Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 49 (6) (2020), 2104 – 2118
DOI : 10.15672/hujms.629192

Research Article

An estimation of Phi divergence and its application in testing normality

Mahsa Tavakoli*[1], Hadi Alizadeh Noughabi[2], Gholam Reza Mohtashami Borzadaran[1]

[1] Department of Statistics, Ferdowsi University of Mashhad, Iran
[2] Department of Statistics, University of Birjand, Birjand, Iran

Abstract

In this article, a new goodness of fit test for normality is introduced based on Phi divergence. The test statistic is estimated using spacing and the consistency of the test is proved. Then with replacing some special cases of Phi divergence, the efficiency of each test statistic is analyzed by Monte Carlo simulation against some competitors (based on Phi divergence using kernel density function and also some classical competitors). It is shown that each special case of Phi divergence based test is the most powerful in each group of alternatives (depending on symmetry or support).

Mathematics Subject Classification (2010). 62G10, 62G20

Keywords. Phi divergence measure, normality test, spacing, Monte Carlo simulation, test power

1. Introduction

The normal distribution is the most widely known and used of all distributions. Because the normal distribution approximates many natural phenomena. So, it plays a predominant role in various sciences and many applications.
Since Pearson [29] proposed the chi-squared goodness of fit test for normality, considerable attention has been given to the problem of goodness of fit test and various tests were proposed, like the moment-based tests, regression-based tests and entropy-based tests. Some other popular tests for normality founded by Cramer [13], Von Mises [38], Kolmogorov [22], Smirnov [34], Shapiro-Wilk [33], Anderson-Darling [7], Kuiper [23] and Watson [40]. Let f denote the density of a given population, then the null hypothesis of normality is stated formally as

$$H_0 : f(x) = g(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp\{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2\},$$

for some $(\mu, \sigma) \in \Theta = \mathbb{R} \times \mathbb{R}^+$. Also the alternative hypotheses is

$$H_1 : f(x) \neq g(x; \mu, \sigma),$$

*Corresponding Author.
Email addresses: mahsa_tavakoli88@yahoo.com (M. Tavakoli), alizadehhadi@birjand.ac.ir (H. Alizadeh Noughabi), grmohtashami@um.ac.ir (G.R. Mohtashami Borzadaran)
Received: 08.10.2019; Accepted: 05.10.2020

for any $(\mu, \sigma) \in \Theta$, where μ and σ are unspecified.
For the first time, Vasicek [37] introduced a normality test based on entropy. The entropy of X with distribution function $F(x)$ and a continuous density function $f(x)$ is defined by Shannon [32] as

$$H(f) = -\int_{-\infty}^{\infty} f(x) \log f(x) dx. \tag{1}$$

The problem of estimating Shannon entropy has been considered by many authors such as Ahmad and Lin [1], Vasicek [37], Dudewicz and Van der Meulen [14], Joe [20], Van Es [36], Ebrahimi *et al.* [17], Correa [12], Wieczorkowski and Grzegorzewski [41], Yousefzadeh and Arghami [42] and Alizadeh [3]. Among these various entropy estimators, Vasicek's sample entropy has been the most widely used in developing entropy-based statistical procedures. Using $F(x) = p$, Vasicek [37] expressed equation (1) as

$$H(f) = \int_0^1 \log\{\frac{d}{dp} F^{-1}(p)\} dp, \tag{2}$$

and by replacing the distribution function F by the empirical distribution function F_n and using a difference operator instead of the differential operator, the estimator is given as

$$H_{n,m} = \frac{1}{n} \sum_{i=1}^{n} \log \left\{ \frac{n}{2m} (X_{(i+m)} - X_{(i-m)}) \right\}, \tag{3}$$

where $X_{(1)} \leq \cdots \leq X_{(n)}$ are the order statistics and m is a positive integer, $m \leq n/2$ and $X_{(i)} = X_{(1)}$ if $i < 1$, $X_{(i)} = X_{(n)}$ if $i > n$.
The sample entropy was considered in establishing a goodness of fit test statistic for uniformity by Dudewicz and Van der Meulen [14], for exponentiality by Ebrahimi *et al.* [16] and for normality by Arizono and Ohta [8] and Vasicek [37], which introduced the test statistic based on the property of maximum entropy of normal distribution as

$$K_{n,m} = \frac{\exp(H_{n,m})}{\hat{\sigma}} = \frac{n}{2m\hat{\sigma}} \left(\prod_{i=1}^{n} [X_{(i+m)} - X_{(i-m)}] \right)^{1/n}, \tag{4}$$

where

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2}.$$

Some other works related to the goodness of fit test are as follows.
Park [28] provided a test of normality based on the sample entropy of order statistics. Choi [11]improved Vasicek's entropy test. Esteban *et al.* [18] compared four tests of normality using four statistics based on different entropy estimators namely, Vasicek [37], Van Es [36], Correa [12] and Wieczorkowski and Grzegorzewski [41]. Similar to this article Alizadeh [3] compared the four mentioned tests with a test based on a new estimator of entropy using kernel density function. In the other article, Alizadeh and Arghami [4] compared seven different tests of normality namely Kolmogorov-Smirnov, Anderson- Darling [7], Kuiper [23], Jarque-Bera [19], Cramer-von Mises [13], Shapiro-Wilk [33] and Vasicek [37]. Lequesne [25] introduced a test based on entropy to pareto distribution. Moreover, Lequesne [26] proposed a test of Student distribution based on Renyi information. Lee [24] investigated the Vasicek test for a composite hypothesis. Lequesne and Regnault [27] investigated the details of the mathematical justification of Vasicek test and Song [35] test. Alizadeh Noughabi and Balakrishnan [6] introduced a general goodness of fit test based on Phi-divergence. They estimated the test statistic using kernel density estimation and computed the power of the tests for 20 alternatives divided to four groups. Alizadeh and Arghami [5], presented an exponentiality test based on characterizations of the exponential distribution. Karadag and Aktas [21] introduced goodness of fit tests for generalized gamma distribution.

The rest of the article is organized as follows:
In Section 2, eight test statistics based on various divergence measures (special cases of Phi divergence) which are estimated using spacing are introduced for normality. Also in this section invariant and consistency of proposed tests are proved. In Section 3, using Monte Carlo simulation, it is demonstrated that the proposed test has the greatest power among the competitors.

2. Test statistic

Phi divergence measure between two density functions $f(x)$ and $g(x)$ is defined as

$$D_\Phi(f,g) = \int_{-\infty}^{\infty} \Phi\Big(\frac{g(x)}{f(x)}\Big) f(x)dx. \tag{7}$$

where Φ is a convex functions such that $\Phi(1) = 0$ and $\Phi''(1) > 0$.
Some Special cases of Phi divergence are:
1) Kullback-Leibler divergence measure ($\Phi(t) = t\log(t)$)

$$D_{KL}(f,g) = \int_{-\infty}^{\infty} \log\big(\frac{f(x)}{g(x)}\big) f(x)dx;$$

2) Pearson divergence measure ($\Phi(t) = (t-1)^2$)

$$D_P(f,g) = \int_{-\infty}^{\infty} \frac{(f(x)-g(x))^2}{f(x)}dx;$$

3) Hellinger divergence measure ($\Phi(t) = \frac{1}{2}(\sqrt{t}-1)^2$)

$$D_H(f,g) = \int_{-\infty}^{\infty} \frac{1}{2}\big(\sqrt{f(x)} - \sqrt{g(x)}\big)^2 dx;$$

4) Triangular divergence measure ($\Phi(t) = \frac{(1-t)^2}{1+t}$)

$$D_T(f,g) = \int_{-\infty}^{\infty} \frac{(f(x)-g(x))^2}{f(x)+g(x)}dx;$$

5) Lin-Wang divergence measure ($\Phi(t) = t\log\frac{2}{1+t}$)

$$D_{LW}(f,g) = \int_{-\infty}^{\infty} f(x)\log\frac{2f(x)}{f(x)+g(x)}dx;$$

6) Jeffreys divergence measure ($\Phi(t) = (t-1)\log(t)$)

$$D_J(f,g) = \int_{-\infty}^{\infty} \log\big(\frac{f(x)}{g(x)}\big) f(x)dx + \int_{-\infty}^{\infty} \log\big(\frac{g(x)}{f(x)}\big) g(x)dx;$$

7) Total variation divergence measure ($\Phi(t) = |t-1|$)

$$D_{TV}(f,g) = \int_{-\infty}^{\infty} |f(x)-g(x)|dx;$$

8) Balakrishnan-Sanghvi divergence measure [9] ($\Phi(t) = (\frac{t-1}{t+1})^2$)

$$D_{BS}(f,g) = \int_{-\infty}^{\infty} \Big(\frac{f(x)-g(x)}{f(x)+g(x)}\Big)^2 f(x)dx.$$

It is obvious that the above divergence measures are non negative and they equal to zero if and only if $f(x) = g(x)$. So it motivates us to use them as a test statistic for normality. Taking

$$g(x) = f_0(x;\mu,\sigma) = (1/\sqrt{2\pi\sigma^2})\exp\{-(x-\mu)^2/2\sigma^2\},$$

in (7), we obtain

$$D_\Phi(f,g) = \int_{-\infty}^{\infty} f(x)\Phi\Big(\frac{(1/\sqrt{2\pi\sigma^2})\exp\{-(x-\mu)^2/2\sigma^2\}}{f(x)}\Big)dx. \qquad (8)$$

Now, similar to Vasicek's method for estimating the entropy, by using $F(x) = p$, equation (8) is expressed as

$$D_\Phi(f,g) = \int_0^1 \Phi\Big(\frac{(1/\sqrt{2\pi\sigma^2})\exp\{-(F^{-1}(p)-\mu)^2/2\sigma^2\}}{(\frac{dF^{-1}(p)}{dp})^{-1}}\Big)^2 dp. \qquad (9)$$

With replacing F by F_n and using difference operator in place of differential operator, we get an estimator V_Φ of $D_\Phi(f,g)$ as

$$V_\Phi = \frac{1}{n}\sum_{i=1}^{n}\Phi\Big(\frac{(\frac{1}{\sqrt{2\pi\hat{\sigma}^2}})\exp\{\frac{-(X_{(i)}-\hat{\mu})^2}{2\hat{\sigma}^2}\}}{\frac{2m}{n(X_{(i+m)}-X_{(i-m)})}}\Big), \qquad (10)$$

where $X_{(1)} <= ... <= X_{(n)}$ are the order statistics and m is a positive integer, $m <= n/2$ and $X_{(i)} = X_{(1)}$ for $i < 1$ and $X_{(i)} = X_{(1)}$ for $i > n$. Also, $\hat{\mu} = \bar{X}$ and $\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ are the MLEs of μ and σ^2 under the null hypothesis.
It is obvious that V_Φ is invariant with respect to scale and location transformations since

$$V_\Phi(cx+d) = \frac{1}{n}\sum_{i=1}^{n}\Phi\Big(\frac{(\frac{1}{\sqrt{2\pi c\hat{\sigma}^2}})\exp\{\frac{-(cx_{(i)}+d-c\hat{\mu}-d)^2}{2c\hat{\sigma}^2}\}}{\frac{2m}{n(cx_{(i+m)}+d-cx_{(i-m)}-d)}}\Big) = V_\Phi(x).$$

Now, we prove the test based on V_Φ is consistent, that is the power of the test under the alternative hypothesis increaes to one as $n \to \infty$.

Theorem 2.1. *Let F be an unknown continuous distribution on real line and f_0 be the normal distribution with unspecified parameters. Then the test based on V_Φ is consistent.*

Proof of Theorem 2.1. As $n, m \to \infty$ and $m/n \to 0$, we have

$$\begin{aligned}\frac{2m}{n} &= F_n(X_{(i+m)}) - F_n(X_{(i-m)}) \simeq F(X_{(i+m)}) - F(X_{(i-m)})\\ &\simeq \frac{f(X_{(i+m)}) + f(X_{(i-m)})}{2}(X_{(i+m)} - X_{(i-m)}),\end{aligned}$$

where $F_n(a) = \#(X_i \le a)/n = (1/n)\sum I_{(-\infty,X_i]}(a)$ and I is the indicator function. Therefore noting that $\hat{\mu}$ and $\hat{\sigma^2}$ are consistent, we have

$$\begin{aligned}V_\Phi &= \frac{1}{n}\sum_{i=1}^{n}\Phi\Big(\frac{(\frac{1}{\sqrt{2\pi\hat{\sigma}^2}})\exp\{\frac{-(X_{(i)}-\hat{\mu})^2}{2\hat{\sigma}^2}\}}{\frac{2m}{n(X_{(i+m)}-X_{(i-m)})}}\Big)\\ &\simeq \frac{1}{n}\sum_{i=1}^{n}\Phi\Big(\frac{(\frac{1}{\sqrt{2\pi\sigma^2}})\exp\{\frac{-(X_{(i)}-\mu)^2}{2\sigma^2}\}}{\frac{F(X_{(i+m)})-F(X_{(i-m)})}{X_{(i+m)}-X_{(i-m)}}}\Big)\\ &\longrightarrow E\Big\{\Phi\Big(\frac{(\frac{1}{\sqrt{2\pi\sigma^2}})\exp\{\frac{-(X-\mu)^2}{2\sigma^2}\}}{\frac{F(X_{(i+m)})-F(X_{(i-m)})}{X_{(i+m)}-X_{(i-m)}}}\Big)\Big\}\end{aligned}$$

where the last limit holds by the law of large numbers. Also, $X_{(i-m)}$ and $X_{(i+m)}$ belong to an interval in which $f(x)$ is positive and continuous, then there exsists a value $X'_i \in (X_{(i-m)}, X_{(i+m)})$ such that

$$\frac{F(X_{(i+m)}) - F(X_{(i-m)})}{X_{(i+m)} - X_{(i-m)}} = f(X'_i).$$

Therefore $V_\Phi \to D_\Phi(f, f_0)$. So, the test based on V_Φ is consistent. □

3. Simulation study

A simulation study is investigated to analyze the behavior of the proposed test. For this purpose, some special cases of Phi divergence are considered to construct some test statistics that are, the tests based on Kullback-Leibler measure (V_{KL}) (Vasicek test with different m), Pearson measure (V_P), Triangular measure (V_T), Jeffreys measure (V_J), Balakrishnan-Sanghvi measure (V_{BS}), Hellinger measure (V_H), Lin-wang measure (V_{LW}) and Total Variation measure(V_{TV}). Also, the critical values are determined using Monte Carlo simulation with 10,000 replicates at the significance level 0.05 (Table 2).

The values of m in (10) are suggested that attained maximum power. These values for each test statistic are given in Table 1.

Table 1. The values of m

n	V_J	V_{BS}	V_{LW}	V_{TV}	V_P	V_H	V_{KL}	V_T
5 − 9	2	2	2	1	2	2	2	1
10 − 19	3	3	1	2	2	4	2	2
20 − 29	5	5	2	5	3	7	7	5
30 − 49	8	7	2	8	5	11	11	8
50 − 79	14	11	2	15	8	18	13	15
80 − 100	14	20	2	15	16	18	13	15

Table 2. Critical values of the proposed statistics at the significance level 0.05

n	V_J	V_{BS}	V_{LW}	V_{TV}	V_P	V_H	V_{KL}	V_T
5	0.791	0.260	0.266	0.784	0.396	0.184	0.899	0.457
6	0.788	0.251	0.239	0.778	0.421	0.183	0.840	0.447
7	0.796	0.247	0.218	0.778	0.452	0.184	0.792	0.436
8	0.785	0.240	0.200	0.774	0.475	0.182	0.751	0.428
9	0.758	0.227	0.186	0.770	0.493	0.178	0.697	0.419
10	0.634	0.196	0.181	0.624	0.515	0.138	0.652	0.300
11	0.614	0.187	0.172	0.618	0.517	0.135	0.612	0.289
12	0.598	0.182	0.166	0.616	0.522	0.131	0.588	0.283
13	0.596	0.176	0.161	0.611	0.537	0.131	0.564	0.283
14	0.564	0.167	0.154	0.604	0.526	0.125	0.526	0.273
15	0.557	0.162	0.148	0.602	0.537	0.124	0.509	0.270
20	0.460	0.135	0.113	0.475	0.413	0.108	0.468	0.192
25	0.422	0.118	0.098	0.463	0.409	0.099	0.389	0.176
30	0.395	0.109	0.088	0.426	0.317	0.097	0.411	0.161
40	0.344	0.091	0.077	0.405	0.300	0.084	0.314	0.142
50	0.350	0.089	0.069	0.392	0.258	0.088	0.275	0.144
60	0.316	0.079	0.063	0.375	0.250	0.079	0.228	0.131
70	0.296	0.072	0.060	0.363	0.237	0.073	0.195	0.123
80	0.274	0.083	0.057	0.349	0.230	0.068	0.173	0.115
90	0.259	0.076	0.054	0.341	0.224	0.065	0.154	0.108
100	0.246	0.071	0.052	0.334	0.218	0.061	0.139	0.103

To comparing the proposed test with competitors, 20 alternatives are considered that can be divided into four groups, based on the support and shape of their densities. These alternatives were used by Alizadeh and Balakrishnan [6] and Esteban *et al.* [18] which are, as follows:

Group I: Support $(-\infty, \infty)$, symmetric.

1) Students t with 1 degree of freedom (i.e. the standard Cauchy),
2) Students t with 3 degrees of freedom,
3) Standard logistic,
4) Standard double exponential (Laplace);

Group II: Support $(-\infty, \infty)$, asymmetric.

5) Gumbel with parameters $\alpha = 0$ (location) and $\beta = 1$ (scale),
6) Gumbel with parameters $\alpha = 0$ (location) and $\beta = 2$ (scale),
7) Gumbel with parameters $\alpha = 0$ (location) and $\beta = 1/2$ (scale);

Group III: Support $(0, -\infty)$.

8) Exponential with mean 1,
9) Gamma with parameters $\beta = 1$ (scale) and $\alpha = 2$ (shape),
10) Gamma with parameters $\beta = 1$ (scale) and $\alpha = 1/2$ (shape),
11) Lognormal with parameters $\mu = 0$ (scale) and $\sigma = 1$ (shape),
12) Lognormal with parameters $\mu = 0$ (scale) and $\sigma = 2$ (shape),
13) Lognormal with parameters $\mu = 0$ =0 (scale) and $\sigma = 1/2$ (shape),
14) Weibull with parameters $\beta = 1$ (scale) and $\alpha = 1/2$ (shape),
15) Weibull with parameters $\beta = 1$ (scale) and $\alpha = 2$ (shape);

Group IV: Support (0,1).

16) Uniform,
17) Beta (2,2),
18) Beta (0.5,0.5),
19) Beta (3,1.5),
20) Beta (2,1).

We regard the entropic based tests using Kernel method and some classical tests such as Cramer Von Mises, Kolmogorov, Kuiper and Anderson-Darling as the competitors. Alizadeh and Balakrishnan [6] have concluded that the most powerful tests among all mentioned in the last line, are five following tests (tests 1 to 5). Moreover, in regards to Romao *et al.* [31], Shapiro-Wilk test was concluded as one of the most powerful tests among 30 normality tests, we consider, Shapiro-Wilk test as the competitor (test 6th), as well.

1) K_{KL}: Test based on Kullback-Leibler measure using kernel density estimation,
2) K_H: Test based on Hellinger measure using kernel density estimation,
3) K_J: Test based on pearson measure using kernel density estimation,
4) K_P: Test based on Jeffreys measure using kernel density estimation,
5) A: Anderson-Darling test,
6) W: Shapiro-Wilk test.

So, we compare the treatment of the proposed test with these five powerful tests as the competitors.

Under each alternative, 10,000 samples of size 10, 20, 30 and 50 were generated. For each sample, the test statistics A, K_{KL}, K_P, K_H and K_J were evaluated and the power of the corresponding test was estimated by the frequency of the event "the statistic is in the critical region".

Tables 3 to 7 show the estimated power of the proposed tests and those of the competing tests, at the significance level $\alpha = 0.05$ based on 10,000 iterations with sample sizes 10, 20, 30 and 50.

Table 3. The power comparisons at $\alpha = 0.05$ under alternatives from group I

Alternative	n	K_{KL}	K_J	K_P	K_H	A	V_J	V_{BS}	V_{LW}	V_{TV}	V_P	V_H	V_{KL}	V_T	W
t (1)	10	0.623	0.554	0.453	0.553	0.616	0.524	0.484	0.435	0.619	**0.630**	0.503	0.429	0.579	0.599
	20	0.886	0.849	0.729	0.846	0.878	0.868	0.800	0.584	0.873	**0.899**	0.794	0.465	0.858	0.859
	30	0.970	0.951	0.897	0.951	0.967	0.964	0.946	0.746	0.968	**0.975**	0.922	0.425	0.960	0.961
	50	0.997	0.995	0.988	0.995	0.997	0.997	0.996	0.913	0.997	**0.998**	0.994	0.729	0.995	0.996
t (3)	10	**0.212**	0.203	0.141	0.203	0.183	0.165	0.138	0.095	0.192	0.208	0.175	0.093	0.168	0.186
	20	0.378	0.354	0.172	0.347	0.326	0.351	0.223	0.087	0.339	**0.396**	0.292	0.088	0.321	0.332
	30	0.498	0.453	0.208	0.439	0.422	0.491	0.317	0.092	0.472	**0.537**	0.391	0.049	0.438	0.463
	50	0.688	0.643	0.368	0.621	0.607	0.711	0.564	0.118	0.684	**0.738**	0.626	0.058	0.649	0.639
Logistic	10	0.097	0.098	0.076	0.097	0.080	0.082	0.071	0.052	0.086	**0.097**	0.086	0.052	0.078	0.0829
	20	0.135	0.125	0.058	0.123	0.103	0.122	0.071	0.040	0.110	**0.138**	0.107	0.044	0.106	0.114
	30	0.178	0.155	0.055	0.146	0.130	0.174	0.087	0.032	0.160	**0.203**	0.128	0.025	0.140	0.151
	50	0.233	0.201	0.061	0.178	0.159	0.248	0.135	0.026	0.222	**0.276**	0.192	0.015	0.197	0.194
Double exponential	10	**0.192**	0.169	0.112	0.168	0.162	0.127	0.103	0.073	0.171	0.185	0.146	0.066	0.137	0.151
	20	0.313	0.267	0.116	0.259	0.270	0.269	0.144	0.052	0.271	**0.346**	0.193	0.043	0.241	0.267
	30	0.417	0.345	0.141	0.334	0.368	0.392	0.246	0.051	0.390	**0.484**	0.271	0.019	0.342	0.362
	50	0.594	0.513	0.272	0.494	0.542	0.599	0.490	0.060	0.568	**0.676**	0.470	0.011	0.517	0.521

Table 4. The power comparisons at $\alpha = 0.05$ under alternatives from group II

Alternative	n	K_{KL}	K_J	K_P	K_H	A	V_J	V_{BS}	V_{LW}	V_{TV}	V_P	V_H	V_{KL}	V_T	W
Gumbel (0,1)	10	0.162	0.168	0.145	**0.169**	0.143	0.158	0.146	0.101	0.152	0.158	0.167	0.100	0.142	0.153
	20	0.286	0.320	0.243	0.322	0.271	0.305	0.257	0.158	0.298	0.243	**0.335**	0.186	0.306	0.317
	30	0.415	0.457	0.325	0.458	0.398	0.453	0.367	0.215	0.457	0.361	**0.487**	0.185	0.451	0.461
	50	0.620	0.666	0.512	0.665	0.603	0.634	0.550	0.338	0.648	0.491	**0.672**	0.293	0.662	0.671
Gumbel (0,2)	10	0.156	0.159	0.139	0.160	0.141	0.150	0.141	0.102	0.147	0.151	**0.160**	0.105	0.139	0.155
	20	0.289	0.321	0.249	0.324	0.276	0.311	0.264	0.166	0.303	0.248	**0.337**	0.193	0.311	0.311
	30	0.410	0.452	0.335	0.455	0.400	0.445	0.369	0.216	0.455	0.356	**0.479**	0.190	0.449	0.465
	50	0.628	0.678	0.516	0.676	0.607	0.639	0.548	0.336	0.652	0.499	0.676	0.291	0.665	**0.688**
Gumbel (0,1/2)	10	0.161	0.163	0.146	**0.165**	0.145	0.151	0.140	0.099	0.148	0.155	0.163	0.102	0.139	0.153
	20	0.285	0.318	0.240	0.320	0.270	0.307	0.255	0.158	0.300	0.245	**0.332**	0.184	0.306	0.321
	30	0.412	0.450	0.323	0.451	0.393	0.450	0.370	0.214	0.454	0.360	**0.481**	0.183	0.449	0.463
	50	0.620	0.673	0.514	0.670	0.605	0.633	0.551	0.341	0.652	0.497	**0.674**	0.300	0.666	0.669

According to Table 3, in group I, the most amount of powers is related to V_P test for most of the alternatives, especially for large sample sizes. According to Table 4, in group II, the most amount of powers is related to V_H test for most of the alternatives. Also, according to Tables 5 and 6, in group III, the most amount of powers is related to V_{BS} and V_H tests and eventually, according to Table 7 in group IV, the most amount of powers is related to V_{KL} test for most of the alternatives.

Remark 3.1. It should be noted that for a sample size of less than or equal to 10, K_{KL} is also suitable for group I because for 2 of the 4 alternatives, it gained the highest power. This is true about K_H for group II, and V_{LW} for group IV.

Table 5. The power comparisons at $\alpha = 0.05$ under alternatives from group III

Alternative	n	K_{KL}	K_J	K_P	K_H	A	V_J	V_{BS}	V_{LW}	V_{TV}	V_P	V_H	V_{KL}	V_T	W
Exponential	10	0.382	0.379	0.406	0.385	0.412	0.489	**0.494**	0.371	0.429	0.406	0.479	0.420	0.462	0.446
	20	0.722	0.799	0.773	0.809	0.771	0.863	0.882	0.753	0.829	0.658	**0.885**	0.822	0.871	0.829
	30	0.908	0.943	0.915	0.948	0.931	0.973	**0.982**	0.934	0.971	0.874	0.979	0.926	0.979	0.970
	50	0.995	0.998	0.994	0.999	0.997	0.999	**1.000**	0.999	1.000	0.974	**1.000**	0.997	1.000	0.999
Gamma (2)	10	0.230	0.233	0.229	0.235	0.225	0.254	0.250	0.182	0.229	0.226	**0.261**	0.190	0.234	0.236
	20	0.457	0.518	0.470	0.529	0.470	0.551	0.546	0.379	0.531	0.374	**0.595**	0.448	0.565	0.525
	30	0.641	0.716	0.629	0.725	0.656	0.752	0.747	0.556	0.761	0.560	**0.800**	0.536	0.773	0.746
	50	0.895	0.928	0.864	0.932	0.897	0.924	0.944	0.830	0.947	0.739	**0.955**	0.823	0.955	0.953
Gamma (1/2)	10	0.623	0.603	0.635	0.611	0.702	0.811	**0.820**	0.711	0.722	0.693	0.771	0.786	0.788	0.732
	20	0.937	0.968	0.966	0.971	0.970	0.992	**0.995**	0.982	0.985	0.930	0.993	0.988	0.993	0.986
	30	0.994	0.998	0.996	0.999	0.998	1.000	**1.000**	0.999	1.000	0.994	**1.000**	0.999	1.000	1.000
	50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Lognormal (0,1)	10	0.558	0.545	0.558	0.551	0.585	0.646	**0.650**	0.524	0.602	0.577	0.644	0.568	0.620	0.601
	20	0.885	0.920	0.902	0.924	0.905	0.944	0.949	0.866	0.932	0.846	**0.953**	0.914	0.947	0.928
	30	0.978	0.989	0.974	0.990	0.985	0.993	0.994	0.972	0.993	0.968	**0.995**	0.970	0.994	0.992
	50	1.000	1.000	0.999	1.000	1.000	1.000	**1.000**	0.999	1.000	0.997	**1.000**	1.000	1.000	1.000

Remark 3.2. Because of the importance of normal distribution, in this paper goodness of fit test for normality was studied. However, it could be done for any distribution such as exponential, uniform, etc in a similar way.

Table 6. The power comparisons at $\alpha = 0.05$ under alternatives from group III

Alternative	n	K_{KL}	K_J	K_P	K_H	A	V_J	V_{BS}	V_{LW}	V_{TV}	V_P	V_H	V_{KL}	V_T	W
Lognormal (0,2)	10	0.864	0.831	0.825	0.837	0.908	0.946	**0.951**	0.906	0.916	0.901	0.931	0.936	0.939	0.919
	20	0.997	0.998	0.998	0.999	0.999	0.999	1.000	0.999	0.999	0.996	1.000	0.999	1.000	0.999
	30	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Lognormal (0,1/2)	10	0.246	0.255	0.235	0.257	0.234	0.257	0.246	0.170	0.241	0.243	**0.271**	0.177	0.237	0.248
	20	0.463	0.520	0.440	0.526	0.459	0.518	0.481	0.318	0.505	0.400	**0.555**	0.375	0.520	0.512
	30	0.645	0.700	0.585	0.706	0.643	0.711	0.665	0.462	0.721	0.576	**0.747**	0.432	0.719	0.729
	50	0.876	0.908	0.821	0.911	0.871	0.890	0.870	0.706	0.915	0.760	**0.918**	0.688	0.917	0.923
Weibull (1/2)	10	0.812	0.781	0.786	0.786	0.881	0.938	**0.945**	0.893	0.892	0.871	0.914	0.934	0.930	0.895
	20	0.993	0.997	0.996	0.997	0.998	1.000	**1.000**	0.998	0.999	0.993	**1.000**	0.999	1.000	0.998
	30	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Weibull (2)	10	0.082	0.079	0.081	0.080	0.082	0.084	0.084	0.079	0.081	0.079	**0.085**	0.075	0.081	0.081
	20	0.127	0.156	0.150	0.161	0.130	0.159	0.159	0.132	0.150	0.095	**0.177**	0.145	0.167	0.153
	30	0.175	0.218	0.196	0.225	0.181	0.224	0.234	0.170	0.227	0.133	**0.265**	0.176	0.239	0.238
	50	0.293	0.369	0.328	0.380	0.304	0.332	0.378	0.285	0.377	0.158	**0.424**	0.295	0.420	0.413

Table 7. The power comparisons at $\alpha = 0.05$ under alternatives from group IV

Alternative	n	K_{KL}	K_J	K_P	K_H	A	V_J	V_{BS}	V_{LW}	V_{TV}	V_P	V_H	V_{KL}	V_T	W
Uniform	10	0.046	0.024	0.072	0.025	0.077	0.085	0.110	**0.173**	0.078	0.047	0.050	0.168	0.114	0.083
	20	0.099	0.115	0.333	0.138	0.171	0.214	0.356	0.445	0.174	0.065	0.133	**0.471**	0.273	0.198
	30	0.190	0.352	0.562	0.395	0.292	0.307	0.604	0.669	0.367	0.109	0.274	**0.764**	0.469	0.382
	50	0.471	0.746	0.850	0.773	0.570	0.511	0.904	0.925	0.643	0.155	0.609	**0.975**	0.806	0.749
Beta (2,2)	10	0.032	0.022	0.041	0.023	0.046	0.042	0.050	**0.080**	0.040	0.027	0.031	0.073	0.050	0.042
	20	0.030	0.034	0.118	0.041	0.056	0.050	0.106	0.150	0.047	0.017	0.042	**0.153**	0.070	0.056
	30	0.046	0.075	0.178	0.091	0.081	0.044	0.150	0.218	0.066	0.014	0.051	**0.250**	0.089	0.076
	50	0.084	0.189	0.329	0.215	0.133	0.031	0.241	0.362	0.076	0.006	0.067	**0.464**	0.133	0.161
Beta (1/2,1/2)	10	0.134	0.045	0.157	0.047	0.259	0.322	0.378	0.487	0.285	0.196	0.148	**0.507**	0.407	0.294
	20	0.429	0.517	0.785	0.565	0.623	0.768	0.874	0.910	0.705	0.458	0.663	**0.931**	0.820	0.737
	30	0.717	0.897	0.957	0.913	0.857	0.935	0.986	0.989	0.945	0.718	0.924	**0.996**	0.972	0.937
	50	0.980	0.998	0.999	0.999	0.992	0.998	1.000	1.000	0.999	0.924	1.000	**1.000**	1.000	0.999
Beta (3,1/2)	10	0.471	0.459	0.514	0.466	0.557	0.688	**0.701**	0.584	0.587	0.550	0.637	0.680	0.665	0.597
	20	0.848	0.913	0.918	0.920	0.923	0.976	**0.984**	0.952	0.952	0.827	0.977	0.966	0.978	0.953
	30	0.977	0.989	0.982	0.990	0.990	0.998	**1.000**	0.997	0.997	0.972	0.998	0.999	0.999	0.997
	50	1.000	1.000	1.000	1.000	1.000	1.000	**1.000**	1.000	1.000	0.999	1.000	**1.000**	1.000	1.000
Beta (2,1)	10	0.096	0.085	0.129	0.087	0.116	0.153	0.169	0.166	0.124	0.100	0.127	**0.177**	0.151	0.133
	20	0.179	0.238	0.363	0.259	0.257	0.344	0.438	0.412	0.305	0.128	0.360	**0.458**	0.385	0.307
	30	0.305	0.439	0.552	0.472	0.417	0.516	0.681	0.626	0.536	0.204	0.571	**0.662**	0.604	0.508
	50	0.617	0.774	0.828	0.797	0.723	0.754	0.925	0.905	0.827	0.280	0.835	**0.931**	0.885	0.843

4. Illustration with a real data set

In this section, two real data sets are used to illustrate the application of the proposed tests.

Example 4.1. We choose the leghorn chicks data set given in Bliss [10]. It contains the weights in grams of n = 20 twenty-one-day-old leghorn chicks. The data (Table 9) was also analyzed in Wang *et al.* [39], which observed that the normal distribution fits these data at level 5%. Figure 1 shows the histogram of the kilos of chicks.

Table 8. The kilos of the chicks data

156	162	168	182	186	190	190	196	202	210
214	220	226	230	230	236	236	242	246	270

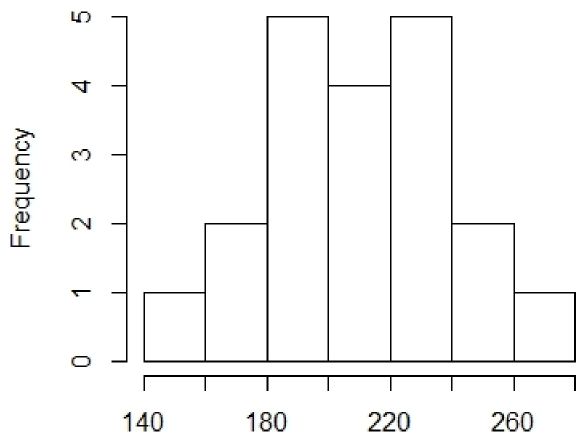

Figure 1. Histogram for data set in Example 1

Now, using the proposed tests, we test whether the data come from a normal distribution. For the normal assumption, the values of the proposed test statistics (the critical values at level 5%) are:

$V_J = 0.282(0.460)$, $V_{BS} = 0.092(0.135)$, $V_{LW} = 0.563(0.113)$, $V_{TV} = 0.315(0.475)$,
$V_P = 0.1495(0.413)$, $V_H = 0.0764(0.108)$, $V_{KL} = 0.339(0.468)$, $V_T = 0.120(0.192)$.
It is clear that the values of the proposed statistics are all smaller than the corresponding critical values and so the normal model is not rejected at the significance level of 0.05 and this agrees with the previous conclusions. In addition, the Shapiro-Wilk test does not reject the normality assumption for these data, as its $p-value$ is 0.866.

Example 4.2. The second example is related to 100 breaking strengths presented by Duncan [15]. Table 9 shows the data set and Figure 2 indicates the histogram of the breaking strengths of yarn.
Recently, these data were reviewed and analyzed by Alizadeh [2], who concluded that a normal distribution does not fit well for this data set. We also examine here the appropriateness of the normal distribution for this data. The values of the proposed tests are as follows:
$V_J = 0.581(0.246)$, $V_{BS} = 0.116(0.071)$, $V_{LW} = 0.03(0.052)$, $V_{TV} = 0.59(0.334)$,
$V_P = 0.63(0.218)$, $V_H = 0.132(0.061)$, $V_{KL} = 0.163(0.139)$, $V_T = 0.239(0.103)$.

Table 9. The breaking strengths of yarn data

66	117	132	111	107	85	89	79	91	97	138	103	111	86	78
96	93	101	102	110	95	96	88	122	115	92	137	91	84	96
97	100	105	104	137	80	104	104	106	84	92	86	104	132	94
99	102	101	104	107	99	85	95	89	102	100	98	97	104	114
111	98	99	102	91	95	111	104	98	98	102	109	88	91	103
94	105	103	96	100	101	98	97	97	101	102	98	94	100	98
99	92	102	87	99	62	92	100	96	98					

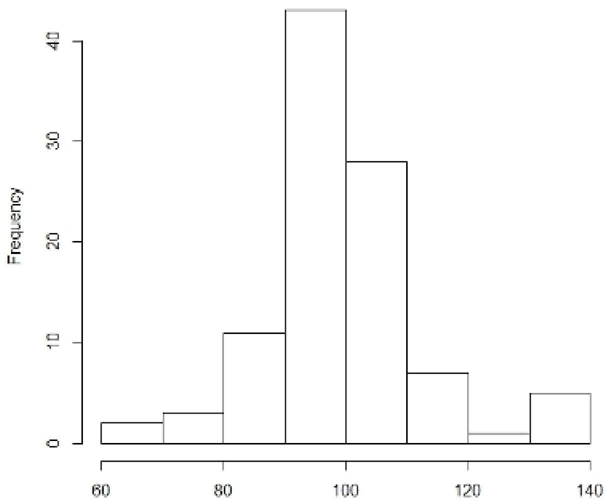

Figure 2. Histogram for data set in Example 2

It can be seen that the values of the test statistics for all tests (with the exception of V_{LW}) are higher than their critical values, which means that the normal distribution for this data set is rejected, consistent with the result of Puig and Stephenes [30]. Moreover, the $p - value$ of the Shapiro-Wilk test for this data set is 0.000. So, Shapiro-Wilk test rejects the assumption of normality as well.

5. Conclusions

We introduced a goodness of fit test for normality based on Phi divergence using spacings. The test statistic was estimated using spacing and the consistency of the test was proved. Then with replacing some special cases of Phi divergence (like Kullback-Leibler measure, Pearson measure, Triangular measure, Jeffreys measure, Hellinger measure, Total variation measure, Lin-Wang measure and Balakrishnan-Sanghvi measure), the efficiency of each test statistic was analyzed by Monte Carlo simulation against some competitors based on Phi divergence using kernel density function and also some classical competitors. It was shown, for a symmetric alternative with support $(-\infty, \infty)$, the most powers were related to V_P. For a asymmetric alternative with support $(-\infty, \infty)$, the most powers were related to V_H. Also, for an alternative with support $(0, \infty)$, the most powerful tests are V_{BS} and V_H, and at last for an alternative with support (0,1), the most powerful test is V_{KL} test.

Acknowledgment. The authors would like to thank the editor and anonymous reviewers for their constructive comments which led to the improvement of the paper.

References

[1] I.A. Ahmad and P.E. Lin, *A nonparametric estimation of the entropy of the absolutely continuous distribution*, IEEE Trans. Inform. Theor. **22**, 327-375, 1976.

[2] H. Alizadeh Noughabi, *A new estimator of Kullback–Leibler information and its application in goodness of fit tests*, J. Stat. Comput. Simul. **89** (10), 1914-1934, 2019.

[3] H. Alizadeh Noughabi, *A new estimator of entropy and its application in testing normality*, J. Stat. Comput. Simul. **86**, 1151-1162, 2010.

[4] H. Alizadeh Noughabi and N.R. Arghami, *Monte Carlo comparison of seven normality tests*, J. Stat. Comput. Simul. **8**, 965-972, 2011a.

[5] H. Alizadeh Noughabi and N.R. Arghami, *Testing exponentiality based on characterizations of the exponential distribution*, J. Stat. Comput. Simul. **81**, 1641-1651, 2011b.

[6] H. Alizadeh Noughabi and N. Balakrishnan, *Tests of goodness of fit based on Phidivergence*, J. Appl. Stat. **43** (3), 412-429, 2016.

[7] T.W. Anderson and D.A. Darling, *A test of goodness of fit*, J. Amer. Statist. Assoc. **49**, 765-769, 1954.

[8] I. Arizono and H. Ohta, *A test for normality based on Kullback Leibler information*, Amer. Statist. **43**, 20-23, 1989.

[9] V. Balakrishnan and L.D. Sanghvi, *Distance between populations on the basis of attribute*, Biometrics **24**, 859-865, 1968.

[10] C.I. Bliss, *Statistics in Biology: Statistical methods for research in the natural sciences*, McGrawHill Book Company, New York, 1967.

[11] B. Choi, *Improvement of goodness of fit test for normal distribution*, J. Stat. Comput. Simul. **78**, 781-788, 2008.

[12] J.C. Corea, *A new estimator of entropy*, Comm. Statist. Theory Methods **24**, 2439-2449, 1995.

[13] H. Cramer, *On the composition of elementary errors*, Scand. Actuar. J. **1**, 1374, 1928.

[14] E.S. Dudewicz and E.C. Van der Meulen, *Entropy-based tests of uniformity*, J. Amer. Statist. Assoc. **76**, 967-974, 1981.

[15] A.J. Duncan, *Quality Control and Industrial statistics*, Homewood (IL), Irwin, 1974.

[16] N. Ebrahimi, M. Habibullah and E.S. Soofi, *Testing exponentiality based on Kullback-Leibler information*, J. R. Stat. Soc. Ser. B. Stat. Methodol. **54**, 739-748, 1992.

[17] N. Ebrahimi, K. Pflughoeft and E.S. Soofi, *Two measures of sample entropy*, Statist. Probab. Lett. **20**, 225-234, 1994.

[18] M.D. Esteban, M.E. Castellanos, D. Morales and I. Vajda, *Monte Carlo comparison of four normality tests using different entropy estimates*, Comm. Statist. Simulation Comput. **30**, 761-785, 2001.

[19] C.M. Jarque and A.K. Bera, *A test normality of observations and regression residuals*, Int. Stat. Rev. **55**, 163-172, 1987.

[20] H. Joe, *Estimation of entropy and other functionals of a multivariate density*, Ann. Inst. Statist. Math. **41**, 683-697, 1989.

[21] O. Karadag and S. Aktas, *Goodness of fit tests for generalized gamma distribution*, International Conference of Numerical Analysis and Applied Mathematics 2015, AIP Conference Proceedings, **1738**, 2016.

[22] A.N. Kolmogorov, *Sulla determinazione empiricadi une legga di distribuzione*, Giornale dell Istituto Italiano degli Attuari **4**, 83-91, 1993.

[23] N.H. Kuiper, *Tests concerning random points on a circle*, Proc. K. Ned. Akad. Wet. **63**, 38-47, 1960.

[24] S. Lee Bull, *Entropy-based goodness of fit test for a composite hypothesis*, Bull. Korean Math. Soc. **53**(2), 351-363, 2016.

[25] J. Lequesne, *Entropy-based goodness-of-fit test: Application to the Pareto distribution*, AIP Conf. Proc. of the International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, **1553**, 155-62, 2013.

[26] J. Lequesne, *A goodness-of-fit test of student distributions based on Renyi entropy*, AIP Conf. Proc. of the International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, **1641**, 487-94, 2015.

[27] J. Lequesne and P. Regnault, *Goodness-of-fit tests based on entropy: R package KL-goftest*, Workinprogress, 2017.

[28] S. Park, *A goodness-of-fit test for normality based on the sample entropy of order statistics*, Statist. Probab. Lett. **44** (4), 359-363, 1999.

[29] K. Pearson, *On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling*, Philos. Mag. Lett. **50**, 157-175, 1900.

[30] P. Puig and M.A.Stephens, *Tests of fit for the Laplace distribution with applications*, Technometrics **4**, 417424, 2000.

[31] X. Romao, R. Delgado and A. Costa, *An empirical power comparison of univariate goodness-of-fit tests for normality*, J. Stat. Comput. Simul. **80** (5),545-591, 2010.

[32] C.E. Shannon, *Mathematical theory of communications*, Bell Syst. tech. **27**, 379-423, 623-656, 1948.

[33] S.S. Shapiro and M.B. Wilk, *An analysis of variance test for normality (Complete Sample)*, Biometrika **52**, 591-611,1965.

[34] N. Smirnov, *Table for estimating the goodness of fit of empirical distributions*, Ann Math Stat. **19** (2), 279281, 1948.

[35] K.S. Song, *Goodness-of-fit tests based on KullbackLeibler discrimination information*, IEEE Trans. Inf. Theory **48** (5),110317, 2002.

[36] B. Van Es, *Estimating functional related to a density by a lass of statistic based on spacings*, Scand. J. Stat. **19**, 61-72, 1992.

[37] O. Vasicek, *A Test for normality based on sample entropy*, J. R. Stat. Soc. Ser. B. Stat. Methodol. **38**, 730-737, 1976.

[38] R.E. Von Mises, *Wahrscheinlichkeit, Statistik und Wahrheit*, Julius Springer, 1928.

[39] X. Wang, Y. Liu and B. Han, *Goodness-of-fit tests based on Bernstein distribution estimator*, J. Nonparametr. Stat., 2018.

[40] G.S. Watson, *Goodness of fit tests on a circle*, Biometrika **48**, 109-114, 1961.

[41] P. Wieczorkowski and P. Grzegorzewsky, *Entropy estimators improvements and comparisons*, Comm. Statist. Simulation Comput. **28**, 541-567, 1999.

[42] F. Yousefzadeh and N.R. Arghami, *Testing exponentiality based on type II censored data and a New cdf estimator*, Comm. Statist. Simulation Comput. **37**, 1479-1499, 2008.