Application of Non-parametric Analysis Technique amongst Postgraduate Education Research: A Survey of South African Universities

Anass BAYAGA*

Liile Lerato LEKENA**

Abstract

The objective of this research was to determine factors that influence application of non-parametric analysis technique. The data emanated from research done by postgraduate students over a ten year period (1995-2004) and archived by the project in postgraduate education research (PPER).

A Survey of three South African universities was conducted. The classification of researches from chosen prominent universities were made by research title, research topic, target population, data collection method, and other diversity titles which were used to map the position of non parametric analysis. The sample in the three (3) universities included four hundred and twenty-one (421) sampled researches.

The first finding indicated that the data presentation chapters of the sampled researches were all analysed using descriptive analysis without application of non-parametric technique. Thus, no sampled research applied non-parametric analysis technique. Secondly, the findings suggested that there was a relationship between research title and data analysis technique. Thirdly, there was association between research titles and target populations, which consequently influence choice of data analysis. Lastly, the dominant themes amongst the sampled researches were age, inclusive education and education**.**

Key Words: Non-parametric analysis, parametric analysis, Research Methods, Data analysis. Postgraduate Research

1.0 Backround of Study

Research methods and particularly data analysis topic are among rather difficult subjects tackled by postgraduate researchers (Yin, 2008). The reason being that many research problems call for analysis and prediction of a dichotomous[1] outcome/responses. Traditionally, such research responses were addressed by either ordinary least squares (OLS) regression or linear discriminant function analysis (Hair, 2006; Stern, Frohna & Gruppen, 2005).

However, both techniques are akin to parametric analysis. This means that they depend on strict statistical assumptions, which most researchers fail to grasp (Tabachnick & Fidell, 2001). These assumptions[2] include normality of independent variables, linearity of relationships, lack of multicollinearity among independent variables, equal dispersion matrices for discriminant analysis (Ott & Longnecker, 2008; Yin, 2008; Tabachnick & Fidell, 2001).

Introduction of non-parametric analysis (NPA) appeared as an alternative in the late 1960s and early 1970s (Johnson & Wichern, 2007; Neuman, 2003; Tabachnick & Fidell, 2001). NPA technique, like other data analysis methods, started to find a prominent place in the statistical software packages statistical techniques, most particularly NPA and parametric test (Creswell, 2008; Harrell, 2001; Tabachnick & Fidell, 2001).

Regardless of availability of such software, recent international and South African studies suggest that postgraduate researchers in education, social and even exact (physical) sciences in data analysis courses do not understand many of the concepts they study in research methods, particularly data analysis (North & Zewotir, 2006; Stohl, 2005; DeWet, 2002). There are three main reasons attributed to the above, as literature suggested.

---

* Correspondence author; University of Fort Hare, South Africa.

** University of the Witwatersrand, South Africa.

[1] Male/female, yes/no, user/nonuser, satisfied/unsatisfied, etc.

[2] For details cf. Tabachnick & Fidell, (2001).

Firstly, besides lack of understanding of such strict assumptions, postgraduate researchers often tend to respond to problems involving data analysis in general by falling into a number crunching mode, plugging quantities into a tables and graphs or procedure without forming an internal representation of the data (Stokes, Davis & Koch, 2008).

Secondly, postgraduate researchers may be able to memorise formulas and the steps to follow in familiar, well-defined problems, but only seldom appear to get much sense of what the rationale is or how data can be applied in new situations. Consequently, within the conceptual underpinnings, the details they have learned or memorised, for whatever use they might be, soon fade. This is particularly true with (Creswell, 2008; Powers & Xie, 2000; Stokes, Davis & Koch, 2008).

Thirdly, in statistical inference in particular, or hypothesis testing, the traditional tests are called parametric tests because, as aforementioned, they depend on the specification of a probability distribution[3] (cf. assumptions) (Huck, Cormier & Bounds, 2004; Tabachnick & Fidell, 2001). Thus, parametric tests are said to depend on distributional assumptions, but studies suggest that most research students do not even understand these assumptions, not alone apply them (Stohl, 2005; Russo & Passannante, 2001).

Nonparametric analysis (NPA), on the other hand and as noted earlier on, does not require any strict distributional assumptions. Therefore, the method is also sometimes (and more appropriately) called parameter-free methods or distribution-free methods. Even if the data are distributed normally, nonparametric methods are often almost as powerful as parametric methods, but always misapplied or not applied by postgraduate researchers at all due to the aforementioned reasons. Although NPA procedures (cf. section 1.1 for details) are specifically targeted for distribution free data, Tabachnick & Fidell (2001) have established that postgraduate researchers still find difficulty during data analysis, even without necessarily applying statistical test of assumptions. Tabachnick & Fidell's (2001) lamentation forms the motive of this current paper, which was applied project in postgraduate education research (PPER) in three South African Universities. The study reports on part of the 4th national progress report on the Project for Postgraduate Education Research (PPER) in South Africa (1995-2004). It is however the first report that offers a preliminary reading of data that is produced in Higher Education Institutions in the Western Cape Province. It uses the data gathered in 2007 and 2008 which formed part of phase 1 and 2 respectively. The current study takes the 4th national progress report (Balfour et al, 2008) as its reference. While, the 4th National progress report represents the preliminary report of the whole project, the 1st regional report is part of the National Research Foundation (NRF) request to provide a preliminary analysis of the data that includes phase 1 and 2 developments. The next section explores the usage of NPA.

1.1 Usage of Non parametric Analysis (NPA)

One of the applications of NPA arises, when a researcher has say two samples that need to be compared concerning their mean value for some variable of interest. In such instance, a researcher would use the *t*-test for independent samples-associated with parametric analysis/test. Nonparametric alternatives for this test are the Wald runs test, the Mann-Whitney U test, and the Kolmogorov-Smirnov two-sample test (Tabachnick & Fidell, 2001).

If, there exist multiple groups, then a researcher would use analysis of variance or multivariate analysis of variance (ANOVA/MANOVA). Again the nonparametric equivalents to this method are the Kruskal-Wallis analysis of ranks and the Median test. This is what is termed differences between independent groups (Tabachnick & Fidell, 2001).

Alternatively, differences between dependent groups arises if a researcher wants to compare two variables measured in the same sample, in this regard, a researcher would customarily use the t-test for dependent samples[4]. Nonparametric alternatives to this test are the Sign test and Wilcoxon's matched pairs test (cf. section 3.2 for such application). However, if the variables of interest are dichotomous in nature (for instance "pass" or "fail"), then McNemar's chi-square test is appropriate.

---

[3] Except for a set of free parameters
[4] In basic statistics for eaxmple, if one wants to compare students' math skills at the beginning of the semester with their skills at the end of the semester.

If there are more than two variables that were measured in the same sample, then a researcher would customarily use repeated measures ANOVA. Nonparametric alternatives to this method are Friedman's two-way analysis of variance and Cochran Q test (if the variable was measured in terms of categories for instance "passed" or "failed"). Noting that Cochran Q is particularly useful for measuring changes in frequencies (proportions) across time.

In addition to the above test or analysis technique for significant levels, there are other three types of commonly used nonparametric correlation coefficients, which are used to test relationship between variables. These include (1) Spearman R (2) Kendall Tau and (3) Gamma coefficient; these would be explained shortly. Noting that chi-square statistic computed for two-way frequency tables, also provides a careful measure of a relation between two (tabulated) variables, and unlike the correlation measures listed, it could be used for variables that are measured on a nominal scale (Tabachnick & Fidell, 2001).

With reference to NPA correlation measures such as Spearman R, it assumes that the variables under consideration were measured on at least an ordinal (rank order) scale that is, individual observations can be ranked into two ordered series. Spearman R could be thought of as the regular pearson product moment corrlation coefficient, that is, in terms of proportion of variability accounted for, except that Spearman R is computed from ranks (Tabachnick & Fidell, 2001).

Kendall Tau is equivalent to Spearman R with regard to the underlying assumptions. It is also comparable in terms of its statistical power. However, Spearman R and Kendall Tau are usually not identical in magnitude because their underlying logic as well as their computational formulas are very different.

The Gamma statistic is preferable to Spearman R or Kendall Tau when the data contain many tied observations. In terms of the underlying assumptions, Gamma is equivalent to Spearman R or Kendall Tau; in terms of its interpretation and computation, it is more similar to Kendall Tau than Spearman R (Tabachnick & Fidell, 2001). In short, Gamma is also a probability, specifically, it is computed as the difference between the probabilities that the rank ordering of the two variables agree minus the probability that they disagree, divided by 1 minus the probability of ties. Thus, Gamma is basically equivalent to Kendall Tau, except that ties are explicitly taken into account.

Conclusively, with regards to (1) the contestations elucidated in the background and (2) its application in general context, the current paper intents to explore via the under listed research questions and hypothesis to investigate application of non-parametric analysis technique of PPER among South African Universities.

Research Questions

The following major research question emanated from the background of the study.

Major research question

What factors influence the choice of data analysis technique amongst postgraduate researchers?

Sub-research question

What diversity variables predominates postgraduate (PhD and M.Ed in 1995-2004 and 2000-2004) research output in South Africa that influence choice of data analysis?

Following the major research question, it was deemed necessary to pose the below hypothesis

Hypothesis 1

Ho= there is no relationship between research title and data analysis method.

Ha= there is a relationship between research title and data collection method.

Significance of the Study

The environment for both academic and applied researchers has changed dramatically in the last ten years as a result of a series of developments in computer technology/software programs, information, and

statistical techniques (Tabachnick & Fidell, 2001). The revolution of the computers and continuous development of new computerised techniques have contributed to a greater analytical capability.

These developments, however, have been confronted with the complexity of the research topics problems and response types. In this context, NPA techniques have emerged as useful tools allowing researchers to engage in developing and testing their theoretical models and coping with complex theoretical and practical problems. Suggesting that NPA techniques are of outmost significance to practitioners who formulate strategies to be applied in different segments of institutions with varied characteristics.

2. Method

The study surveyed and analysed research trends by using the PPER database from 1995-2004. Although the Western Cape Province of South Africa has four universities, namely the University of Cape Town (UCT), University of the Western Cape (UWC), Stellenbosch University (SUN) and the Cape Peninsula University of Technology (CPUT) as the number of universities used. The data that is discussed below excluded the Cape Peninsula University of Technology, since the data for this institution was not available at the time of data collection. Hence, the total universities involved were three (3) with a sample of four hundred and twenty-one (421) researches between 1995 and 2004 (cf. section 3 for details).

However, it proved to be difficult as most research often focused on multiple and non-mutually exclusive topics. To do the classification, main topics were classified according to major areas by researchers (cf. table 3.5). Then, the final categories were constructed by the comparison of individual findings. Accordingly, it is observed that research conducted in the selected research addressed diverse set of topics (cf. section 3.1).

3. Results and Discussion of Findings

The distribution of sample was by period of completion of research study in various researches. The period was split into two halves (1995-1999 and 2000-2004) for both Doctor of Philosophy (PhD) and Master of Education (MEd) researchers. The first part of the results detailed specific diversity variables viz; language/medium of research, year of publication, race, gender and class. The intent was geared towards addressing sub research question. The second section addressed the research questions/hypothesis posed, with the objective of addressing the major research question.

3.1. Descriptive Statistics: Institutional Demography

This section deals with the total number of theses per institution and by degree per institution, generated between the years 1995 and 2004. The distribution as revealed in table 3.1 suggest that University of Cape Town (UCT) predominated (41%:171 out of 421). SUN produced a little over one-third (35%). Amongst the three research, UWC generated the least number of theses (24%:105 out of 421) which was almost one-half the number of UCT.

Table 3.1: Percentage of research per institution: Western Cape 1995-2004

Institution	% of research out put
UWC	24
SUN	35
UCT	41

In terms of distribution of doctoral and masters theses per institution, table 3.2 revealed that while UWC generated a total of 105, this accounted for 91 out of 105 M.Eds, while, the rest accounted for doctoral studies. Of the three institutions, it could be inferred from the table that the bulk of the theses produced were from the masters' students for both SUN and UCT. In a sharp contrast, the total number of doctoral students generated by SUN (34 out of 145) out numbered the sum of the other two research by 7, which was quite substantial in terms research output.

Figure 3.2: Masters and Doctoral Theses: Western Cape 1995-2004

|	MED	Doctorate	Total
UWC	91	14	105
SUN	**111**	34	145
UCT	**158**	13	171
TOTAL	360	61	421

Suggesting that with reference to the ratio of masters' theses to that of doctoral, the number of theses exceeded the number of doctoral degrees by 6:1 at UWC; 3:1 at SUN and 11:1 at UCT. This indicates a higher number of doctoral students at SUN and UWC than at UCT. On average, for every doctoral thesis, there are 5 masters' theses being produced. Suggesting the depth of research undertaken between the two levels.

With regards to language of dissertation, English predominates as the language of dissertation in the Western Cape. Regionally, English was (and still) is used in 89% of the theses and Afrikaans, 11%. The one thesis that was outside of the two language groups was French, which was archived at UCT.

With regards to the indexes of language of dissertation, the distribution between English and Afrikaans was in accordance with the historical context of the university. Afrikaans was the predominant home language[5] of the Coloured race. Noting that about one-tenth (12%) of Afrikaans theses produced in the Western Cape could have been written by Coloureds. The number of Afrikaans theses produced at Stellenbosch was more than one third of the total number of theses produced by institution (49 out of 96). These may have been produced by members of the White Afrikaans community.

On the other hand, the distribution in figure 3.2.1 of theses per institution showing year on year progression suggested a decline. The time span for the studies was divided in to two parts to facilitate a clearer picture of the progression of research output through the years. Noting that in this study, output is the end product of research, thus research would have begun, on average, 2 to 4 years earlier than the output shows. Figure 3.2.1 shows that research output remained fairly static in the first 5 year even dipping by approximately 10% in 1999 (from 48 to 39 theses).

**Figure 3.2.1 Distribution of theses per institution - Western Cape 1995-1999**

---
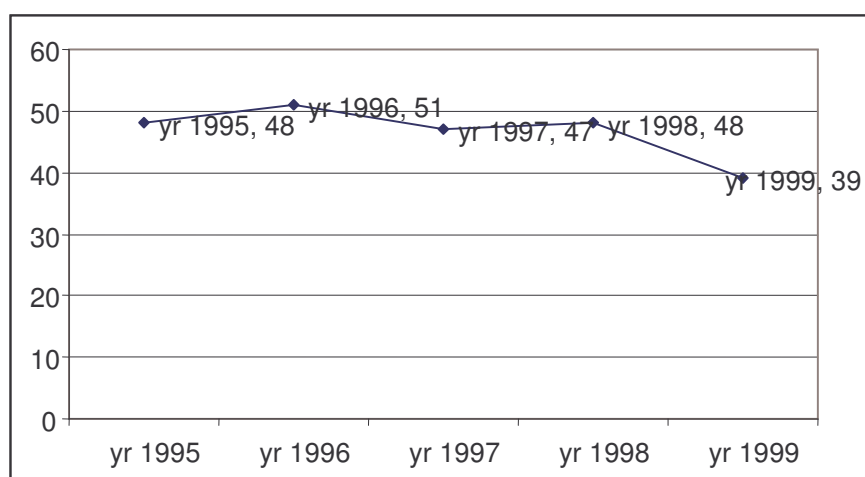
[5] There are 11 official languages in South Africa.

Following the Education Act of 1996 the number of postgraduate studies in education increased dramatically. These studies would have started a year or two earlier. An interesting comparison as a recommendation would be to see what was studied in which year. In fact, policy changed at a rapid rate, making some research invalid even before it was complete. For example, the policy of developmental appraisal (DAS) of educators (a policy widely researched, but which was changed by the unions within a matter of two years, additionally the implementation of Curriculum 2005 saw policy change even as research on the policy of 1997 – 2000 was being completed by individual post graduate student researchers**.**

Noting that the output at UCT showed the greatest increase over the first five years. It is commonly known that UCT has impressive international standing (it is rated amongst the best universities in Southern Africa) and with open access to higher education institutions post-1994, this would have been the University of Choice for many previously disadvantaged students. The next section looks at period of researches from 2000-2004.

This subsection elaborates on the distribution of theses per institution showing progression from 2000-2004. Figure 3.2.2 shows that the output in research in the Western Cape has more than doubled between 2000 and 2004, moving from 42 theses in 2000 to 79 theses in 2004.

Figure 3.2.2: Distribution of theses per year 2000-2004: Western Cape
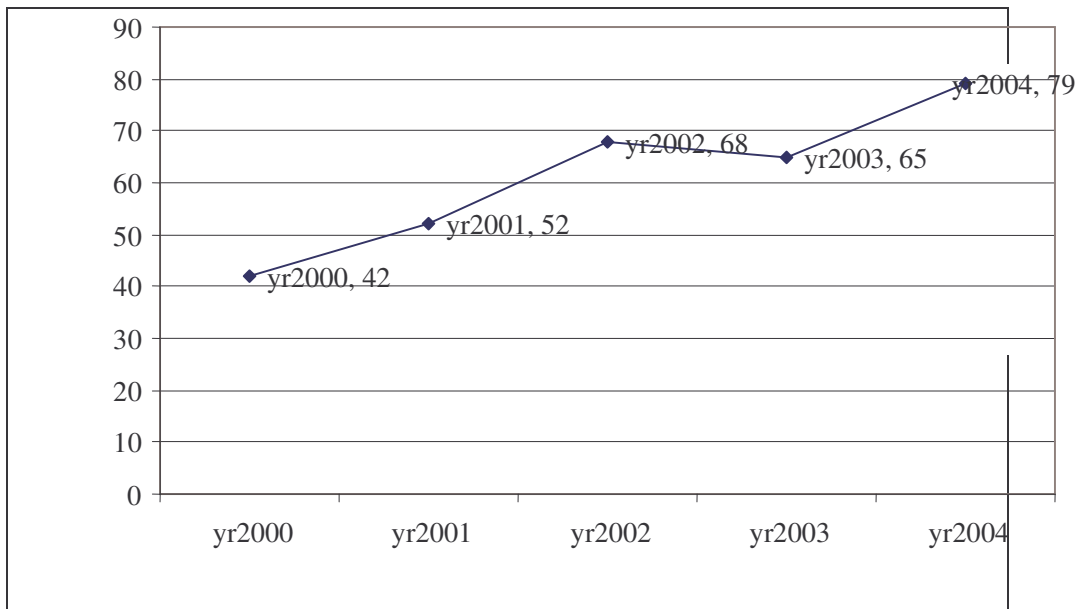
The increment in research output is exponentially greater at UWC than at the other two institutions, that is the number of theses produced has more than doubled. SUN has also shown a marked increase with an approximately 90% output.

Meanwhile, over the entire period (2000-2004), the number of theses produced has increased in a gradual but steady upward trend at each institution with minor depressions in certain years (cf. fig. 3.3 for details).

Figure 3.3: Distribution of theses per institution 2000-2004: Western Cape

YEAR	UWC	SUN	UCT	TOTAL
2000	12	19	11	42
2001	12	20	20	52
2002	12	33	24	68
2003	12	29	24	65
2004	28	34	17	64
TOTAL	**76**	**135**	**86**	**297**

Other variables of importance are as shown in table 3.3, which categorises the diversity concepts found in the theses per institution. Up to twelfth (12) categories namely: class, race, gender, age, religion, special education, language, culture, identity, feminism, ethnicity and HIV/AIDS, were identified in relation to trends in specific diversity concepts.

Of the 171 theses from UCT, they was 46(27%) studies in diversity issues, SUN had 37(26%) out of 145 and UWC had 30 (29%) out of 105. It was interesting to note that in terms of percentage, even though UWC has the least number of theses, it seems to be doing more research in diversity followed by UCT, while SUN has the least percentage (cf. table 3.4).

Table 3.4 suggest that majority of diversity research at UCT is on language issues (15), followed by age (14), race (6), class (5). The others are religion (3), identity (3), gender (2), inclusive/ special education (1), feminism (1) and culture (1). Nothing was found on ethnicity, HIV/AIDS and sexual orientation. Inclusive/special education research dominated at SUN with 16, followed by age with 15, culture 5, language 4, HIV/AIDS 1, identity 1. The dominant type of research at UWC was found to be in inclusive/special education with 11, followed by language with 7, age 3, race and gender and HIV/AIDS tie with 2 each, ethnicity and identity both have 1 and the rest have nothing. The cummulative results of the three univesities are as shown in table 3.4, in which case the most dominate diversity issue were age, inclusive education and language.

Table 3.4: Trends in specific diversity concepts in theses

| Diversity Variable	UCT	SUN	UWC
Class	5	0	0
Race	6	0	2
Gender	3	0	2
Age	**14**	**15**	**3**
Sexual Orientation	0	0	0
Religion	4	0	0
Inclusive/special Education	**2**	**16**	**11**
Language	**15**	**4**	**7**
Culture	1	5	4
HIV/AIDS	0	1	2
Identity	3	1	1
Feminism	1	0	0
Ethnicity	0	0	1

Other variable considered was race of authors as revealed in figure 3.5. This was because, desegregation of research occurred at the turn of the century, which was half way through the decade, specified in the study. This has impacted on the distribution of race per institution with a predominance of White authors at all institutions.

Figure 3.5: Distribution of Race of Authors: Western Cape 1995-2004

	Blacks	Coloureds	indians	Whites
Seri 1	27	6	1	66
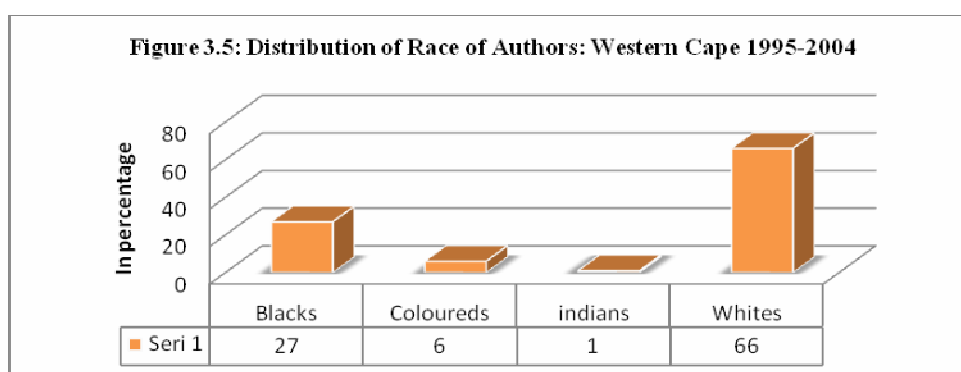
It should be noted that several names that appear 'white' may, in fact, be names of Coloured authors, especially at UWC – a historically black institution. The indexes in figure 3.5 show a broad trend in distribution of race. In figure 3.6, the distributions of gender of author are shown.

The line graph in figure 3.6 suggest that there are more female researchers than male researchers overall in the region. The male to female ratio at UWC is almost 50:50, but there are a greater number of females at both SUN and UCT.

Figure 3.6: Distribution of Gender of Author: Western Cape 1995-2004
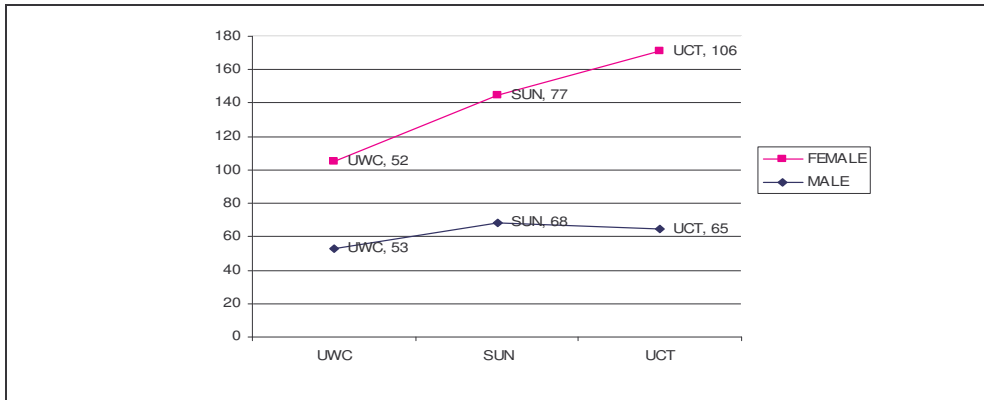
The preponderance of females was at the masters level, for instance 100 out of 158 at UCT were written by females at masters level and 8 at doctoral level, this quoted indexes contrasted sharply with that of males (cf. fig 3.5: where 33 males at master level and 5 males at doctoral level). The next section deals with inferential analysis.

3.2. Inferential analysis: non parametric analysis amongst postgraduate students

To examine the trend of application of non parametric data analysis usage by researchers over time, linear curve estimation technique was employed. The reviewed four hundred and twenty-one (421) researches spread over a time span of nine (9) years. In the analysis, time was accepted as the independent variable, while number of researches as the dependent variable. The descriptive statistics (R^2=0.98, F=21.960, p>.05) shows that the regression model is significant and explains 98 percent of the variance. This reflected a linear downward trend in the number of non parametric data analysis researches over time.

On the other hand, over seventy per cent of the researches were prepared with the data gathered by interviews, face-to-face and observations. According to of correspondence analysis and related chi-square value (121.41), there is a significant association between the research title and data analysis method. Hence, the null hypothesis was rejected accepting the alternate hypothesis. In fact, the inertia value is 0.624 and the first dimension explains 39.5 per cent of the total variance, thus, it is concluded that the analysis is significant and has high explanatory power.

Additionally, a Man-Whitney test was used to test diversity variable. The test indicated that Inclusive/special education was greater for females doctoral course (Mdn = 5) than for males who were doing doctoral programme (Mdn = 4), U = 60.5, p = .031, r = .34. Wilcoxon Signed-ranks Test used for two related samples was also applied. A Wilcoxon Signed-ranks test indicated that language was preferred more in female (Mdn = 0.87) than in male (Mdn = 0.66), Z = 3.20, $p < .001$, r = .69. The research was also examined according to their target populations {parents; 33.9 percent, managers (principals); 27.2 percent, and students; 27.2 percent}. Correspondence analysis points to a significant association between research titles and target populations (Chi square=81.454; inertia value=0.548; explained variance=51.8 percent).

It is evident from table 3.7 that interviews (22) were chosen as the main data collection technique. Other three following techniques are: face-toface (18*),* secondary data (10) and observation (12).

Table 3.7: Distribution of research by data collection methods

Face to face/administration	**18**
Mail survey	5
Experimental survey	0
Secondary data	**10**
Telephone Interview	1
Interview	**22**
Observation	**12**
E-mail	1
Panel data	3

It can evidently be seen also in table 3.8 that nonparametric data analysis was not the chosen as main analysis technique.

Table 3.8: Nonparametric Test Statistics

Spearman R	**0**
Man-Whitney test	0
Wilcoxon Signed-ranks test	0
Kendall Tau	0
Gamma coefficient	0

With reference to the hypothesis test, two constraints regarding application of NPA could be raised. It could be suggested foremost that sample size (cf. table 3.4) is a necessary and sufficient factor.

Firstly, for large data sets, nonparametric methods could be problematic. This is because, nonparametric methods are most appropriate when sample sizes are small. When the data set is large (for instance n > 100) it often makes little sense to use nonparametric statistics at all. Thus, when the samples become very large, then the sample means will follow the normal distribution even if the respective variable is not normally distributed in the population, or is not measured very well. Thus, parametric methods, which are usually much more sensitive (i.e., have more statistical power) are in most cases appropriate for large samples. However, the tests of significance of many of the nonparametric statistics described here are based on asymptotic (large sample-i.e 421)theory; therefore, meaningful tests can often not be performed if the sample sizes become too small as in the case of researches of PPER (cf. table 3.4).

Secondly, data gathered from students and employees (managers) were mainly used in. it is thus it is not easy to give simple advice concerning the use of nonparametric procedures among postgraduate students in this study. Each nonparametric procedure has its peculiar sensitivities and blind spots. For example, the Kolmogorov-Smirnov two-sample test is not only sensitive to differences in the location of distributions (for example, differences in means), but is also greatly affected by differences in their shapes. The Wilcoxon matched pairs test assumes that one can rank order the magnitude of differences in matched observations in a meaningful manner. None was found in the findings of the research of the postgraduate studies.

4. Summary and Conclusion

Two sets of findings emerged. Firstly, there are many general concepts used in the study of diversity. The most popular concepts studied in the three researches were found to be inclusive/special education for SUN and UWC and for UCT it was age which coincidentally is second for SUN. For UCT and UWC the second most popular concept studied was language and it was third most popular at SUN. Sexual orientation was not studied in any of the research. This is an interesting trend as it mirrors the fact that the South African community was still not open to talking about issues of sexual orientation. That can be attributed to the fact that, by and large, the community embraces Christian morals which romanticises the idea of one sexual orientation-heterosexuality.

Secondly, with regards to the major research question, classification of all the reviewed researches in terms of main research topics/areas, target populations, and data collection methods, has mapped the position of nonparametric analysis. Studies that did not use this analysis appeared most frequently (all researches in this case). It was also observed that researches conducted in the selected universities addressed a diverse set of topics. The data collection methods employed in the reviewed analysis applications included face to face, observations and interviews.

Also, the study discovered non usage trend of NPA by curve estimation in the period of 9 years. It was also observed that the analysis (NPA) was not successfully used for probing many concepts such as the diversity based studies. As a result non parametric has not been well used for analysis among postgraduate education researchers in the period of 1995-2004 in Western Cape of South African universities.

REFERENCES

BALFOUR R; MOLETSANE R; KARLSSON J.; PILLAY G; RULE, P.; NKAMBULE T; DAVEY B.; LEKENA L; MOLEFE S.; MADIYA N.; BENGESAI A.; GOBA B. (2008). *Project Postgraduate Educational Research in Education: Issues, Trends, and Reflections (1995-2004): A Progress Report (3)*: July, 2008. Faculty of Education-UKZN: KwaZulu-Natal.

CRESWELL J.W. (2008). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Pub.

HAIR J.F (2006). *Multivariate Data Analysis,* Prentice Hall.

HARRELL F. E. (2001). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression and Survival Analysis,*UK, Springer.

HEDEKER D. (2003). A Mixed-Effects Multinomial Logistic Regression Model. UK, John Wiley & Sons, Heights, MA: Allyn & Bacon.

HENDRICKX J. (2000). *Special Restrictions in Multinomial Logistic*, Stata Technical Bulletin Regression, 56, 18-26.

HOSMER D.W. AND LEMESHOW, S. (2000). *Applied Logistic Regression*. US, Wiley-Interscience.

HUCK S.W.; CORMIER W.H.; BOUNDS W.G. (2004). *Reading Statistics and Research*, Pearson.

JOHNSON R. A.; WICHERN D. W. (2007). *Applied Multivariate Statistical Analysis*. Upper Saddle Ltd, *Middle School, 6(6), 370 – 376.*

NEUMAN W.L. (2003). Social Research Methods: Qualitative And Quantitative Approaches, Pearson Education.

OTT L. AND LONGNECKER, M. (2008). *An Introduction to Statistical Methods and Data Analysis*. Duxbury Pr.

POWERS D.A. AND XIE, Y. (2008). *Statistical Methods for Categorical Data Analysis*, Emerald Group Publishing .

RUSSO L. M; PASSANNANTE M. R. C. (2001). Statistics Fever, *Mathematics Teaching*, (?).

SAUNDERS M.; THORNHILL A.; LEWIS P. (2009). "Research Methods for Business Students", *Financial Times/Prentice Hall.*

STERN D.T.; FROHNA A.Z.; GRUPPEN L.D (2005). "The Prediction of Professional Behaviour", *Medical education 39(1), pp 75-82.*

STOHL H. (2005(. *Probability in Teacher Education and Development,* in G. Jones (Ed) Exploring Probability İn Schools: Challenges for Teaching and Leaning (Pp. 345-366) New York: Springer.

RIVER NJ: Prentice Hall.

STOKES M.; DAVIS C.; KOCH G (2000). *Categorical Data Analysis Using the SAS*, SAS Publishing.

TABACHNICK B. G.; FIDELL L. S. (2001). *Using Multivariate Statistics*, (4th ed.). Needham.

YIN R. K. (2008). *Case Study Research: Design and Methods*, Sage Publications, Inc.