Statistics and Probability Letters 90 (2014) 129–135

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Non-parametric estimation of the generalized past entropy function with censored dependent data

R. Maya [a], E.I. Abdul-Sathar [a,*], G. Rajesh [b]

[a] Department of Statistics, University of Kerala, Thiruvananthapuram - 695 581, India

[b] Department of Statistics, DB Parumala College, Pampa - 689 626, India

ARTICLE INFO

Article history:
Received 23 August 2013
Received in revised form 18 March 2014
Accepted 18 March 2014
Available online 28 March 2014

Keywords:
Generalized past entropy function
Past entropy function
Residual entropy function
Kernel estimate
α-mixing
Residual life

ABSTRACT

The generalized past entropy function introduced by Gupta and Nanda (2002) is viewed as a dynamic measure of uncertainty in past life. This measure finds applications in modeling past life time data. In the present work we provide non-parametric kernel-type estimator for the generalized past entropy function based on censored data. Asymptotic properties of the estimator are established under suitable regularity conditions. Simulation studies are carried out using the Monte Carlo method.

1. Introduction

Shannon's entropy (see, Shannon, 1948) has been extensively used as a useful tool for measuring the uncertainty associated with a random phenomenon. For a continuous non-negative random variable X admitting an absolutely continuous distribution with survival function $\overline{F}(x)$ and the probability density function $f(x)$, Shannon's entropy (also differential entropy) is defined as

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

Ebrahimi and Pellerey (1995) and Ebrahimi (1996) defined the residual entropy function as Shannon's entropy associated with the random variable $(X - t)$ conditional on $X > t(\geq 0)$ and is given by

$$H(f; t) = -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \log \frac{f(x)}{\overline{F}(t)}\, dx. \tag{1.2}$$

Belzunce et al. (2001) proposed the kernel-type estimation of the residual entropy function in the case of independent complete data sets and Belzunce et al. (2004) has established that $H(f; t)$ determines the distribution uniquely if $H(f; t)$ is increasing in t.

It is reasonable to imagine 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

* Corresponding author. Tel.: +91 9495366416.
E-mail address: sathare@gmail.com (E.I. Abdul-Sathar).

http://dx.doi.org/10.1016/j.spl.2014.03.012
0167-7152/

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 over $(0, t)$. They have discussed the necessity of the past entropy, its relation with residual entropy and many interesting results. 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

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

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

It can also be used in Forensic Sciences where the exact time of failure (death in the case of human being) is important when at some time t the unit is found to be in the failure state. This measure also deserves applications in Actuarial Science (see, Sachlas and Papaioannou, 2014). Di Crescenzo and Longobardi (2004) studied a measure of discrimination based on past entropy. 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.

Gupta and Nanda (2002) defined the generalized past entropy function of order β for the past life time as first kind past entropy of order β and second kind past entropy of order β and is given respectively as

$$\overline{H}^{*\beta}(f;t) = \frac{1}{(\beta-1)} \left\{ 1 - \int_0^t \left(\frac{f(x)}{F(t)} \right)^{\beta} dx \right\}, \quad \text{for } \beta \neq 1,\ \beta > 0 \tag{1.4}$$

and

$$\overline{H}^{\beta}(f;t) = \frac{1}{(1-\beta)} \left\{ \log \int_0^t \left(\frac{f(x)}{F(t)} \right)^{\beta} dx \right\}, \quad \text{for } \beta \neq 1,\ \beta > 0. \tag{1.5}$$

Note that, as $\beta \to 1$, (1.4) and (1.5) reduce to $\overline{H}(f;t)$. Again, as $\beta \to 1$ and $t \to \infty$, (1.4) and (1.5) reduce to $H(X)$.

The relation between $\overline{H}^{*\beta}(f;t)$ and $\overline{H}^{\beta}(f;t)$ is given by

$$\overline{H}^{*\beta}(f;t) = \frac{1}{(\beta-1)} \left(1 - \exp\left\{ (1-\beta)\overline{H}^{\beta}(f;t) \right\} \right).$$

It is clear that for practical purposes, we need to develop some inference techniques about this measure. In the present work, we provide the non-parametric estimator for $\overline{H}^{\beta}(f;t)$ defined in (1.5) using the kernel-type estimation for censored data. We consider only situations where the data under study are dependent, i.e., the underlying lifetimes are assumed to be α-mixing (see, Rosenblatt, 1956) and whose definition is given below.

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

$$\alpha(n) = \sup_{k \geq 1} \{ |P(A \cap B) - P(A)P(B)|;\ A \epsilon \mathfrak{F}_1^k,\ B \epsilon \mathfrak{F}_{k+n}^{\infty} \}, \tag{1.6}$$

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 estimator for $\overline{H}^{\beta}(f;t)$ given in (1.5) using censored samples and examine the asymptotic properties of the estimator. In Section 3, a simulation study to illustrate the behavior of the estimator is undertaken.

2. Estimation

In reliability and life testing, due to time constraints or cost consideration the experimenter is forced to terminate the experiment after a specific period of time or after a 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.