Application of Ranked Set Sampling to Normality Tests

A. Erem and G. Arslan*

Izmir University of Economics, Department of Mathematics, Turkey

Abstract: The normality assumption is used in many statistical analyses and is also a fundamental concept in statistics. Because of this there are many statistical tests for testing the normality assumption. Two of the most primitive ones are the *R* and *Z* tests. The main aim of this study is to investigate the application of ranked set sampling to these tests. Therefore, by using the idea of ranked set sampling, modifications of the *R* and *Z* tests are considered. Using simulations the results of these new tests are discussed.

Keywords: Simple random sampling, ranked set sampling, normality tests, R test, Z test.

1. INTRODUCTION

In many applications of statistics and statistical tests the normality assumption is of vital importance. One of the most primitive but conceptually simple univariate normality tests is the *R* test. It combines two fundamentally easy concepts: the probability plot and the correlation coefficient. Filiben [1] introduced the probability plot correlation coefficient test statistic. Although it is an old test, it is widely used because the logic behind this test is simple. Another important test for testing normality, suggested by Tiku [2, 3] is a test based on generalized spacings. It was shown that Tiku's test based on generalized spacings is more powerful than the *R* test against skew alternatives but less powerful against symmetric alternatives [2].

The aim of this study is to investigate the performance of these tests when ranked set sampling is used instead of simple random sampling. Ranked set sampling (RSS) is a data collection technique that results in a collection of measurements that are more likely to span the range of values in the population compared to a simple random sample [4]. This approach is especially useful and appropriate when ranking is cheap compared to actual measurement of observations.

The main objective of this study is to use this fact and apply it in goodness-of-fit testing for the normal distribution. Since ranked set sampling is expected to represent the population more accurately than simple random sampling it seems natural to investigate whether the performance of the *R* and *Z* tests may be improved by using ranked set sampling.

*Address correspondence to these authors at the Izmir University of Economics, Department of Mathematics; Tel: +90 232 4889858; Fax: +90 232 2792626; E-mail; guvenc.arslan@gmail.com

The paper is organized as follows. In Section 2 and 3 we present the basic idea and concepts of the R and Z tests, respectively. After presenting the R and Z statistic in detailed, in Section 4 the basic idea about ranked set sampling is given. Then simulation results that are composed in R programme are introduced in Section 5 and 6.

2. THE PROBABILITY PLOT CORRELATION COEFFICIENT TEST FOR NORMALITY (R TEST)

Let $Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n}$ denote the order statistics of a sample from a distribution whose cumulative distribution function (cdf) is of the form $F\left[(y-\mu)/\sigma\right]$,where μ and σ are the location and scale parameters, respectively. To construct a probability plot, the sample order statistic, $y_{i:n}$, is plotted (usually on the vertical axis) against $x_{i:n} = F^{-1}(p_i)$ (usually on the horizontal axis), where p_i is an estimate of $F\left[(y_{i:n}-\mu)/\sigma\right]$. This estimate is called the plotting position [5].

There are lots of plotting positions but the most common ones are given below (see [6, 7] and [8] respectively)

$$p_i = (i-0.5)/n \tag{1}$$

$$p_i = i/(n+1) \tag{2}$$

$$p_i = (i-0.375)/(n+0.25) \tag{3}$$

Looney S.W. and Gulledge T.R (see [5]) showed that the Blom' s plotting position yields a correlation coefficient test that is more powerful than Filiben's, Hazen' s and Weibull' s plotting positions.

The correlation statistic is simply

$$R = 1-\hat{\rho}^2 \qquad (0 < R < 1), \tag{4}$$

where $\hat{\rho}$ is the ordinary correlation coefficient between $y_{i:n}$ and $\mu_{i:n}$ (see [1] and [9]). Here, the $Y_{i:n}$ $(1 \le i \le n)$ are the order statistics of a random sample of size n from the normal distribution and $Z_{i:n} = (Y_{i:n} - \mu)/\sigma$ are the corresponding standardized normal order statistic. $\mu_{i:n}$, $(1 \le i \le n)$ are the expected values of the standardized normal order statistics, i.e, $\mu_{i:n} = E[Z_{i:n}]$. The correlation coefficient statistics can be used for testing any assumed density of type $(1/\sigma) f((y-\mu)/\sigma)$. For ease of computation, in several studies, values of $\mu_{i:n}$ are obtained by the population quantiles

$$\mu_{i:n} = F_0^{-1}(i/n+1), \ (1 \le i \le n),$$

where $F_0(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{z^2}{2}} dz$ is the cumulative distribution function of the standard normal distribution. The expected values of the standardized normal statistics $(\mu_{i:n})$ for some particular i and n values can be obtained from Harter [10]. Details on how the $\mu_{i:n}$ are calculated in this study are given in section 5. Note that the null distribution of R in general is not known for a given density f_0. Its percentage points have to be determined empirically by Monte Carlo simulation. The critical values for $\alpha = 0.1$ and some values of n are given in Table **1**.

Table 1: Critical Values for the *R* Test (α=0.1)

n	Simpson's Rule	Blom's Plotting Position
10	0.933	0.934
20	0.960	0.960
50	0.981	0.981

3. THE Z TEST

Let $Y_{1:n}, Y_{2:n}, ..., Y_{n:n}$ be the order statistics of a random sample of size n from an exponential distribution with parameters μ and σ, i.e., $Y_i \sim E(\mu, \sigma)$: $f(y) = \frac{1}{\sigma} e^{-\frac{(y-\mu)}{\sigma}}$, $\mu < y < \infty$. It is well known that the spacings

$$D_i = (n-i+1)(y_{i+1:n} - y_{i:n}), \quad 1 \le i \le n-1. \tag{5}$$

are also exponential. Using this fact, Tiku [2,3] defined the statistic

$$Z_E = \frac{2\sum_{i=1}^{n-1}(n-1-i)D_i}{(n-2)\sum_{i=1}^{n-1}D_i}. \tag{6}$$

To test whether a data set comes from an exponential population, he showed that the null distribution of $Z_E/2$ is exactly the same as the distribution of the mean of $n-2$ independent and identically distributed (iid) uniform $(0,1)$ random variates. Therefore

$$E(Z_E) = 1 \text{ and } V(Z_E) = \frac{1}{3(n-2)}, \tag{7}$$

and the null distribution of Z_E converges to a normal distribution very quickly.

The basic idea of the *Z* test is based on generalization of the statistic given in (6). Let $Y_{1:n}, Y_{2:n}, ..., Y_{n:n}$ be the order statistic from a population with cumulative distribution function F. The *Z* statistic is given by

$$Z = \frac{2\sum_{i=1}^{n-1}(n-1-i)G_i}{(n-2)\sum_{i=1}^{n-1}G_i}, \ 0 < Z < \infty, \tag{8}$$

where

$$G_i = \frac{y_{i+1:n} - y_{i:n}}{\mu_{i+1:n} - \mu_{i:n}}, \ 1 \le i \le n-1 \tag{9}$$

are the generalized spacings. Similar to *R* the *Z* test is used for testing location scale family distributions. For large n (effectively, $n \ge 10$) the null distribution of *Z* is normal [11].

The asymptotic distribution of G_i is identical with the distribution of D_i [12] and the asymptotic distribution of Z is also normal. But for small sample size ($n < 100$), the values of V are not well approximated by $\frac{1}{3(n-2)}$. Because of this reason the values of V are obtained by simulation [11]. In this study, critical values are calculated by using simulation for any sample size. The critical values for $\alpha = 0.1$ and some values of n are given in Table **2**.

4. RANKED SET SAMPLING

Ranked set sampling (RSS) was first proposed by Mc Intire [13]. He used this model for estimating the mean of pasture yields. This design appeared as a useful technique for improving the accuracy of the estimated means [14]. The RSS procedure is based on the selection of independent samples, not necessarily of the same size, by using simple random sampling (SRS). The sampled units are ranked and only the selected units are measured and used in statistical inference.

Table 2: Critical Values for the *Z* Test (α=0.1)

	Simpson' s Rule		Tiku	
n	**5%**	**95%**	**5%**	**95%**
10	0.6265	1.1484	0.71	1.29
20	0.7601	1.1256	0.81	1.19
50	0.8681	1.0924	0.88	1.12

An RSS can be described as follows. Let $X_1, X_2, ..., X_n, ...$ be independent and identically distributed random variables with cdf F. Consider r independent sets of sample sizes $n_1, n_2, ..., n_r,$ from this distribution, where $r \le n_r$. The n_i, $i = 1,...,r$ represent the set sizes for each independent collection of random variables. From the first set of n_1 independent we select the smallest $\left(X_{1:n_1}^{(1)}\right)$, while from the second set we select the second smallest $\left(X_{2:n_2}^{(2)}\right)$. In this way we continue to select independent random variables until we have selected r representative random variables denoted here by $X_{[1,n_1]}, X_{[2,n_2]}, ..., X_{[r,n_r]}$. The notation $X_{[i,n_j]}$, $1 \le j \le r$ is used to express the fact that each ordered random variable is selected from independent sets as described. In this way a set of independent order statistics is obtained. We note here also that the basic idea in RSS is to rank the observations in each set without actual measurement. This process has been summarized as follows [15].

$$\begin{array}{cccccccc} \underline{X_{1:n_1}^{(1)}} & \underline{X_{2:n_1}^{(1)}} & \cdots & \underline{X_{n_1:n_1}^{(1)}} & \to & X_{[1,n_1]} & \sim & F_{1:n_1}(x) \\ \underline{X_{1:n_2}^{(2)}} & \underline{X_{2:n_2}^{(2)}} & \cdots & \underline{X_{n_2:n_2}^{(2)}} & \to & X_{[2,n_2]} & \sim & F_{2:n_2}(x) \\ \cdots & \cdots & \cdots & \cdots & \to & & \cdots & \\ \underline{X_{1:n_r}^{(r)}} & \underline{X_{2:n_r}^{(r)}} & \cdots & \underline{X_{n_r:n_r}^{(r)}} & \to & X_{[n_r,n_r]} & \sim & F_{r:n_r}(x) \end{array}$$

The random variables $X_{[1,n_1]}, X_{[2,n_2]}, ..., X_{[n_r,n_r]}$ obtained in this way represent one cycle in a RSS. This process may be repeated K times to obtain more random variables. In this study, we assume that all $n_i = n$ for $i = 1,...,r,$ which corresponds to a balanced RSS.

5. APPLICATION OF RANKED SET SAMPLING

Let $X_1, X_2, ..., X_n$ be a random sample from a population with cumulative distribution function $F(x)$. In goodness of fit test we are interested to test hypotheses of the form:

$H_0 : F = F_0$ against $H_1 : F \ne F_0$,

where $F_0(x)$ is a specific distribution. In this study we assume that F_0 represents a normal distribution.

In the *R* and *Z* normality tests the null hypothesis H_0 is tested by using a simple random sample. In this section, it is explained how ranked set sampling is applied to these tests. Therefore, instead of simple random samples ranked set samples with cycles of 1 through 5 are used. Then for different alternative distributions the power of these tests are calculated using simulations for several sample sizes.

It should be noted that the *R* and *Z* tests both are based on comparisons of the order statistics $Y_{i:n}$ from a given population with the theoretical expectations $\mu_{i:n}$ of the ordered data from the standard normal distribution. Therefore, finding accurate estimates for the order statistics, $Y_{i:n}$, when calculating critical values of these tests is an important step in applying these tests.

Let $Y_{1:n}, Y_{2:n}, ..., Y_{n:n}$ be the order statistics from a given population with cdf F. In the *R* test the correlation between $Y_{1:n}, Y_{2:n}, ..., Y_{n:n}$ and is used as a measure of to conclude whether the data come from a normal population. In this study, we propose to use ranked set samples $Y_{[1,n]}, Y_{[2,n]}, ..., Y_{[n,n]}$ instead of simple random samples $Y_{1:n}, Y_{2:n}, ..., Y_{n:n}$. Note that

$$\mu_{i:n} = i\binom{n}{i}\int_{-\infty}^{\infty} x\left[F_0(x)\right]^{i-1}\left[1 - F_0(x)\right]^{n-i} f_0(x)\,dx \tag{10}$$

where $\mu_{i:n} = E\left(Z_{[i:n]}\right)$. Therefore, we will use the correlation between $Y_{[1,n]}, Y_{[2,n]}, ..., Y_{[n,n]}$ and $\mu_{1:n}, \mu_{2:n}, ..., \mu_{n:n}$ as a measure for normality.

Table 3: Critical values for *R* and *R test (α=0.1)**

n	5	10	20	30	50
SRS	**0.903**	**0.933**	**0.959**	**0.971**	**0.980**
1 cycle	0.684	0.847	0.922	0.947	0.968
2 cycle	0.836	0.918	0.959	0.973	0.984
3 cycle	0.886	0.946	0.973	0.982	0.989
4 cycle	0.917	0.959	0.979	0.986	0.991
5 cycle	0.934	0.967	0.983	0.989	0.993

Table 4: Critical Values for the Z and Z^* Tests

N	10		20		30		50	
	5%	95%	5%	95%	5%	95%	5%	95%
SRS	0.627	1.148	0.760	1.126	0.817	1.114	0.868	1.092
1 cycle	0.608	1.168	0.795	1.099	0.862	1.071	0.910	1.049
2 cycle	0.693	1.092	0.840	1.056	0.889	1.041	0.922	1.032
3 cycle	0.730	1.052	0.861	1.032	0.903	1.028	0.939	1.019
4 cycle	0.746	1.027	0.870	1.022	0.912	1.021	0.946	1.015
5 cycle	0.760	1.008	0.879	1.015	0.916	1.014	0.943	1.011

Remark 1: Since the integrand formula (10) is very close to 0 for $|x| > 7.6$ [10] the trapezoidal rule for numerical integration can be applied to evaluate these values. Instead of using plotting positions, the $\mu_{i:n}$ values are calculated by using numerical integration in this way.

Suppose that *K* cycles are used. Let $Y^{(j)} = \left(Y^{(j)}_{[1:n]}, Y^{(j)}_{[2:n]}, \ldots, Y^{(j)}_{[n:n]}\right)$ denote the ranked set sample corresponding to the $j-th$ cycle, where 1≤*j*≤*K* and $\mu = \left(\mu_{1:n}, \mu_{2:n}, \ldots, \mu_{n:n}\right)$. The modified *R* test will be denoted by $R^* = 1 - \bar{\rho}^2$, where $\hat{\rho}_j = Cor\left(y^{(j)}, \mu\right)$ and $\bar{\rho} = \frac{1}{k} \sum_j \hat{\rho}_j$.

To find the critical values of the *R** statistic, ranked set samples from the standard normal distribution are generated for each cycle $j = 1, 2, \ldots, K$. Then for each cycle $\hat{\rho}_j$ is calculated. The results are given Table **3** for some values of n and cycle sizes of one through five.

In a similar way the modified *Z* test is denoted by *Z**, which is defined as

$Z^* = \frac{1}{K} \sum_{j=1}^{K} Z^*_{(j)}$, where

$$Z^*_{(j)} = \frac{2\sum_{i=1}^{n-1}\left(n-1-i\right)G^*_{i,j}}{\left(n-2\right)\sum_{i=1}^{n-1} G^*_{i,j}},\quad G^*_{i,j} = \frac{y^{(j)}_{[i+1:n]} - y^{(j)}_{[i:n]}}{\mu_{i+1:n} - \mu_{i:n}}.$$

Using simulation critical points for the *Z** are computed for 5% and 95% percentage point. The results are given in Table **4** for some values of n and cycle sizes of one through five.

6. SIMULATION RESULTS

Using the critical values as obtained in Tables **3** and **4**, for different alternative distributions in the alternative hypothesis, the power values are calculated. But before presenting the power results the Type I error probabilities for both of the *R** and *Z** tests are summarized in Table **5** for a significance level of α=0.10.

Table **6** shows the simulation results of the *R** test for a sample size of *n*=10. It should be noted that for ranked set sampling the sample size is chosen such that $K.m = n$, where K is the cycle size and m is the

Table 5: Type I Errors for the R^* and Z^*

Method	n=5	n=10	n=20	n=30	n=50
The R^* Test					
SRS	0.101	0.097	0.096	0.099	0.104
1 cycle RSS	0.101	0.088	0.105	0.104	0.111
2 cycle RSS	0.098	0.097	0.092	0.105	0.107
3 cycle RSS	0.088	0.103	0.105	0.110	0.106
4 cycle RSS	0.098	0.103	0.093	0.096	0.106
5 cycle RSS	0.101	0.098	0.091	0.112	0.085
The Z^* Test					
SRS	0.096	0.102	0.093	0.099	0.101
1 cycle RSS	0.102	0.108	0.096	0.101	0.101
2 cycle RSS	0.101	0.100	0.094	0.100	0.069
3 cycle RSS	0.109	0.102	0.107	0.09	0.100
4 cycle RSS	0.097	0.098	0.100	0.105	0.095
5 cycle RSS	0.095	0.104	0.093	0.096	0.067

Table 6: Power of the *R* and *R for Symmetric and Skewed Distributions**

	1 cycle	2 cycle	SRS
Sample Size *n*	10	5	10
Symmetric Distributions			
Student $t_{(2)}$	0.332	0.163	0.417
Student $t_{(4)}$	0.185	0.120	0.231
Logistic	0.129	0.105	0.152
Uniform	0.094	0.110	0.110
Skew Distributions			
$\chi^2_{(1)}$	0.687	0.304	0.783
$\chi^2_{(2)}$	0.404	0.177	0.519
$\chi^2_{(4)}$	0.228	0.135	0.333
Lognormal	0.609	0.279	0.681
Weibull (2)	0.121	0.102	0.140
Beta (2,1)	0.125	0.113	0.187

set size. Similarly, Table **7** and Table **8** show the simulation results for sample sizes of $n = 20$ and *n*=50 are presented. It can be seen that for sample size of *n*=20 and *n*=50 the power of the *R** test, in general, is better than for the *R* test for one cycle. For a sample size of *n*=10 there seems to be no improvement in using ranked set sampling.

Table **9** shows the simulations results of the *Z** test for a sample size of *n*=10. In contrast to the *R** there seems to be improvement even for a small sample size of *n*=10 for symmetric alternatives. At this point we note that Tiku [2] showed that the *Z* test is more powerful than the *R* test, except for testing Uniform, Normal and Logistic (symmetric distributions) against symmetric alternatives.

In general, we can say that using ranked set sampling is worth investigating in goodness-of-fit tests. In this study we only considered small modifications of

Table 7: Power of the *R* and *R for Symmetric and Skewed Distributions**

	1 cycle	2 cycle	3 cycle	4 cycle	SRS
Sample Size *n*	**20**	**10**	**7**	**5**	**20**
Symmetric Distributions					
Student $t_{(2)}$	0.701	0.458	0.357	0.223	0.640
Student $t_{(4)}$	0.362	0.207	0.175	0.124	0.352
Logistic	0.209	0.136	0.125	0.101	0.203
Uniform	0.097	0.096	0.112	0.107	0.175
Skew Distributions					
$\chi^2_{(1)}$	1.000	0.932	0.813	0.886	0.986
$\chi^2_{(2)}$	0.927	0.658	0.526	0.573	0.866
$\chi^2_{(4)}$	0.638	0.376	0.303	0.267	0.606
Lognormal	0.987	0.850	0.736	0.547	0.941
Weibull (2)	0.186	0.143	0.156	0.110	0.205
Beta (2,1)	0.223	0.157	0.163	0.127	0.327

Table 8: Power of the *R* and *R for Symmetric and Skewed Distributions**

	1 cycle	2 cycle	3 cycle	4 cycle	5 cycle	SRS
Sample Size *n*	**50**	**25**	**17**	**12**	**10**	**50**
Symmetric Distributions						
Student $t_{(2)}$	0.990	0.921	0.849	0.727	0.682	0.914
Student $t_{(4)}$	0.775	0.556	0.453	0.340	0.298	0.592
Logistic	0.418	0.274	0.226	0.176	0.156	0.314
Uniform	0.744	0.372	0.300	0.199	0.197	0.640
Skew Distributions						
$\chi^2_{(1)}$	1.000	1.000	1.000	1.000	1.000	1.000
$\chi^2_{(2)}$	1.000	1.000	0.999	1.000	1.000	1.000
$\chi^2_{(4)}$	0.999	0.984	0.940	0.969	0.943	0.959
Lognormal	1.000	1.000	1.000	0.998	0.999	1.000
Weibull (2)	0.618	0.407	0.362	0.279	0.952	0.455
Beta (2,1)	0.983	0.790	0.660	0.449	0.421	0.815

Table 9: Power of the Z and Z^* for Symmetric and Skewed Distributions

	1 cycle	2 cycle	SRS
Sample Size n	**10**	**5**	**10**
Symmetric Distributions			
Student $t_{(2)}$	0.365	0.416	0.358
Student $t_{(4)}$	0.227	0.242	0.214

(Table 9...Continue)

	1 cycle	2 cycle	SRS
Sample Size n	**10**	**5**	**10**
Logistic	0.152	0.082	0.146
Uniform	0.034	0.085	0.076
Skew Distributions			
$\chi^2_{(1)}$	0.725	0.450	0.822
$\chi^2_{(2)}$	0.537	0.268	0.590
$\chi^2_{(4)}$	0.351	0.174	0.378
Lognormal	0.663	0.399	0.722
Weibull (2)	0.146	0.113	0.153
Beta (2,1)	0.133	0.104	0.198

Table 10: Power of the *Z* and *Z for Symmetric and Skewed Distributions**

	1 cycle	2 cycle	3 cycle	4 cycle	SRS
Sample Size n	**20**	**10**	**7**	**5**	**20**
Symmetric Distributions					
Student $t_{(2)}$	0.503	0.416	0.368	0.270	0.472
Student $t_{(4)}$	0.300	0.239	0.201	0.147	0.265
Logistic	0.188	0.148	0.133	0.115	0.171
Uniform	0.006	0.025	0.040	0.059	0.072
Skew Distributions					
$\chi^2_{(1)}$	0.996	0.929	0.840	0.671	0.989
$\chi^2_{(2)}$	0.959	0.792	0.648	0.472	0.886
$\chi^2_{(4)}$	0.797	0.562	0.434	0.286	0.665
Lognormal	0.980	0.876	0.784	0.622	0.960
Weibull (2)	0.330	0.219	0.180	0.129	0.253
Beta (2,1)	0.492	0.246	0.187	0.132	0.380

the classical *R* and *Z* tests. For the considered alternatives, in testing normality, *Z** generally (except for the Uniform distribution) is more powerful than the Z test, except for a small sample size of *n*=10 with skewed alternative distributions.

In parameter estimation problems it has been shown that RSS provides an unbiased estimator for the population mean with smaller variance than for simple random sampling (for example [16, 17]). Although this study shows that also in goodness-of-fit testing the use

Table 11: Power of the *Z* and *Z for Symmetric and Skewed Distributions**

	1 cycle	2 cycle	3 cycle	4 cycle	5 cycle	SRS
Sample Size *n*	**50**	**25**	**17**	**12**	**10**	**50**
Symmetric Distributions						
Student $t_{(2)}$	0.642	0.614	0.571	0.521	0.495	0.587
Student $t_{(4)}$	0.396	0.355	0.336	0.298	0.271	0.338
Logistic	0.219	0.208	0.184	0.168	0.168	0.206
Uniform	0.001	0.004	0.008	0.015	0.026	0.069
Skew Distributions						
$\chi^2_{(1)}$	1.000	1.000	1.000	1.000	0.998	1.000
$\chi^2_{(2)}$	1.000	1.000	0.999	0.992	0.980	1.000
$\chi^2_{(4)}$	1.000	0.996	0.980	0.935	0.875	0.975
Lognormal	1.000	1.000	1.000	1.000	0.995	1.000
Weibull (2)	1.000	0.775	0.590	0.444	0.397	0.587
Beta (2,1)	1.000	0.954	0.870	0.675	0.619	0.771

of ranked set sampling may improve the power more detailed analysis and simulations should be done.

REFERENCES

[1] Filiben JJ. The probability plot correlation coefficient test for normality. Technometrics 1975; 17: 111-117. http://dx.doi.org/10.1080/00401706.1975.10489279

[2] Tiku M. Goodness of fit statistics based on the spacing of complete or censored samples. Australian Journal of Statistics 1980; 22: 260-275. http://dx.doi.org/10.1111/j.1467-842X.1980.tb01175.x

[3] Tiku M. Order statistics in goodness of fit tests. Communications in Statistics-Theory and Methods 1988; 17: 2369-2387. http://dx.doi.org/10.1080/03610928808829751

[4] Wolfe W. Ranked set sampling: an approach to more efficient data collection. Statistical Science. A Review Journal of the Institute of Mathematical Statistics 2004; 19: 636-643.

[5] Looney S. Use of the correlation coefficient with normal probability plots. The American Statistician 1985; 39: 75-79.

[6] Hazen A. Storage to be provided in the impounding reservoirs for municipal water supply. Transactions of the American Society of Civil Engineers 1914; 77: 1539-1640.

[7] Weibull W. The Phenomenon of Rupture in Solids. Ingeniors Vetenskaps Akademien Handlingar 1939; 153: 1-55.

[8] Blom G. Statistical Estimates and Transformed Beta Variables. John Wiley 1958.

[9] Smith RB. Correlation type goodness of fit statistics with censored sampling. Communications in Statistics-Theory and Methods 1976; 5: 119-132. http://dx.doi.org/10.1080/03610927608827337

[10] Harter HL. Expected values of normal order statistics. Biometrica 1961; 48: 151-165. http://dx.doi.org/10.1093/biomet/48.1-2.151

[11] Sürücü B. Goodness-of-Fit Tests for multivariate distributions. Communications in Statistics--Theory and Methods 2006; 35: 1319-1331. http://dx.doi.org/10.1080/03610920600628999

[12] Pyke R. Spacings (with discussion). Journal of the Royal Statistical Society Series B 1965; 27: 395-449

[13] McIntire G. A method of unbiased selective sampling using ranked sets. J Agric Res 1952; 3: 385-390. http://dx.doi.org/10.1071/AR9520385

[14] Bouza C. Sampling Using Ranked Sets: Concepts, Results and Perspectives. Revista Investigacion Operacional 2005; 26: 275-293.

[15] Arslan G. A new characterization of the power distribution. Journal of Computational and Applied Mathematics 2014, 260: 99-102. http://dx.doi.org/10.1016/j.cam.2013.09.068

[16] Dell TT. Ranked set sampling theory with order statistics background. Biometrics 1972; 28: 545-555. http://dx.doi.org/10.2307/2556166

[17] Takashi KW. On unbiased estimates of the population mean based on the sample stratified by means of ordering. Ann Inst Statist Math 1988; 20: 1-31. http://dx.doi.org/10.1007/BF02911622

Received on 30-12-2014 Accepted on 12-01-2015 Published on 02-04-2015

DOI: http://dx.doi.org/10.15377/2409-5761.2014.02.01.2