South African Statistical Journal
Vol. 53, No. 2, 65–72

KERNEL ESTIMATION OF RESIDUAL EXTROPY FUNCTION UNDER α-MIXING DEPENDENCE CONDITION

R. Maya
Department of Statistics, University College, Thiruvananthapuram, India
e-mail: *mayasekharan@gmail.com*

M. R. Irshad
Department of Statistics, Cochin University of Science and Technology, Kochi, India
e-mail: *irshadm24@gmail.com*

As in the context of introducing the concept of residual entropy in the literature, Qiu and Jia (2018b) introduced the concept, residual extropy to measure the residual uncertainty of a random variable. In this work, we propose a nonparametric estimator for the residual extropy, where the observations under consideration are exhibiting α-mixing (strong mixing) dependence condition. Asymptotic properties of the estimator is derived under suitable regular conditions. A Monte Carlo simulation study is carried out to evaluate the performance of the estimator using the mean squared errors.

Key words: α-mixing, Entropy, Extropy, Kernel estimator, Residual extropy.

1. Introduction

The notion of entropy emerged in conceptually distinct contexts. From a statistical point of view, it is a measure of uncertainty associated with a random variable (see Shannon, 1948). Let X be a non-negative random variable possessing an absolutely continuous distribution function $F(x)$ and with probability density function (pdf) $f(x)$. Then the Shannon's entropy associated with X is defined as

$$H(X) = -\int_0^\infty f(x)\log f(x)dx.$$

Similar to the notion of entropy, Martinas (1998) introduced the concept of extropy in physics. It provides a calculable physical measure of the human impact on environment and it formulates the physical limits of economic growth. Following the work of Martinas (1998), Furuichi and Mitroi (2012) and Vontobel (2013) discussed the same extropy, and they found applications of extropy in thermodynamics and astrophysics, etc.

Recently, Lad, Sanfilippo and Agro (2015) defined statistically the term extropy as a potential measure of uncertainty. Really it is an alternative measure of Shannon's entropy. For a random variable X, its extropy is defined as

$$J(X) = -\frac{1}{2}\int_0^\infty f^2(x)dx.$$

[1]Corresponding author.
MSC2010 subject classifications. 62G05, 62G20.

For statistical applications of extropy, one can refer to Gneiting and Raftery (2007). The extropy of order statistics and record values are studied by Qiu (2017). Again Qiu and Jia (2018a) provided some estimators for extropy with applications in testing uniformity.

A serious difficulty involved in the application of Shannon's entropy is that it is not applicable to a system which has survived for some units of time. In this situation, Ebrahimi (1996) proposed the concept of residual entropy. As in the scenario of introducing the concept of residual entropy, Qiu and Jia (2018b) introduced residual extropy to measure the residual uncertainty of a random variable.

For a random variable X, the residual extropy is defined as (see Qiu and Jia, 2018b)

$$J(f;t) = \frac{-1}{2R^2(t)} \int_t^\infty f^2(x)dx,$$

where $R(t) = 1 - F(t)$ is the survival function of X. Also in the same paper, the authors elaborately elucidated its various characterizing properties. Even though the practical utility of the above measure is considered, there is a necessity to develop some inferential aspects. Hence in this work our prime objective is to develop nonparametric estimator for the residual extropy function. Here we consider the observations which are exhibiting some mode of dependence because for practical data it seems more realistic to drop independence and replace it by some mode of dependence. For the details of the inference problems, one may refer Masry (1986), Castellana and Leadbetter (1986) and Castellana (1989). Among several mixing conditions found in the literature, α-mixing (strong mixing) has various practical applications (see Rosenblatt, 1956).

Let $(\Omega, \mathfrak{F}, P)$ be a probability space and $\mathfrak{F}_i^k$ be the σ-algebra of events obtained by the random variables $\{X_j; i \le j \le k\}$. The stationary process $\{X_j\}$ is said to satisfy the α-mixing (strong mixing) condition if

$$\sup_{\substack{A\in\mathfrak{F}_{-\infty}^k \\ B\in\mathfrak{F}_{i+k}^\infty}} |P(AB) - P(A)P(B)| = \alpha(k) \downarrow 0$$

as $k \to \infty$. In particular, this means that the random variables X_i and X_{i+k} become asymptotically independent as k tends to infinity. The coefficient $\alpha(k)$ is referred to as the α-mixing (strong mixing) coefficient (see Rosenblatt, 1956).

The remaining part of the paper is organized as follows. In Section 2, we propose a nonparametric estimator for the residual extropy function. Asymptotic properties of the proposed estimator is elucidated in Section 3. Section 4 is devoted to a simulation study to illustrate the performance of the proposed estimator.

2. Estimation

In this section, we propose a nonparametric estimator for the residual extropy function.

Suppose $\{X_i; 1 \le i \le n\}$ is a sequence of identically distributed random variables representing the lifetimes for n components. Note that the X_i need not be mutually independent, that is, the lifetimes are assumed to be α-mixing (strong mixing).

A simple nonparametric estimator of $J(f;t)$ is written as

$$J_n^*(f;t) = \frac{-1}{2} \left\{ \frac{n^{-1} \sum_{i=1}^n f_n(X_i) I_{(X_i \ge t)}}{R_n^2(t)} \right\},$$

where

$$f_n(X_i) = \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{1}{b_j} K\left(\frac{X_i - X_j}{b_j}\right)$$

is the kernel estimator obtained from the sample without X_i (see Hall and Morton, 1993) and

$$R_n(t) = \frac{1}{n} \sum_{i=1}^{n} I_{(X_i \geq t)}, \tag{1}$$

where

$$I_{(X_i \geq t)} = \begin{cases} 1 & \text{if } X_i \geq t, \\ 0 & \text{otherwise} \end{cases}$$

is the empirical survival function under dependence condition (see Roussas, 1990).

A kernel estimator for $J(f;t)$ is defined as

$$J_n(f;t) = \frac{-1}{2} \left\{ \frac{\int_t^{\infty} f_{n,\tau}^2(x) dx}{R_n^2(t)} \right\}, \tag{2}$$

where $f_{n,\tau}(x)$ is a nonparametric estimator of density function and $R_n(t)$ is defined in (1).

Ruiz and Guillamon (1996) proposed the following estimator for the density function:

$$f_{n,\tau}(x) = \frac{1}{\sum_{j=1}^{n} b_j^{\tau}} \sum_{i=1}^{n} b_i^{\tau} K\left(\frac{x - X_i}{b_i^{\tau}}\right), \tag{3}$$

where $K(x)$ is a kernel of order s satisfying the following conditions: $K(x)$ is bounded, non-negative, symmetric, $K_i(x) = b_i^{-1} K(x/b_i)$, $\int K(x)\, dx = 1$, $\int xK(x)\, dx = 0$, τ is a positive real number, and $\{b_i\}$ is a sequence of real numbers satisfying the requirements $\lim_{n\to\infty} b_n = 0$, $\lim_{n\to\infty} nb_n = \infty$, and $\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} (b_i/b_n)^j = \beta_j < \infty$, $j = 1, 2, \ldots, s+1$.

It is a recursive density estimator and it can be calculated recursively by

$$f_{n,\tau}(x) = \frac{S_{n-1}}{S_n} f_{n-1,\tau}(x) + \frac{b_n^{\tau}}{S_n} K\left(\frac{x - X_n}{b_n^{\tau}}\right),$$

where $S_n = \sum_{j=1}^{n} b_j^{\tau}$. This property is particularly useful for large sample sizes.

Under α-mixing dependence condition, the expression for the bias and variance of $f_{n,\tau}(x)$ is (see Ruiz and Guillamon, 1996)

$$Bias\,(f_{n,\tau}(x)) \simeq \frac{b_n^s}{s!} \frac{\beta_{\tau+s}}{\beta_{\tau}} c_s f^{(s)}(x) \tag{4}$$

and

$$Var\,(f_{n,\tau}(x)) \simeq \frac{1}{nb_n} \frac{\beta_{2\tau-1}}{\beta_{\tau}^2} C_K f(x), \tag{5}$$

where $c_s = \int_{-\infty}^{\infty} u^s K(u)\, du$ and $C_K = \int_{-\infty}^{\infty} K^2(u)\, du$. Under the α-mixing dependence condition, the expression for the bias and variance of $R_n(t)$ is (see Roussas, 1990)

$$Bias\,(R_n(t)) = 0 \tag{6}$$

and

$$Var\,(R_n(t)) \simeq \frac{1}{n}\Big\{R(t)(1-R(t)) + D^n\Big\}, \tag{7}$$

where D^n is a remainder term due to the dependence involved.

3. Recursive property and asymptotic results

In this section, we establish a recursive property and some asymptotic results of $J_n(f;t)$.

Theorem 3.1. Let $J_n(f;t)$ be a nonparametric estimator of $J(f;t)$ as defined in (2), where $f_{n,\tau}(x)$ and $R_n(t)$ are given in (3) and (1) respectively. Then $J_n(f;t)$ satisfies the recursive property,

$$\begin{aligned} J_n(f;t) = \frac{1}{R_n^2(t)S_n^2}\Bigg\{ & S_{n-1}^2 R_{n-1}^2(t) J_{n-1}(f;t) \\ & - b_n^\tau \Bigg\{\int_t^\infty K\left(\frac{x-X_n}{b_n^\tau}\right)\left(S_{n-1} f_{n-1,\tau}(x) + \frac{b_n^\tau}{2} K\left(\frac{x-X_n}{b_n^\tau}\right)\right) dx\Bigg\}\Bigg\}. \end{aligned} \tag{8}$$

Proof. We have,

$$J_n(f;t) = \frac{-1}{2R_n^2(t)}\int_t^\infty f_{n,\tau}^2(x)dx$$

and

$$J_{n-1}(f;t) = \frac{-1}{2R_{n-1}^2(t)}\int_t^\infty f_{n-1,\tau}^2(x)dx.$$

Now

$$\begin{aligned} -2R_n^2(t)J_n(f;t) = {} & \frac{S_{n-1}^2}{S_n^2}\left(-2R_{n-1}^2(t)J_{n-1}(f;t)\right) \\ & + \frac{b_n^\tau}{S_n^2}\int_t^\infty K\left(\frac{x-X_n}{b_n^\tau}\right)\left(2S_{n-1}f_{n-1,\tau}(x) + b_n^\tau K\left(\frac{x-X_n}{b_n^\tau}\right)\right)dx. \end{aligned} \tag{9}$$

Simplifying (9), we get (8). ∎

Theorem 3.2. Suppose $J_n(f;t)$ is a nonparametric estimator of $J(f;t)$ as defined in (2), where $f_{n,\tau}(x)$ and $R_n(t)$ are given in (3) and (1) respectively. Then the estimator $J_n(f;t)$ is a consistent estimator of the residual extropy function $J(f;t)$. That is, as $n \to \infty$,

$$J_n(f;t) = \frac{-1}{2}\left\{\frac{\int_t^\infty f_{n,\tau}^2(x)dx}{R_n^2(t)}\right\} \xrightarrow{p} \frac{-1}{2}\left\{\frac{\int_t^\infty f^2(x)dx}{R^2(t)}\right\} = J(f;t).$$

Proof. By using Taylor's series expansion, we get

$$\int_t^\infty f_{n,\tau}^2(x)dx \simeq \int_t^\infty f^2(x)dx + 2\int_t^\infty \Big(f_{n,\tau}(x) - f(x)\Big)f(x)dx. \tag{10}$$

Using (4), (5) and (10), the expression for the bias and the variance of $\int_t^\infty f_{n,\tau}^2(x)dx$ is given by

$$Bias\left(\int_t^\infty f_{n,\tau}^2(x)dx\right) \simeq \frac{2b_n^s}{s!}\frac{\beta_{\tau+s}}{\beta_\tau}c_s\int_t^\infty f^{(s)}(x)f(x)\,dx$$

and

$$Var\left(\int_t^\infty f_{n,\tau}^2(x)dx\right) \simeq \frac{4}{nb_n}\frac{\beta_{2\tau-1}}{\beta_\tau^2}C_k\int_t^\infty f^3(x)dx.$$

The corresponding MSE is obtained by

$$MSE\left(\int_t^\infty f_{n,\tau}^2(x)dx\right) \simeq \left(\frac{2b_n^s}{s!}\frac{\beta_{\tau+s}}{\beta_\tau}c_s\int_t^\infty f^{(s)}(x)f(x)\,dx\right)^2 + \frac{4}{nb_n}\frac{\beta_{2\tau-1}}{\beta_\tau^2}C_k\int_t^\infty f^3(x)dx. \tag{11}$$

From (11), as $n \to \infty$

$$\text{MSE}\left(\int_t^\infty f_{n,\tau}^2(x)dx\right) \to 0.$$

Therefore, the estimator $\int_t^\infty f_{n,\tau}^2(x)dx$ is consistent (in the probability sense). That is,

$$\int_t^\infty f_{n,\tau}^2(x)dx \xrightarrow{p} \int_t^\infty f^2(x)dx.$$

By using Taylor's series expansion, we get

$$R_n^2(t) \simeq R^2(t) + 2\Big(R_n(t) - R(t)\Big)R(t). \tag{12}$$

Using (6), (7) and (12), the expression for the bias and the variance of $R_n^2(t)$ is given by

$$Bias\left(R_n^2(t)\right) \simeq 0$$

and

$$Var\left(R_n^2(t)\right) \simeq \frac{4R^2(t)}{n}\Big\{R(t))(1 - R(t) + D^n\Big\}.$$

The corresponding MSE is obtained by

$$MSE\left(R_n^2(t)\right) \simeq \frac{4R^2(t)}{n}\Big\{R(t)(1 - R(t)) + D^n\Big\}. \tag{13}$$

From (13), as $n \to \infty$

$$\text{MSE}\left(R_n^2(t)\right) \to 0.$$

Therefore, the estimator $R_n^2(t)$ is consistent (in the probability sense). That is,

$$R_n^2(t) \xrightarrow{p} R^2(t).$$

Then by Slutsky's theorem, we have as $n \to \infty$ that

$$J_n(f;t) = \frac{-1}{2}\left\{\frac{\int_t^{\infty} f_{n,\tau}^2(x)dx}{R_n^2(t)}\right\} \xrightarrow{p} \frac{-1}{2}\left\{\frac{\int_t^{\infty} f^2(x)dx}{R^2(t)}\right\} = J(f;t).$$

That is, $J_n(f;t)$ is a consistent estimator of $J(f;t)$. ■

Theorem 3.3. Suppose $J_n(f;t)$ is a nonparametric estimator of $J(f;t)$ as defined in (2), where $f_{n,\tau}(x)$ and $R_n(t)$ are given in (3) and (1) respectively. Then

$$(nb_n)^{\frac{1}{2}}\left\{\frac{J_n(f;t)-J(f;t)}{\sigma_{J_t}}\right\}$$

has a standard normal distribution as $n \to \infty$, with

$$\sigma_{J_t}^2 \simeq \frac{1}{R^2(t)}\frac{\beta_{2\tau-1}}{\beta_\tau^2}C_K\int_t^{\infty} f^3(x)dx.$$

Proof. Observe that

$$\begin{aligned}
&(nb_n)^{\frac{1}{2}}\left(J_n(f;t)-J(f;t)\right)\\
&= (nb_n)^{\frac{1}{2}}\left\{\frac{\int_t^{\infty} f_{n,\tau}^2(x)dx}{-2R_n^2(t)} - \frac{\int_t^{\infty} f^2(x)dx}{-2R^2(t)}\right\}\\
&= \frac{-(nb_n)^{\frac{1}{2}}}{2R_n^2(t)}\left\{\int_t^{\infty} f_{n,\tau}^2(x)dx - \frac{R_n^2(t)}{R^2(t)}\int_t^{\infty} f^2(x)dx\right\}\\
&= \frac{-(nb_n)^{\frac{1}{2}}}{2R_n^2(t)}\left\{\int_t^{\infty}\left(f_{n,\tau}^2(x)-f^2(x)\right)dx\right\} - \frac{(nb_n)^{\frac{1}{2}}}{2R_n^2(t)}\frac{\int_t^{\infty} f^2(x)dx}{R^2(t)}\left(R^2(t)-R_n^2(t)\right)\\
&\simeq \frac{-(nb_n)^{\frac{1}{2}}}{R_n^2(t)}\left\{\int_t^{\infty}\left(f_{n,\tau}(x)-f(x)\right)f(x)dx\right\} - \frac{2(nb_n)^{\frac{1}{2}}}{R_n^2(t)}J(f;t)R(t)\left(R_n(t)-R(t)\right).
\end{aligned}$$

By using the asymptotic normality of $f_{n,\tau}(x)$ established by Ruiz and Guillamon (1996), the asymptotic normality of $R_n(t)$ established by Roussas (1990), and by using Slutsky's theorem, the proof is immediate. ■

4. Simulation study

A Monte Carlo simulation study is carried out to compare the kernel estimators $J_n(f;t)$ and $J_n^*(f;t)$ in terms of the mean squared error. For that we considered the normal distribution with parameters $\mu = 5$ and $\sigma = 3$ and generated $\{X_i\}$ from an AR(1) process with correlation coefficient $\rho = 0.3$. The Gaussian kernel is used as the kernel function for the estimation. The biases and mean squared errors of $J_n(f;t)$ and $J_n^*(f;t)$ are computed for various values of $t \in (0.1, 1.5)$, $\tau = 0.01$, and sample sizes 100 and 300. The results are given in Table 1. From the table we can say that the mean squared error of $J_n(f;t)$ is small compared to the mean squared error of $J_n^*(f;t)$.

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

Table 1. Bias and mean squared error (in brackets) of $J_n(f;t)$ and $J_n^*(f;t)$ for $\tau = 0.01$.

	Bias and MSE of $J_n(f;t)$		Bias and MSE of $J_n^*(f;t)$	
t	$n = 100$	$n = 300$	$n = 100$	$n = 300$
0.1	0.00230 (0.000014)	0.00368 (0.000015)	-0.01055 (0.000133)	-0.00570 (0.000037)
0.2	0.00210 (0.000012)	0.00369 (0.000016)	-0.01072 (0.000136)	-0.00572 (0.000037)
0.3	0.00213 (0.000012)	0.00371 (0.000016)	-0.01073 (0.000133)	-0.00575 (0.000037)
0.4	0.00161 (0.000011)	0.00394 (0.000018)	-0.01111 (0.000143)	-0.00560 (0.000036)
0.5	0.00123 (0.000012)	0.00399 (0.000019)	-0.01146 (0.000155)	-0.00562 (0.000036)
0.6	0.00097 (0.000011)	0.00415 (0.000020)	-0.01175 (0.000163)	-0.00556 (0.000036)
0.7	0.00110 (0.000009)	0.00426 (0.000021)	-0.01168 (0.000160)	-0.00551 (0.000034)
0.8	0.00086 (0.000012)	0.00421 (0.000019)	-0.01198 (0.000172)	-0.00560 (0.000035)
0.9	0.001142 (0.000012)	0.00423 (0.000020)	-0.011879 (0.000169)	-0.00564 (0.000035)
1.0	0.00119 (0.000015)	0.00427 (0.000020)	-0.012005 (0.000177)	-0.00570 (0.000036)
1.1	0.00101 (0.000015)	0.00425 (0.000021)	-0.012246 (0.000183)	-0.00581 (0.000038)
1.2	0.00111 (0.000015)	0.00450 (0.00002)	-0.01223 (0.000182)	-0.00566 (0.000036)
1.3	0.00097 (0.000015)	0.00448 (0.000023)	-0.01240 (0.000185)	-0.00574 (0.000037)
1.4	0.00137 (0.000019)	0.00473 (0.000025)	-0.01220 (0.000183)	-0.00566 (0.000036)
1.5	0.00142 (0.000018)	0.00478 (0.000026)	-0.01229 (0.000183)	-0.00570 (0.000036)

References

Castellana, J. V. (1989). Integrated consistency of smoothed probability density estimators for stationary sequences. *Stochastic Processes and their Applications*, **33**, 335–346.

Castellana, J. V. and Leadbetter, M. R. (1986). On smoothed probability density estimation for stationary processes. *Stochastic processes and their applications*, **21**, 179–193.

Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution. *Sankhyā: The Indian Journal of Statistics, Series A*, **58**, 48–56.

Furuichi, S. and Mitroi, F.-C. (2012). Mathematical inequalities for some divergences. *Physica A: Statistical Mechanics and its Applications*, **391**, 388–400.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, **102**, 359–378.

Hall, P. and Morton, S. C. (1993). On the estimation of entropy. *Annals of the Institute of Statistical Mathematics*, **45**, 69–88.

Lad, F., Sanfilippo, G., and Agro, G. (2015). Extropy: complementary dual of entropy. *Statistical Science*, **30**, 40–58.

Martinas, K. (1998). Thermodynamics and sustainability a new approach by extropy. *Periodica Polytechnica Chemical Engineering*, **42**, 69–83.

Masry, E. (1983). Probability density estimation from sampled data. *IEEE Transactions on Information Theory*, **29**, 696–709.

Qiu, G. (2017). The extropy of order statistics and record values. *Statistics & Probability Letters*, **120**, 52–60.

Qiu, G. and Jia, K. (2018a). Extropy estimators with applications in testing uniformity. *Journal of Nonparametric Statistics*, **30**, 182–196.

Qiu, G. and Jia, K. (2018b). The residual extropy of order statistics. *Statistics & Probability Letters*, **133**, 15–22.

Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the United States of America*, **42**, 43–47.

Roussas, G. G. (1989). Some asymptotic properties of an estimate of the survival function under dependence conditions. *Statistics & Probability Letters*, **8**, 235–243.

Ruiz, J. and Guillamón, A. (1996). Nonparametric recursive estimator for mean residual life and vitality function under dependence conditions. *Communications in Statistics – Theory and Methods*, **25**, 1997–2011.

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

Vontobel, P. O. (2012). The Bethe permanent of a nonnegative matrix. *IEEE Transactions on Information Theory*, **59**, 1866–1901.

Manuscript received 2018-08-04, revised 2018-12-17, accepted 2019-02-05.