Revista Colombiana de Estadística
July 2017, Volume 40, Issue 2, pp. 279 to 290
DOI: http://dx.doi.org/10.15446/rce.v40n2.60375

Goodness of Fit Tests for Rayleigh Distribution Based on Phi-Divergence

Pruebas de bondad de ajuste para distribución Rayleigh basadas en Divergencia Phi

Ehsan Zamanzade[1,a], M. Mahdizadeh[2,b]

[1]Department of Statistics, University of Isfahan, Isfahan, Iran
[2]Department of Statistics, Hakim Sabzevari University, Sabzevar, Iran

Abstract

In this paper, we develop some goodness of fit tests for Rayleigh distribution based on Phi-divergence. Using Monte Carlo simulation, we compare the power of the proposed tests with some traditional goodness of fit tests including Kolmogorov-Smirnov, Anderson-Darling and Cramer von-Mises tests. The results indicate that the proposed tests perform well as compared with their competing tests in the literature. Finally, the new procedures are illustrated via two real data sets.

***Key words*:** Goodness of fit test, Monte Carlo simulation, Phi-divegence, Rayleigh distribution.

Resumen

En este artículo desarrollamos pruebas de bondadn de ajuste para distribución Rayleigh basados en divergencia Phi. Usando simulaciones de Monte Carlo, comparamos el poder de las pruebas propuestas con algunas pruebas tradicionales incluyendo Kolmogorov-Smirnov, Anderson-Darling y Cramer von-Mises. Los resultados indican que la prueba propuesta funciona mejor que las otras pruebas reportadas en literatura. Finalmente, los procedimientos nuevos son ilustrados sobre dos conjuntos de datos reales.

***Palabras clave*:** distribución Rayleigh, Divergencia Phi, Pruebas de bondad de ajuste, Simulaciones Monte Carlo.

[a]PhD. E-mail: e.zamanzade@sci.ui.ac.ir; ehsanzamanzadeh@yahoo.com
[b]PhD. E-mail: mahdizadeh.m@live.com

1. Introduction

A continuous random variable X is said to have Rayleigh distribution with location parameter $\mu \in \mathbb{R}$, and scale parameter $\sigma > 0$, denoted by $X \sim Ra(\mu, \sigma)$, if its probability distribution function (pdf) is given by

$$f_0(x, \mu, \sigma) = \frac{x-\mu}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \geq \mu, \tag{1}$$

with the corresponding cumulative distribution function (cdf)

$$F_0(x, \mu, \sigma) = 1 - e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{2}$$

for $x \in [\mu, \infty)$.

Figure 1 presents the pdf and cdf of Rayleigh distribution for $\mu = 0$ and different values of σ.

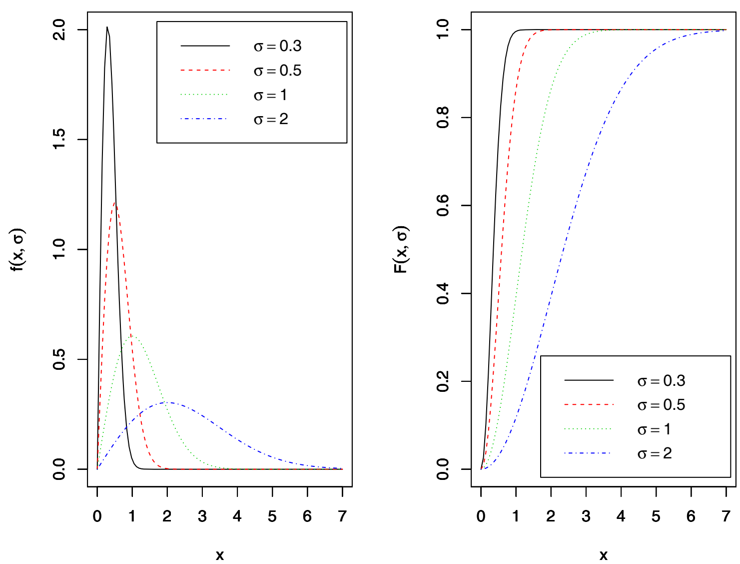

FIGURE 1: Probability density function (left panel) and cumulative distribution function (right panel) of Rayleigh distribution for $\mu = 0$ and different values of σ. This figure appears in color in the electronic version of this paper.

It is evident from this figure that the Rayleigh distribution is right skewed, and therefore it provides a useful tool for modelling right-skewed data. The Rayleigh distribution was firstly motivated with a problem in acoustics, and has been utilized for modelling the distribution of the distance between two individuals in a Poisson process. This distribution typically arises when overall size of a vector is related to its directional components. In fact, it is well-known that if Z and W are two independent and identical random variables from a standard normal distribution with mean zero and variance σ^2, then $X = \sqrt{Z^2 + W^2}$ follows a Rayleigh distribution with location parameter $\mu = 0$ and scale parameter σ. For more information about the applications and properties of the Rayleigh distribution, we refer the interested reader to Siddiqui (1962) and Johnson, Kotz & Balakrishnan (1994).

This paper introduces some goodness of fit tests for Rayleigh distribution when the location parameter μ is known and the scale parameter σ is unknown. Without loss of generality, throughout this paper we assume that μ is known to be zero. If $\mu \neq 0$, then one can simply apply the proposed procedures in this paper after subtracting the parameter μ from data. Dey, Dey & Kundu (2014) discussed the problem of estimating the parameters of Rayleigh distribution, and they proposed a numerical method for obtaining the maximum likelihood estimators of the Rayleigh parameters.

In Section 2, we develop some goodness of fit tests for Rayleigh distribution based on Phi-divergence. We compare the proposed tests with some well-known goodness of fit tests for Rayleigh distribution via Monte Carlo simulation in Section 3. We show that the proposed tests are generally more powerful than their competitors in the literature. In Section 4, we illustrate the proposed procedures by using two real data examples. Section 5 contains some concluding remarks.

2. Tests of Fit for Rayleigh Distribution Based on Phi-Divergence

Let P_1 and P_2 be two probability measures over measurable space Ω. If P_1 is absolutely continuous with respect to P_2, then Phi-divergence of P_1 from P_2 is denoted by $D_\phi\left(P_1||P_2\right)$ and is given by

$$D_\phi\left(P_1||P_2\right) = \int_\Omega \phi\left(\frac{dP_1}{dP_2}\right) dP_2, \tag{3}$$

where $\phi(\cdot)$ is a convex function such that $\phi(1) = 0$ and $\phi''(1) > 0$.

It is important to note that the Phi-divergence measure enjoys the non-negativity and joint convexity properties. This means that $D_\phi\left(P_1||P_2\right) \geq 0$ and the equality happens if and only if $P_1 = P_2$ almost surely, and also for a fixed $\gamma \in [0, 1]$, we have:

$$D_\phi\left(\gamma P_1 + (1-\gamma) P_1||\gamma P_2 + (1-\gamma) P_2\right) \leq \gamma D_\phi\left(P_1||P_2\right) + (1-\gamma) D_\phi\left(P_1||P_2\right),$$

We refer the interested reader to Csiszar (1967), and Liese & Vajda (2006) and references therein for more details about the Phi-divergence.

Let f_0 and f be the pdf under null and alternative hypotheses, respectively. Some well-known divergence measures in statistics can be obtained by appropriate choice of $\phi(\cdot)$ function in (3) as follows:

- Kullback-Liebler (KL) distance is obtained by replacing $\phi(t)$ by $-\log(t)$ in (3) as

$$KL = \int f(x) \log\left(\frac{f(x)}{f_0(x)}\right) dx = E_f\left(\log\left(\frac{f(x)}{f_0(x)}\right)\right).$$

- Hellinger (H) distance is obtained by using $\frac{1}{2}\left(1-\sqrt{t}\right)^2$ as the $\phi(t)$ in (3) as

$$H=\frac{1}{2}\int f(x)\left(1-\sqrt{\frac{f_0(x)}{f(x)}}\right)^2 dx=\frac{1}{2}E_f\left(\left(1-\sqrt{\frac{f_0(x)}{f(x)}}\right)^2\right).$$

- Jeffreys (J) distance is obtained by replacing $\phi(t)$ by $(t-1)\log(t)$ in (3) as

$$\begin{aligned}J&=\int f(x)\left(\frac{f_0(x)}{f(x)}-1\right)\log\left(\frac{f_0(x)}{f(x)}\right)dx\\&=E_f\left(\left(\frac{f_0(x)}{f(x)}-1\right)\log\left(\frac{f_0(x)}{f(x)}\right)\right).\end{aligned}$$

- Total variation (TV) distance is obtained by replacing $\phi(t)$ by $|t-1|$ in (3) as

$$TV=\int f(x)\left|\frac{f_0(x)}{f(x)}-1\right|dx=E_f\left(\left|\frac{f_0(x)}{f(x)}-1\right|\right).$$

- Chi-square (χ^2) distance is obtained by replacing $\phi(t)$ by $\frac{1}{2}\left(1-t\right)^2$ in (3) as

$$\chi^2=\frac{1}{2}\int f(x)\left(1-\frac{f_0(x)}{f(x)}\right)^2 dx=\frac{1}{2}E_f\left(\left(1-\frac{f_0(x)}{f(x)}\right)^2\right).$$

Alizadeh Noughabi & Balakrishnan (2016) utilized above distances to introduce goodness of fit tests for normal, exponential, uniform and Laplace distributions.

Let $x_1,\ldots,x_n$ be a simple random sample of size n from the population of interest with pdf $f(x)$. Suppose that we are interested in testing the null hypothesis $H_0: f(x)=f_0(x,0,\sigma)$ for some $\sigma>0$, where $f_0(x,0,\sigma)$ is defined in (1) against the alternative hypothesis $H_1: f(x)\neq f_0(x,0,\sigma)$ for all $\sigma>0$. Let

$$D_n\left(f_0(x,0,\sigma)||f(x)\right)=\frac{1}{n}\sum_{i=1}^{n}\phi\left(\frac{f_0(x_i,0,\sigma)}{f(x_i)}\right),$$

be the sample estimate of $D\left(f_0(x,0,\sigma)||f(x)\right)$. Let $\hat{f}(x)$ be a nonparametric density estimator and $\hat{\sigma}$ be an appropriate estimator of σ under H_0. Then $D_n\left(f_0(x,0,\hat{\sigma})||\hat{f}(x)\right)$ can be considered as a test statistic for testing H_0 versus H_1, and the null hypothesis H_0 should be rejected for large enough values of $D_n\left(f_0(x,0,\hat{\sigma})||\hat{f}(x)\right)$.

We propose to estimate $f(x_i)$ by using kernel density estimator, i.e.

$$\hat{f}(x_i)=\frac{1}{nh}\sum_{j=1}^{n}k\left(\frac{x_i-x_j}{h}\right), \tag{4}$$

where $k(\cdot)$ is pdf of a standard normal distribution, $h=\left(4s^2/(3n)\right)^{1/5}$ is the bandwidth suggested by Silverman (1986), and $s^2=\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2/(n-1)$ is the sample variance. We also replace the parameter σ by its maximum likelihood estimator, i.e.

$$\hat{\sigma} = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} x_i^2}. \tag{5}$$

Therefore, the null hypothesis that the population distribution follows a Rayleigh distribution is rejected if

$$\hat{D}_n\left(\hat{f}_0(x,0,\sigma)||\hat{f}(x)\right) = \frac{1}{n}\sum_{i=1}^{n}\phi\left(\frac{\hat{f}_0(x_i,0,\sigma)}{\hat{f}(x_i)}\right) \geq d_\phi(n,\alpha), \tag{6}$$

where $d_\phi(n,\alpha)$ is $(1-\alpha)$ quantile of $\hat{D}_n$ under the null hypothesis.

In what follows, the test statistic obtained by using $\phi(\cdot)$ function corresponding to Kullback-Liebler, Hellinger, Jeffreys, Total variation and Chi-square distances are denoted by KL_n, H_n, J_n, TV_n and χ_n^2, respectively.

We also report the critical values of the proposed test statistics based on Monte Carlo simulation at significance level $\alpha = 0.01, 0.05$ and 0.1. For this purpose, we have generated 50,000 random samples from Rayleigh distribution with location parameter $\mu = 0$ and scale parameter $\sigma = 1$ using VGAM package of R statistical software. We then estimated the $(1-\alpha)$ quantiles of the proposed test statistics by using their $(1-\alpha)$ sample quantiles. It is important to note that all proposed test statistics are invariant under scale transformation, and therefore the critical values of the tests do not depend on the unknown parameter σ. The critical values are reported in Table 1.

Another way to estimate $D\left(f_0(x,0,\sigma)||f(x)\right)$ is to use spacing technique for estimation of $f(x)$. Alizadeh Noughabi, Alizadeh Noughabi & Ebrahimi Moghaddam Behdadi (2014) and Jahanshahi, Habibi Rad & Fakoor (2016) used this technique to develop a goodness of fit test based for Rayleigh distribution based on Kullback-Liebler and Hellinger distances, respectively. However, the drawback of estimation of $D\left(f_0(x,0,\sigma)||f(x)\right)$ based on spacing is that it is not applicable whenever there are ties in data. We also refer the interested reader to Mahdizadeh (2012) and Zamanzade & Mahdizadeh (2017) for more information on this topic.

3. Power Comparison

In this section, we compare the powers of the proposed tests with those of the competitor tests in the literature. Let $x_{(1)}, \ldots, x_{(n)}$ be an ordered random sample corresponding to simple random sample of $x_1, \ldots, x_n$. Let $F_0\left(x,0,\sigma\right)$ and $\hat{\sigma}$ be as defined in (2) and (5). The competing test statistics considered in this section are

- Kolmogorov-Smirnov (KS) test statistic which has the form

$$KS = \max\left\{\max_{i=1,\ldots,n}\left\{\frac{i}{n} - F_0\left(x_{(i)},0,\hat{\sigma}\right)\right\}, \max_{i=1,\ldots,n}\left\{F_0\left(x_{(i)},0,\hat{\sigma}\right) - \frac{i-1}{n}\right\}\right\}.$$

TABLE 1: Values of $d_\phi(n,\alpha)$ for different tests when $n = 10, 20, 50$ and $\alpha = 0.01, 0.05, 0.1$.

n	α	KL_n	H_n	J_n	TV_n	χ^2_n
	0.01	0.412	0.040	0.327	0.471	0.141
10	0.05	0.242	0.026	0.211	0.383	0.094
	0.1	0.169	0.020	0.162	0.340	0.075
	0.01	0.208	0.025	0.213	0.365	0.092
20	0.05	0.125	0.018	0.143	0.305	0.063
	0.1	0.089	0.014	0.116	0.274	0.052
	0.01	0.144	0.020	0.166	0.320	0.072
30	0.05	0.088	0.014	0.116	0.268	0.050
	0.1	0.063	0.011	0.094	0.242	0.041
	0.01	0.111	0.017	0.139	0.290	0.059
40	0.05	0.067	0.012	0.097	0.242	0.042
	0.1	0.048	0.010	0.079	0.219	0.034
	0.01	0.086	0.014	0.116	0.267	0.050
50	0.05	0.055	0.010	0.084	0.225	0.036
	0.1	0.039	0.008	0.069	0.203	0.030
	0.01	0.049	0.009	0.078	0.214	0.033
100	0.05	0.028	0.007	0.056	0.180	0.024
	0.1	0.020	0.006	0.047	0.162	0.020

- Anderson-Darling (AD) test statistic which has the from

$$AD = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left\{\log\left(F_0\left(x_{(i)},0,\hat{\sigma}\right)\right) + \log\left(1 - F_0\left(x_{(i)},0,\hat{\sigma}\right)\right)\right\}.$$

- Cramer-von Mises (CM) test statistic which has the form

$$\frac{1}{12n} + \sum_{i=1}^{n}\left(\frac{2i-1}{n} - F_0\left(x_{(i)},0,\hat{\sigma}\right)\right)^2.$$

All above tests reject the null hypothesis that the population distribution is Rayleigh if their observed values are greater than the corresponding critical points at a given significance level.

In order to compare the power of different goodness of fit tests, we generated 50,000 simple random samples of sizes $n = 10, 20$ and 50 under 14 alternative distributions considered by Best et al. (2010). We then estimated the power of the aforementioned tests by dividing the number of test statistics values are greater than their critical points at significance level $\alpha = 0.05$ to number of the repetitions in the simulation study, i.e. 50,000. The alternative distributions as follows:

- Weibull distribution with scale parameter 1 and shape parameter θ, denoted by $W(\theta)$, for $\theta = 1, 2, 3$.

- Gamma distribution with scale parameter 1 and shape parameter θ, denoted by $G(\theta)$, for $\theta = 1.5, 2$.
- Lognormal distribution with pdf $\frac{1}{\theta x\sqrt{2\pi}} \exp\{-(\log x)^2/(2\theta^2)\}$, denoted by $LN(\theta)$, for $\theta = 0.8$.
- Gompertz distribution with cdf $1 - \exp\{(1 - e^x)/\theta\}$, denoted by $GO(\theta)$, for $\theta = 0.5, 1.5$.
- Power distribution with pdf $(x^{\frac{1}{\theta}-1})/\theta$ for $0 \leq x \leq 1$, denoted by $PW(\theta)$, for $\theta = 1, 2$.
- Exponential power distribution with cdf $1 - \exp\left\{1 - \exp(x^\theta)\right\}$, denoted by $EP(\theta)$, for $\theta = 1, 2$.
- Poisson-exponential distribution defined as $\sum_{i=1}^{N} X_i$, where X_i's are independent random variables from standard exponential distribution and N follows a poisson distribution with parameter θ, denoted by $PE(\theta)$, for $\theta = 3, 4$.

The simulation results are presented in Tables 2-4. Table 2 presents the results for sample size $n = 10$. It is clear from this table that none of the tests can dominate the others in terms of power. However, the tests based on Jeffreys and Hellinger distances have generally higher powers, and the difference between their performances is negligible in most considered cases. In fact, the J_n and H_n are the most powerful tests for 8 out 14 alternative distributions considered in our simulation study. The AD test also has good powers, and it is the best test when the alternative distributions are $W(3), G(1.5)$ and $LN(0.8)$. It is worth noting that all considered goodness of fit tests have the highest powers for $W(1)$ distribution and the lowest powers for $W(2)$ distribution, which is a special case of Rayleigh distribution.

Table 2: Estimated powers for different tests of size 0.05 when $n = 10$.

Dist.	KS	AD	CM	KL_n	H_n	J_n	TV_n	χ_n^2
$W(1)$	0.571	0.771	0.619	0.428	0.770	0.770	0.727	0.755
$W(2)$	0.050	0.051	0.051	0.051	0.051	0.051	0.050	0.050
$W(3)$	0.210	0.186	0.249	0.344	0.179	0.176	0.190	0.172
$G(1.5)$	0.322	0.474	0.361	0.128	0.454	0.455	0.403	0.435
$G(2)$	0.180	0.258	0.203	0.040	0.229	0.231	0.187	0.211
$LN(0.8)$	0.387	0.454	0.422	0.132	0.378	0.381	0.277	0.338
$GO(0.5)$	0.310	0.540	0.349	0.211	0.551	0.551	0.506	0.532
$GO(1.5)$	0.134	0.305	0.146	0.108	0.333	0.334	0.289	0.311
$PW(1)$	0.102	0.260	0.113	0.131	0.287	0.287	0.244	0.272
$PW(2)$	0.582	0.871	0.620	0.707	0.883	0.883	0.842	0.870
$EP(1)$	0.195	0.400	0.218	0.139	0.418	0.419	0.374	0.399
$EP(2)$	0.141	0.121	0.165	0.225	0.138	0.135	0.153	0.143
$PE(3)$	0.318	0.660	0.359	0.488	0.669	0.670	0.586	0.629
$PE(4)$	0.170	0.402	0.194	0.230	0.414	0.415	0.335	0.371

TABLE 3: Estimated powers for different tests of size 0.05 when $n = 20$.

Dist.	KS	AD	CM	KL_n	H_n	J_n	TV_n	χ^2_n
$W(1)$	0.863	0.956	0.899	0.803	0.967	0.967	0.945	0.957
$W(2)$	0.051	0.051	0.049	0.050	0.052	0.052	0.052	0.053
$W(3)$	0.395	0.439	0.479	0.583	0.290	0.281	0.336	0.334
$G(1.5)$	0.576	0.733	0.638	0.358	0.760	0.760	0.677	0.714
$G(2)$	0.324	0.430	0.372	0.120	0.445	0.448	0.335	0.380
$LN(0.8)$	0.664	0.726	0.708	0.393	0.706	0.709	0.552	0.635
$GO(0.5)$	0.558	0.784	0.612	0.469	0.830	0.829	0.777	0.798
$GO(1.5)$	0.224	0.459	0.248	0.216	0.551	0.549	0.493	0.510
$PW(1)$	0.153	0.395	0.187	0.283	0.527	0.520	0.547	0.534
$PW(2)$	0.872	0.983	0.891	0.945	0.991	0.990	0.986	0.989
$EP(1)$	0.348	0.603	0.388	0.298	0.678	0.677	0.619	0.640
$EP(2)$	0.242	0.250	0.294	0.366	0.224	0.216	0.284	0.270
$PE(3)$	0.567	0.879	0.626	0.764	0.900	0.900	0.818	0.853
$PE(4)$	0.298	0.609	0.344	0.431	0.656	0.657	0.512	0.563

TABLE 4: Estimated powers for different tests of size 0.05 when $n = 50$.

Dist.	KS	AD	CM	KL_n	H_n	J_n	TV_n	χ^2_n
$W(1)$	0.998	1.000	0.999	0.997	1.000	1.000	1.000	1.000
$W(2)$	0.054	0.052	0.051	0.052	0.053	0.053	0.053	0.051
$W(3)$	0.818	0.901	0.898	0.926	0.652	0.637	0.721	0.743
$G(1.5)$	0.931	0.977	0.954	0.813	0.981	0.981	0.956	0.970
$G(2)$	0.663	0.771	0.732	0.382	0.794	0.796	0.651	0.717
$LN(0.8)$	0.958	0.973	0.972	0.843	0.972	0.973	0.925	0.952
$GO(0.5)$	0.925	0.985	0.950	0.885	0.991	0.991	0.981	0.986
$GO(1.5)$	0.508	0.775	0.543	0.511	0.862	0.859	0.807	0.829
$PW(1)$	0.358	0.748	0.471	0.743	0.908	0.902	0.924	0.910
$PW(2)$	0.998	1.000	0.999	1.000	1.000	1.000	1.000	1.000
$EP(1)$	0.721	0.907	0.762	0.677	0.946	0.945	0.908	0.927
$EP(2)$	0.557	0.632	0.663	0.729	0.513	0.496	0.629	0.598
$PE(3)$	0.927	0.996	0.949	0.982	0.997	0.997	0.986	0.994
$PE(4)$	0.631	0.899	0.696	0.783	0.929	0.929	0.806	0.870

Simulation results for sample size $n = 20$ are given in Table 3. The pattern of the performance of goodness of fit tests is somewhat similar to that of Table 2. The highest possible power occurs when $W(1)$ is the alternative distribution and the lowest powers occur when $W(2)$ is the alternative distribution. However, as one expects, the powers for sample size $n = 20$ are higher than those for sample size $n = 10$. We also observe that the tests based on Jeffreys and Hellinger distances have the best performance, and they are the most powerful tests for 8 out of 12 considered alternative distributions. However, the tests based on KL and TV distances and AD are the best tests in certain cases.

Table 4 gives the simulation results for sample size $n = 50$. It is clear that the powers in this table are higher than those in Tables 2-3.

Finally, we would like to mention that $W(2)$ distribution is a special case of Rayleigh distribution. Therefore, it is clear from Tables 2-4 that all considered tests control type **I** error well.

4. Real Data Examples

In this section, we applied our proposed procedures to two real data sets for illustration purpose.

Example 1. The data set used for this example is known as average wind speed data (AWSD) and can be found in Best, Rayner & Thas (2010). The AWSD includes 30 average daily wind speeds (in km/h) for the month of November 2007 recorded at Elanora Heights, a northeastern suburb of Sydney, Australia. The data are as follows:

```
2.7, 3.2, 2.1, 4.8, 7.6, 4.7, 4.2, 4.0, 2.9, 2.9,
4.6, 4.8, 4.3, 4.6, 3.7, 2.4, 4.9, 4.0, 7.7, 10.0,
5.2, 2.6, 4.2, 3.6, 2.5, 3.3, 3.1, 3.7, 2.8, 4.0.
```

Suppose that we are interested in testing the null hypothesis that the average wind speed follows a Rayleigh distribution and we also assume that the location parameter is known and $\mu = 1.5$. Then, the estimated value of σ based on the transformed data (after subtracting the location parameter) is given by $\hat{\sigma} = 2.231$. The Q-Q plot and the corresponding histogram along with fitted Rayleigh density function is presented in Figure 2.

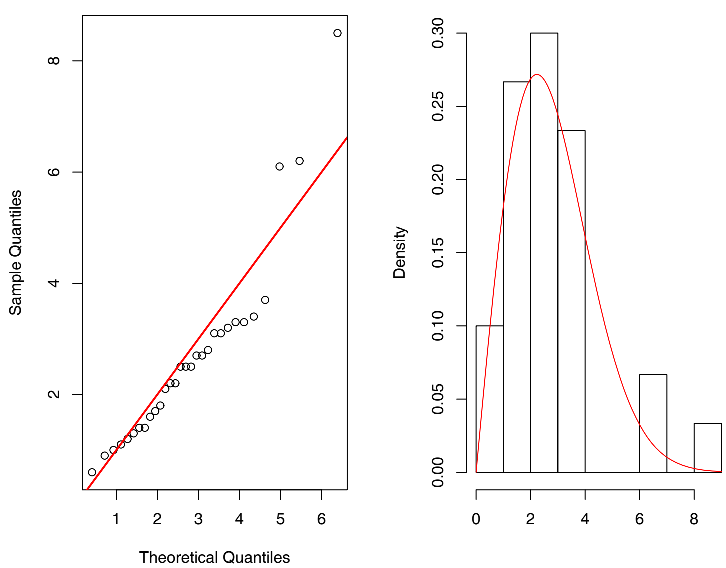

Figure 2: The Rayleigh Q-Q plot of the transformed AWSD (left panel) and the corresponding histogram along with fitted Rayleigh density (right panel). This figure appears in color in the electronic version of this paper.

The values of the proposed test statistics for Rayleigh distribution are $KL_n = 0.068$, $H_n = 0.009$, $J_n = 0.083$, $TV_n = 0.078$, $\chi^2_n = 0.016$. Comparing above test statistics values with their corresponding critical values at significance level

$\alpha = 0.05$ for sample size $n = 30$, we observe that the null hypothesis is not rejected. This is consistent with what is reported by Best et al. (2010).

Example 2. The second data set we use for illustrative purpose is survival times for lung cancer patients (in days) which is collected by Lawless (1982). This data set was also analyzed by Soliman & Al-Aboud (2008). The data are as follows:

```
6.96, 9.30, 6.96, 7.24, 9.30, 4.90, 8.42, 6.05,
10.18, 6.82, 8.58, 7.77, 11.94, 11.25, 12.94, 12.94.
```

Here, we assume that the location parameter is known and equals to $\mu = 3.5$ and we apply our procedures on after subtracting location parameter from them. The estimated value of the parameter of σ is given by $\hat{\sigma} = 3.499$. Figure 3 shows Q-Q plot and the corresponding histogram along with fitted Rayleigh density for this data set.

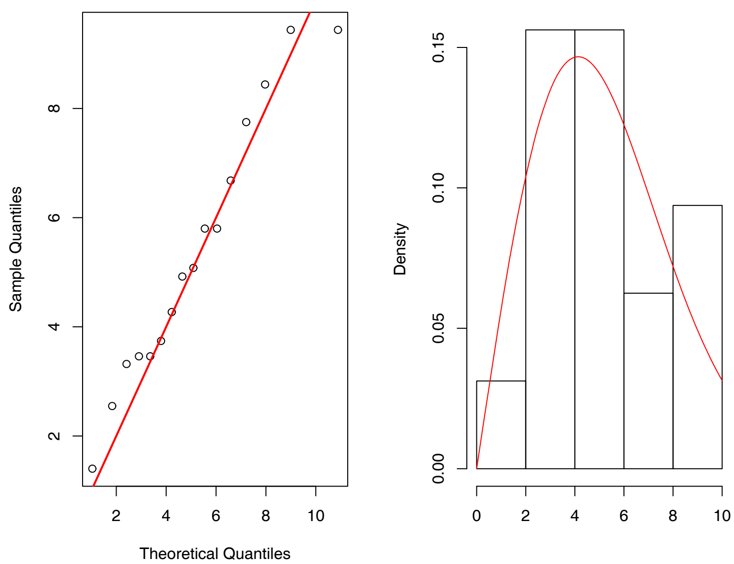

FIGURE 3: The Rayleigh Q-Q plot of the transformed survival times for lung cancer patients (left panel) and the corresponding histogram along with fitted Rayleigh density (right panel). This figure appears in color in the electronic version of this paper.

The values of the proposed test statistics for Rayleigh distribution are $KL_n = 0.009$, $H_n = 0.003$, $J_n = 0.028$, $TV_n = 0.122$, $\chi^2_n = 0.013$. On the other hand, the simulated critical values for KL_n, H_n, J_n, TV_n and χ^2_n for $n = 16$ and $\alpha = 0.05$ are $0.152, 0.020, 0.164, 0.329$ and 0.072, respectively. Thus, the null hypothesis is not rejected.

5. Conclusion

This paper deals with constructing goodness of fit tests for Rayleigh distribution based on Phi-divergence. Several goodness of fit tests are proposed, and their powers are compared with some well-known goodness of fit tests based on empirical distribution function. It is found that the introduced tests have good performance as compared with their competitors.

Acknowledgements

The authors are grateful to the reviewers for providing valuable comments and suggestions.

[Received: September 2016 — Accepted: April 2017]

References

Alizadeh Noughabi, H. & Balakrishnan, N. (2016), 'Tests of goodness of fit based on Phi-divergence', *Journal of Applied Statistics* **43**(3), 412–429.

Alizadeh Noughabi, R., Alizadeh Noughabi, H. & Ebrahimi Moghaddam Behdadi, A. (2014), 'An entropy test for the rayleigh distribution and power comparison', *Journal of Statistical Computation and Simulation* **84**(1), 151–158.

Best, D., Rayner, J. & Thas, O. (2010), 'Easily applied tests of fit for the Rayleigh distribution', *Synkhya, B* **72**, 254–263.

Csiszar, I. (1967), 'On topology properties of f-divergences', *Studia Scientiarum Mathematicarum Hungarica* **2**, 329–339.

Dey, S., Dey, T. & Kundu, D. (2014), 'Two parameter Rayleigh distribution: Different methods of estimation', *merican Journal of Mathematical and Management Sciences* **33**(1), 55–74.

Jahanshahi, S., Habibi Rad, A. & Fakoor, V. (2016), 'A goodness-of-fit test for Rayleigh distribution based on Helinger distance', *Annals of Data Sciences* **3**(4), 401–411.

Johnson, N., Kotz, S. & Balakrishnan, N. (1994), *Continuous univariate distributions*, Wiley.

Lawless, J. F. (1982), *Statistical model and methods for lifetime data*, Wiley, New York.

Liese, F. & Vajda, I. (2006), 'On divergences and informations in statistics and information theory', *IEEE Transactions on Information Theory* **52**, 4394–4412.

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.

Siddiqui, M. (1962), 'Some problems connected with Rayleigh distribution', *Journal of Research of the National Bureau of Standards* **66D**, 167–174.

Soliman, A. & Al-Aboud, F. M. (2008), 'Bayesian inference using record values from Rayleigh model with application', *European Journal of Operational Research* **185**, 659–672.

Zamanzade, E. & Mahdizadeh, M. (2017), 'Entropy estimation from ranked set samples with application to test of fit', *Revista Colombiana de Estadística* **40**(2), 223–241.