937) has shown that, when the number of rows and/or columns is large enough, the following statistics is approximately distributed as X^2 with k – 1 degree of freedom.

$$x^2 = \left(\frac{12}{NK(N+1)} \sum_{i=1}^{k} (R_1)^2 \right) - 3N(k-1)$$

Where R_1 is the sum of the ranks for row 1: this is Friedman's Statistics for two-way analysis of variance by ranks. Kendall's coefficient of concordance can be obtained by transforming the Friedman's X^2 statistics as follows.

$$W = \frac{x^2}{K(N-1)} \Leftrightarrow x^2 = K(N-1)W$$

The above expression can be expressed as

$$W = \frac{12 \sum_{i=1}^{k} \left(R_i - \overline{R}\right)^2}{K^2(N^3 - N)}$$

Kendall's W statistics is simply the variance of the row sums of ranks R: divided by the maximum possible value that variance can take: this occurs when all variables are in total agreement. Hence, the Kendall coefficient of concordance, W, provides an indication of the association between several rankings of different parameters. W ranges between 0 and 1 with 1 designating perfect concordance, and 0 indicating no agreement or independence of samples. The Friedman two-

way analysis of variance and the Kendall Coefficient of concordance are linearly related Siegel and Castellan, (1988) and as such provide, similar results. They address hypothesis concerning the same data and they use the same x^2 Statistics for testing. They differ only in the formulation of their respective hypothesis. An extension of the Friedman test does however allow multiple comparisons of different ranks, thus enabling the user to identify significant differences between the important of paranieters ranked. As in the case for the Kendall, statistics data have to be converted to rank prior to analysis and modifications exist for tied ranks. The null hypothesis (Ho) subjected to testing is that the row sums R; are equal or, in other words, that the K sets of ranks are independent of one another. The X^2 statistics is compared to a X^2_α value read in a table of critical values of $X^2$for N- 1 degree of freedom. When X^2 is small than the critical value X^2 (that is to say probability larger than α) the null hypothesis that the row sums Ri are equal cannot be rejected. This leads to the conclusion that the K variables are independent and differ in the way they rank the N objects. On the contrary, $X^2 \geq X^2_\alpha$ (that is probability smaller than or equal to α). Indicates good agreement among the variables in the way they rank the objects.

3.10 Cochran Q Test

The Cochran Q test for K related samples, provides a method for testing whether more than one sets of frequencies or proportions differ significantly among themselves. The test is particularly suitable when the data are in a nominal scale or are dichotomized ordinal information. If the

data of the research can be arranged in two way table consisting of N rows and K columns. It is possible to test the null hypothesis that the proportion (or frequency) of responses of a particular kind is the same in each column, except for chance differences. If the null hypothesis is true, that is if there is no difference in the probability of success under each condition the statistics is

$$Q = \frac{K^{(k-1)} \sum_{i=1}^{k} \left(Z_1 - \overline{Z}\right)^2}{K \sum_{i=1}^{N} L_1 - \sum_{i=1}^{N} L_1^2}$$

which is distributed as X^2_{k-1}

Where Z_j is the total number of "success" in i^{th} column,
$\overline{Z}$ is the mean of the Z_j

Li is the total number of 'success" in i^{th} row. The above formula is equivalent to

$$Q = \frac{K^{(k-1)} \sum_{i=1}^{k} Z_j^2 - \left(\sum_{i=1}^{k} - \overline{Z}_j\right)^2}{K \sum_{i=1}^{N} L_1 - \sum_{i=1}^{N} L_1^2}$$

The probability associated with the occurrence under the null hypothesis of values as large as an observed Q may be determined by reference to tabular. If the observed value of Q is equal to or greater than that of tabular.

Then for α level and a value of df= K-1, the null hypothesis is rejected in favour to the alternative hypothesis.

CHAPTER FOUR

ANALYSIS

4.01 Hoteling's-T^2-Test

Hoteling T^2 test of independence

Table 4.1 Analysis of Variance for Hoteling T^2 test of independence

Source of Variation	Sum of Sq	DF	Mean Square	F
Between People	10787.8279	37	291.5629	
Within People	128271.8686	266	482.2251	
Between Measures	102106.0608	7	14586.5801	144.3840 .0000
Residual	26165.8078	259	101.0263	
Total	139059.6965 303	458.9429		

Table 1

Hotelling's T-Squared 1342.9892 F = 160.7439 Prob. = .0000

Degrees of Freedom: Numerator = 7 denominator = 31

From Table 1, the null hypothesis, that is variables are statistically independent, is rejected at-α= 0.05. This mean that the indicators:

HDI, EL, LEI, *GDPIlndex,* HPI, PNUIWS, AIR, PBNS are not independent meaning that they are correlated

4.02 Wilks' Lambda test

Table 4.2: Wilks' Lambda test of independence for human development index report

Multivariate Tests of Significance (S = 4, M = 2, N = 14)

Test Name	Value	Approx. F	Hypoth. DF	Error DF	Sig. of F
Pillais	2.14185	9,50962	16,00	132,00	000
Hotellings	53,25485	94,86020	16,00	114,00	,000
Wilks	,00137	44,11116	16,00	92,29	,000
Roys	,97601				

The null hypothesis is Ho:variables are independent against the alternative. The result from Table 2 indicates that small value ^ of 0.00137 implies that the null hypothesis is rejected at α = 0.05 level. That is variables are not independent. The interpretation of this is that the indicators HDI, LEI, El, GDP, HPI, PNUIWS, AIR, PBNS are interrelated.

4.03 Canonical correlation analysis

Table 4.3a: Canonical correlation test of independence for human development report analysis

Correlations for Set-i

	HDI	El	LEI	GDP
HDI	1.0000	.8030	.8174	.8957
El	.8030	1.0000	.3918	.6074
LEI	.8174	.3918 1	.0000	.6881
GDP	.8957	.6074	.6881	1.0000

Table 4.3b: Canonical correlation test of independence for human development report analysis

	HPI	PNUIWS	AIR	PBNS
HPI	1.0000	.6098	.8368	.6652
PNUIWS	.6098 1	.0000	.3260	.5053
AIR	.8368	.3260	1.0000	.2166
PBNS	.6652	.5053	.2166	1.0000

Correlations Between table 4.3a and 4.3b

	HPI	PNUIWS	AIR	PBNS
HDI	-.9155	-.6161	-.6770	-.7958
El	-.8772	-.3836	-.9516	-.3531
LEI	-.6805	-.5085	-.2584	-.9852
GDP	-.7372	-.6642	-.4655	-.6860

Canonical Correlations results for the analysis

1 .988

2 .962

3 .491

4 .013

Test that remaining correlations are zero:

	Wilks	Chi-SQ	DF	Sig.
1	.001	214.193	16.000	.000
2	.057	92.967	9.000	.000
3	.759	8.972	4.000	.062
4	1.000	.006	1.000	.939

From the canonical correlation results of the analysis, 0.988 and 0.962 are very close to one. Therefore, an overall conclusion can be made, and this can be due to the strong relationship between the first two variables (HDI and El). On the other

hand the last two canonical correlation *0.491* and 0.013 are low, this implies that the variables corresponding have less impact (LEI and GDP) on the relationship. Also the comparison of the obtained and the tabulated values for the chi-square at $\alpha = 0.05$ level indicate that only the last two canonical variate are significant. This means that the null hypothesis is accepted at these two variables with canonical correlation not different from zero. The zero canonical correlation implies independence between the sets of variables. So, the two sets of variables are not fairly independent. The conclusion is to reject the null hypothesis that the two sets of variables are independent.

4.04 The Wilcoxon signed-ranks sum test

Table 4.4a: Wilcoxon signed ranks test of independence for human development report

		N	Mean Rank	Sum of Ranks
HPI WILCOXON -	Negative Ranks	2 [a]	1.50	3.00
HDI WILCOXON	Positive Ranks	149[b]	77.00	11473.00
	Ties	O[c]		
	Total	151		

a. HPI WILCOXON < HDI WILCOXON

b. HPI WILCOXON> HDI WILCOXON

c. HDI WILCOXON = HPI WILCOXON

Table 4.4b: Wilcoxon signed ranks test of independence for human development report

	HPI WILCOXON HDI WILCOXON
Z	-10.654[B]
Asymp.Sig.(2-table)	.000

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

From Table 4.4a, the Z value is -10.654 and the p-value associated from the tabular (2-tailed) is less than that $\alpha = 0.05$. So, the null hypothesis is rejected. This mean that variables are not independent, thus we can conclude that Human development index and Human poverty index are related to each other.

4.05 The Friedman two way analysis of variance test

Table 4.5a: Friedman multivariate test of independence for human development report analysis

Ranks

HDI	2.53
El	3.30
LEI	2.30
GDP/lndex	2.08
HPI	6.55
PNUIWS	6.29
AIR	6.82
PBNS	6.13

Table 4.5b: Friedman multivariate test of independence for human development report analysis

Test-Statistics[a]

N	38
Chi-square	199.932
df	7
asymp.Sig.	.000

Result from Table 4.5a indicates that the null hypothesis is rejected at $\propto =$ 0.05 level, meaning that the variables are not independent, the interpretation of this is that the 8 indicators are related.

4.06 Kendall's W test

Table 4. 6: Kendall's W test of independence for human development report

Test Statistics

N	38
Kendall's	752
Chi-square	199.932
Df	7
Asymp.sig	000

a. Kendall's Coefficient of Concordance

From Table 4.6, the null hypothesis that is variables are independent is, rejected at α= 0.05 level of significance. This implies that the variables are not independent, meaning that all the indicators of interest are correlated.

4.07 Cochran Q test

Table 4.7: Cochran Q test of independence for human development report

Analysis of Variance for human development report

Source of Variation	Sum of Sq.	DF	Mean Square	Q	Prob.
Between people	10787.8279	37	291.5629		
Within People	28271.8686	266	482.2251		
Between Measures	102106.0608	7	14586.5801	211.7394	.0000
Residual	26165.8078	259	101.0263		
Total	139059.6965	303	458.9429		

Table 4.7 indicates that the null hypothesis is rejected at $\alpha = 0.05$ level. Meaning that the variables are not independent so indicators are related to each other.

CHAPTER FIVE

5.01 SUMMARY AND CONCLUSION

5.02 Summary

In this research work multivariate parametric and non parametric tests for independence were performed using multivariate data collected from World Bank on Human Development Report. The null hypothesis (Ho: variables are independent) is rejected for both the selected parametric (Hotelling's T^2, Wilks Lambda, Canonical correlation analysis) and non parametric (Friedman, Kendall W. Wilcoxon and Cochran Q) tests.

5.03 Conclusion

Multivariate parametric and non parametric are two statistical methods of inference. Multivariate parametric methods depend upon the assumption of a specific distributional form for example an approximate multivariate normal distribution. And data for these method will be interval or ratio scales. On the other hand the multivariate non parametric method is referred to as distribution free method using ordinal or normal data and even interval and ratio data when the distributional assumption is unspecified. The analysis of the data from human development report using both methods yields the results. From the parametric test of independence, Hotelling's T^2,

Wilks Lambda and Canonical correlation analysis the null hypothesis that is variable are independent and is rejected in favour of the alternative. From the nonparametric test, Wilcoxon signed-ranks, Friedman, Kendall and Cochran, the null hypothesis is also rejected at the same levels of significant so both methods yield the same result. But a multivariate parametric method for example, Canonical correlation analysis have some times an advantage over nonparametric method. That is, if the null hypothesis is rejected, it shows which of the components hypothesis led to the rejection of the null hypothesis. As well as knowing that the null hypothesis should be rejected, one could enquire which specific linear combination led to its rejection. This work has equally shown us the similarities and differences between parametric and non parametric method in analyzing specific data.

5.04 Recommendation

The essence of this work is to expose the relationship between component of human development reports. The result shows that the education index and poverty index are closely related. Hence, assistance need to be given to Countries with high poverty indices in terms of their educational institutions.

References

Bruce, F. and Arusha, C. (2004), parametric and nonparametric test for RIP among the G7 Nations. School of Economics. Discussion paper (unpublished).

Chatfield, C, and Collins, A. J., (1980), Introduction to multivariate analysis. Chapman and Hall. London. NewYork. Tokyo. Melbourne. Madras.

Chernoff, S. and Hofges, L. (1958), results for Wilks' test of Multivariate independence. Paper pre-printed submitted to Elsevier science (unpublished).

Conover, W. J., (1 999), A practical nonparametric statistics. (Third edition), John Wiley and Sons, New York.

Genest, C. Jean, F.Q. and Bruno, R., (2002) tests of serial independence based on Kendall's process. The Canadian journal of Statistics vol.30.No.3, page 1-21.

Genest, C and \/erret F. (2003). Locally most powerful rank tests of independencefor copied models (unpublished).

Hallin, M. and Paindaveine, D. (2002), Multivarate signed-ranks: Randles interdirection on Tyler's Angles (Unpublished).

Harvey, M., (1 995), Intuitive Biostatistics: Choosing a statistical test. Chapter 37. Copyright (c) by Oxford University press Inc.

Horowitz,J.L., (2004), Testing a parametric model against a nonpararnetric alternative with identification through instrumental.

Human Development Report 2000. New York.

James, R.S.,(2004), Testing for independence in high dimensions. University of Central Florida, orlando, Florida 328 16- 2370, USA. Jschott(d ucf.edu).

Kai, C.C., et al, (2003), A statistical approach to testing mutual independence of ICA Recovered sources. 4th International Symposium on independent component analysis and Blind signal separation.

Mardia, K.V., Kent, J.T.and Bibby, J.M. (1979), Multivariate analysis. London New York. Toronto Sydney. San Francisco. A subsidiary of Harcourt Brace jovanvich, publishers.

McGraw l-l.K., (1956), Nonparametric statistics: For behavioural sciences.

International student edition. mc, Publisher, Library of Congress catalog card No.56-8 185.

Oja, H. and Paindaveine, D., (2004), Optimal signed-rank tests based on hyperplanes. (unpublished).

Paindaveine, D., (2003), Elementary derivations of multivariate non-admissibility results. Technical report 0329. hhp://www.stat.ucl.ac.be/IA P.

Park, H. et al, (2000), Nonparametric One-sided tests for multivariate data. Sankhya: the Indian Journal of statistics, vol. 63 series 13 Pt 3. pp 286-297.

Piater, A.P.C. and Sauvageot. J.L., (1975). Statistique 2: Elements d' analyse et Applications. Presses Universitaires de France.

Sara, T. Oja, H. and Ronald, H.R. (2003), Multivariate nonparametric test of independence (unpublished).

Ronald, E.W and Raymond, H.M (1978) Probability and statistics for Engineers and Scientists. Macmillan publishing Co. Inc. New York.

Siegel, and Castellan, N.J. (1988), Nonparametric statistic for the behavioural sciences (2nd ed.), New York: Mc Graw. Hill.

Kai, C.C., et al, (2003), A statistical approach to testing mutual independence of ICA Recovered sources. 4th International Symposium on independent component analysis and Blind signal separation.

McGraw l-I.K., (1956), Nonparametric statistics: For behavioural sciences. International student edition. mc, Publisher, Library of Congress catalog card No. *56-8* 185.

Oja, H. and Paindaveine, D., (2004), Optimal signed-rank tests based on hyperplanes. (unpublished).

Paindaveine, D., (2003), Elementary derivations of multivariate non- admissibility results. Teachnical report 0329. hhp://www.stat.ucl .ac.be/l AP.

Park, H. Jong, H. and Desu, M.M., (2000), Nonpararnetric One-sided tests for multivariate data. Sankhya: the Indian Journal of statistics, vol. 63 series 13 pt 3. pp 286-297

Piater. A.P.C. and Sauvageot. J.L.. *(1975),* Statistique 2: Elements d' analyse et Applications. Presses Universitaires de France.

Sara, T., Oja, H. and Ronald, H.R. (2003), Multivariate nonparametric test of independence (unpublished).

Ronald, E.W and Raymond, H.M (1978) Probability and statistics for Engineers and Scientists. Macmillan publishing Co. Inc. New York.

Subhashish Chakravarty (2001), Testing independence for multivariate data using Bootstrap (unpublished).

Taskinen, S., Kankainen, A. and Oja H., (2003), Rank score tests of multivariate independence. (unpublished).

APPENDIX A

HUMAN DEVELOPMENT INDEX

Countries	HDI	EI	LEI	GDP/INDEX	HPI	PNUIWS	AIR	PENS
Mauritius	0.75	0.77	0.77	0.76	16	0	15.8	5.4
Libyan J	0.77	0.83	0.75	0.72	16.7	28	20.9	6.4
Cape verde	0.7	0.75	0.74	0.63	20.9	26	26.4	10.4
South Afri	0.7	0.87	0.48	0.75	18.7	14	15.1	24.4
Algeria	0.69	0.69	0.74	0.66	23.5	6	33.4	10.5
Egypt	0.63	0.62	0.7	0.59	31.7	5	45.4	10.3
Morocco	0.59	0.49	0.7	0.59	36.4	18	52	11.8
Zimbabwe	0.55	0.8	0.3	0.56	36.2	15	12	51.6
Ghana	0.54	0.61	0.53	0.49	29.1	36	29.7	27
Lesotho	0.54	0.75	0.38	0.49	25.8	9	17.1	35.4
Kenya	0.51	0.71	0.44	0.39	31.8	51	18.5	34.6
Comoros	0.51	0.51	0.57	0.44	29.9	4	40.8	20.6
Cameroon	0.5	0.64	0.42	0.46	31.1	38	25.2	36.2
Congo	0.5	0.74	0.44	0.5	30.7	49	20.5	34.8
Togo	0.48	0.58	0.44	0.44	38.3	46	43.7	34.1
Madagascar	0.46	0.59	0.45	0.35	38.6	53	34.3	31.6
Nigeria	0.45	0.57	0.44	0.36	36.1	43	37.4	33.7
Djibouti	0.44	0.5	0.32	0.53	34.7	0	36.6	42.3
Sudan	0.43	0.49	0.51	0.32	34.8	25	43.1	27.3
Mauritania	0.43	0.41	0.43	0.46	47.8	63	58.4	33.1
Tanzania	0.43	0.61	0.44	0.27	32.4	46	25.3	33.3
Uganda	0.43	0.59	0.3	0.41	41	50	33.9	48.4
Congo	0.42	0.51	0.43	0.35	40	55	39.7	34.7
Zambia	0.42	0.68	0.27	0.34	40	36	22.8	53.6
Cote d'ivoire	0.42	0.43	0.38	0.47	42.9	23	54.3	40.2
Senegal	0.42	0.36	0.47	0.44	45.9	22	63.6	28.5
Benin	0.42	0.41	0.48	0.37	45.8	37	61	29.7
Eritrea	0.41	0.44	0.45	0.36	44	54	47.3	31.7
Gambia	0.39	0.39	0.35	0.46	49.6	38	64.3	40.5
Malawi	0.39	0.64	0.26	0.3	43.4	43	40.8	50.4
Rwanda	0.39	0.57	0.25	0.36	44.2	59	34.2	51.9
Mali	0.37	0.36	0.44	0.34	47.8	35	60.2	38.5
Central Africa	0.37	0.38	0.32	0.41	46.1	40	54.6	45.3
Chad	0.35	0.38	0.34	0.36	53.1	73	59	41
Guinea Biss	0.33	0.37	0.33	0.32	49.6	51	62.3	42.2
Mozambique	0.32	0.36	0.25	0.36	48.3	40	56.8	49.2
Ethiopia	0.32	0.34	0.32	0.31	57.2	76	62.6	43.6
Niger	0.27	0.15	0.33	0.34	63.6	41	84.7	41.4

Source: Note on statistics in the human development report: World Bank (March, 2005)

GDP = Gross Domestic Product Index

EI = Education Index

LEI = Life Expectancy Index

HDI = Human Development Index

HPI-1 = Human Poverty Index (%)

PNUIWS = Population Not Using Improved Water Sources (%)

PBN = Probability at birth of not surviving to age 40 (%)

AIR = Adult Illiteracy rate

Appendix-B

Data for Wilcoxon

HDI	HPI
0.75	16
0.77	16.7
0.7	20.9
0.7	18.7
0.69	23.5
0.63	31.7
0.59	36.4
0.55	36.2
0.54	29.1
0.54	25.8
0.51	31.8
0.51	29.9
0.5	31.1
0.5	30.7
0.48	38.3
0.46	38.6
0.45	36.1
0.44	34.7
0.43	34.8
0.43	47.8
0.43	32.4
0.43	41
0.42	40
0.42	40
0.42	42.9
0.42	45.9
0.42	45.8
0.41	44
0.39	49.6
0.39	43.4
0.39	44.2
0.37	47.8
0.37	46.1
0.35	53.1
0.33	49.6
0.32	48.3
0.32	57.2
0.27	63.6
0.77	0
0.83	28
0.75	26
0.87	14
0.69	6
0.62	5
0.49	18
0.8	15
0.61	36
0.75	9

0.71	51
0.51	4
0.64	38
0.74	49
0.58	46
0.59	53
0.57	43
0.5	0
0.49	25
0.41	63
0.61	46
0.59	50
0.51	55
0.68	36
0.43	23
0.36	22
0.41	37
0.44	54
0.39	38
0.64	43
0.57	59
0.36	35
0.38	40
0.38	73
0.37	51
0.36	40
0.34	76
0.15	41
0.77	15.8
0.75	20.9
0.74	26.4
0.48	15.1
0.74	33.4
0.7	45.4
0.7	52
0.3	12
0.53	29.7
0.38	17.1
0.44	18.5
0.57	40.8
0.42	25.2
0.44	20.5
0.44	43.7
0.45	34.3
0.44	37.4
0.32	36.6
0.51	43.1
0.43	58.4
0.44	25.3
0.3	33.9
0.43	39.7
0.27	22.8

0.38	54.3
0.47	63.6
0.48	61
0.45	47.3
0.35	64.3
0.26	40.8
0.25	34.2
0.44	60.2
0.32	54.6
0.34	59
0.33	62.3
0.25	56.8
0.32	62.6
0.33	84.7
0.76	5.4
0.72	6.4
0.63	10.4
0.75	24.4
0.66	10.5
0.59	10.3
0.59	11.8
0.56	51.6
0.49	27
0.49	35.4
0.39	34.6
0.44	20.6
0.46	36.2
0.5	34.8
0.44	34.1
0.35	31.6
0.36	33.7
0.53	42.3
0.32	27.3
0.46	33.1
0.27	33.3
0.41	48.4
0.35	34.7
0.34	53.6
0.47	40.2
0.44	28.5
0.37	29.7
0.36	31.7
0.46	40.5
0.3	50.4
0.36	51.9
0.34	38.5
0.41	45.3
0.36	41
0.32	42.2
0.36	49.2
0.31	43.6
0.34	41.4