International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:3, No:6, 2009

On the Comparison of Several Goodness of Fit tests under Simple Random Sampling and Ranked Set Sampling

F. Azna A. Shahabuddin, Kamarulzaman Ibrahim, and Abdul Aziz Jemain

Abstract**—**Many works have been carried out to compare the efficiency of several goodness of fit procedures for identifying whether or not a particular distribution could adequately explain a data set. In this paper a study is conducted to investigate the power of several goodness of fit tests such as Kolmogorov Smirnov (KS), Anderson-Darling(AD), Cramer- von- Mises (CV) and a proposed modification of Kolmogorov-Smirnov goodness of fit test which incorporates a variance stabilizing transformation (FKS). The performances of these selected tests are studied under simple random sampling (SRS) and Ranked Set Sampling (RSS). This study shows that, in general, the Anderson-Darling (AD) test performs better than other GOF tests. However, there are some cases where the proposed test can perform as equally good as the AD test.

Keywords**—**Empirical distribution function, goodness-of-fit, order statistics, ranked set sampling.

I. Introduction

GOODNESS of fit tests have been applied in many areas of research. Goodness of fit tests (GOF) measure the degree of agreement between the distribution of an observed sample data and a theoretical statistical distribution. The problems involve a comparison of the empirical distribution function (EDF) for a set of ordered observations of size *n*, say $F_n(x_{(i:n)})$, with a particular theoretical distribution with known parameters, denoted as $F_0(x_{(i:n)})$. The problem can be formulated under the test of hypothesis involving $H_0 : F(x) = F_0(x_{(i:n)})$ where F_0 is the hypothesized continuous cumulative distribution function (cdf) with known parameters against $H_1 : F(x) \neq F_0(x_{(i:n)})$. The standard practice of GOF test is that the observations are sampled based on a simple random sampling (SRS) procedure.

Another good and efficient sampling procedure which has received numerous attention in the current statistics literature is ranked set sampling. This sampling procedure was first introduced by McIntyre in 1952 [1] in his effort to find an efficient method to estimate the yield of pastures [2].

From his study, he found that RSS was more efficient and cost effective than the commonly used simple random sampling, particularly when visual ordering of sample units can be done easily and cheaply but the actual measurement of the sample units are expensive and difficult. After his first investigation, there is almost no application of RSS by any researchers until it was rediscovered by Halls and Dale [3]. They found that the estimator for the population mean based on RSS is more efficient that SRS. Takahashi and Wakimoto [4] was the first to provide the mathematical theory of RSS. They found that when ranking is perfect the sample mean based on RSS is an unbiased estimator of the population mean. The same result was later obtained by Dale and Clutter [5]. RSS has received a lot of interest from various researchers and recently there have been many development in the theories and methodologies of RSS, see for example Chen [6], Stokes & Sager [7], Patil [8], Patil et.al [9], Bai & Chen[10] and Jemain et.al [11].

In this paper, the performance of several goodness of fit tests such as Kolmogorov Smirnov (KS), Anderson-Darling(AD), Cramer-von-Mises (CV) are investigated. In addition a modified KS goodness of fit test which incorporates a variance stabilizing transformation (FKS) is proposed. The performances of these selected tests are studied under two sampling techniques which are Simple Random Sampling (SRS) and Ranked Set Sampling (RSS).

II. Ranked Set Sampling Procedures

The RSS procedure as suggested by McIntyre [1] is as follows:

To obtain a RSS sample of size *k*, *k* sets of SRS samples each of size *k* are selected from the target population. The units within each set are then rank with respect to a variable of interest by visual judgment or any other inexpensive ranking mechanism but not involving actual measurements of the variable. From the first sample, the smallest ranked unit is selected and the actual measurement is made on the variable of interest denoted as $X_{(1:k)}$. Then the second SRS of size *k* is selected from the population and ranked without actual measurement as before. From this sample, the second smallest rank unit is selected and actual measurement is made on the variable of interest denoted as $X_{(2:k)}$. The process is continued until from the *k*-th sample, the *k*-th ranked unit is selected and measurement is taken, denoted as

F. Azna A.Shahabuddin is with the School of Mathematical Sciences, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia. (603-8921 5893, fax:603-8925 4519; e-mail: fazna99@gmail.com).

Kamarulzaman Ibrahim is with the School of Math. Sciences, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia. (e-mail: kamarulz@ukm.my).

Abdul Aziz Jemain is also with the School of Math.Sciences, Universiti Kebangsaan Malaysia,43600 Bangi, Selangor, Malaysia. (e-mail: azizj@ukm.my).

The authors acknowledge the financial support received from UKM research grants, UKM-GUP-BTT-07-25-168 and UKM-GUP-TK-08-16-061.

$X_{(k:k)}$. In a cycle of getting a ranked set sample of size *k*, sample units of size k^2 have actually been considered. The cycle can be repeated many times, and if the cycle is repeated *m* times, the final data sets for RSS would be of size *n=mk*, given as follows:

$$\begin{array}{l} X_{(1:k)1}\ X_{(2:k)1} \ldots X_{(k:k)1} \\ X_{(1:k)2}\ X_{(2:k)2} \ldots X_{(k:k)2} \\ \ldots \quad \ldots \quad \ldots \\ X_{(1:k)m}\ X_{(2:k)m} \ldots X_{(k:k)m} \end{array}$$

The random variables in each row are the order statistics associated with the SRS observations, which can be written as $X_{(1:k)\,i} \le X_{(2:k)\,i} \le \ldots \le X_{(k:k)\,i}$ for $i = 1, 2, \ldots m$. The probability density function (pdf) of the *r*- th order statistic $X_{(r:k)}$ for a SRS of size *k* is denoted as :

$$f_{(r:k)}(x) = \frac{k!}{(r-1)!(k-r)!}\left[F(x)\right]^{r-1}[1-F(x)]^{k-r} f(x) \quad (1)$$

where $f(x)$ and $F(x)$ are the probability density function and cumulative distribution function for a random sample $X_1, X_2, ..., X_k$ respectively. We can easily verify the relationship between $f_{(r:k)}(x)$ and $f(x)$, see for example [2], as follows:

$$f(x) = \frac{1}{k}\sum_{r=1}^{k} f_{(r:k)}(x) \quad (2)$$

Also the following fundamental equality holds for all *x*:

$$F(x) = \frac{1}{k}\sum_{r=1}^{k} F_{(r:k)}(x) \quad (3)$$

where

$$F_{(r:k)}(x) = \int_0^{F(x)} \frac{\Gamma(k+1)}{\Gamma(r)\Gamma(k-r+1)} t^{r-1}(1-t)^{k-r}\,dt \quad (4)$$

III. Goodness of Fit Procedures

In order to test whether a ranked set sample, comes from a particular distribution, for example a standard normal $N(0,1)$, the null hypothesis $H_0 : F(x) = N(0,1)$ is tested against the alternative hypothesis $H_1 : F(x) \neq N(0,1)$. The GOF procedure is described as the following:

The RSS for the *i*-th cycle $X_{(r:k)i}$, where *r=1,2,…,k* and *i=1,2,…,m* is generated assuming a particular cdf of *F(x)*. These observations are then ordered from smallest to largest, denoted as $y_{(1:n)}, y_{(2:n)}, ..., y_{(n:n)}$ where *n=mk*. The empirical distribution function (EDF) is defined as:

$$F_n(y_{(i:n)}) = \frac{i}{n+1} \quad (5)$$

where $i = 1, 2, \ldots n$. The EDF values will be compared to the theoretical distribution of ordered observations based on:

$$F(y_{(i:n)}) = \frac{1}{k}\left(F_{(1:k)}(y_{(i:n)}) + F_{(2:k)}(y_{(i:n)}) + ... + F_{(k:k)}(y_{(i:n)})\right) \quad (6)$$

for all $i = 1, 2, \ldots n$. To study the degree of discrepancies between the EDF and the theoretical distribution, there are various GOF statistics in the literature that have been used. The GOF tests that are of particular interest in this study include the Kolmogorov-Smirnov (KS), given by Kolmogorov and Smirnov[12], Anderson-Darling (AD), Cramer-von-Mises (CV) and the proposed modification of KS which incorporates variance stabilizing transformation (FKS). The popular KS test is defined as:

$$KS = \max(D^+, D^-) \quad (7)$$

where $D^+ = \max\left(F(y_{(i:n)}) - \frac{i}{n}\right)$ and

$D^- = \max\left(F(y_{(i:n)}) - \frac{i-1}{n}\right)$.

The respective Anderson-Darling and Cramer-von-Mises tests are define as:

$$AD = -\frac{n}{2}\sum_{i=1}^{n}\left[\begin{array}{l}(i-0.5)\log F\left(y_{(i:n)}\right) + \\ (n-i-0.5)\log\left(1-F\left(y_{(i:n)}\right)\right)\end{array}\right] - n \quad (8)$$

and

$$CV = \sum_{i=1}^{n}\left[F\left(y_{(i:n)}\right) - \frac{i-0.5}{n}\right]^2 + \frac{1}{12n} \quad (9)$$

Another alternative for GOF test is offered beside the statistics shown in equations (7) to (9). Variance stabilizing transformation is incorporated to the original modified KS of Green and Hegazy [14]. The proposed modified statistics which is called FKS is defined as:

$$FKS = \max\left|\sin^{-1}(F(y_{(i:n)})) - \sin^{-1}(\frac{i}{n+1})\right| \quad (10)$$

IV. Results

A simulation study was carried out to test the hypothesis $H_0 : F(x) = N(0,1)$ against $H_1 : F(x) \neq N(0,1)$ under SRS and RSS. The following alternative hypotheses are considered under both sampling techniques against the null hypothesis to allow for differences in locations and variances in the contending distributions.

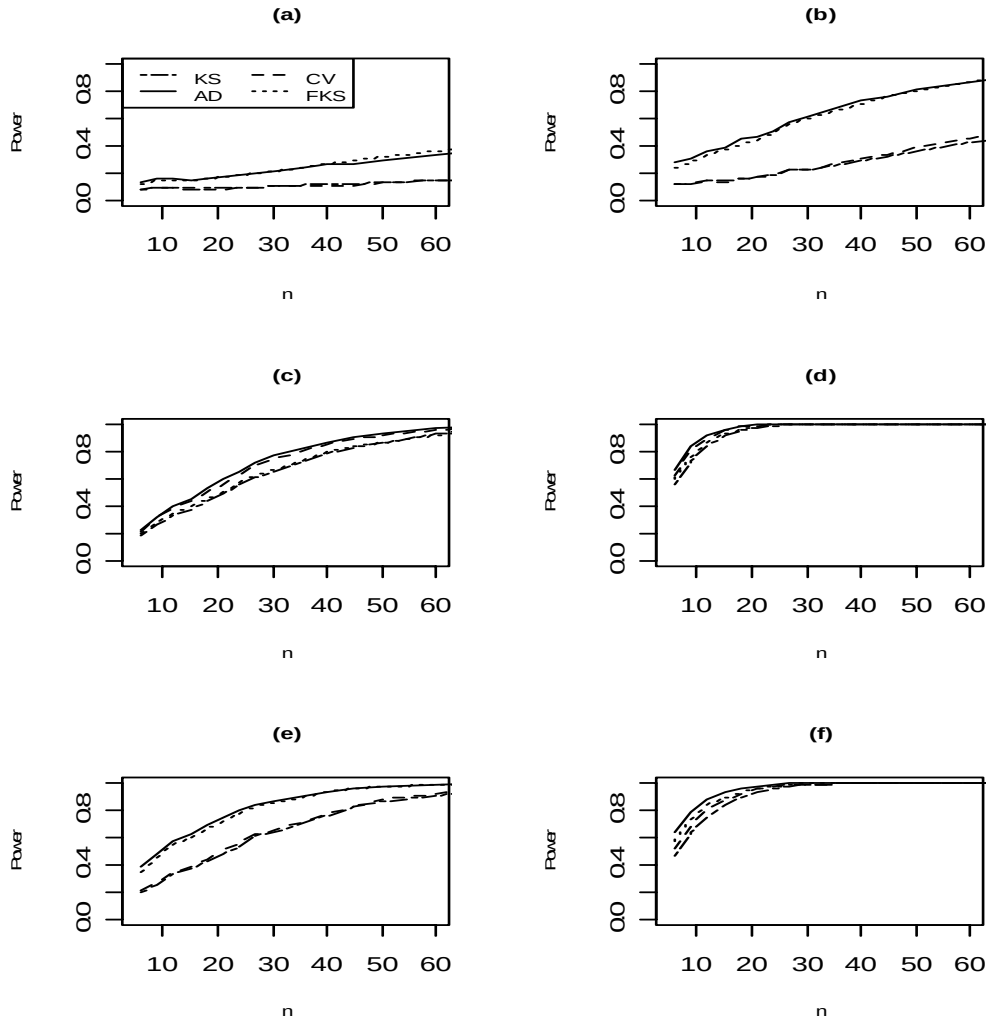

Fig. 1 Power comparison for normality test for various alternative hypotheses under SRS

The alternative hypotheses that are considered are as follows:

(a) N(0, 1.25)
(b) N(0, 1.5)
(c) N(0. 5,1)
(d) N(1,1)
(e) N(0.5,1.5)
(f) N(1,1.25)

The following results as shown in figures Fig. 1 (a) to Fig. 1 (f) are found based on simulation study under SRS. In the case of same location but differing in variances as opposed to the null hypothesis as shown in Fig. 1 (a) and Fig. 1 (b), the performance of FKS is found to be almost as powerful as AD test. The same results are observed in the cases of Fig. 1(e) and Fig. 1 (f), where the differences are due to the different values of both the mean and variance. AD is found to be most powerful and this is quite closely followed by FKS. In the case of allowing for the difference in location but same variance, the results as shown in Fig.1(c) and Fig. 1 (d) indicate that AD is slightly more powerful but is again followed closely by FKS.

For the purpose of comparison, simulation results under RSS, are obtained by setting k=3 and m=2, 3, 4, 5, 6, 7, 8, 9, 10, 15 and 20 to obtain the sample sizes $n = mk$ = 6, 9, 12, 15, 18, 21, 24, 27, 30, 45 and 60. The results are shown in figures Fig.2(a) to Fig.2(f). Under RSS, the Anderson-Darling test has a better performance in all cases investigated. However FKS still remains as the second best test when compared to other tests.

V. Conclusion

In this paper a new GOF tests which incorporates variance stabilizing transformation is introduced. The proposed modified GOF test, i.e. FKS is found to perform almost as powerful as the well known AD in SRS. Under RSS, AD performs better than FKS. However, in all the cases considered, FKS always outperform the original KS and CV.

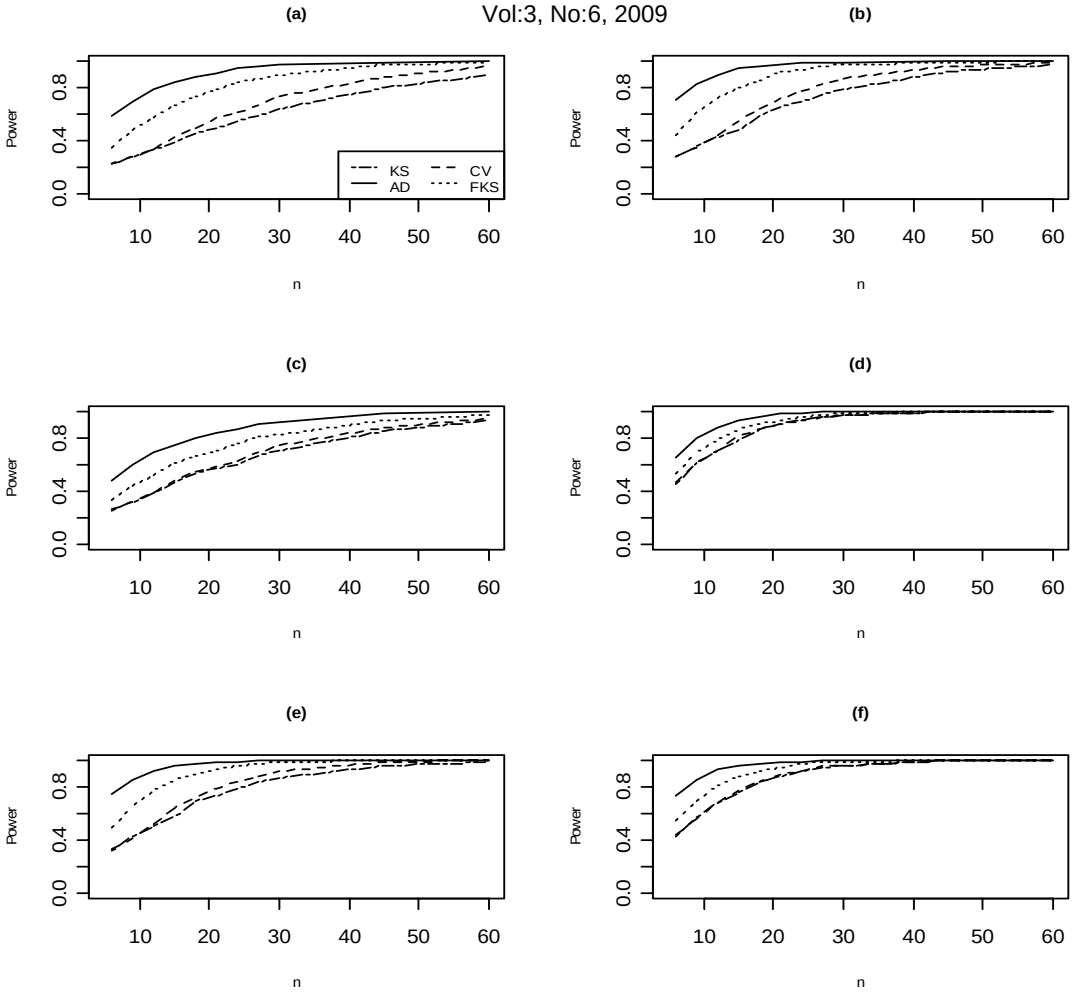

Fig. 2 Power comparison for normality test for various alternative hypotheses under RSS

REFERENCES

[1] G.A. McIntyre, "A method for unbiased selective sampling using ranked sets", *Australian Journal of Agricultural Research*, No.3, 1952, pp.385-390.

[2] Z. Chen, Z. Bai & B.K. Sinha. *Ranked Set Sampling: Theory and Applications*. Springer. 2004.

[3] L.S.Hall and T.R.Dell, "Trial of ranked set sampling for forage yields". *Forest Science*, No. 22, 1966, pp. 22-26.

[4] K. Takahashi and K. Wakimoto, "On unbiased estimates of the population mean based on the sample stratified by means of ordering", *Annals of the Institute of Statistical Mathematics*, No. 20, 1968, pp.1-31.

[5] T.R. Dale and J.L. Clutter, "Ranked set sampling theory with order statistics background", *Biometrika*, No.28, 1972, pp.545-555.

[6] S.H. Chen, "Ranked set sampling theory with selective probability vector", *Journal of Statistical Planning and Inference*, No. 8, 1983, pp. 161-174.

[7] S.L. Stokes and T.W. Sager, "Characterization of a ranked set sample with application to estimating distribution functions", *Journal of the American Statistical Association*, No. 83, 1988, pp. 374-381.

[8] G.P. Patil. Editorial: "Ranked set sampling", *Environmental and Ecological Statistics*, No.2, 1995, pp.271-285.

[9] G.P. Patil, A. K. Sinha, and C. Tallie, "Ranked set sampling: a bibliography", *Environmental and Ecological Statistics*, No.6, 1999, pp.91-98.

[10] Z.D. Bai and Z. Chen, "On the theory of ranked set sampling and its ramifications". *Journal of Statistical Planning and Inference*. No. 109, 2003, pp. 81-99.

[11] A.A. Jemain, A. Al-Omari, and K. Ibrahim, "Modified ratio estimator for the population mean using double median ranked set sampling", *Pakistan Journal of Science*, No. 24(3), 2008, pp. 217-226.

[12] X12. R.B. D'Agostino and M.A. Stephens, *Goodness of fit techniques*, Marcel Dekker, New York, 1986.

[13] J. Zhang, "Powerful goodness of fit tests based on the likelihood ratio". *Journal of Royal Statist. Soc. B*, No. 64, Part 2, 2002, pp. 281-294.

[14] J.R. Green and Y.A.S. Hegazy, "Powerful Modified Goodness of fit tests", *Journal of the American Statistical Association*, Vol. 71, No.353, 1976, pp. 204-209.