Computational Statistics and Data Analysis 88 (2015) 1–14

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Local linear estimation of residual entropy function of conditional distributions

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

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

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

ARTICLE INFO

Article history:
Received 17 October 2013
Received in revised form 2 February 2015
Accepted 3 February 2015
Available online 12 February 2015

Keywords:
Residual life
Local linear estimate
ρ-mixing
Residual entropy function
Conditional residual entropy function

ABSTRACT

Local linear estimators for the conditional residual entropy function in the case of complete and censored samples are proposed. The resulting estimators are shown to be consistent and asymptotically normally distributed under certain regularity conditions. The performance of the estimator is compared by using a real data set and simulation studies are carried out by using the Monte-Carlo method.

1. Introduction

The concept of differential entropy was introduced by Shannon (1948). Entropy is a useful measure of uncertainty and dispersion and has been widely used in many applications such as statistical communication theory, quantization theory, statistical decision theory and contingency table analysis. For a non-negative random variable X having an absolutely continuous distribution function F with pdf f, the entropy of the random variable is defined as

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

$H(X)$ measures the uncertainty inherent in the distribution of X.

Identification of joint distribution using conditional densities has been an important problem considered by many authors. This method of finding a bivariate density using the conditionals is called the conditional specification of joint distribution (see, Arnold et al., 1999). These models are very useful in two component systems where the operational status of one component is known. For more recent works on conditionally specified models, we refer to Sunoj and Sankaran (2005), Kotz et al. (2007), Navarro et al. (2014), Navarro and Sarabia (2013) and the references therein.

Joint entropy and conditional entropy are simple extensions of (1.1) that measure the uncertainty in the joint distribution and conditional distribution of a pair of random variables respectively. If (X, Y) is a pair of non-negative random variables with joint distribution function $F(x, y)$ and joint probability density function $f(x, y)$, then the entropy for the conditional distribution of Y given $X = x$ is defined as

$$H(Y|X = x) = -\int_0^\infty f(y|x)\log f(y|x)\,dy, \tag{1.2}$$

where $f(y|x)$ is the conditional probability density function of Y given $X = x$.

* Corresponding author.
E-mail addresses: rajeshg75@yahoo.com (G. Rajesh), sathare@gmail.com (E.I. Abdul-Sathar), mayasekharan@gmail.com (R. Maya).

http://dx.doi.org/10.1016/j.csda.2015.02.002
0167-9473/

Now the conditional entropy is defined as

$$H(Y|X) = \int_0^\infty H(Y|X = x)f(x)\,dx. \tag{1.3}$$

Conditional entropy of Y conditional on X refers to the average entropy of Y conditional on the value of X, averaged over possible values of X. This is different than conditioning on X taking one particular value of X, though we can write the other as well. Additional information never increases entropy.

That is

$$H(Y|X) \leq H(Y).$$

The chain rule for joint entropy states that the total uncertainty about the value of X and Y is equal to the uncertainty about X plus the uncertainty about Y, once you know X. That is

$$H(X, Y) = H(X) + H(Y|X).$$

If we consider X as the lifetime of a new unit, then $H(X)$ can be useful for measuring the associated uncertainty. However, for a used unit $H(X)$ is no longer useful for measuring uncertainty about the remaining lifetime of the unit. Ebrahimi (1996) and Ebrahimi and Pellerey (1995) have introduced the concept of residual entropy function, which is the Shannon's entropy associated with the residual life of the unit. For a non-negative random variable X, representing the lifetime of a unit, the residual entropy function is defined as

$$H(f; t) = -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \log \frac{f(x)}{\overline{F}(t)}\,dx, \tag{1.4}$$

where $\overline{F}(t) = Pr(X > t)$ denotes the survival function. 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), Asadi and Ebrahimi (2000) and the references therein.

Rajesh and Nair (2000) defined the concept of residual entropy function for conditional distributions and called it as conditional residual entropy function. Let (X, Y) be a pair of non-negative random variables admitting an absolutely continuous distribution function $F(x, y)$ with density function $f(x, y)$, survival function $\overline{F}(x, y)$, marginal density of $X, f(x)$ and marginal density of $Y, f(y)$. Then the conditional residual entropy function of Y given $X = x$ as

$$H(y|x) = -\int_y^\infty \frac{f(t|x)}{\overline{F}(y|x)} \log\left(\frac{f(t|x)}{\overline{F}(y|x)}\right) dt. \tag{1.5}$$

The above equation simplifies to

$$H(y|x) = \log \overline{F}(y|x) - \frac{1}{\overline{F}(y|x)} \int_y^\infty f(t|x) \log f(t|x)\,dt, \tag{1.6}$$

where $\overline{F}(y|x)$ is the conditional survival function of Y given $X = x$.

Several proposals have been made to estimate Shannon's entropy given in (1.1). Dmitriev and Tarasenko (1973) and Ahmad and Lin (1976) proposed estimators of the entropy using kernel-type estimators for the density. Vasicek (1976) proposed an entropy estimator based on spacings. A detailed review work for the non-parametric estimators of the Shannon's entropy is given in Beirlant et al. (1997). Non-parametric estimation of reliability measures of independent and identically distributed random variables are studied by several authors. Belzunce et al. (2001) proposed kernel-type estimation of the residual entropy function in the case of independent complete data sets. From a practical point of view, it seems more realistic to drop independence and replace it by some mode of dependence. For example, consider a system consisting of n identical components. When the component size is small, the failure of a single component affects the performance of the remaining items of that system. But when the number of components in a system is large, the effect of failure of an item is small compared to the first case. Maya et al. (2014) proposed non-parametric estimators for the Rényi's residual entropy function of order α under dependence condition.

Rajesh and Nair (2000) studied several properties of conditional residual entropy function. But till now, no work is seen carried out to study the estimation aspects of this function. Hence in this paper, we develop local linear estimators for the conditional residual entropy function using complete and censored data. In both situations, the underlying lifetimes are assumed to be ρ-mixing (see, Kolmogorov and Rozanov, 1960) and whose definition is given below.

Definition 1. Let $\mathfrak{F}_i^k$ be the σ-algebra of events generated by the random variables $\{X_j, Y_j,\ i \leq j \leq k\}$ and $L_2\left(\mathfrak{F}_i^k\right)$ be the collection of all $\mathfrak{F}_i^k$-measurable square integrable random variables. Let

$$\rho(k) = \sup_{U \epsilon L_2\left(\mathfrak{F}_{-\infty}^0\right), V \epsilon L_2\left(\mathfrak{F}_k^\infty\right)} \frac{|\mathrm{cov}(U, V)|}{\mathrm{var}^{\frac{1}{2}}(U)\mathrm{var}^{\frac{1}{2}}(V)} \tag{1.7}$$

denote the ρ-mixing coefficient. The sequence is said to be ρ-mixing, if the mixing coefficient $\rho(k) \to 0$ as $k \to \infty$.