Chilean Journal of Statistics
Volume 13, Number 2 (December 2022), Pages 155–163
DOI: 10.32372/chjs.13-02-02

Nonparametric statistics
Research Paper

Nonparametric log kernel estimator of extropy function

Muhammed Rasheed Irshad[1,*] and Radhakumari Maya[2]

[1]Department of Statistics, Cochin University of Science and Technology, Kerala, India,
[2]Department of Statistics, Govt. College for Women, University of Kerala, Kerala, India

(Received: 26 January 2021 · Accepted in final form: 21 December 2021)

Abstract

About seventy years ago, Shannon (1948) introduced the concept of entropy, a measure of uncertainty associated with the random variable X. Lad et al. (2015) identified that Shannon's entropy function has a complementary dual function, named as extropy. For developing inferential aspects of this measure, in this work, we propose a log kernel estimator for the new measure of uncertainty, extropy. Asymptotic properties of the estimator is proved under suitable regularity conditions. A Monte Carlo simulation study is explored to illustrate the performance of the proposed estimator. The methods are illustrated using real data sets.

Keywords: Data analysis · Entropy · Nonparametric kernel estimator.

Mathematics Subject Classification: 94A17 · 62G07.

1. Introduction

Shannon (1948), made his remarkable contribution in statistics by introducing the concept of entropy, a measure of disorder in a probability distribution. Starting from the pioneering work of Shannon, different researchers have shown advantages of entropy in different fields. Apart from statistics and thermodynamics, application of Shannon entropy varies over diverse fields such as information theory, economics, finance, psychology, wavelet analysis, image recognition, computer sciences, fuzzy sets and so on. Inspite of its well known applications, Shannon entropy possesses some demerits that have been suitably modified by different researchers. Now one may use modified entropies which serve the same purpose as that of using Shannon entropy. Because the modified entropies are improvement over Shannon entropy in some sense, it is expected that the tests developed (or distribution estimated) based on the modified entropies will be better, which may be in terms of power of the test or anything alike. Thus the literature on entropy has been developing for the last seven decades. An excellent review of various variants on entropy and its inferential aspects can be found in the book by Cover and Thomas (2006).

Surprisingly, after seventy years, Lad et al. (2015) discovered that the entropy measure of a probability distribution has a dual measure, a complementary companion named as

*Corresponding author. Email: irshadm24@gmail.com

ISSN: 0718-7912 (print)/ISSN 0718-7920 (online)

http://www.soche.cl/chjs

extropy. The authors motivation for finding such a measure came from their interest in the use of proper scoring rules for assessing the quality of alternative probability distributions asserted as forecasts of observable quantities of interest. Also they observed that the entropy and the extropy of a binary distribution are identical and as in the case of entropy, the maximum extropy distribution is the uniform distribution. As companions these two measures associate as do the positive and negative images of a photographic film and they contribute together to characterizing the information in a distribution in much the same way. Following the work of Lad et al. (2015), the study of extropy has gained momentum in theoretical perspective as well as in terms of its applications. Qiu (2017) developed some interesting properties of extropy, including some characterization results using order statistics and record values. Qiu and Jia (2018) proposed extropy estimators with applications in testing uniformity. Qiu et al. (2018) used extropy to compare the uncertainties of two random variables and from applied perspective, Becerra et al. (2018) used extropy in the methods of automatic speech recognition. For various variants of extropy one can also refer Lad et al. (2018). While considering the inferential aspects of extropy, there is a necessity to develop some reasonably good estimators for extropy. Hence in this work, our aim is to develop a nonparametric estimator for extropy function using kernel type estimation. For the survey of nonparametric estimation using kernel type estimation, one may refer Azevedo and Oliveira (2011), Maya and Irshad (2019) and Maya and Irshad (2021).

The remaining part of the paper is organized as follows. In Section 2, we provide a nonparametric log kernel estimator for extropy function and studied its asymptotic properties. The empirical illustration of the proposed estimator is given in Section 3. The performance of the proposed estimator is validated through a simulation study and are discussed here. The conclusion is given in Section 4.

2. Log kernel estimator of extropy function

In the following subsection, we present the definition of entropy and extropy functions.

2.1 Entropy and extropy

Let X be a non-negative random variable admitting an absolutely continuous cumulative distribution function (CDF) $F_X(x)$ and with probability density function (PDF) $f_X(x)$. Then the Shannon entropy associated with X is defined as

$$H(X) = -\int_0^\infty f_X(x)\log f_X(x)\mathrm{d}x.$$

For a non-negative absolutely continuous random variable X with PDF $f_X(x)$, extropy is defined as

$$J(X) = -\frac{1}{2}\int_0^\infty f_X^2(x)\mathrm{d}x.$$

In the next subsection, we look into some basics about the log-kernel density estimator (log-KDE).

2.2 Log kernel density estimator

Let X be a random variable having an unknown PDF $f_X(x)$. Assume that, A_1: Support of X is on R and A_2: $f_X(x)$ is continuously differentiable. Suppose $\{X_i; \quad 1 \leq i \leq n\}$ be a sequence of independently and identically distributed (iid) random variables. Then under the assumptions A_1 and A_2, the most common kernel density estimator (KDE) of $f_X(x)$ is given by (Parzen (1962) and Rosenblatt (1956)),

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right),$$

where $K(x)$ is the kernel function satisfying the conditions, B_1: $\int_{\mathrm{R}} K(x)\mathrm{d}x = 1$, B_2: $\int_{\mathrm{R}} xK(x)\mathrm{d}x = 0$, B_3: $\int_{\mathrm{R}} x^2K(x)\mathrm{d}x = 1$ and B_4: $\int_{\mathrm{R}} K^2(x)\mathrm{d}x < \infty$ and $h = h_n$ is a bandwidth sequence of positive numbers such that C_1: $h_n \to 0$ and C_2: $nh_n \to \infty$ as $n \to \infty$. KDE are inevitable tools for nonparametric estimation of PDF for real valued data sets. When we deal with positive data, these usual KDEs do not provide bonafide PDFs. A log-transformation methodology can be applied to produce a nonparametric estimator that is appropriate and yields proper PDFs over positive supports, see Charpentier and Flachaire (2015). Nguyen et al. (2018) called the KDE obtained using this transformation as log-KDE. Nguyen et al. (2018) provided biases, variances and mean-squared errors (MSEs), mean integrated squared error (MISE) and asymptotic MISE results of log-KDE and demonstrate the log-KDEs methodology via R package, logKDE.

The procedure developed for obtaining log-KDE by Nguyen et al. (2018) is explained in the following steps. Here the assumption A_1 is replaced by A_1^*, where A_1^*: Support of X is on $(0, \infty)$. Let $Y = \log X$, $Y_i = \log X_i, i = 1, 2, \cdots, n$ and $f_Y(y)$ be the PDF of Y. As a result X is supported on $(0, \infty)$, then the support of $f_Y(y)$ satisfies A_1.

A log-KDE is given by (Nguyen et al. (2018))

$$\hat{f}_{\log}(x) = x^{-1}\hat{f}_Y\ \log(x) = \frac{1}{nh}\sum_{i=1}^{n} x^{-1}K\left(\frac{\log\, x - \log\, X_i}{h}\right) = \frac{1}{n}\sum_{i=1}^{n} L(x; X_i, h), \tag{2.1}$$

where $L(x; z, h) = (xh)^{-1}K\left(\log[(x/z)^{1/h}]\right)$ is the log-kernel function with bandwidth $h > 0$, at location parameter z. For any $z, h \in (0, \infty)$, $L(x; z, h)$ satisfies the conditions D_1: $L(x; z, h) \geq 0$ for all $x \in (0, \infty)$ and D_2: $\int_0^\infty L(x; z, h)\mathrm{d}x = 1$ and by using the property D_2, we get $\int_0^\infty f_X(x)\mathrm{d}x = 1$, thus making Equation (2.1) a bonafide PDF on $(0, \infty)$. Under the assumptions A_1 and A_2 regarding $f_Y(y)$, A_1^* and A_2 regarding $f_X(x)$ and B_1 - B_4 regarding $K(y)$, for $y \in R$, we have the bias and variance of $\hat{f}(y)$ as (Nguyen et al. (2018))

$$\mathrm{Bias}[\hat{f}(y)] = \mathrm{E}[\hat{f}(y)] - f_Y(y) = \frac{1}{2}h^2 f_Y^{(2)}(y) + O(h^2),$$
$$\mathrm{Var}[\hat{f}(y)] = \frac{1}{nh}f_Y(y)\int_{\mathrm{R}} K^2(z)\mathrm{d}z + O\Big((nh)^{-1}\Big),$$

where $a_n = o(b_n)$ as $n \to \infty$ if and only if $\lim_{n\to\infty}|a_n/b_n| = 0$. Therefore, for any $x \in (0, \infty)$, the bias and variance of $\hat{f}_{\log}(x)$ is given by

$$\mathrm{Bias}[\hat{f}_{\log}(x)] = \mathrm{E}[\hat{f}_{\log}(x)] - f_X(x) = \frac{h^2}{2}\{f_X(x) + 3xf_X^{(1)}(x) + x^2 f_X^{(2)}(x)\} + O(h^2),$$
$$\mathrm{Var}[\hat{f}_{\log}(x)] = \frac{1}{nhx}f_X(x)\int_{\mathrm{R}} K^2(z)\mathrm{d}z + O\Big((nh)^{-1}\Big).$$

2.3 Log kernel estimator for extropy function and some asymptotic results

In this subsection, we propose a nonparametric log-kernel estimator (log-KE) for extropy function.

As in the line of proposing Equation (2.1), here we propose a nonparametric log-KE for extropy function and is given by

$$\hat{J}_n(X) = -\frac{1}{2}\int_0^\infty \hat{f}_{\log}^2(x)\mathrm{d}x. \tag{2.2}$$

In the following theorem, we study the consistency of our estimator $\hat{J}_n(X)$.

Theorem 2.1 Suppose $\hat{J}_n(X)$ is a log-KE of $J(X)$ as defined in Equation (2.2). Then, it can be concluded that $\hat{J}_n(X)$ is a consistent estimator of $J(X)$. That is, as $n \to \infty$

$$\hat{J}_n(X) = \frac{-1}{2}\left\{\int_0^\infty \hat{f}_{\log}^2(x)\mathrm{d}x\right\} \xrightarrow{p} \frac{-1}{2}\left\{\int_0^\infty f_X^2(x)\mathrm{d}x\right\} = J(X).$$

Proof By using Taylor's series expansion, the expressions for the bias and the variance of $\int_0^\infty \hat{f}_{\log}^2(x)\mathrm{d}x$ are given by

$$\text{Bias}\left(\int_0^\infty \hat{f}_{\log}^2(x)\mathrm{d}x\right) = h^2\int_0^\infty f_X(x)\left\{f_X(x) + 3xf_X^{(1)}(x) + x^2 f_X^{(2)}(x)\right\}\mathrm{d}x + O(h^2), \tag{2.3}$$

$$\text{Var}\left(\int_0^\infty \hat{f}_{\log}^2(x)\mathrm{d}x\right) = \frac{4}{nh}C_k\int_0^\infty \frac{f_X^3(x)}{x}\mathrm{d}x + O\Big((nh)^{-1}\Big), \tag{2.4}$$

where $C_k = \int_{\mathrm{R}} K^2(z)\mathrm{d}z$.

The corresponding MSE is given by

$$\begin{aligned}\text{MSE}\left(\int_0^\infty \hat{f}_{\log}^2(x)\mathrm{d}x\right) &= \left(h^2\int_0^\infty f_X(x)\left\{f_X(x) + 3xf_X^{(1)}(x) + x^2 f_X^{(2)}(x)\right\}\mathrm{d}x\right)^2 \\ &\quad + \frac{4}{nh}C_k\int_0^\infty \frac{f_X^3(x)}{x}\mathrm{d}x + O(h^4) + O\Big((nh)^{-1}\Big).\end{aligned} \tag{2.5}$$

From Equation (2.5), as $n \to \infty$, $\text{MSE}\left(\int_0^\infty \hat{f}_{\log}^2(x)\mathrm{d}x\right) \to 0$. Therefore,

$$\hat{J}_n(X) = \frac{-1}{2}\left\{\int_0^\infty \hat{f}_{\log}^2(x)\mathrm{d}x\right\} \xrightarrow{p} \frac{-1}{2}\left\{\int_0^\infty f_X^2(x)\mathrm{d}x\right\} = J(X).$$

That is, $\hat{J}_n(X)$ is a consistent estimator of $J(X)$. Thus the theorem is proved. $\square$

The expressions for the bias and variance of the estimator $\hat{J}_n(X)$ is given in the following theorem.

THEOREM 2.2 Suppose $\hat{J}_n(X)$ is a log-KE of $J(X)$ as defined in Equation (2.2). Then

$$\text{Bias}\left(\hat{J}_n(X)\right) = -\frac{h^2}{2}\int_0^\infty f_X(x)\left\{f_X(x) + 3xf_X^{(1)}(x) + x^2 f_X^{(2)}(x)\right\}\mathrm{d}x + O(h^2), \tag{2.6}$$

$$\text{Var}\left(\hat{J}_n(X)\right) = -\frac{2}{nh}C_k\int_0^\infty \frac{f_X^3(x)}{x}\mathrm{d}x + O\Big((nh)^{-1}\Big). \tag{2.7}$$

The proof of the theorem is easily obtained using Equation (2.3) and (2.4) and hence omitted. □

In the following theorem, we prove that $\hat{J}_n(X)$ is integratedly uniformly consistent in quadratic mean estimator of $J(X)$.

THEOREM 2.3 Suppose $\hat{J}_n(X)$ is a log-KE of $J(X)$ as defined in Equation (2.2). Then, $\hat{J}_n(X)$ is integratedly uniformly consistent in quadratic mean estimator of $J(X)$.

PROOF Consider the MISE of the estimator $\hat{J}_n(X)$. That is

$$\begin{aligned}\text{MISE}(\hat{J}_n(X)) &= \mathrm{E}\int_0^\infty \left[\hat{J}_n(X) - J(X)\right]^2 \mathrm{d}x = \int_0^\infty \mathrm{E}\left[\hat{J}_n(X) - J(X)\right]^2 \mathrm{d}x \\ &= \int_0^\infty \left[\text{Var}(\hat{J}_n(X)) + [\text{Bias}(\hat{J}_n(X))]^2\right]\mathrm{d}x = \int_0^\infty \text{MSE}(\hat{J}_n(X))\mathrm{d}x.\end{aligned}$$

Using Equation (2.6) and (2.7), we have

$$\begin{aligned}\text{MISE}(\hat{J}_n(X)) = &\int_0^\infty \left(\frac{-h^2}{2}\int_0^\infty f_X(x)\left[f_X(x) + 3xf_X^{(1)}(x) + x^2 f_X^{(2)}(x)\right]\mathrm{d}x\right)^2 \mathrm{d}x \\ &+ \int_0^\infty \left(\frac{-2}{nh}C_k\int_0^\infty \frac{f_X^3(x)}{x}\mathrm{d}x\right)\mathrm{d}x + O\Big((nh)^{-1}\Big) + O(h^4).\end{aligned} \tag{2.8}$$

We have, as $n \to \infty$,

$$\text{MSE}\left(\hat{J}_n(X)\right) \to 0.$$

Therefore, from Equation (2.8), we have

$$\text{MISE}\left(\hat{J}_n(X)\right) \to 0, \text{as } n \to \infty. \tag{2.9}$$

From Equation (2.9), we can say that $\hat{J}_n(X)$ is integratedly uniformly consistent in quadratic mean estimator of $J(X)$ (Wegman (1972)).

Thus the theorem is proved.

□

2.4 Optimal band width selection

In this section, we derive the expression for the bandwidth. The expression for MISE of the log-KE $\hat{J}_n(X)$ is given in Equation (2.8). By ignoring higher order terms, we get asymptotic-MISE (A-MISE). By minimizing A-MISE($\hat{J}_n(X)$) with respect to the parameter h, we get the optimal bandwidth h_0. By the definition,

$$\begin{aligned}\text{A-MISE}(\hat{J}_n(X)) = & \frac{h^4}{4}\int_0^\infty \left(\int_0^\infty f_X(x)\left[f_X(x)+3xf_X^{(1)}(x)+x^2 f_X^{(2)}(x)\right]\mathrm{d}x\right)^2 \mathrm{d}x \\ & +\frac{1}{nh}\int_0^\infty \left(-2C_k\int_0^\infty \frac{f_X^3(x)}{x}\mathrm{d}x\right)\mathrm{d}x.\end{aligned}$$

$\frac{\partial \text{A-MISE}(\hat{J}_n(X))}{\partial h} = 0 \Rightarrow$

$$h^3\int_0^\infty \left(\int_0^\infty f_X(x)\left[f_X(x)+3xf_X^{(1)}(x)+x^2 f_X^{(2)}(x)\right]\mathrm{d}x\right)^2 \mathrm{d}x = \frac{1}{nh^2}\int_0^\infty \left(-2C_k\int_0^\infty \frac{f_X^3(x)}{x}\mathrm{d}x\right)\mathrm{d}x.$$

$\Rightarrow$

$$h^5 = \frac{\frac{1}{n}\int_0^\infty \left(-2C_k\int_0^\infty \frac{f_X^3(x)}{x}\mathrm{d}x\right)\mathrm{d}x}{\int_0^\infty \left(\int_0^\infty f_X(x)\left[f_X(x)+3xf_X^{(1)}(x)+x^2 f_X^{(2)}(x)\right]\mathrm{d}x\right)^2 \mathrm{d}x}.$$

Therefore,

$$\begin{aligned}h_0 &= \left[\frac{\int_0^\infty \left(-2C_k\int_0^\infty \frac{f_X^3(x)}{x}\mathrm{d}x\right)\mathrm{d}x}{\int_0^\infty \left(\int_0^\infty f_X(x)\left[f_X(x)+3xf_X^{(1)}(x)+x^2 f_X^{(2)}(x)\right]\mathrm{d}x\right)^2 \mathrm{d}x}\right]^{\frac{1}{5}} n^{-\frac{1}{5}} \\ &= O\left(n^{-\frac{1}{5}}\right).\end{aligned}$$

3. Empirical illustration and simulation study

In this section, we provide empirical and simulation study in order to evaluate the performance of the proposed estimator.

3.1 Empirical illustration

In this section, we illustrate the estimator of the extropy function using the real data sets reported by Bjerkedal (1960) and it represents the survival times of guinea pigs injected with different doses of tubercle bacilli. Guinea pigs are known to have high susceptibility to human tuberculosis. Even an infection initiated with a few virulent tubercle bacilli will lead to progressive disease and death. We consider the data sets obtained under regimens $1.4, 2.4, 3.4, 4.4$ and 6.4 and there were 79 observations each. Using this data, the estimator $\hat{J}_n(X)$ is calculated using the log-gaussian kernel and is given in Table 1. From the table it is inferred that, the extropy is minimum corresponding to regimen 6.4 and it is maximum corresponding to regimen 1.4. All these calculations were done using the software *Wolfram Mathematica10* with standard computation time.

Table 1. Values of $\hat{J}_n(X)$

Regimen	$\hat{J}_n(X)$
1.4	-1.07303×10^{-38}
2.4	-7.37904×10^{25}
3.4	-1.63298×10^{45}
4.4	-1.73375×10^{55}
6.4	-1.21396×10^{67}

3.2 Simulation study

A Monte Carlo simulation study is carried out to compare the kernel estimator $\hat{J}_n(X)$ in terms of the MSE. We consider the normal distribution with parameters $\mu = 5$ and $\sigma = 2$. The log-gaussian kernel is used as the kernel function for the estimation. The bias and MSE of $\hat{J}_n(X)$ for various sample sizes $10, 20, 50, 100, 150, 200, 250, 300, 350, 400, 450$ and 500 are calculated and are given in Table 2.

Table 2. Bias and MSE of $\hat{J}_n(X)$

	Bias and MSE of $\hat{J}_n(X)$	
Sample size	Bias	MSE
n=10	0.281	0.0386
n=20	0.212	0.0355
n=50	0.186	0.0289
n=100	0.155	0.0192
n=150	0.113	0.0138
n=200	0.098	0.0130
n=250	0.090	0.0126
n=300	0.083	0.0093
n=350	0.081	0.0090
n=400	0.078	0.0072
n=450	0.075	0.0067
n=500	0.069	0.0056

4. Conclusion

In this work, we propose a nonparametric estimator for extropy function using log kernel type estimation. Certain asymptotic properties of the proposed estimator is established. A simulation study is conducted to find the MSE of the estimator and it shows that it decreases with increasing sample size. The performance of the estimator is elucidated using real life data sets.

Author contributions Conceptualization, M.R.I., R.M.; methodology, M.R.I., R.M.; software, M.R.I., R.M.; validation, M.R.I., R.M.; formal analysis, M.R.I., R.M.; investigation, M.R.I., R.M.; data curation, M.R.I., R.M.; writing-original draft preparation, M.R.I., R.M.; writing-review and editing, M.R.I., R.M.; visualization, M.R.I., R.M.; supervision, M.R.I., R.M. All authors have read and agreed the published version of the paper.

Acknowledgements The authors express their gratefulness for the constructive criticism of the referees which helped to improve considerably the revised version of the paper.

Funding The authors received no financial support for the research, authorship, and/or publication of this article.

Conflicts of interest The authors declare no conflict of interest.

References

Azevedo, C. and Oliveira, P.E., 2011. On the kernel estimation of a multivariate distribution function under positive dependence. Chilean Journal of Statistics, 2, 99-113.

Becerra, A., de la Rosa, J.I., González, E., Pedroza, A.D., and Escalante, N.I., 2018. Training deep neural networks with non-uniform frame-level cost function for automatic speech recognition. Multimedia Tools and Applications, 77, 27231-27267.

Bjerkedal, T., 1960. Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. American Journal of Hygiene, 72, 130-148.

Charpentier, A. and Flachaire, E., 2015. Log-transform kernel density estimation of income distribution. L'Actualite economique, Revue d' Analyse Economique, 91, 141-159.

Cover, T.M. and Thomas, J.A., 2006. Elements of information theory, 2nd Edition. John Wiley, New York.

Lad, F., Sanfilippo, G., and Agró, G., 2015. Extropy: Complementary dual of entropy. Statistical Science, 30, 40-58.

Lad, F., Sanfilippo, G. and Agró, G., 2018. The Duality of entropy/extropy, and completion of the Kullback information complex. entropy, 20, 1-11.

Maya, R. and Irshad, M.R., 2019. Kernel estimation of residual extropy function under α-mixing dependence condition. South African Statistical Journal, 53, 65-72.

Maya, R. and Irshad, M.R., 2021. Kernel estimation of Mathai-Haubold entropy and residual Mathai-Haubold entropy functions under α-mixing dependence condition. American Journal of Mathematical and Management Sciences, 1-13.

Nguyen, H.D., Jones, A.T., and Mclachlan, G.J., 2018. Positive data kernel density estimation via the LogKDE Package for R. 16^{th} Australasian Conference, AusDM 2018, Bahrurst, NSW, Australia.

Parzen, E., 1962. On estimation of a probability density function and mode. The Annals of Mathematical Statistics, 33, 1065-1076.

Qiu, G., 2017. The extropy of order statistics and record values. Statistics and Probability Letters, 120, 52-60.

Qiu, G. and Jia, K., 2018. Extropy estimators with applications in testing uniformity. Journal of Nonparametric Statistics, 30, 182-196.

Qiu, G., Wang, L., and Wang, X., 2018. On extropy properties of mixed systems. Probability in the Engineering and Informational Sciences, 33, 471-486.

Rosenblatt, M., 1956. Remarks on some nonparametric estimates of a density function. The Annals of Mathematical Statistics, 27, 832-837.

Shannon, C.E., 1948. A mathematical theory of communication. Bell System Technical Journal, 27, 379-423.

Wegman, E.J., 1972. Nonparametric probability density estimation: I.Asummary of available methods. Technometrics, 14, 533-546.