Different Goodness of Fit Tests for Rayleigh Distribution in Ranked Set Sampling

Amer Ibrahim Al-Omari
Department of Mathematics, Faculty of Science
Al al-Bayt University, P.O. Box 130095, Mafraq 25113, Jordon
alomari_amer@yahoo.com

Ehsan Zamanzade
Department of Statistics, University of Isfahan, Isfahan 81746-73441, Iran
ehsanzamanzadeh@yahoo.com; e.zamanzade@sci.ui.ac.ir

Abstract

In this paper, different goodness of fit tests for the Rayleigh distribution are considered based on simple random sampling (SRS) and ranked set sampling (RSS) techniques. The performance of the suggested estimators is evaluated in terms of the power of the tests by using Monte Carlo simulation. It is found that the suggested RSS tests perform better than their counterparts in SRS.

Keywords: Goodness of fit tests; Simple random sampling; Ranked set sampling; Kullback-Leibler information; Critical values; Rayleigh distribution.

MSC: 94A17

1. Introduction

Let X be a continuous random variable with probability density function $f(x)$ and cumulative distribution function (cdf) $F(x)$. The entropy is known as a measure of uncertainty and dispersion and it is defined by Shanon (1948) for the random variable X as

$$H(f) = -\int_{-\infty}^{+\infty} f(x)\log f(x)dx.$$

As it was shown by Vasicek (1976), the above estimator can be expressed as

$$H(f) = \int_0^1 \log\left(\frac{\mathrm{d}}{\mathrm{d}p}F^{-1}(p)\right)dp.$$

Let $X_1, X_2, ..., X_n$ be a simple random sample of size n from a population with cdf $F(x)$, and let $X_{(1)}, X_{(2)}, ..., X_{(n)}$ be the order statistics of this sample. Vasicek (1976)'s estimator of entropy is defined as

$$HV_{mn} = \frac{1}{n}\sum_{i=1}^{n} \mathrm{Log}\left\{\frac{n}{2m}\left(X_{(i+m)} - X_{(i-m)}\right)\right\}, \tag{1}$$

where the window size m is a positive integer less than $n/2$, $X_{(i)} = X_{(n)}$ if $i > n$ and $X_{(i)} = X_{(1)}$ if $i < 1$. Vasicek (1976)'s estimator converges in probability to $H(f)$ as $n, m \to \infty$, and $m/n \to 0$.

Ebrahimi et al. (1994) suggested an estimator of entropy which is given by

$$HE_{mn} = \frac{1}{n}\sum_{i=1}^{n} \text{Log}\left\{ \frac{n}{c_i m}\left(X_{(i+m)} - X_{(i-m)} \right) \right\},$$

where $c_i = 1 + \frac{i-1}{m}$ if $1 \le i \le m$, $c_i = 2$ if $m+1 \le i \le n-m$, and $c_i = 1 + \frac{n-i}{m}$ if $n-m+1 \le i \le n$.

Ebrahimi et al. (1994) utilized the weight $\frac{n}{2m}$ in Vasicek (1976) estimator to have a smaller weights and they proved that this estimator converges in probability to $H(f)$ as $n \to \infty$, $m \to \infty$ and $m/n \to 0$.

Correa (1995) suggested an estimator of entropy to have a smaller mean square error given by

$$HC_{mn} = -\frac{1}{n}\sum_{i=1}^{n} \log\left[\frac{\sum_{j=i-m}^{i+m} (j-i)\left(X_{(j)} - \bar{X}_{(i)} \right)}{n \sum_{j=i-m}^{i+m} \left(X_{(j)} - \bar{X}_{(i)} \right)^2} \right],$$

where $\bar{X}_{(i)} = \frac{1}{2m+1}\sum_{j=i-m}^{i+m} X_{(j)}$, and $X_{(i-m)} = X_{(1)}$ for $i \le m$ and $X_{(i+m)} = X_{(n)}$ for $i \ge n-m$.

Recently, Al-Omari (2014) suggested new estimators of entropy based simple random sampling, ranked set sampling (RSS) and double ranked set sampling (DRSS) methods in the form

$$AHM_{mn} = \frac{1}{n}\sum_{i=1}^{n} \text{Log}\left\{ \frac{n}{c_i m}\left(X^{M}_{(i+m)} - X^{M}_{(i-m)} \right) \right\}, \ M = SRS, RSS, DRSS,$$

where $c_i = 1 + \frac{1}{2}$ if $1 \le i \le m$, $c_i = 2$ if $m+1 \le i \le n-m$, and $c_i = 1 + \frac{1}{2}$ if $n-m+1 \le i \le n$, and $X_{(i-m)} = X_{(1)}$ for $i \le m$ and $X_{(i+m)} = X_{(n)}$ for $i \ge n-m$.

This paper is laid out as follows. Rayleigh distribution and its properties are discussed in Section 2. Section 3, is devoted to the suggested goodness of fit tests based on SRS and RSS methods. Section 4 is considered for power comparison using Mote Carlo simulations. Finally, Section 5 is the conclusion.

2. Rayleigh distribution

The Rayleigh distribution is suggested by Rayleigh (1880), and later Siddiqui (1962) discussed its properties. The Rayleigh is commonly used when two orthogonal components have an absolute value. As an example, direction and wind velocity may be combined to produce a wind speed. Also, it is used in clinical studies and life testing experiments.

Here, we investigate the powers of different goodness of fit tests for Rayleigh distribution with probability density function (PDF) given by

$$f_0(x;\theta) = \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}} \qquad x \geq 0, \tag{2}$$

where $\theta > 0$ is an unknown parameter. Let X follows a Rayleigh distribution, then the cumulative distribution function of X is defined as

$$F_0(x;\theta) = 1 - e^{-\frac{x^2}{2\theta^2}} \qquad x \geq 0. \tag{3}$$

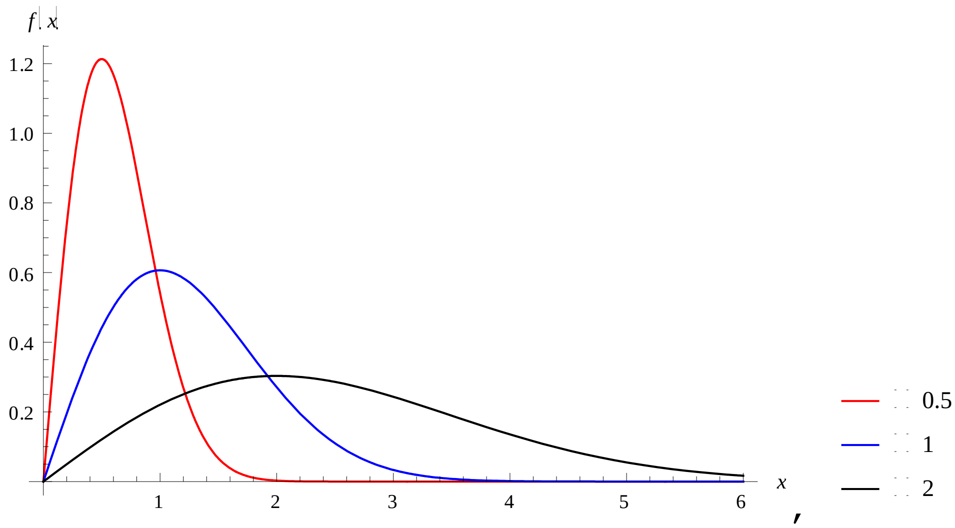

Figure 1: The PDF of the Rayleigh distribution for different values of θ

The mean and variance of X are $\mu = \theta\sqrt{\frac{\pi}{2}}$ and $\sigma^2 = \frac{4-\pi}{2}\theta^2$, and its entropy is $1 + \ln\left(\frac{\theta}{\sqrt{2}}\right) + \frac{\lambda}{2}$, where λ is the Euler's constant. The Rayleigh distribution is skewed to the right, where its skewness is $\frac{2\sqrt{\pi}(\pi-3)}{\sqrt{(4-\pi)^3}}$. Therefore, it can be used efficiently to analyze skewed data. The survival function is $S(x) = P(X \geq x) = Exp\left(-\frac{x^2}{\theta}\right)$ and the hazard function is $h(x) = \frac{f(x)}{S(x)} = \frac{2x}{\theta}$. Also, its inverse distribution function is $F^{-1}(\upsilon) = \sqrt{-\theta \ln(1-\upsilon)}$, $0 < \upsilon < 1$, see Hirano (1986) for more details about the properties of Rayleigh distribution.

3. Test statistics

In this section, we will discuss the suggested goodness of fit tests based on SRS and RSS methods.

3.1 Using SRS

In SRS, the goodness of fit tests which are considered here are:

- The Kullback-Leibler based on Vasicek (1976)'s entropy estimator, which is given by:

$$KL_{mn} = -HV_{mn} - \frac{1}{n}\sum_{i=1}^{n}\log\left[f_0\left(x_i, \hat{\theta}_{SRS}\right)\right], \tag{6}$$

  where the distribution of KL_{mn} is free of θ. Noughabi et al. (2013) showed under the null hypothesis that KL_{mn} approach to zero in probability as $N \to \infty, m \to \infty$, and $\frac{m}{n} \to 0$.

- The Kolmogorov-Smirnov test statistics :

$$KS = Max\left\{ \underset{1\le i\le n}{Max}\left[\frac{i}{n} - F_0\left(x_{(i)}, \hat{\theta}_{SRS}\right)\right], \underset{1\le i\le n}{Max}\left[F_0\left(x_{(i)}, \hat{\theta}_{SRS}\right) - \frac{i-1}{n}\right]\right\} \tag{7}$$

- The Anderson-Darling (1954) test statistics:

$$A^2 = -2\sum_{i=1}^{n}\left\{\left(i-\frac{1}{2}\right)\log\left[F_0\left(x_{(i)}, \hat{\theta}_{SRS}\right)\right] + \left(n-i+\frac{1}{2}\right)\log\left[1-F_0\left(x_{(i)}, \hat{\theta}_{SRS}\right)\right]\right\} - n. \tag{8}$$

- The Cramer-von Mises test statistics:

$$W^2 = \sum_{i=1}^{n}\left[F_0\left(x_{(i)}, \hat{\theta}_{SRS}\right) - \frac{2i-1}{2n}\right]^2 + \frac{1}{12n}. \tag{9}$$

- The Zhang (2002) test statistics:

$$Z_K = \underset{1\le i\le n}{Max}\left\{\left(i-\frac{1}{2}\right)\log\left[\frac{i-\frac{1}{2}}{nF_0\left(x_{(i)}, \hat{\theta}_{SRS}\right)}\right] + \left(n-i+\frac{1}{2}\right)\log\left[\frac{n-i+\frac{1}{2}}{n\left[1-F_0\left(x_{(i)}, \hat{\theta}_{SRS}\right)\right]}\right]\right\}, \tag{10}$$

$$Z_A = -\sum_{i=1}^{n}\left\{\frac{\log\left[F_0\left(x_{(i)}, \hat{\theta}_{SRS}\right)\right]}{n-i+\frac{1}{2}} + \frac{\log\left[1-F_0\left(x_{(i)}, \hat{\theta}_{SRS}\right)\right]}{i-\frac{1}{2}}\right\}, \tag{11}$$

$$Z_C = \sum_{i=1}^{n}\left[\log\left(\frac{F_0\left(x_{(i)}, \hat{\theta}\right)^{-1} - 1}{\left(n-\frac{1}{2}\right)\Big/\left(i-\frac{3}{4}\right) - 1}\right)\right], \tag{12}$$

where $HV_{mn} = \frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{n}{2m}\left(x_{(i+m)} - x_{(i-m)}\right)\right]$, $F_0\left(.,\theta\right)$ is the distribution function of Rayleigh distribution and $\hat{\theta}_{SRS} = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} x_i^2}$, in all above test statistics.

3.2 Using RSS

McIntyre (1952) was the first who suggested the ranked set sampling design for estimating the population mean. The RSS involves randomly selecting k^2 units from the population. These units are randomly allocated into *k* sets, each of size *k*. The *k* units within each sample are ranked visually or by any inexpensive method with respect to the variable of interest from smallest to largest Then the smallest ranked unit is measured from the first set of *k* units. The second smallest ranked unit is measured from the second set of *k* units. The process is continued in this way until the largest ranked unit is measured from the *k*th set of *k* units. The whole process can be repeated *h* times to get a sample of size *n=hk.*

For more works in RSS see Al-Omari (2011, 2014), Mahdizadeh (2012), Zamanzade et al. (2014), Mahdizadeh and Arghami (2010), Haq et al. (2013).

Let $X_{(i)j}$, $i = 1, 2, \cdots, k, \ j = 1, 2, \cdots, h,$ be a ranked set sample of size $n = hk$ from the population of interest, where $X_{(i)j}$ is the ith order statistic from the jth sample, and let $Z_{(1)} \le Z_{(2)} \le \cdots \le Z_{(n)}$ be corresponding ordered statistics values of the ranked set sample. We develop the goodness of fit tests described above based on ranked set sampling method. Thus, above goodness of tests counterparts for RSS are

- The test based on Kullback- Leibler distance and Vasicek (1976)'s entropy estimator counterpart in RSS, which has the form

$$KL_{mn}^{RSS} = -HV_{mn}^{RSS} - \frac{1}{n}\sum_{i=1}^{n}\log\left[f_0\left(z_{(i)}, \hat{\theta}_{RSS}\right)\right]. \tag{13}$$

- Kolmogorov-Smirnov counterpart test statistics in RSS:

$$KS = Max\left\{\underset{1\le i\le n}{Max}\left[\frac{i}{n} - F_0\left(z_{(i)}, \hat{\theta}_{RSS}\right)\right], \underset{1\le i\le n}{Max}\left[F_0\left(z_{(i)}, \hat{\theta}_{RSS}\right) - \frac{i-1}{n}\right]\right\}. \tag{14}$$

- Anderson-Darling counterpart test statistics in RSS:

$$A^2 = -2\sum_{i=1}^{n}\left\{\left(i - \frac{1}{2}\right)\log\left[F_0\left(z_{(i)}, \hat{\theta}_{RSS}\right)\right] + \left(n - i + \frac{1}{2}\right)\log\left[1 - F_0\left(z_{(i)}, \hat{\theta}_{RSS}\right)\right]\right\} - n. \tag{15}$$

- Cramer-von Mises test statistics:

$$W^2 = \sum_{i=1}^{n}\left[F_0\left(z_{(i)}, \hat{\theta}_{RSS}\right) - \frac{2i-1}{2n}\right]^2 + \frac{1}{12n}. \tag{16}$$

- Zhang (2002) counterpart test statistics in RSS:

$$Z_K = \underset{1\le i\le n}{Max}\left\{\left(i-\frac{1}{2}\right)\log\left(\frac{i-\frac{1}{2}}{nF_0\left(z_{(i)},\hat{\theta}_{RSS}\right)}\right)+\left(n-i+\frac{1}{2}\right)\log\left(\frac{n-i+\frac{1}{2}}{n\left[1-F_0\left(z_{(i)},\hat{\theta}_{RSS}\right)\right]}\right)\right\}, \quad (17)$$

$$Z_A = -\sum_{i=1}^{n}\left\{\frac{\log\left[F_0\left(z_{(i)},\hat{\theta}_{RSS}\right)\right]}{n-i+\frac{1}{2}}+\frac{\log\left[1-F_0\left(z_{(i)},\hat{\theta}_{RSS}\right)\right]}{i-\frac{1}{2}}\right\}, \quad (18)$$

$$Z_C = \sum_{i=1}^{n}\left[\log\left(\frac{F_0\left(z_{(i)},\hat{\theta}_{RSS}\right)^{-1}-1}{\left(n-\frac{1}{2}\right)\Big/\left(i-\frac{3}{4}\right)-1}\right)\right], \quad (19)$$

where $HV_{nm}^{RSS} = \frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{n}{2m}\left(z_{(i+m)}-z_{(i-m)}\right)\right]$, $F_0\left(.,\theta\right)$ is the distribution function of Rayleigh distribution and $\hat{\theta}_{RSS} = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} z_{(i)}^2}$, in all above test statistics.

4. Monte Carlo Study

In this section, we carry out a simulation study to investigate the performance of the suggested goodness of fit tests.

4.1 Critical values

In all considered tests the null hypothesis is rejected when the test statistics are larger than the corresponding critical values at a given significance level.

The performance is evaluated based on 100,000 replications of random samples of sizes $n=10$, 20 and 40 from the Rayleigh distribution using SRS and RSS designs for different significance levels. The problem of choosing the optimal values of $m<(n/2)$ which maximizes the powers subject to *n* is still open in the field of entropy estimation. The window size corresponding to a given sample size should be selected in advance. Therefore, in our simulations, we have used Grzegorzewski and Wieczorkowski (1999)'s heuristics formula for choosing *m* as: $m=\left[\sqrt{n}+0.5\right]$. The critical values are reported in Tables 1-3 for $\alpha=0.01$, 0.05, 0.1, respectively.

4.2 Power comparison

One of the main goodness of fit tests is based on the empirical distribution function (EDF). The widely important EDF tests are the Anderson-Darling (1954), Kolmogorov-Smirnov test, and Cramer-von Mises. Powers of tests based on KL_{mn}, and KL_{mn}^{RSS} depend on the value of *m*, and m is obtained using $m=\left[\sqrt{n}+0.5\right]$. A total of 100,000 random samples of sizes $n=10,20,40$ are generated from each of the alternatives.

In order to compare the powers of goodness of fit tests in SRS and RSS designs, we have considered the eight following distributions as alternative distributions:

- the Exponential distribution with mean 1, denoted by Exp(1),
- the Uniform distribution on (0,1), denoted by U(0,1),
- the Beta distribution with parameters 0.5 and 0.5, denoted by Beta(0.5, 0.5),
- the Beta distribution with parameters 2 and 1, denoted by Beta(2,1),
- the Gamma distribution with scale parameter 1 and shape parameter 0.5, denoted by G(0.5),
- the Gamma distribution with scale parameter 1 and shape parameter 2, denoted by G(2),
- the Weibull distribution with scale parameter 1 and shape parameter 0.8, denoted by W(0.8),
- the Weibull distribution with scale parameter 1 and shape parameter 1.4, denoted by W(1.4).

The power of the tests at significance level $\alpha=0.05$ are summarized in Tables 4,5,6, for $n=10,20,40$, respectively.

In RSS scheme, the value of *k* (set size) is taken to be 2 and 5, so we can observe the effect of increasing sample size while set size being fixed, and the effect of increasing set size while sample is being fixed.

Based on simulation results it can be observed that:

- The suggested RSS tests are more powerful as compared to their SRS counterparts for all cases considered in this study.
- The bold fonts in Tables 4-6 are the maximum power values for the test statistics for each distribution. However, among all alternatives, the test Anderson-Darling has the greatest powers in most cases.
- The power of the goodness of fit tests increase in the sample size. As an example, based on RSS with $k=2$ for the Beta distribution, $Beta(2,1)$, the power values of Z_C are 0.234, 0.671, 0.984, respectively with $n=10,20,40$.
- Power simulations report that the tests statistics are most powerful when the alternative distribution is gamma distribution with shape parameter 0.5.

Table 1: Critical values of different tests of Rayleigh distribution for different values of (n,k) in RSS design, at significance level $\alpha = 0.01$

		Test Statistic						
n	*k*	KL^a	KS	A^2	W^2	Z_K	Z_A	Z_C
2	2	1.920	0.543	18.924	0.271	2.073	1.112	3.610
	3	1.274	0.453	39.794	0.267	2.383	0.700	2.663
	4	0.957	0.389	68.224	0.253	2.500	0.500	2.104
	5	0.793	0.344	104.270	0.238	2.516	0.387	1.684
3	2	1.340	0.464	40.651	0.289	2.524	0.715	2.775
	3	0.913	0.379	87.012	0.277	2.769	0.446	1.968
	4	0.722	0.326	150.555	0.259	2.857	0.321	1.553
	5	0.601	0.286	231.422	0.243	2.863	0.249	1.242
4	2	1.047	0.412	70.436	0.296	2.834	0.518	2.276
	3	0.743	0.334	152.024	0.281	3.024	0.325	1.611
	4	0.588	0.285	264.818	0.261	3.085	0.235	1.220
	5	0.491	0.250	408.705	0.245	3.072	0.183	0.997
5	2	0.881	0.373	108.118	0.302	3.051	0.404	1.922
	3	0.633	0.302	235.160	0.283	3.227	0.255	1.377
	4	0.498	0.257	411.110	0.262	3.263	0.184	1.040
	5	0.422	0.225	635.853	0.246	3.236	0.144	0.842
6	2	0.763	0.343	153.650	0.300	3.188	0.329	1.669
	3	0.552	0.278	336.427	0.286	3.367	0.209	1.172
	4	0.441	0.236	589.345	0.263	3.393	0.152	0.901
	5	0.367	0.205	912.858	0.245	3.391	0.119	0.731
7	2	0.681	0.320	207.506	0.304	3.374	0.278	1.495
	3	0.498	0.258	455.525	0.285	3.492	0.176	1.036
	4	0.391	0.219	799.458	0.264	3.538	0.129	0.797
	5	0.330	0.192	1240.526	0.248	3.487	0.101	0.644
8	2	0.614	0.300	268.935	0.303	3.470	0.239	1.352
	3	0.452	0.242	592.014	0.283	3.587	0.153	0.927
	4	0.361	0.206	1041.470	0.265	3.590	0.112	0.705
	5	0.297	0.180	1617.511	0.248	3.576	0.088	0.582
9	2	0.562	0.285	338.596	0.307	3.553	0.210	1.216
	3	0.410	0.230	747.402	0.288	3.713	0.135	0.850
	4	0.329	0.195	1316.814	0.267	3.734	0.099	0.659
	5	0.275	0.170	2045.018	0.248	3.680	0.078	0.530
	2	0.523	0.271	416.434	0.307	3.660	0.188	1.122

0.230

Table 2: Critical values of different tests of Rayleigh distribution for different values of (n,k) in RSS design, at significance level $\alpha = 0.05$

		Test Statistic						
n	*k*	KL[a]	KS	A^2	W^2	Z_K	Z_A	Z_C
2	2	1.422	0.470	16.630	0.191	1.341	1.009	2.308
	3	0.991	0.389	36.759	0.186	1.567	0.643	1.688
	4	0.778	0.333	64.488	0.174	1.665	0.468	1.314
	5	0.654	0.294	100.010	0.165	1.711	0.366	1.064
3	2	1.038	0.400	37.233	0.201	1.673	0.654	1.817
	3	0.741	0.325	82.280	0.189	1.854	0.417	1.288
	4	0.593	0.278	144.879	0.177	1.936	0.304	0.997
	5	0.507	0.244	225.183	0.167	1.962	0.239	0.806
4	2	0.831	0.351	65.698	0.202	1.889	0.478	1.478
	3	0.608	0.285	145.867	0.191	2.059	0.307	1.062
	4	0.493	0.243	257.256	0.179	2.118	0.224	0.802
	5	0.409	0.214	400.388	0.168	2.145	0.177	0.659
5	2	0.711	0.318	102.225	0.204	2.042	0.376	1.268
	3	0.528	0.257	227.406	0.192	2.206	0.242	0.898
	4	0.416	0.219	401.672	0.179	2.242	0.178	0.685
	5	0.355	0.192	625.399	0.168	2.263	0.140	0.558
6	2	0.621	0.291	146.805	0.205	2.171	0.309	1.114
	3	0.459	0.236	326.976	0.193	2.311	0.199	0.782
	4	0.374	0.201	578.201	0.180	2.373	0.147	0.603
	5	0.308	0.176	900.470	0.168	2.372	0.116	0.489
7	2	0.567	0.271	199.172	0.204	2.275	0.262	0.987
	3	0.421	0.219	444.286	0.192	2.396	0.169	0.694
	4	0.330	0.186	786.372	0.180	2.455	0.125	0.535
	5	0.281	0.164	1225.597	0.169	2.456	0.099	0.433
8	2	0.512	0.254	259.677	0.204	2.351	0.227	0.894
	3	0.378	0.205	579.718	0.192	2.472	0.147	0.627
	4	0.306	0.175	1026.723	0.181	2.514	0.109	0.479
	5	0.253	0.154	1600.645	0.169	2.526	0.086	0.390
9	2	0.467	0.241	328.377	0.206	2.441	0.201	0.820
	3	0.345	0.195	733.432	0.193	2.559	0.130	0.572
	4	0.278	0.165	1299.204	0.181	2.590	0.096	0.438
	5	0.237	0.145	2025.973	0.170	2.598	0.076	0.358
10	2	0.431	0.230	404.924	0.207	2.506	0.180	0.760
	3	0.318	0.185	905.167	0.195	2.620	0.117	0.531
	4	0.255	0.157	1603.536	0.181	2.626	0.086	0.403
	5	0.217	0.138	2500.957	0.168	2.643	0.068	0.327

[a]The value of m is taken to be $\left[\sqrt{n}+0.5\right]$, where $N = nk$ is total sample size.

Table 3: Critical values of different tests of Rayleigh distribution for different values of (*n*,*k*) in RSS design, at significance level $\alpha = 0.1$

		Test Statistic						
n	k	KL^a	KS	A^2	W^2	Z_K	Z_A	Z_C
2	2	1.216	0.434	15.770	0.157	1.043	0.959	1.865
	3	0.873	0.356	35.495	0.151	1.236	0.619	1.352
	4	0.703	0.305	62.997	0.142	1.325	0.454	1.059
	5	0.593	0.270	98.291	0.134	1.375	0.357	0.868
3	2	0.913	0.366	35.842	0.162	1.322	0.627	1.465
	3	0.667	0.297	80.361	0.152	1.481	0.404	1.040
	4	0.536	0.254	142.592	0.143	1.561	0.297	0.813
	5	0.463	0.224	222.543	0.135	1.590	0.235	0.660
4	2	0.743	0.321	63.827	0.162	1.494	0.462	1.197
	3	0.548	0.261	143.302	0.155	1.647	0.299	0.858
	4	0.449	0.222	254.186	0.144	1.718	0.220	0.662
	5	0.372	0.196	396.885	0.136	1.754	0.174	0.545
5	2	0.637	0.289	99.888	0.163	1.630	0.365	1.024
	3	0.481	0.235	224.116	0.154	1.775	0.237	0.729
	4	0.377	0.200	397.782	0.145	1.828	0.175	0.562
	5	0.325	0.176	621.042	0.136	1.858	0.138	0.460
6	2	0.558	0.266	143.942	0.164	1.735	0.301	0.903
	3	0.416	0.215	323.054	0.154	1.870	0.196	0.636
	4	0.341	0.184	573.448	0.145	1.934	0.145	0.496
	5	0.281	0.161	895.243	0.136	1.950	0.115	0.403
7	2	0.515	0.247	195.908	0.163	1.820	0.256	0.805
	3	0.386	0.200	439.883	0.155	1.950	0.167	0.569
	4	0.301	0.170	780.996	0.145	2.008	0.123	0.441
	5	0.258	0.150	1219.515	0.136	2.027	0.098	0.358
8	2	0.465	0.232	255.904	0.163	1.891	0.222	0.731
	3	0.346	0.188	574.699	0.155	2.015	0.145	0.515
	4	0.282	0.160	1020.532	0.145	2.065	0.108	0.396
	5	0.232	0.141	1593.739	0.137	2.091	0.085	0.323
9	2	0.425	0.220	324.117	0.165	1.971	0.197	0.672
	3	0.315	0.178	727.710	0.155	2.083	0.128	0.471
	4	0.255	0.151	1292.254	0.145	2.128	0.095	0.363
	5	0.219	0.133	2018.068	0.137	2.152	0.076	0.296
10	2	0.390	0.210	400.098	0.165	2.025	0.176	0.623
	3	0.290	0.169	898.702	0.157	2.146	0.115	0.437
	4	0.235	0.144	1595.768	0.146	2.168	0.085	0.335
	5	0.200	0.126	2492.068	0.136	2.201	0.068	0.272

[a]The value of *m* is taken to be $\left[\sqrt{n}+0.5\right]$, where $N = nk$ is total sample size.

Table 4: Power estimates of different goodness of fit tests in SRS and RSS designs for $n = 10$ and $\alpha = 0.05$

Sampling scheme	Alternative distribution	KL $m = 3$	KS	A^2	W^2	Z_K	Z_A	Z_C
SRS	Exponential(1)	0.484	0.566	**0.767**	0.613	0.728	0.715	0.717
	Uniform(0,1)	0.178	0.103	0.263	0.112	0.275	0.245	**0.293**
	Beta (0.5,0.5)	0.682	0.334	0.732	0.387	0.730	0.732	**0.763**
	Beta(2,1)	**0.475**	0.280	0.293	0.361	0.192	0.320	0.327
	Gamma (0.5)	0.921	0.885	**0.978**	0.908	0.973	0.974	0.975
	Gamma(2)	0.072	0.181	**0.257**	0.200	0.205	0.174	0.170
	Weibull(0.8)	0.763	0.777	**0.921**	0.812	0.903	0.900	0.901
	Weibull(1.4)	0.104	0.200	**0.332**	0.227	0.290	0.255	0.263
RSS $k = 2$	Exponential(1)	0.535	0.613	**0.806**	0.662	0.764	0.762	0.755
	Uniform(0,1)	0.180	0.093	0.268	0.106	0.280	0.254	**0.299**
	Beta (0.5,0.5)	0.694	0.333	0.758	0.403	0.757	0.760	**0.787**
	Beta(2,1)	0.484	0.278	0.278	**0.362**	0.176	0.315	0.324
	Gamma (0.5)	0.941	0.910	**0.986**	0.931	0.981	0.983	0.982
	Gamma(2)	0.080	0.202	**0.286**	0.226	0.226	0.198	0.186
	Weibull(0.8)	0.801	0.812	**0.939**	0.847	0.921	0.923	0.920
	Weibull(1.4)	0.115	0.224	**0.366**	0.253	0.316	0.287	0.285
RSS $k = 5$	Exponential(1)	0.656	0.726	**0.904**	0.786	0.870	0.874	0.850
	Uniform(0,1)	0.202	0.092	0.328	0.119	0.339	0.315	**0.340**
	Beta (0.5,0.5)	0.774	0.398	0.861	0.501	0.854	0.858	**0.864**
	Beta(2,1)	**0.562**	0.339	0.354	0.455	0.208	0.369	0.404
	Gamma (0.5)	0.983	0.963	**0.997**	0.976	0.996	0.996	0.996
	Gamma(2)	0.100	0.264	**0.374**	0.305	0.294	0.274	0.230
	Weibull(0.8)	0.900	0.896	**0.981**	0.927	0.972	0.975	0.969
	Weibull(1.4)	0.139	0.291	**0.474**	0.340	0.407	0.389	0.346

Table 5: Power estimates of different goodness of fit tests in SRS and RSS designs for $n = 20$ and $\alpha = 0.05$

Sampling scheme	Alternative distribution	KL $m=4$	KS	A^2	W^2	Z_K	Z_A	Z_C
SRS	Exponential(1)	0.845	0.864	**0.956**	0.900	0.935	0.935	0.938
	Uniform(0,1)	**0.469**	0.156	0.399	0.189	0.410	0.397	0.462
	Beta (0.5,0.5)	**0.966**	0.598	0.930	0.693	0.919	0.944	0.950
	Beta(2,1)	**0.845**	0.527	0.627	0.672	0.376	0.684	0.632
	Gamma (0.5)	0.998	0.996	**1**	0.997	**1**	**1**	**1**
	Gamma(2)	0.179	0.329	**0.432**	0.375	0.335	0.298	0.311
	Weibull(0.8)	0.979	0.976	**0.996**	0.985	0.994	0.994	0.995
	Weibull(1.4)	0.255	0.368	**0.534**	0.420	0.463	0.424	0.449
RSS $k=2$	Exponential(1)	0.869	0.891	**0.967**	0.922	0.949	0.950	0.950
	Uniform(0,1)	0.468	0.150	0.410	0.192	0.421	0.413	**0.474**
	Beta (0.5,0.5)	**0.968**	0.622	0.943	0.721	0.936	0.953	0.959
	Beta(2,1)	**0.892**	0.551	0.664	0.715	0.372	0.73	0.671
	Gamma (0.5)	0.999	0.997	**1**	0.998	**1**	**1**	**1**
	Gamma(2)	0.193	0.357	**0.466**	0.410	0.360	0.330	0.336
	Weibull(0.8)	0.984	0.982	**0.997**	0.989	0.995	0.996	0.996
	Weibull(1.4)	0.271	0.400	**0.575**	0.457	0.497	0.465	0.480
RSS $k=5$	Exponential(1)	0.934	0.951	**0.992**	0.971	0.985	0.985	0.982
	Uniform(0,1)	0.521	0.165	0.508	0.235	0.505	0.494	**0.530**
	Beta (0.5,0.5)	**0.987**	0.730	0.983	0.838	0.977	0.984	0.985
	Beta(2,1)	**0.961**	0.669	0.836	0.868	0.484	0.876	0.821
	Gamma (0.5)	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	Gamma(2)	0.216	0.437	0.575	0.509	0.45	0.418	0.391
	Weibull(0.8)	0.996	0.995	**1**	0.998	0.999	0.999	0.999
	Weibull(1.4)	0.314	0.500	**0.703**	0.579	0.611	0.579	0.557

Table 6: Power estimates of different goodness of fit tests in SRS and RSS designs for significance level $n = 40$ and $\alpha = 0.05$

Sampling scheme	Alternative distribution	KL $m = 6$	KS	A^2	W^2	Z_K	Z_A	Z_C
SRS	Exponential(1)	0.990	0.991	**0.999**	0.996	0.997	0.998	0.998
	Uniform(0,1)	**0.871**	0.283	0.651	0.377	0.64	0.737	0.744
	Beta (0.5,0.5)	**1**	0.897	0.998	0.958	0.996	0.999	0.999
	Beta(2,1)	**0.998**	0.854	0.950	0.953	0.934	0.99	0.964
	Gamma (0.5)	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	Gamma(2)	0.38	0.571	**0.686**	0.639	0.557	0.527	0.555
	Weibull(0.8)	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	Weibull(1.4)	0.520	0.638	0.796	0.708	0.722	0.698	0.722
RSS $k = 2$	Exponential(1)	0.993	0.994	**0.999**	0.997	0.998	**0.999**	**0.999**
	Uniform(0,1)	**0.874**	0.285	0.668	0.393	0.649	0.747	0.755
	Beta (0.5,0.5)	**1**	0.917	0.998	0.967	0.997	1	0.999
	Beta(2,1)	**1**	0.889	0.975	0.976	0.969	0.998	0.984
	Gamma (0.5)	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	Gamma(2)	0.398	0.605	**0.719**	0.677	0.586	0.561	0.581
	Weibull(0.8)	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	Weibull(1.4)	0.539	0.673	**0.829**	0.744	0.750	0.731	0.747
RSS $k = 5$	Exponential(1)	0.999	0.999	**1**	**1**	**1**	**1**	**1**
	Uniform(0,1)	**0.923**	0.348	0.790	0.508	0.738	0.828	0.820
	Beta (0.5,0.5)	**1**	0.971	**1**	0.992	**1**	**1**	**1**
	Beta(2,1)	**1**	0.964	0.999	0.999	0.999	**1**	0.999
	Gamma (0.5)	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	Gamma(2)	0.442	0.707	0.826	**0.784**	0.691	0.658	0.648
	Weibull(0.8)	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	Weibull(1.4)	0.611	0.785	**0.920**	0.855	0.849	0.831	0.823

5. Conclusion

The goodness of fit tests for the Rayleigh distribution using SRS and RSS methods is considered in this paper. Monte Carlo simulation are devoted to study the power of the tests. It is found that the tests based on RSS are more powerfull than thier SRS counterparts. Also, the test Anderson-Darling is the most powerful among all tests in most cases.

References

1. Abd-Elfattah, A.M. (2011). Goodness of fit test for the generalized Rayleigh distribution with unknown parameters. *Journal of Statistical Computation and Simulation*, 81(3), 357-366.
2. Al-Omari, A.I. (2014). Estimation of entropy using random sampling. *Journal of Computational and Applied Mathematics*, 261, 95-102.
3. Al-Omari, A.I. (2011). Estimation of mean based on modified robust extreme ranked set sampling. *Journal of Statistical Computation and Simulation*, 81(8), 1055-1066.
4. Chakravarti, Laha, and Roy, (1967). *Handbook of Methods of Applied Statistics, Volume I*, John Wiley and Sons, 392-394.
5. Correa, J.C. (1995). A new estimator of entropy. *Communication in Statistics-Theory Methods*, 24(10), 2439–2449.
6. Darling, D.A. (1957). The Kolmogrov-Smirnov, Gramer-Von Mises tests. *The Annals of Mathematical Statistics*, 823-838.
7. David, H.A. and Nagaraja, H.N. (2003). Order Statistics, Third Edition, John Wiley & Sons, Inc., Hoboken, New Jersey.
8. Ebrahimi, N., Pflughoeft, K. and Soofi, E. (1994). Two measures of sample entropy. *Statistics & Probability Letters*, 20, 225–234.
9. Haq, A., Brown, J., Moltchanova, E. and Al-Omari, A.I. (2013). Partial ranked set sampling design. *Environmetrics*, 24(3): 201-207.
10. Hirano, K. (1986). Rayleigh Distributions, Wiley, New York, 1986.
11. Kullback, S. and Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, 22, 79-86.
12. Mahdizadeh, M. (2012). On the use of ranked set samples in entropy based test of fit for the Laplace distribution. *Revista Colombiana de Estadística*, 35(3): 443-455.
13. Mahdizaheh, M. and Arghami, N.R. (2010). Efficiency of ranked set sampling in entropy estimation and goodness-of-fit testing for the inverse Gaussian law. *Journal of Statistical Computation and Simulation*, 80(7), 761-774.
14. McIntyre, G.A. (1952). A method for unbiased selective sampling using ranked sets. *Australian Journal Agricultural Research*. 3, 385-390.

15. Noughabi, R.A., Noughabi, H.A., and Behabadi, A.E.M. (2014). An entropy test for the Rayleigh distribution and power comparison. *Journal of Statistical Computation and Simulation*, 84(1), 151-158.

16. Rayleigh, J. (1980). On the resultant of a large number of vibrations of the same pitch and of arbitrary phase, Philos. Mag., 10, 73–78.

17. Shannon, C.E. (1948). A mathematical theory of communications. *Bell System Technical Journal*, 27, 379–423, 623–656.

18. Siddiqui, M.M. (1962). Some problems connected with Rayleigh distributions, J. Res. Nat. Bur. Stand, 60D, 167–174.

19. Song, K.S. (2002). Goodness of fit tests based on Kullback-Leibler discrimination. *IEEE Transactions on Information Theory*, 48(5), 1103-1117.

20. Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons, *Journal of the American Statistical Association*, 69, 730-737.

21. Stephens, M. A. (1979). Tests of Fit for the Logistic Distribution Based on the Empirical Distribution Function, *Biometrika*, 66, 591-595.

22. Yamane, T. (1998). Application of the Rayleigh distribution to size selectivity of small prawn pots for the oriental river prawn, Macrobrachium nipponense, Fisheries Res. 36(1): 27–33.

23. Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of the Royal Statistical* Society. B 38, 54–59.

24. Wieczorkowski, R. and Grzegorzewsky, P. (1999). Entropy estimators improvements and comparisons. *Communication in Statistics-Simulation and Computation*, 28(2), 541–567.

25. Zamanzade, E., Arghami, N.R., and Vock, M. (2014). A parametric test of perfect ranking in balanced ranked set sampling. *Communications in Statistics - Theory and Methods*, 43(21), 4589-4611.

26. Zhang, J. (2002). Powerful goodness-of-fit tests based on the likelihood ratio. *Journal of the Royal Statistical Society: Series B*, 64(2), 281-294.