

Constructing a Criterion-Referenced Test for Measuring the Statistical Competencies of the Postgraduate Students in Education Colleges in Yemeni Universities

By : Mamun Ali Naji Qasem - Abdulrhman Altrairy - Punita Govil

Publication date / 21/ 12/ 2012

Abstract

This research has aimed at constructing Criterion Referenced Test to measure the statistical competencies of the Post-graduate Students in Education Colleges in Yemeni Universities, at examining the validity of the test's grades (the descriptive validity and the Domain Selection Validity), at examining the test's grades Reliability according to (Livingston). In addition, it aims at determining the adequate Cut-off – Scores for every statistical Competency through which one can identify those students who are skilful or unskillful of this Competency, and finally to know to what extent the Post-graduate Students in Education Colleges in the Yemeni Universities are in full command of using these statistical Competencies which are included in this test.

The researcher has used a method called Evaluation researches. The sample of the research has consisted of (157) students in the MA program. The researcher has constructed a test after following these stages (Analysis, test construction, try out, and print out the test)

After following these stages the study has produced a Criterion Referenced Test to measure the statistical competencies of the Postgraduate Students in Education Colleges in the Yemeni Universities. This test includes (77) items which measure seven statistical Competencies which seem to be adequate: (Determining the basic statistical concepts and idioms- using and inferring the descriptive statistics- determining and explaining the connection element- using the parametric statistical methods- using the nonparametric statistical methods- reading and explaining the inferred results from the statistical program SPSS- and selecting the adequate statistical style). It has also been concluded that this test has a descriptive validity as well as Domain Selection Validity. It has been found that the test has a high Reliability according to Livingston which has reached at (0.99). The adequate cutoff – Scores is (0.63). There was also a clear decrease in the statistical Competencies of the students of the Postgraduate Students in Education Colleges in the Yemeni Universities.

Introduction

With the growing recognition of the importance of statistics and its multiple uses in the various fields of theoretical and applied knowledge, and in response to the movement of overall change that targeted all stages of the educational process in most Universities, statistics was scheduled as a mandatory subject for most Faculty students as part of their program requirements; in order to enable students to analyze the raw data by converting it into figures that can be easily handled and analyzed, using the appropriate statistical tool or test. Thus, students get used to dealing with such programs and become fully prepared to interact with the statistical aspects of working life outside university classrooms, as well. (Rian, 2008)

In fact, it has become necessary for each postgraduate student to have the competencies of basic statistics that help him cope with the fast developments and rapid changes of the scientific works; especially in the current era. In the past, research studies relied mainly on the traditional tests in which much interest was given to comparing the individual performance to his peers without giving much attention to identifying the knowledge and competencies that students can't actually achieve. Although most of the Yemeni universities have been newly established, they made much effort to qualifying their teaching staff by means of running postgraduate programs in some of their faculties of education.

Objectives of Research

This research aims at:

- 1- Constructing Criterion Referenced Test to measure the statistical competencies of the Postgraduate Students in Education Colleges in the Yemeni Universities.
- 2- Examining the validity of the test's grades (the descriptive validity and a behavioral Domain Selection validity).
- 3- Examining the test's grades Reliability according to (Livingston)
- 4- Determining the adequate cutoff – Scores for every statistical Competency through which one can identify those students who are skillful or unskillful of this Competency
- 5- Knowing to what extent the Postgraduate Students in Education Colleges in the Yemeni Universities are in full command of using these statistical Competencies which are included in this test.

Method of Research

The researcher has used a method called Evaluation researches because it has seemed to be the most adequate method to fulfill the aims of the research.

Population and Sample of Research

The sample of the research has consisted of (157) students who have been taught some statistical courses in the first semester of the academic year (2010/2011) in four Yemeni Universities which include the MA program in the faculty of education. These universities are Sana'a, Aden, Tamar, and Ibb.

Tool of Research: The researcher has constructed a test after following these stages

First: Analysis Process: This process includes these steps:

- **Determining the behavioral Domain of the competencies.** This has been done by analyzing the description and the special aims of the courses of statistics which are presented to the MA students in the Yemeni Universities and also by seeking the assistance of some experts and specialists in teaching the course of psychological and educational statistics throughout distributing an open question about the statistical competencies which must be available with the students of higher studies in the Faculty of Education as viewed by the specialist, then by having a look at the permanent references and literature reviews of the psychological and educational statistics.

- **Analyzing the behavioral Domain (to determine the statistical competencies).** This has been done and resulted in seven statistical competencies which are: (determining the basic statistical concepts and idioms- using and inferring the descriptive statistics- determining and explaining the coefficient correlation- using the parametric statistical methods- using the imparametric statistical methods- reading and explaining the inferred results from the statistical program Spss- and selecting the adequate statistical style).
- **Determining the main aims** for each statistical Competency as each one has been expressed by one general main aim.
- **Translating the general main aims** into measurable and behavioral aims and organize them in accordance with the priority of each aim for the sake of achieving the main Competency. The behavioral aims are (83) distributed to the seven statistical competencies included in the test.

Second: the Constructing Process: it comprises three sub steps:

- Determine the bases of constructing the test which has been determined by eight bases.
- Wording some experimental items in order to measure the measurable and behavioral aims. Some multiple four-choice alternative questions have been used. (Choosing the correct answer).
- Judging the items. The experimental items have been shown to some experts and jurists in measurement, estimation and the educational statistics to judge the correspondence of the item with the behavioral aim and its capability of being estimated and presented for this behavioral aim. Those experts were (28).

Third: The Experimental Process: it includes the following steps:

- Applying the test for a sample in the form of a questionnaire. The research sample has consisted of (120) students, males and females.
- Analyzing the items of the test by finding out the corresponding point between the item and the behavioral aim which it estimates then to examine the difficulty of the item.
- Checking the validity of the test's grades by using the descriptive validity through the sincerity of the jury and the linguistic determination of how the experimental items are related to its behavioral aims then to find the coefficient correlation of the (individual) item by means of the whole grade of the Domain of this item. This happens by using (p.p) coefficient correlation and by counting the coefficient correlation of each Domain with the whole grade of the test by using the coefficient correlation (Person).
- The descriptive validity of the test's grades has been examined by using the indication of comparing two separate samples. This happened by determining the high group and the low group.

- The Domain Selection Validity has been examined by presenting a question to some professors who are majoring in teaching statistics course about the statistical Competency (which in case it is performed efficiently by the students it could be possible to issue a decision that he/she is very skillful at using statistics).
- The Reliability of the test's grades This has been achieved by using Livingston coefficient. It has been found that the test has a high Reliability according to Livingston coefficient which has reached at (0.99).
- Determining the adequate cutoff – Scores The researcher has used Nedlesky method to determine the adequate cutoff – Scores for every statistical Competency through which one can identify those students who are skillful or unskillful of this Competency. The adequate cutoff – Scores were (0.63)

Fourth: to print out the final test and apply it to a sample of (157) students in the four Yemeni Universities which have been mentioned above.

Results of Research

After following these steps to construct the test and after analyzing the data from the First and the last application by using adequate statistical methods, it has been concluded that.

- 1- The study has produced a Criterion Referenced Test to measure the statistical competencies of the Postgraduate Students in Education Colleges in the Yemeni Universities. This test includes (77) items which measure seven statistical competencies which seem to be adequate: (Determining the basic statistical concepts and idioms- using and inferring the descriptive statistics- determining and explaining the coefficient correlation- using the parametric statistical methods- using the imparmetric statistical methods- reading and explaining the inferred results from the statistical program Spss- and selecting the adequate statistical style).
- 2- The test has a descriptive validity as well as a behavioral Domain Selection valid based on some indications which have been used by the researcher.
- 3- The test has a high Reliability according to Livingston coefficient which has reached at (0.99).
- 4- The adequate cutoff – Scores have been found as that seen by the experts (0.63).
- 5- There was a clear decrease in the statistical competencies of the Postgraduate Students in Education Colleges in the Yemeni universities let alone in the parametric and imparmetric statistics and using the statistical program SPSS.

Recommendations

- 1) The professors, whose major is statistics in the Yemeni Universities, ought to update and develop their curriculums so that they could include all the important statistical subjects which are needed by the students of higher studies.



- 2) The researcher recommends that all the statistics professors ought to use the test which has been conducted in this current study in order to be able to diagnose how far their students are skillful at the statistical competencies
- 3) The researcher recommends the deanships of the Education Faculties to use this test as an admission test for enrolling the (PhD) program in any major in the Faculty of Education since these faculties prepare researchers and educational professors.

References

- 1- Abalone, R. M. (2004). *Research Methodologies in Psychological and Educational Sciences* (4th Ed.). Cairo: Universities Press Publishers.
- 2- AbuHashem, A. M. (2004). *The Statistical Guide in Analyzing Data Using SPSS*. Riyadh: Al-rushed Bookstore Publishers .
- 3- Abulabdah, S. M. (1996). *Principles of Psychological Evaluation and Educational Evaluation* (4th Ed.). Amman: Association of Cooperative Workers Press.
- 4- AbuSaleh, M. S. (2009). *Statistical Methods*. Amman: Dar Al-yazory for Publishing.
- 5- Abuzianah, F. K. (1998). *Basics of Measurement and Evaluation in Education* (2nd Ed.). Kuwait: Al-Falah Bookstore for Publishing and Distribution.
- 6- A'ddas, A. (1987). *Principles of Analytic Statistics* (Vol. II). Amman: Dar Alfekr.
- 7- A'ddas, A. (1997). *Principles of Statistics in Education and Psychology: Analytic Statistics* . Amman : Dar Alfekr for Publishing and Distribution
- 8- Al'assaf, S. A. (1995). *An Introduction to Research in Behavioral Sciences* . Riyadh: Alobaikan Bookstore.
- 9- Al-dawsry, E. M. (2001). *A Reference Guide for Educational Evaluation* (3rd Ed.). Riyadh: Arabian Education Department of Gulf Countries.
- 10- Alkassas, M. M. (2007). *Principles of Statistics and Social Measurement* . Almansorah: Almansorah University Press, Eyygpt.
- 11- A'llam, S. M. (2000). *Analyzing Data of Psychological, Educational and Social Researches*. Cairo : Arabian Dar Alfekr.
- 12- A'llam, S. M. (2005). *Criterion Diagnosis Tests in Educational, Psychological and Training Fields* . Cairo: Arabian Dar Alfekr.
- 13- A'llam, S. M. (2005). *Parametric and Non-Parametric Statistical Methods* .: Arabian Dar Alfekr .
- 14- A'llam, S. M. (2006). *Educational and Psychological Tests and Measures* (1st Ed.). Cairo: Arabian Dar Alfekr.
- 15- Anger, A. O. (1422 AH). The problem of Identifying the Suitable Statistical Method in Humanities Studies and Researches by Teaching Staff at King Faisal University. *King Saud University Journal of Educational Sciences and Islamic Studies* , 545-585.

- 16- Altarery, A. S. (1997). *Psychological and Educational Measurement: Theories, Baiscs and Applications* (1st Ed.). Riyadh: Alroshd Bookstore.
- 17- Alyan, R. M. (2001). *Scientific Research: Baiscs, Methodologies, Methods and Procedures* . Amman: House of International Ideas .
- 18- Assayed, F. A. (2005). *Statistical Psychology and Meausing the Human Mind* . Cairo: Dar Al-Fikr Al-Araby.
- 19- Audah, A. (2010). *Measurement and Evaluation in Teaching Process* (2nd Ed.). Amman: Dar Al'amal (hope publishing).
- 20- Audio, A., & Alkhalily, K. Y. (1988). *Statistics for a Researcher in Education and Humanities Sciences* . Amman : Dar Alfekr Publishers .
- 21- Carlson. B. (2002). Preparing Workers for the 21ST Century: The Importance of Statistical Competencies. United Nations Economic Commission for Latin America and the Caribbean (ECLAC)
- 22- Greene. J & Oliveira. M. (One thousand nine hundred ninety-nine). Learning to Use Statistical Tests in psychology. [Http://www.openup.co.uk](http://www.openup.co.uk)
- 23- Luis M. Lozano. (2008). Effect of the Number of Response Categories on the Reliability and Validity of Rating Scales. *Methodology*. www.eric.com
- 24- Messick, S. (1994) The Interplay of Evidence and Consequences in the Validation of Performance Assessment. *Educational Researcher*, 23,2 13-23
- 25- Rian, A. A. (2008). *Tests and Measurements in Psychological and Educational Sciences* . Cairo : Modern Dar Alketab.
- 26- Taylor. S. (1993). An exploratory study of the statistical competencies of certain future elementary and middle school teachers. Dissertation (Doctor of philosophy) Texas A. M University
- 27- Brian. D, (2009). Criterion Referenced Test Is not A mastery test. www.mountainmeasurement.com
- 28- - Beggs. D & Lewis. (1975). *Measurement and Evaluation in the school* . Educational and Psychological Measurement, 60 (3), pp. 361- 370
- 29- Carlson. B .(2002) . Preparing Workers For The 21ST Century: The Importance of Statistical Competencies. United Nations Economic Commission for Latin America and the Caribbean (ECLAC)
- 30- Greene . J &Oliveira. M. (1999). Learning to Use Statistical Tests in psychology. <http://www.openup.co.uk>
- 31- Chase. C. (1987). *Measurement in Educational Evaluation*. London: Addison Wesley Publishing. Co., 2nd Ed www.eric.com
- 32- Daniel. L & Michalis. p. (2005). Setting Cut Scores for Collega Placement. www.collegeboard.com
- 33- Hambleton. P & Simon , A , (2000). Steps for constructing Criterion referenced test, Laboratory of psychometric and Evaluative Research, www.eric.com
- 34- Hambleton. P & Novick. M, (1978). Toward an integration of theory and Method for Criterion Referenced Tests , *Educational Testing and Assessment: Lessons From the Past Directions for the Future*, International Journal of Research, 27 (5), 379-393

- 35- Halpin , g. & Sigmon g. (1983). Minimum Competency Standard set of three Divergent Groups of raters using three Judgmental procedures : Implications for Validity. Educational & Psychological Measurement. 43: PP185-196.
- 36- Keith. A, & Giert. J. (2000). Automated Test assembly Procedures for Criterion Referenced Testing using optimization heuristics Paper presented at the annual meeting of the AERA. New Orleans Louisiana, USA
- 37- Kumazawa. T (2009). Criterion Referenced Test Administration designs and analyses. Tohoku Bunka Gakuan. University. pp 65-74
- 38- Murphy .K. & David. C, (2001). Psychological Testing: Principles and Applications, New Jersey: Prentice Hall.
- 39- Mylonas. A & wagon. C. (2004). The Journey to a Criterion Referenced as University, Teaching and Learning Support Support Services Queensland University of Technology
- 40- Nicely. L (2009). Cut Off Scores: THE Basic Angoff Method and The Item Response Theory Method. International Journal of Educational. 1(1). P 39-47
- 41- Patricia, M . (2003). What Competencies Should be in Program: American Association of colleges for teacher education, www.eric.com
- 42- Patricia. C& Barbara. H (1987). Norm-Referenced and Criterion- Referenced Test. Journal of measurement. 67(12) 873-898
- 43- Pester. M. (2003) Language Intervention Effects on Norm referenced. Criterion referenced test Scores. Submitted in the Partial Fulfillment of The Requirements for The degree of Master of Arts. Department of Speech Pathology and Audiology. The faculty of Miami University, Oxford. Ohio
- 44- Peter. S & Carolyn. A: (2000) Cut Scores: Results May Vary , the National Board on Educational Testing and Public Policy, v 1, n 1, p1- 31
- 45- Popham. W. (2003) .Test Better Teach Better, International Journal of Testing , 9 (2), pp. 3–19
- 46- Popham. W. & Husek. T. (1969). The implication of Criterion referenced measurement. Journal of Educational measurement. 6(1).
- 47- Popham, W. J. (1978). Criterion Referenced Measurement. www.eric.com
- 48- Sharon , A. &William , C. (2007) .Criterion Referenced Test development : technical and legal guidelines for corporate training. 3rd, www.eric.com
- 49- Susan. L. (2011): Evaluating the Bookmark Standard Setting Method: The Impact of Random Item Ordering. International Journal of Testing. 11 (1), p 24- 37
- 50- Taylor. S .(1993). An exploratory study of the statistical competencies of certain future elementary and middle school teachers. Dissertation (Doctor of philosophy) Texas A. M University