Journal of the Kerala Statistical Association
Vol.24, December 2013, 12-36 *ISSN 2249-4553*

Kernel Estimation of the Past Entropy Function with dependent data

R. Maya*
Department of Statistics
University of Kerala, Thiruvananthapuram - 695 581.

ABSTRACT

The past entropy function, introduced by Di Crescenzo and Longobardi (2002), is viewed as a dynamic measure of uncertainty in past life. This measure find applications in modeling life time data. In the present work we provide non-parametric kernel type estimators for the past entropy function based on complete and censored data. Asymptotic properties of the estimators are established under suitable regularity conditions. Monte-Carlo simulation studies are carried out to compare the performance of the estimators using the mean-squared error. The methods are illustrated using real data sets.

Key words and Phrases: *Past entropy function, Residual entropy function, Kernel estimate, α-mixing, Residual life.*

1 Introduction

Recently, many researchers have shown a keen interest in the measurement of uncertainty associated with a probability distribution of particular interest in probability and statistics is the notion of entropy, introduced by Shannon (1948). If X is a random variable having an absolutely continuous distribution function F with

*email: mayasekharan@gmail.com

probability density function f, then the entropy of the random variable X is defined as

$$H(X) = H(f) = -\int_0^\infty f(x)\log f(x)dx. \tag{1.1}$$

If we consider X as the lifetime of a new unit then $H(f)$ can be viewed as a useful tool for measuring the associated uncertainty. Ebrahimi and Pellerey (1995) and Ebrahimi (1996) have introduced the concept of residual entropy in terms of a conditional measure. For a non-negative random variable X, representing the life time of a component, the residual entropy function is the Shannon's entropy associated with the random variable X given $X > t$, and is defined as

$$\begin{aligned} H(f;t) &= -\int_t^\infty \frac{f(x)}{\overline{F}(t)}\log\frac{f(x)}{\overline{F}(t)}dx \\ &= 1 - \frac{1}{\overline{F}(t)}\int_t^\infty f(x)\log h(x)dx,\ \overline{F}(t) > 0, \end{aligned} \tag{1.2}$$

where $\overline{F}(t) = P(X > t)$ denotes the survival function and $h(x) = \frac{f(x)}{\overline{F}(x)}$ is the hazard function of X. Belzunce et al. (2004) have established that if $H(f;t)$ is increasing in t then $H(f;t)$ determines the distribution uniquely. Given that an item has survived up to time t, $H(f;t)$ measures the uncertainty in its remaining life. For a discussion of the properties and applications of residual entropy we refer to Ebrahimi and Kirmani (1996), Nair and Rajesh (1998) and Asadi and Ebrahimi (2000).

It is reasonable to presume that in many realistic situations uncertainty is not necessarily related to the future but can also refer to the past. For instance, consider a system whose state is observed only at certain preassigned inspection times. If at time t, the system is inspected for the first time and it is found to be 'down', then the uncertainty relies on the past, i.e. on which instant in (0, t) it has failed. It thus seems natural to introduce a notion of uncertainty that is dual to the residual entropy, in the sense that it refers to past time and not to future time. Based on this idea Di Crescenzo and Longobardi (2002) have studied the past entropy. Further, they discussed the necessity of the past entropy, its relation with residual entropy and many interesting results. In Di Crescenzo and Longobardi (2004) a measure of discrimination based on past entropy has been studied. If X denotes the lifetime of a component/system or of living organism, then the past entropy of X at time t is defined as

$$\begin{aligned} \overline{H}(f;t) &= -\int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} \, dx \\ &= 1 - \frac{1}{F(t)} \int_0^t f(x) \log T(x) \, dx, \end{aligned} \tag{1.3}$$

where $T(x) = \frac{f(x)}{F(x)}$ is the reversed hazard rate of x.

It can also be used in Forensic Sciences where exact time of failure (death in case of human being) is important when at some time t the unit is found to be in failure state. Gupta and Nanda (2002) have defined generalized uncertainty of lifetime distribution by truncating the distributions above some point t. Nanda and Paul (2006) have studied some properties and applications of past entropy. Gupta (2009) established that the past entropy determines the distribution uniquely, under certain conditions.

To make valid decisions regarding the extent of uncertainty in past data, one requires a reasonable estimate of the past entropy. Motivated by this in the present paper we provide nonparametric estimators for $\overline{H}(f;t)$ using kernel type estimation for complete as well as censored data. We consider only situations where the data under study are dependent. In both situations, the underlying lifetimes are assumed to be α-mixing (see, Rosenblatt (1956)).

Definition 1.

Let $\{X_i; i \geq 1\}$ denote a sequence of random variables. Given a positive integer n, set

$$\alpha(n) = \sup_{k\geq 1}\{|P(A \cap B) - P(A)P(B)|; A\epsilon\mathfrak{F}_1^k, B\epsilon\mathfrak{F}_{k+n}^\infty \tag{1.4}$$

where $\mathfrak{F}_i^k$ denote the σ- field of events generated by $\{X_j; i \leq j \leq k\}$. The sequence is said to be α-mixing (strong mixing) if the mixing coefficient $\alpha(n) \to 0$ as $n \to \infty$. Among various mixing conditions, α-mixing is reasonably weak and has many practical applications.

The rest of the paper is organized as follows. In Section 2, we present estimators for $\overline{H}(f;t)$, given in (1.3), using complete and censored samples. In Section 3, we examine the consistency and asymptotic normality of the estimators. In Section 4, we evaluate the estimators for real data sets and in Section 5 a simulation study to illustrate the behavior of the estimators is undertaken.

2 Estimation

In this section, we propose nonparametric estimators for the past entropy function for complete as well as censored data sets.

2.1 Complete Samples

Let $\{X_i; 1 \leq i \leq n\}$ be a sequence of identically distributed random variables. Note that ${X_i}'s$ need not be mutually independent. A simple nonparametric estimator of $\overline{H}(f;t)$ is given as

$$\overline{H}^*(f;t) = \frac{-1}{n}\sum_{i=1}^{n}\log\left(\frac{f_n(X_i)}{F_n(t)}\right) I_{(X_i \leq t)}, \tag{2.1}$$

where $f_n(X_i) = \frac{1}{(n-1)}\sum\limits_{j\neq i}^{n}\frac{1}{b_n}K\left(\frac{X_i - X_j}{b_n}\right)$ is the kernel estimator obtained from the sample without X_i, $F_n(t)$ is either an empirical or a kernel estimator for the distribution function and

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

A kernel estimator of $\overline{H}(f;t)$ for the above sample is defined as

$$\overline{H}_n(f;t) = -\int_0^t \frac{f_n(x)}{F_n(t)}\log\frac{f_n(x)}{F_n(t)}dx. \tag{2.2}$$

(2.2) can also be written as

$$\overline{H}_n(f;t) = \log F_n(t) - \frac{1}{F_n(t)}\int_0^t f_n(x)\log f_n(x)dx, \tag{2.3}$$

where $f_n(x)$ is a nonparametric estimator of $f(x)$ and $F_n(t) = \int_0^t f_n(x)\,dx$ is a nonparametric estimator of distribution function $F(t)$.
The most common nonparametric density estimator of $f(x)$ is the kernel estimator given by (see, Parzen (1962), Rosenblatt (1970))

$$f_n(x) = \frac{1}{nb_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{b_n}\right), \tag{2.4}$$

where $K(x)$ is a kernel of order s with compact support satisfying the conditions

i) $K(x) \geq 0$ for all x

ii) $\int K(x)\,dx = 1$

iii) K(.) is symmetric about zero

iv) $K_n(x) = \frac{1}{b_n} K(\frac{x}{b_n})$ where $\{b_n\}$ is a bandwidth sequence of positive numbers such that $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$ and

v) $K(.)$ satisfies Lipschitz condition, namely there exist a constant M such that $|K(x) - K(y)| \leq M|x - y|$.

Under α-mixing dependence conditions, expressions for the bias and variance of $f_n(x)$ are

$$Bias\,(f_n(x)) \backsimeq \frac{b_n^s c_s}{s!} f^{(s)}(x) \tag{2.5}$$

and

$$Var\,(f_n(x)) \backsimeq \frac{1}{nb_n} f(x) C_K, \tag{2.6}$$

where $c_s = \int\limits_{-\infty}^{\infty} u^s K(u)\ du$ and $C_K = \int\limits_{-\infty}^{\infty} K^2(u)\ du$.

2.2 Censored Samples

In reliability and life testing, due to time constraints or cost considerations the experimenter is forced to terminate the experiment after a specific period of time or after the failure of a specified number of units. In this context the underlying data will be censored. In the context of right censoring only, the lower bounds on life time will be available for some individuals and in the context of left censoring data will be recorded as the upper bound of life time for some individuals. Another common type of censoring is random censoring.

Let $\{X_i; 1 \leq i \leq n\}$ be a sequence of non-negative random variables representing the life times for n components/devices. The random variables are not assumed to be mutually independent. However they have a common unknown continuous marginal distribution function $F(x)$ with a probability density function $f(x) = F^{'}(x)$. Let the random variable X_i be censored on the right by the random variable Y_i. In this random censorship model, the censoring times Y_i are assumed to be independently and identically distributed and they are also assumed to be independent of X_i. The censoring times Y_1, Y_2, ..., Y_n have the common distribution function $P(y)$. This scheme is very common in clinical trials. In such experiments, patients enter into the study at random time points, while the experiment itself is terminated at a prespecified time. Let $Z_i = min(X_i, Y_i)$ and $\delta_i = I(X_i \leq Y_i)$, where I(.) denotes the indicator function of the event specified in parentheses. The actually observed $Z_i's$ have a distribution function G satisfying

$$1 - G(t) = (1 - F(t))(1 - P(t)),\ t\ \epsilon R_+ = [0, \infty).$$

Let $G^*(t) = P(Z_1 \leq t; \delta_1 = 1)$ is the corresponding sub-distribution function for the uncensored observations and $g^*(t) = f(t)(1 - P(t))$ be the corresponding sub-density. A reasonable estimator of f should behave like $\frac{g_n^*(t)}{(1-P(t))}$ where $g_n^*(t) =$

$b_n^{-1} \int\limits_{R^+} K(\frac{t-x}{b_n})\ dG_n^*(x)$ is the kernel estimator pertaining to $G_n^*(t) = \frac{1}{n}\sum\limits_{i=1}^{n} I(Z_i \leq t; \delta_i = 1)$.
A nonparametric estimator for $\overline{H}(f;t)$ based on the censored data is

$$\overline{H}_*^n(f;t) = \frac{-1}{n}\sum_{i=1}^{n} \log\left(\frac{f_n(Z_i)}{F_n^*(t)}\right) I_{(Z_i \leq t)}, \tag{2.7}$$

where $f_n(Z_i) = \frac{1}{(n-1)}\sum\limits_{j\neq i}^{n} \frac{1}{b_n} K\left(\frac{Z_i - Z_j}{b_n}\right)$ is the kernel estimator obtained from the sample without Z_i and $F_n^*(t)$ is a kernel estimator for the distribution function.
A non parametric estimator for past entropy function under $\alpha-$ mixing condition, in the random censorship model, is defined as

$$\overline{H}_n^*(f;t) = \log F_n^*(t) - \frac{1}{F_n^*(t)}\int_0^t f_n^*(x)\log f_n^*(x)dx, \tag{2.8}$$

where

$$f_n^*(t) = \frac{1}{b_n}\int_{R^+} \frac{K(\frac{t-x}{b_n})}{1-P(x)}\ dG_n^*(x), \tag{2.9}$$

is a nonparametric density estimator for $f(x)$ under censoring (see, Cai (1998)), $F_n^*(t) = \int\limits_0^t f_n^*(u)\ du$, $K(x)$ is a kernel of order s with compact support which satisfies the conditions (i)-(v) in Section 2.1 and $\{b_n\}$ is a sequence of real numbers such that $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$.
Under α-mixing dependence conditions, expressions for the bias and variance of $f_n^*(t)$ are given by (see, Cai (1998))

$$Bias\,(f_n^*(t)) \backsimeq \frac{b_n^s c_{s^+}}{s!} f^{(s)}(t) \tag{2.10}$$

and

$$Var\,(f_n^*(t)) \backsimeq \frac{1}{nb_n}\frac{f(t)}{(1-P(t))} C_K, \tag{2.11}$$

where $c_{s^+} = \int\limits_{R^+} u^s K(u)\ du$ and $C_K = \int\limits_{-\infty}^{\infty} K^2(u)\ du$.

3 Asymptotic properties

In this section, we look in to the consistency and asymptotic normality of the estimators (2.3) and (2.8). In order to simplify the notations, define

$V_n(t) = \log F_n(t)$, $V_n^*(t) = \log F_n^*(t)$, $\quad V(t) = \log F(t)$,
$A_n(t) = \int\limits_0^t f_n(x) \log f_n(x) dx$, $A_n^*(t) = \int\limits_0^t f_n^*(x) \log f_n^*(x) dx$
and $\; A(t) = \int\limits_0^t f(x) \log f(x) dx$.

Theorem 3.1. *Let $K(x)$ be a kernel of order s with compact support satisfying the conditions (i)-(v) in Section 2 and $\{b_n\}$ be such that $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$. Then*

a) $\overline{H}_n(f;t)$ is a consistent estimator of $\overline{H}(f;t)$.
b) $\overline{H}_n^*(f;t)$ is a consistent estimator of $\overline{H}(f;t)$.

Proof. a) Under α-mixing dependence conditions, we obtain the expressions for the bias and the variance of $F_n(t)$, $V_n(t)$ and $A_n(t)$ and are given by

$$Bias\,(F_n(t)) \backsimeq \frac{b_n^s c_s}{s!} \int\limits_0^t f^{(s)}(x)\; dx, \tag{3.1}$$

$$Var\,(F_n(t)) \backsimeq \frac{C_K}{nb_n} F(t), \tag{3.2}$$

$$Bias\,(V_n(t)) \backsimeq \frac{b_n^s c_s}{s! F(t)} \int\limits_0^t f^{(s)}(x)\; dx, \tag{3.3}$$

$$Var\,(V_n(t)) \backsimeq \frac{1}{nb_n} \frac{C_k}{F(t)}, \tag{3.4}$$

$$Bias(A_n(t)) \backsimeq \left(\frac{b_n^s}{s!} c_s\right) \int\limits_0^t (1 + log f(x)) f^{(s)}(x) dx \tag{3.5}$$

and

$$Var(A_n(t)) \backsimeq \left(\frac{1}{nb_n}C_k\right)\int\limits_0^t (1+\log f(x))^2 f(x)dx. \tag{3.6}$$

The corresponding MSE's are given by

$$MSE\left(F_n(t)\right) \backsimeq \left(\frac{b_n^s c_s}{s!}\int\limits_0^t f^{(s)}(x)\ dx\right)^2 + \frac{C_K}{nb_n}F(t), \tag{3.7}$$

$$MSE\left(V_n(t)\right) \backsimeq \left(\frac{b_n^s c_s}{s!F(t)}\int\limits_0^t f^{(s)}(x)\ dx\right)^2 + \frac{1}{nb_n}\frac{C_k}{F(t)} \tag{3.8}$$

and

$$\begin{aligned} MSE\left(A_n(t)\right) \backsimeq & \left(\left(\frac{b_n^s}{s!}c_s\right)\int\limits_0^t (1+logf(x))f^{(s)}(x)dx\right)^2 \\ & + \left(\frac{1}{nb_n}C_k\right)\int\limits_0^t (1+\log f(x))^2 f(x)dx. \end{aligned} \tag{3.9}$$

From (3.7), as $n \to \infty$

$$\text{MSE}\left(F_n(t)\right) \to 0.$$

From (3.8), as $n \to \infty$

$$\text{MSE}\left(V_n(t)\right) \to 0.$$

From (3.9), as $n \to \infty$

$$\text{MSE}\left(A_n(t)\right) \to 0.$$

Therefore

$$\overline{H}_n(f;t) = V_n(t) - \frac{A_n(t)}{F_n(t)} \xrightarrow{p} V(t) - \frac{A(t)}{F(t)} = \overline{H}(f;t).$$

That is, $\overline{H}_n(f;t)$is a consistent estimator of $\overline{H}(f;t)$.

b) Under α-mixing dependence conditions, we obtain the expressions for the bias and the variance of $F_n^*(t)$, $V_n^*(t)$ and $A_n^*(t)$ and are given by

$$Bias\left(F_n^*(t)\right) \backsimeq \frac{b_n^s c_{s+}}{s!}\int\limits_0^t f^{(s)}(u)\ du, \tag{3.10}$$

$$Var\left(F_n^*(t)\right) \backsimeq \frac{1}{nb_n} C_K \int\limits_0^t \frac{f(u)}{[1-P(u)]}\ du, \tag{3.11}$$

$$Bias\left(V_n^*(t)\right) \backsimeq \frac{b_n^s}{s!} \frac{c_{s+}}{F(t)} \int\limits_0^t f^{(s)}(u)\ du, \tag{3.12}$$

$$Var\left(V_n^*(t)\right) \backsimeq \frac{1}{nb_n} \frac{C_k}{F^2(t)} \int\limits_0^t \frac{f(u)}{[1-P(u)]} du, \tag{3.13}$$

$$Bias(A_n^*(t)) \backsimeq \left(\frac{b_n^s}{s!} c_s^+\right) \int\limits_0^t (1+logf(u)) f^{(s)}(u) du \tag{3.14}$$

and

$$Var(A_n^*(t)) \backsimeq \left(\frac{1}{nb_n} C_k\right) \int\limits_0^t (1+\log f(u))^2 \frac{f(u)}{[1-P(u)]} du. \tag{3.15}$$

The corresponding MSE's are given by

$$MSE\left(F_n^*(t)\right) \backsimeq \left(\frac{b_n^s c_{s+}}{s!} \int\limits_0^t f^{(s)}(u)\ du\right)^2 + \frac{1}{nb_n} C_K \int\limits_0^t \frac{f(u)}{[1-P(u)]}\ du, \tag{3.16}$$

$$MSE\left(V_n^*(t)\right) \backsimeq \left(\frac{b_n^s}{s!} \frac{c_{s+}}{F(t)} \int\limits_0^t f^{(s)}(u)\ du\right)^2 + \frac{1}{nb_n} \frac{C_k}{F^2(t)} \int\limits_0^t \frac{f(u)}{[1-P(u)]} du \tag{3.17}$$

and

$$\begin{aligned} MSE\left(A_n^*(t)\right) \backsimeq & \left(\left(\frac{b_n^s}{s!} c_s^+\right) \int\limits_0^t (1+logf(u)) f^{(s)}(u) du\right)^2 \\ & + \left(\frac{1}{nb_n} C_k\right) \int\limits_0^t (1+\log f(u))^2 \frac{f(u)}{[1-P(u)]} du. \end{aligned} \tag{3.18}$$

From (3.16), as $n \to \infty$

$$\text{MSE}\left(F_n^*(t)\right) \to 0.$$

From (3.17), as $n \to \infty$

$$\text{MSE}\,(V_n^*(t)) \to 0.$$

From (3.18), as $n \to \infty$

$$\text{MSE}\,(A_n^*(t)) \to 0.$$

Therefore

$$\overline{H}_n^*(f;t) = V_n^*(t) - \frac{A_n^*(t)}{F_n^*(t)} \xrightarrow{p} V(t) - \frac{A(t)}{F(t)} = \overline{H}(f;t).$$

That is, $\overline{H}_n^*(f;t)$ is a consistent estimator of $\overline{H}(f;t)$. □

If $g_n(x)$ is an estimator for $g(x)$, then its MISE (Integrated mean-squared error) is given by:

$$MISE\,(g_n(x)) = E\left\{\int_{-\infty}^{\infty} (g_n(x) - g(x))^2\, w(t)dF(t)\right\},$$

where $w(.)$ is a weight function.
In the following theorem, we give expressions for the MISE of $\overline{H}_n(f;t)$ and $\overline{H}_n^*(f;t)$ as $n \to \infty$.

Theorem 3.2. *Let $K(x)$ be a kernel of order s with compact support satisfying the conditions (i)-(v) in Section 2 and $\{b_n\}$ satisfying the conditions given in Section 2. Then*
a) $\lim\limits_{n\to\infty} MISE\left(\overline{H}_n(f;t)\right) = 0.$
b) $\lim\limits_{n\to\infty} MISE\left(\overline{H}_n^*(f;t)\right) = 0.$

Proof.

$$MISE\left(\overline{H}_n(f;t)\right) = E\left\{\int_{-\infty}^{\infty} \left(\overline{H}_n(f;t) - \overline{H}(f;t)\right)^2 w(t)dF(t)\right\}. \tag{3.19}$$

$$= E\left\{\int_{-\infty}^{\infty} \left[(V_n(t) - V(t)) - \left(\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)}\right)\right]^2 w(t)dF(t)\right\}$$

$$= E\left\{\int_{-\infty}^{\infty} \left(V_n(t) - V(t)\right)^2 w(t)dF(t)\right\}$$

$$+ E\left\{\int_{-\infty}^{\infty} \left(\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)}\right)^2 w(t)dF(t)\right\}$$

$$- 2E\left\{\int_{-\infty}^{\infty} \left(V_n(t) - V(t)\right)\left(\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)}\right) w(t)dF(t)\right\}.$$

Let H_1, H_2 and H_3 are given by

$$H_1 = E\left\{\int_{-\infty}^{\infty} \left(V_n(t) - V(t)\right)^2 w(t)dF(t)\right\}$$

$$= \int_{-\infty}^{\infty} \left[Var(V_n(t)) + Bias^2(V_n(t))\right] w(t)dF(t).$$

$$H_2 = E\left\{\int_{-\infty}^{\infty} \left(\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)}\right)^2 w(t)dF(t)\right\}.$$

Using the approximation $\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)} = \frac{A_n(t) - F_n(t)\frac{A(t)}{F(t)}}{F(t)}\left(1 + o_p(1)\right)$, the above equation simplifies to

$$H_2 = E\left\{\int_{-\infty}^{\infty} \left(\frac{A_n(t) - F_n(t)\frac{A(t)}{F(t)}}{F(t)}\right)^2 w(t)dF(t)\right\}.$$

Using Holder inequality, we get

$$H_2 \le \int_{-\infty}^{\infty} \left\{Var\left(A_n(t)\right) + \left[E\left(A_n(t)\right)\right]^2\right\} \frac{w(t)}{F^2(t)} dF(t)$$

$$+ \int_{-\infty}^{\infty} \left\{Var\left(F_n(t)\right)\frac{A^2(t)}{F^2(t)} + \left[E\left(F_n(t)\right)\right]^2 \frac{A^2(t)}{F^2(t)}\right\} \frac{w(t)}{F^2(t)} dF(t)$$

$$- 2\int_{-\infty}^{\infty} \frac{A(t)}{F(t)} \left[E\left(A_n(t)\right)^2\right]^{\frac{1}{2}} \left[E\left(F_n(t)\right)^2\right]^{\frac{1}{2}} \frac{w(t)}{F^2(t)} dF(t).$$

$$H_3 = 2E\left\{\int_{-\infty}^{\infty} \left(V_n(t) - V(t)\right)\left(\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)}\right) w(t)dF(t)\right\}$$

$$\leq 2\int\limits_{-\infty}^{\infty}\left[E\left(V_n(t)-V(t)\right)^2\right]^{\frac{1}{2}}\left[E\left(\frac{A_n(t)}{F_n(t)}-\frac{A(t)}{F(t)}\right)^2\right]^{\frac{1}{2}} w(t)dF(t).$$

From (3.1), (3.2), (3.3), (3.4), (3.5) and (3.6), it follows that

$$MISE\left(\overline{H}_n(f;t)\right)\rightarrow o \ as \ n\rightarrow\infty. \tag{3.20}$$

b) The proof is similar to that of a). □

In the following theorem, we focus attention on the asymptotic normality of the estimators $\overline{H}_n(f;t)$ and $\overline{H}^*_n(f;t)$.

Theorem 3.3. *Suppose that F be continuous. Assume that K(.) satisfies the assumptions (i)-(v) of Section 2 . Then*

$$a)\,(nb_n)^{\frac{1}{2}}\left\{\frac{(\overline{H}_n(f;t)-\overline{H}(f;t))}{\sigma_{\overline{H}}}\right\} \tag{3.21}$$

has a standard normal distribution as $n\rightarrow\infty$ with
$\sigma^2_{\overline{H}}\backsimeq\frac{C_k}{nb_nF^2(t)}\int\limits_0^t f(x)\left[\frac{A^2(t)}{F^2(t)}+1+(1+\log f(x))^2\right]dx.$

$$b)\,(nb_n)^{\frac{1}{2}}\left\{\frac{(\overline{H}^*_n(f;t)-\overline{H}(f;t))}{\sigma_{\overline{H}^*}}\right\} \tag{3.22}$$

has a standard normal distribution as $n\rightarrow\infty$ with

$\sigma^2_{\overline{H}^*}\backsimeq\frac{C_k}{nb_nF^2(t)}\int\limits_0^t \frac{f(x)}{1-P(x)}\left[\frac{A^2(t)}{F^2(t)}+1+(1+\log f(x))^2\right]dx.$

Proof. a)

$$
\begin{aligned}
&\sqrt{nb_n}\left(\overline{H}_n(f;t)-\overline{H}(f;t)\right)\\
&\quad=\sqrt{nb_n}\left[(V_n(t)-V(t))-\left[\frac{A_n(t)}{F_n(t)}-\frac{A(t)}{F(t)}\right]\right]\\
&\quad\backsimeq\sqrt{nb_n}\left[\left(\frac{F_n(t)-F(t)}{F(t)}\right)-\left(\frac{A_n(t)F(t)-A(t)F_n(t)}{F_n(t)F(t)}\right)\right]\\
&\quad=\sqrt{nb_n}\left[\frac{(F_n(t)-F(t))}{F(t)}-\frac{F(t)\left[A_n(t)-A(t)\right]}{F_n(t)F(t)}\right]\\
&\qquad-\sqrt{nb_n}\left[\frac{A(t)\left[F_n(t)-F(t)\right]}{F_n(t)F(t)}\right].
\end{aligned}
\tag{3.23}
$$

Since $\sup_t |F_n(t)-F(t)| \to 0$ *a.s.*, (3.23) is asymptotically equal to

$$
\begin{aligned}
&\sqrt{nb_n}\left(\overline{H}_n(f;t)-\overline{H}(f;t)\right)\\
&\quad\backsimeq\sqrt{nb_n}\left[\frac{1}{F(t)}-\frac{A(t)}{F^2(t)}\right](F_n(t)-F(t))\\
&\qquad-\sqrt{nb_n}\left(\frac{A_n(t)-A(t)}{F(t)}\right)\\
&\quad\backsimeq\sqrt{nb_n}\left[\frac{1}{F(t)}-\frac{A(t)}{F^2(t)}\right]\int_0^t (f_n(u)-f(u))\;du\\
&\qquad-\frac{\sqrt{nb_n}}{F(t)}\int_0^t (f_n(u)-f(u))\,(1+\log f(u))\;du
\end{aligned}
\tag{3.24}
$$

Note that from Parzen (1962), $\sqrt{nb_n}\,(f_n(x)-f(x))$ is asymptotically normal with mean zero and variance σ_f^2 given in (2.6).
Now from (3.24), it is immediate that

$$
(nb_n)^{\frac{1}{2}}\left\{\frac{(\overline{H}_n(f;t)-\overline{H}(f;t))}{\sigma_{\overline{H}}}\right\}
\tag{3.25}
$$

is asymptotically normal with mean zero. The expression of variance can be obtained from (3.2), (3.4) and (3.6).
b) The proof is similar to that of *a*). □

4 Numerical illustration

Example 4.1.

To illustrate the usefulness of the proposed estimator discussed in Section2.1 with real situations, we consider the times, in months, to the first failure of 20 electric carts used for internal delivery and transportation in a large manufacturing facility (see, Zimmer et al (1998)). We use the bootstrapping procedure to find optimum value of b_n (see, Efron (1981)). The Gaussian kernel $K(z) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right)$ is used as the kernel function for the estimation. At each value of t and b_n we calculate the biases and the mean-squared errors of $\overline{H}_n(f;t)$ using 250 bootstrap samples of size 20. Table 1 presents the bootstrap estimates of the biases and the mean-squared errors, in brackets, for $\overline{H}_n(f;t)$. From the Table 1 it can be seen that the optimum value of b_n is 0.2 for $0.9 < t \leq 12.6$, 0.3 for $12.6 < t \leq 16.3$, 0.4 for $16.3 < t \leq 19.3$, 0.6 for $19.3 < t \leq 22.6$, 0.7 for $22.6 < t \leq 24.8$ and 0.9 for $24.8 < t \leq 53$. In Figure 1 dark line represents the theoretical value $\overline{H}(f;t)$ and starred line represents the estimator $\overline{H}_n(f;t)$. From Figure 1, it is easy to see that for the data set considered the past entropy function is increasing with time.

Example 4.2.

For the illustration of estimation procedure discussed in Section 2.2, we consider the life times (in cycles) of 20 sodium sulphur batteries (see, Ansell and Ansell (1987)). The same method is used for calculating the bootstrap estimate of the bias and the mean-squared errors of $\overline{H}^*_n(f;t)$ as in the case of Example 4.1. The bootstrap estimates of the biases and the mean-squared errors, in brackets, for $\overline{H}^*_n(f;t)$ are given in Table 2. From the Table 2 it can be seen that the optimum value of b_n is 0.1 for $0.76 < t \leq 2.1$, 0.2 for $2.1 < t \leq 7.75$, 0.3 for $7.75 < t \leq 8.14$, 0.5 for $8.14 < t \leq 11.31$, 0.7 for $11.31 < t \leq 14.46$ and 0.9 for $14.46 < t \leq 30.77$. In Figure 2 dark line represents the theoretical value $\overline{H}(f;t)$ and starred line represents the estimator $\overline{H}^*_n(f;t)$. From Figure 2, we can say that for the given data the past entropy function is increasing with time.

5 Simulation studies

A Monte Carlo simulation study is carried out to compare the kernel estimators $\overline{H}_n(f;t)$ and $\overline{H}^*(f;t)$ in the case of complete samples and $\overline{H}^*_n(f;t)$ and $\overline{H}^n_*(f;t)$ in the case of censored samples in terms of the mean-squared error. First we consider the simulation under complete sample. The exponential distribution with parameter $\lambda = 0.6$ is used for the simulation. For the simulation under complete sample, we generated $\{X_i\}$ from AR(1) with correlation coefficient $\rho = 0.3$. The Gaussian kernel is used as the kernel function for the estimation. The estimates for various values of t $(4 < t < 4.8)$, b_n and sample sizes $n = 50$ and $n = 100$ are calculated. The ratios of mean-squared error of $\overline{H}_n(f;t)$ to that of the $\overline{H}^*(f;t)$ are computed and are given in Table 3. From a range of b_n values, we found that $b_n = 0.4$ come close to giving the smallest discrepancy between $\overline{H}_n(f;t)$ and $\overline{H}^*(f;t)$. In Figure 3 dark line represents the theoretical value $\overline{H}(f;t)$, starred line represents the estimator $\overline{H}_n(f;t)$ and dotted line represents the estimator $\overline{H}^*(f;t)$.

For the simulation under right censored sample, we generated $\{X_i\}$ from AR(1) with correlation coefficient $\rho = 0.2$ and the censoring times $\{Y_i\}$ were generated independently from $N(2,1)$. In this case also we used the Gaussian kernel for simulation. The estimates for various values of $t(2 \leq t \leq 2.8)$, b_n and sample sizes $n = 50$ and $n = 100$ are calculated. The ratios of mean-squared error of $\overline{H}^*_n(f;t)$ to that of the $\overline{H}^n_*(f;t)$ for the b_n values between 0.3 - 0.6 are computed and are given in Table 4. The ratios of mean-squared error of $\overline{H}^*_n(f;t)$ to that of the $\overline{H}^n_*(f;t)$ for $b_n = 0.6$ is very small compared to the other b_n values. In Figure 4 dark line represents the theoretical value $\overline{H}(f;t)$, starred line represents the estimator $\overline{H}^*_n(f;t)$ and dotted line represents the estimator $\overline{H}^n_*(f;t)$.

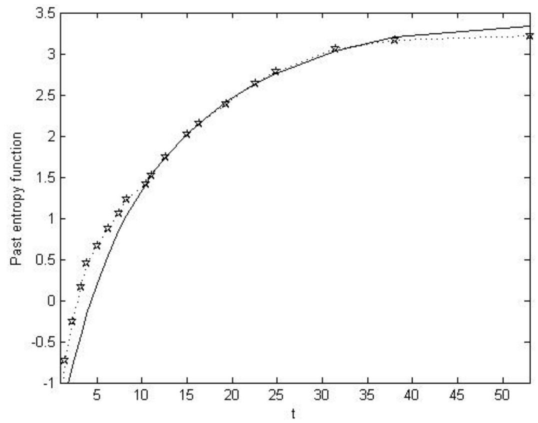

Figure 1: Plots of estimates of past entropy function for the first failure of 20 electric carts

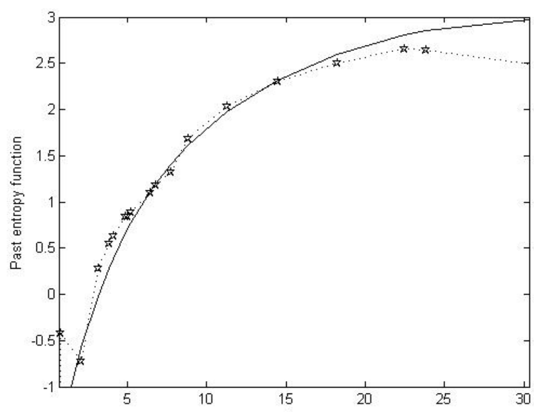

Figure 2: Plots of estimates of past entropy function for the life times of 20 sodium sulphur batteries

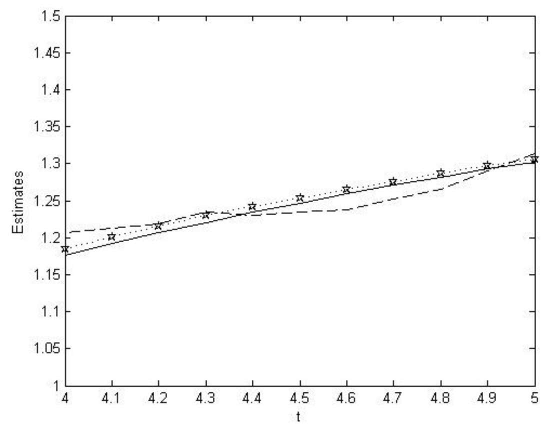

Figure 3: Plots of $\overline{H}_n(f;t)$, $\overline{H}^*(f;t)$ and $\overline{H}(f;t)$ using a simulated sample of size n=100 for $b_n = 0.4$, $\lambda = 0.6$

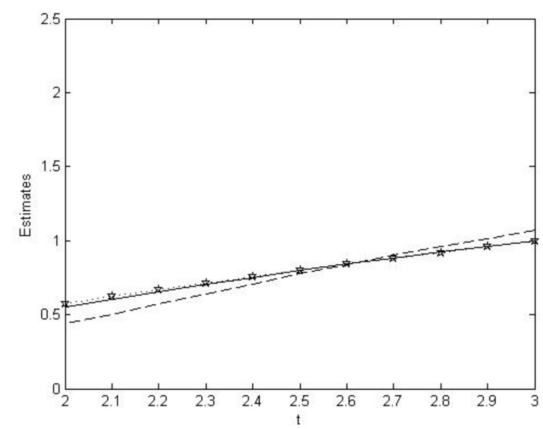

Figure 4: Plots of $\overline{H}_n^*(f;t)$, $\overline{H}_*^n(f;t)$ and $\overline{H}(f;t)$ using a simulated sample of size n=100 for $b_n = 0.6$, $\mu = 2$, $\sigma = 1$

Table 1: Bootstrap bias and mean-squared error estimates of $\overline{H}_n(f;t)$ for a real data set

	b_n							
t	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.9	0.0023	0.5532	0.8844	1.0846	1.2020	1.2609	1.2975	1.3203
	(0.2926)	(0.4403)	(0.8324)	(1.1921)	(1.4853)	(1.6067)	(1.6904)	(1.7463)
1.5	0.6078	0.9093	1.1552	1.2605	1.3597	1.4185	1.4598	1.4882
	(0.5650)	(0.9192)	(1.2871)	(1.6074)	(1.9259)	(2.0498)	(2.1491)	(2.2235)
2.3	0.5346	0.9260	1.1047	1.2274	1.3166	1.3834	1.4340	1.4725
	(0.8725)	(1.1871)	(1.4401)	(1.6560)	(1.8348)	(1.9817)	(2.1013)	(2.1978)
3.2	0.5853	0.9342	1.1175	1.2376	1.3179	1.3728	1.4117	1.4403
	(0.7449)	(1.1748)	(1.4580)	(1.6849)	(1.8515)	(1.9708)	(2.0571)	(2.1219)
3.9	0.6124	0.9964	1.1508	1.2489	1.3135	1.3569	1.3871	1.4092
	(0.6888)	(1.1554)	(1.4555)	(1.6654)	(1.8092)	(1.9068)	(1.9748)	(2.0243)
5	0.46702	0.8723	1.0389	1.1439	1.2116	1.2556	1.3099	1.3239
	(0.4823)	(0.8635)	(1.1541)	(1.3648)	(1.5114)	(1.6105)	(1.7214)	(1.7574)
6.2	0.3402	0.6974	0.8376	0.9278	0.9924	1.0413	1.0798	1.1112
	(0.3020)	(0.5770)	(0.7836)	(0.9316)	(1.0440)	(1.1323)	(1.2041)	(1.2643)
7.5	0.1942	0.5279	0.6910	0.7974	0.8721	0.9269	0.9683	1.0035
	(0.1248)	(0.3397)	(0.5243)	(0.6704)	(0.7856)	(0.9269)	(0.9516)	(1.0116)
8.3	0.1965	0.5232	0.6763	0.7705	0.8343	0.8800	0.9142	0.9407
	(0.1213)	(0.3042)	(0.4798)	(0.6102)	(0.7085)	(0.7842)	(0.8435)	(0.8909)
10.4	-0.0078	0.3083	0.4790	0.5870	0.6596	0.7103	0.7470	0.7743
	(0.0565)	(0.1166)	(0.2452)	(0.3566)	(0.4446)	(0.5123)	(0.5644)	(0.6050)
11.1	-0.0065	0.2857	0.4307	0.5210	0.5838	0.6304	0.6663	0.6947
	(0.0272)	(0.1173)	(0.2165)	(0.2983)	(0.3637)	(0.4166)	(0.4600)	(0.4961)
12.6	0.0065	0.2711	0.4051	0.4812	0.5309	0.5657	0.5910	0.6099

Continued on next page

Table 1 – continued from previous page

	b_n							
t	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	(0.0141)	(0.0875)	(0.1796)	(0.2461)	(0.2953)	(0.3327)	(0.3615)	(0.3839)
15	-0.2251	-0.0113	0.1985	0.2833	0.3422	0.3863	0.4203	0.4468
	(0.0588)	(0.0145)	(0.0548)	(0.0954)	(0.1318)	(0.1637)	(0.1909)	(0.2136)
16.3	-0.2858	-0.0030	0.1477	0.2350	0.2945	0.3379	0.3705	0.3954
	(0.0894)	(0.0245)	(0.0365)	(0.0695)	(0.1004)	(0.1275)	(.1504)	(0.1692)
19.3	-0.4745	-0.1808	-0.0264	0.0694	0.1372	0.1882	0.2275	0.2581
	(0.2357)	(0.0494)	(0.0199)	(0.0238)	(0.0368)	(0.0524)	(0.0677)	(0.0815)
22.6	-0.6330	-0.3336	-0.1726	-0.0706	0.0026	0.0586	0.1026	0.1379
	(0.4080)	(0.1230)	(0.0432)	(0.0181)	(0.0123)	(0.0150)	(0.0214)	(0.0292)
24.8	-0.6913	-0.3843	-0.2147	-0.1062	-0.0288	0.0297	0.0749	0.1106
	(0.4870)	(0.1605)	(0.0605)	(0.0253)	(0.0140)	(0.0130)	(0.0169)	(0.0228)
31.5	-0.8917	-0.5765	-0.3966	-0.2781	-0.1914	-0.1244	-0.0711	-0.0278
	(0.8070)	(0.3483)	(0.1754)	(0.0955)	(0.0544)	(0.0329)	(0.0223)	(0.0181)
38.1	-0.9875	-0.6661	-0.4785	-0.35211	-0.2582	-0.1845	-0.1251	-0.0761
	(0.9878)	(0.4614)	(0.2498)	(0.1461)	(0.0894)	(0.0575)	(0.0399)	(0.0311)
53	-0.9854	-0.7393	-0.5461	-0.4143	-0.3154	-0.2372	-0.1736	-0.1206
	(1.0043)	(0.5629)	(0.3182)	(0.1934)	(0.1226)	(0.0807)	(0.0561)	(0.0420)

Table 2: Bootstrap bias and mean-squared error estimates of $\overline{H}_n^*(f;t)$ for a real data set

	b_n								
t	1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.76	-0.1701	0.5518	0.6866	0.7482	0.9105	1.0117	1.0768	1.1193	1.1472
	(0.0380)	(0.3064)	(0.7417)	(0.9092)	(1.0235)	(1.1234)	(1.2130)	(1.2799)	(1.3299)
0.82	0.0153	0.1799	0.4935	0.7639	0.9252	1.0277	1.0948	1.1393	1.1691
	(0.0094)	(0.8076)	(0.8673)	(0.9611)	(1.0705)	(1.1734)	(1.2606)	(1.3303)	(1.3834)
2.10	-0.1508	0.3183	0.6957	0.9149	1.0504	1.1371	1.1929	1.2281	1.2495
	(0.0587)	(0.2343)	(0.6029)	(0.9263)	(1.1642)	(1.3325)	(1.4478)	(1.5240)	(1.5714)
3.15	-0.3780	0.3057	0.6782	0.7428	0.8810	0.9716	1.0294	1.0636	1.0818
	(0.1944)	(0.1459)	(0.5032)	(0.6026)	(0.8046)	(0.9611)	(1.0724)	(1.1425)	(1.1809)
3.85	-0.4115	0.2665	0.5854	0.6760	0.7807	0.8492	0.8935	0.9212	0.9383
	(0.2279)	(0.1208)	(0.3786)	(0.4939)	(0.6329)	(0.7387)	(0.8137)	(0.8632)	(0.8942)
4.12	-0.3575	0.2448	0.5440	0.6368	0.7329	0.7951	0.8363	0.8639	0.8823
	(0.1871)	(0.1041)	(0.3265)	(0.4451)	(0.5641)	(0.6531)	(0.7176)	(0.7629)	(0.7939)
4.91	-0.4435	0.1661	0.4329	0.5530	0.6176	0.6643	0.6998	0.7268	0.7471
	(0.1871)	(0.1041)	(0.3265)	(0.4451)	(0.5641)	(0.6531)	(0.7176)	(0.7629)	(0.7939)
5.04	-0.4401	0.1309	0.4012	0.5354	0.6015	0.6483	0.6834	0.7098	0.7297
	(0.2643)	(0.0599)	(0.1849)	(0.3236)	(0.3938)	(0.4481)	(0.4911)	(0.5248)	(0.5510)
5.22	-0.3978	0.1180	0.3757	0.5163	0.5832	0.6294	0.6633	0.6885	0.7076
	(0.2091)	(0.0543)	(0.1663)	(0.3041)	(0.3720)	(0.4236)	(0.4638)	(0.4949)	(0.5191)
6.46	-0.5793	-0.0146	0.2737	0.4315	0.4953	0.5332	0.5586	0.5769	0.5906
	(0.3794)	(0.0313)	(0.0933)	(0.2081)	(0.2655)	(0.3036)	(0.3303)	(0.3499)	(0.3647)
6.78	-0.5326	-0.0154	0.2541	0.4052	0.4691	0.5075	0.5331	0.5515	0.5652
	(0.3303)	(0.0314)	(0.0840)	(0.1863)	(0.2398)	(0.2759)	(0.3015)	(0.3203)	(0.3345)
7.75	-0.6148	-0.0811	0.1868	0.3455	0.4048	0.4407	0.4651	0.4827	0.4960

Continued on next page

Table 2 – continued from previous page

	b_n								
t	1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	(0.4281)	(0.0459)	(0.0635)	(0.1332)	(0.1714)	(0.2076)	(0.2293)	(0.2454)	(0.2578)
8.84	-0.7364	-0.1927	0.0787	0.2167	0.3184	0.3555	0.3814	0.4007	0.4157
	(0.6019)	(0.0890)	(0.0472)	(0.0784)	(0.1133)	(0.1382)	(0.1569)	(0.1713)	(0.1828)
11.31	-1.0193	-0.4619	-0.1781	-0.0260	0.0631	0.1204	0.1813	0.2117	0.2357
	(1.0966)	(0.2645)	(0.0733)	(0.0337)	(0.0298)	(0.0342)	(0.0495)	(0.0600)	(0.0691)
14.46	-1.1771	-0.5875	-0.2935	-0.1389	-0.0455	-0.0539	-0.0042	0.0334	0.0633
	(1.4036)	(0.3743)	(0.1146)	(0.0454)	(0.0274)	(0.0232)	(0.0175)	(0.0184)	(0.0176)
18.24	-1.4618	-0.9525	-0.6431	-0.4592	-0.3372	-0.2484	-0.1924	-0.1360	-0.0912
	(1.1969)	(0.9558)	(0.4599)	(0.2542)	(0.1547)	(0.1006)	(0.0518)	(0.0338)	(0.0238)
18.27	-1.5249	-0.9427	-0.6385	-0.4625	-0.3477	-0.2650	-0.1962	-0.1391	-0.0938
	(1.3631)	(0.9211)	(0.4344)	(0.2352)	(0.1386)	(0.0855)	(0.0532)	(0.0346)	(0.0243)
22.48	-1.7304	-0.1148	-0.7847	-0.5811	-0.4389	-0.3332	-0.2742	-0.2029	-0.1472
	(1.4295)	(1.2907)	(0.6567)	(0.3727)	(0.2242)	(0.1411)	(0.0945)	(0.0633)	(0.0460)
23.85	-1.8350	-1.2003	-0.8546	-0.6458	-0.5055	-0.4019	-0.3268	-0.2586	-0.2026
	(1.4978)	(1.3641)	(0.7509)	(0.4353)	(0.2732)	(0.1799)	(0.1223)	(0.0857)	(0.0628)
30.77	-2.6256	-1.6552	-1.2844	-1.0483	-0.8803	-0.7495	-0.6675	-0.5798	-0.5052
	(1.5625)	(1.4242)	(1.2883)	(1.1148)	(1.0687)	(0.8368)	(0.6163)	(0.4981)	(0.4095)

Table 3: Ratios of the mean-squared error of $\overline{H}_n(f;t)$ to that of $\overline{H}^*(f;t)$

		t								
b_n	Sample	4	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8
0.2	n=50	0.6129	0.6451	0.6712	0.6948	0.7161	0.7340	0.7478	0.7592	0.7723
	n=100	0.2785	0.3182	0.3662	0.4235	0.4874	0.5532	0.6165	0.6748	0.7261
0.3	n=50	0.2257	0.2407	0.2560	0.2733	0.2938	0.3189	0.3494	0.3855	0.4271
	n=100	0.1598	0.1805	0.2045	0.2319	0.2629	0.2970	0.3336	0.3716	0.4095
0.4	n=50	0.1528	0.1721	0.1892	0.2061	0.2254	0.2481	0.2748	0.3058	0.3413
	n=100	0.1295	0.1453	0.1633	0.1838	0.2072	0.2335	0.2628	0.2952	0.3306
0.5	n=50	0.1556	0.1724	0.1905	0.2104	0.2323	0.2567	0.2842	0.3150	0.3494
	n=100	0.1497	0.1669	0.1860	0.2073	0.2309	0.2571	0.2858	0.3173	0.3514

Acknowledgements: The author is grateful to Dr. E. I. Abdul Sathar, Assistant Professor, Department of Statistics, University of Kerala for many of his valuable suggestions during this research. The author also wishes to express sincere thanks to Dr. B. V. Rao, Chennai Mathematical Institute and Dr. M. B. Rajarshi, University of Pune for their valuable comments for the advancement of this research. Finally, the financial support from University of Kerala is also gratefully acknowledged.

References

Asadi, M. and Ebrahimi, N. (2000). Residual entropy and its characterizations in terms of hazard function and mean residual life function, Statistics & Probability Letters 49, 263-269.

Ansell, R. O. and Ansell, J. I. (1987). Modeling the reliability of sodium sulphur cells, Reliability Engineering 17, 127-137.

Belzunce, F., Navarro, J., Ruiz, J. M. and Aguila, Y. (2004). Some results on residual entropy function, Metrica 59, 147-161.

Cai, Z. (1998). Kernel density and hazard rate estimation for censored dependent data, Journal of Multivariate Analysis 67, 23-34.

Di Crescenzo, A. and Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions, Journal of Applied Probability 39, 434-440.

Table 4: Ratios of the mean-squared error of $\overline{H}_n^*(f;t)$ to that of $\overline{H}_*^n(f;t)$

		t								
b_n	Sample	2	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8
0.3	n=50	0.1558	0.2402	0.3572	0.5641	0.8525	1.3470	1.9956	1.9176	1.4388
	n=100	0.0596	0.1032	0.1847	0.3232	0.5891	1.4129	2.7669	1.9731	1.0698
0.4	n=50	0.0835	0.1386	0.2176	0.3516	0.5366	0.8639	1.2982	1.1960	0.8841
	n=100	0.0211	0.0424	0.0857	0.1646	0.3136	0.7762	1.4064	1.0303	0.6066
0.5	n=50	0.0367	0.0635	0.1044	0.1734	0.2670	0.4399	0.6886	0.6516	0.5071
	n=100	0.0467	0.0418	0.0586	0.0732	0.1390	0.3537	0.6653	0.5349	0.3381
0.6	n=50	0.0236	0.0309	0.0459	0.0739	0.1127	0.1890	0.3081	0.2913	0.2607
	n=100	0.0191	0.0251	0.0309	0.0651	0.0903	0.1907	0.3423	0.2924	0.1912

Di Crescenzo, A. and Longobardi, M. (2004.) A measure of discrimination between past lifetime distributions, Statistics & Probability Letters 67, 173-182.

Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution, Sankhya A 58, 48-56.

Ebrahimi, N. and Kirmani, S. N. U. A. (1996). Some results on ordering of survival function through uncertainty, Statistics & Probability Letters 29, 167-176.

Ebrahimi, N. and Pellery, F. (1995). New partial ordering of survival functions based on the notion of uncertainty, Journal of Applied Probability 32, 202-211.

Efron, B. (1981). Censored data and the bootstrap, Journal of American Statistical Association 76, 312-319.

Gupta, R. D. and Nanda, A. K. (2002). α- and β-entropies and relative entropies of distributions, Journal of Statistical Theory and Applications 1 (3), 177-190.

Gupta, R. C. (2009). Some characterization results based on residual entropy function, Journal of Statistical Theory and Applications 8 (1), 45-59.

Nair, K. R. M. and Rajesh, G. (1998). Characterization of probability distributions using the residual entropy function, Journal of Indian Satistical Association 36, 157-166.

Nanda, A. K. and Paul, P. (2006). Some properties of past entropy and their applications, Metrika 64, 47-61.

Parzen, E. (1962). On estimation of a probability density function and mode, The Annals of Mathematical Statistics 33(3), 1065-1076.

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.

Rosenblatt, M. (1970). Density estimates and Markov sequences, Nonparametric Techniques in Statistical Inference, ed. M. Puri, Oxford Cambridge University Press, 199-210.

Shannon, C. E. (1948). A mathematical theory of communication, Bell System Technical Journal, 279-423 and 623-656.

Zimmer, W. I., Keats, J. B. and Wang, F. K. (1998). The Burr xii distribution in reliability analysis, Journal of Quality Technology 30, 386-394.