International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2021; 6(6): 01-05

www.mathsjournal.com
Received: 16-07-2021
Accepted: 02-10-2021

Athira NR
Department of Statistics,
Nirmala College, Muvattupuzha,
Kerala, India

ES Jeevanand
Department of Mathematics
Union Christian College, Aluva,
India

Corresponding Author:
Athira NR
Department of Statistics,
Nirmala College, Muvattupuzha,
Kerala, India

Estimation of the residual entropy function of the Finite range distribution using record values

Athira NR and ES Jeevanand

Abstract
Residual entropy has a very significant role in reliability and survival analysis. It is one of the modified form of the Shannon entropy function used to study the uncertainties associated to a non- negative random variable. In this paper, we focus on the Bayesian estimation of the residual entropy function of the Finite range distribution based on record breaking data. The performance of the estimator is evaluated using simulated data sets.

Keywords: Bayes estimation, finite range distribution, record values, residual entropy

Introduction

Entropy, is appropriately associated with lack of information, uncertainty and indefiniteness as one of the most appropriate measure of this probability. Shannon (1948) [5] was the first to introduce entropy, known as Shannon's entropy or Shannon's information measure, into information theory. In the context of information theory, Shannon's entropy plays an important role. Since this entropy is not applicable to a system that has survived for some unit of time, the concept of residual entropy has been developed in Statistics literature. In the reliability context, if X is a random variable representing the life time of a component or a device, a characteristic of special interest in the residual life distribution which is the distribution of the random variable $(X-t)$ truncated at $t(\geq 0)$. A comparison of the residual life distribution and the parent distribution as well as characterization of distributions based on the form of the residual lifetime distributions has received a lot of interest among researchers. The works of Gupta and Gupta (1983), Gupta and Kirmani (1990), Sankaran (1992) focuses attention on this aspect.

Ebrahimi and Pellerey (1995) [6] and Ebrahimi and Kirmani (1996) [7] have used the Shannon's entropy applied to the residual life as a measure of stability of a component. They pointed out that, the measure H does not take into account any information one may have about the current age of the system. Thus, if a unit of life length X is known to have survived to age t, it is the residual entropy of X_t (the remaining lifetime of the system of age $t \geq 0$), rather than that of X, which is relevant. The residual entropy function can advantageously use as a tool at the stage of design and planning in reliability engineering.

For a non-negative random variable X, Ebrahimi (1996) [7] defines the residual entropy function as the Shannon's entropy associated with the random variable $(X-t)$ truncated at $t(\geq 0)$, namely

$$H(f,t) = -\int_{x=t}^{\infty} \frac{f(x)}{\bar{F}(t)} \ln \frac{f(x)}{\bar{F}(t)} dx, \quad \bar{F}(t) \geq 0 \tag{1.1}$$

Where $f(x)$ and $\bar{F}(x)=P(X\geq x)$ are the probability density function and the survival function of X respectively. The residual entropy function can be expressed in terms of the hazard rate through the relation

$$H(f,\ t)=1-\frac{1}{\bar{F}(x)}\int_t^\infty f(x)\ln h(x)\,dx. \tag{1.2}$$

$H(f,\ t)$ Measures the expected uncertainty contained in the conditional density of $(X-t)$ given $X>t$ about the predictability of remaining life time of the component. It may be noticed that $-\infty<H(f,\ t)<\infty$ and that $H(f,\ 0)$ reduces to Shannon's entropy defined over $(0,\infty)$. It is established that $H(f,\ t)$ determines the distribution uniquely. Rajesh (2000) extended this function to the discrete case. He defined the entropy function associated with the discrete random variable X in the support of the set of non-negative integers as

$$H(f,t)=-\sum_{x=t+1}^{\infty}\frac{f(x)}{\bar{F}(t+1)}\ln\frac{f(x)}{\bar{F}(t+1)}$$

$$=\ln\bar{F}(t+1)-\frac{1}{\bar{F}(t+1)}\sum_{x=t+1}^{\infty}f(x)\ln f(x).$$

Also using the relationship $S(x)=\bar{F}(x),\ h(x)=\frac{f(x)}{\bar{F}(x)}$ and

$f(x)=\bar{F}(x)-\bar{F}(x+1),\ H(f,t)$ can be written as

$$H(f,t)=-\ln h(t+1)+\sum_{x=t+1}^{\infty}\frac{\bar{F}(x+1)}{\bar{F}(t+1)}\ln\frac{h(x)}{h(x+1)\left[1-h(x)\right]}. \tag{1.3}$$

After the system or component under consideration has survived up to time t, the entropy function H measures the expected uncertainty contained in the conditional density of $X-t$ given $X\geq t$ about the predictability of the remaining lifetime of the component. The residual entropy function has also been used by Ebrahimi and Kirmani (1996) [7], Asadi and Ebrahimi (2000) [3] and Belzunce *et al*. (2004) [4] to measure ageing and to characterize classify and order lifetime distributions.
The estimation problem entropy function for the different distributions are discussed by Jeevanand, E.S. and Abdul-Sathar E.I. (2009) [9], Mathachan Pathiyil and E.S. Jeevanand (2009a,2009b) [10, 11], E.I.Abdul-Sathar, E.S.Jeevanand and G.Rajesh (2010) [8] Athira, N.R. and E.S.Jeevanand (2021a, 2021b) [1, 2]. Motivated by the relevance, present study aims at estimating the residual entropy function of the finite range distribution using record values.

Residual entropy function of finite range distribution

Let us consider the of the finite range distribution with probability density function given by:

$$f(x;\theta,p)=\frac{p}{\theta^p}(x)^{p-1}p,\theta>0,x>0 \tag{2.1}$$

Where p and θ are the shape and scale parameters of the distribution.

Cumulative distribution function is given by:

$$F(x)=[x/\theta]^p \tag{2.2}$$

The reliability function of the distribution is:

$$R(t)=1-[t/\theta]^p \tag{2.3}$$

The maximum likelihood estimator of p and θ is obtained as

$$\hat{p}=\frac{n}{\log\theta-\sum log x_i}\text{ and }\hat{\theta}=x_{(n)}=\max(x_1,x_2,x_3\ldots\ldots..x_n) \tag{2.4}$$

And then, the residual entropy function associated with the finite range distribution is,

$$H(f,t) = -\int \frac{f(x)}{R(t)} \log \frac{f(x)}{R(t)} dx \;= 1 - (\tfrac{t}{\theta})^p \, p \qquad (2.5)$$

Under the squared loss function and for known θ (since the value of θ is $x_{(n)} = \max(x_1, x_2, x_3 \dots\dots.. x_n)$), the Bayes estimator is estimated as,

$$H^*_G = E(H_G|\underline{x}) \;=\; \int_p H_G f(p|\underline{x}) d_p \qquad (2.6)$$

Cumulative Residual entropy of finite range distribution

The concept of cumulative residual entropy have been proposed by Navarro *et al*. and have established that if H(t) is increasing in t then H(t) determines the distribution uniquely. Even though Shannon's entropy finds applications in many areas of research; recently, Rao *et al*. identified some limitations of the use of Shannon entropy in measuring the randomness of certain systems and introduced an alternative measure of uncertainty called cumulative residual entropy (CRE), through

$$\delta(X) = -\int_0^\infty \bar{F}(x) . log\,(\bar{F}(x)\, dx \qquad (2.7)$$

where $\bar{F}(x)$ is the survival function of the proposed distribution.

Considering the finite range distribution, the corresponding cumulative residual entropy is given by:

$$\delta(X) = -\int_0^\infty (1 - (\tfrac{x}{\theta})^p \; log\,(1 - (\tfrac{x}{\theta})^p \; dx$$

$$= -\frac{3}{4}\left[\frac{1}{(\frac{p}{\theta})(\frac{x}{\theta})^{p-1}}\right] \qquad (2.8)$$

In the following section we obtain the Bayes estimates of using record values from the finite range distribution.

Estimation of the residual entropy and cumulative residual function Using Record Values

Suppose X_1, X_2,...is a sequence of independent and identically distributed random variables with common distribution function F. Then X_j is called a (upper) record value in the sequence if it exceeds all the observations that precede it in the sequence. That is if $X_j > \max(X_1,X_2,...,X_{j-1}), j>1$.

Formally, the nth record in the sequence is denoted by R_n and $R_0 = X_1$ is a trivial or initial record value.

Suppose $\mathbf{R} = (R_0\ R_1\ \dots\ R_n)^T$ is the sequence of the first $(n+1)$ records defined over a random sample from the finite range distribution with pdf (2.1).Then the likelihood function of the record sequence $\mathbf{R}$ is given by

$$l(R|p) = \prod_{i=0}^{n}\left[\frac{f(R_i)}{\prod_{i=0}^{n-1}[1-F(R_i)]}\right], \text{ where } 0 \le R_0 < R_1 < \dots < R_n < \infty.$$

$$= p^{n+1}e^{-pt}, \text{ where } t = \sum \ln\left[\frac{\left(\frac{\theta}{R_i}\right)}{\left(1-\frac{\theta}{R_i}\right)}\right] \qquad (2.9)$$

The maximum likelihood estimator of p and θ is obtained as

$$\hat{p} = \frac{n+1}{\sum \ln\left[\frac{\left(\frac{\theta}{R_i}\right)}{\left(1-\frac{\theta}{R_i}\right)}\right]} \text{ and } \hat{\theta} = x_{(n)} = \max(x_1, x_2, x_3 \dots\dots.. x_n) \quad (2.10)$$

Considering the conjugate prior

$$f(p) = p^m e^{-ps}, p > 0, (m, s) > 0 \qquad (2.12)$$

For p, the posterior density of p is obtained as:

$$f(p|R) = C.p^N e^{-Tp}, \qquad (2.13)$$

where, $C^{-1} = \int_0^\infty p^N e^{-Tp}\, dp$, N = n+m+1 and T=t+ s.

The posterior model is essentially an updated version of our prior knowledge about the sample data. Considering these, the posterior density can be readily derived using the transformation p = $\frac{H}{t-\theta}$. The posterior distribution of H given R is,

$$f(H|R) = C_1.\frac{H}{t-\theta}^N e^{-(T\frac{H}{t-\theta})}, -\infty < H < \infty \quad (2.14)$$

where, $C_1(d) = \int_0^\infty \frac{H}{t-\theta}^N e^{-T\frac{H}{t-\theta}} dp,$ (2.15)

Under squared error loss function, the Bayes estimator of the residual entropy function H defined in (2.1) as

$$H_S = \frac{C_1(1)}{C_1(0)} \quad (2.16)$$

Similarly replacing p in terms of δ by the transformation $p = \frac{1}{1+t}[\frac{3\theta}{4\delta} + t]$, we obtain the posterior density function of δ as

$$f(\delta|R) = C_2 . \frac{1}{1+t}[\frac{3\theta}{4\delta} + t]\, e^{-(T\frac{1}{1+t}[\frac{3\theta}{4\delta}+t])}, -\infty < H < \infty \quad (2.17)$$

where, $C_2(d) = \int_0^\infty \frac{1}{1+t}[\frac{3\theta}{4\delta} + t]\, e^{-(T\frac{1}{1+t}[\frac{3\theta}{4\delta}+t])} dp$ (2.18)

Under squared error loss function, the Bayes estimator of the residual entropy function H defined in (2.1) as

$$H_C = \frac{C_2(1)}{C_2(0)} \quad (2.19)$$

To evaluate 2.15 and 2.18 we seek numerical integration.

Simulation study and discussion

We study the performance of the estimators obtained so far using simulated data. The simulation encompasses different parameter values of the finite range distribution under consideration. The simulated absolute bias (Bias) and (MSE) of the estimators proposed for estimating the residual entropy function of the finite range distribution for 1,000 replications are computed. Then, from these, samples having 6 or more record values (ie. having at least $R_0, R_1, R_2, R_3, R_4, R_5$) are sorted out. Using these short listed sets of records, thus obtained, the estimators of the respective entropy measures are computed. Table 1 gives the estimated values of H_S and H_C using the record values R_1, R_3 and R_5 and for different choices of θ =0.05,0.1,0.3,0.5 and p=0.1,0.2,0.3,0.4,0.5 based on simulated samples from the resulting geometric populations. From the this table, we can conclude that

1. The bias and MSE reduces as the record value increases
2. The bias and MSE reduces as the parameter value increases

Table 1: The estimated values of H_S using the record values R_1, R_3 and R_5

p	θ	H_S	Estimated	Record used for estimation		
				R_1	R_2	R_3
0.1	0.05	3.5640	Mean	3.8147	3.7013	3.5369
			Bias	0.2559	0.2085	0.1556
			MSE	0.1659	0.1033	0.07796
0.2	0.1	3.2456	Mean	3.1250	3.1544	3.1564
			Bias	0.1154	0.1043	0.0746
			MSE	0.0588	0.0489	0.0432
0.3	0.3	2.0456	Mean	1.9231	1.9349	2.0857
			Bias	0.1131	0.1015	0.0396
			MSE	0.0533	0.0485	0.0417
0.4	0.5	1.3845	Mean	1.3524	1.2596	1.3562
			Bias	0.1255	0.1095	0.0681
			MSE	0.211	0.033	0.0307
0.5	0.6	1.1712	Mean	0.8582	1.0672	1.0989
			Bias	0.1643	0.0546	0.0118
			MSE	0.0418	0.0256	0.0221

Table 2: The estimated values of $\mathbf{H_c}$ using the record values R_1, R_3 and R_5

p	θ	H_c	Estimated	Record used for estimation		
				R_1	R_2	R_3
0.1	0.05	2.542	Mean	3.4331	3.4008	3.5512
			Bias	0.2567	0.2134	0.1412
			MSE	0.1765	0.1098	0.7890

0.2	0.1	2.245	Mean	3.1520	3.1324	3.1464
			Bias	0.1254	0.1233	0.0656
			MSE	0.0548	0.0429	0.0402
0.3	0.3	2.009	Mean	2.2826	1.2996	2.3462
			Bias	0.1165	0.1295	0.0781
			MSE	0.2511	0.0453	0.0507
0.4	0.5	1.567	Mean	1.7331	1.5439	2.1237
			Bias	0.0231	0.1455	0.0456
			MSE	0.0673	0.0875	0.0617
0.5	0.6	2.032	Mean	0.7532	1.0452	1.0787
			Bias	0.1533	0.0346	0.0238
			MSE	0.0418	0.0346	0.0251

References

1. Athira NR, Jeevanand ES. On residual entropy function of finite range distribution with type II censoring, Far East Journal of theoretical Statistics 2021a;61(1),61-74. ISSN:0972-0863.
2. Athira NR, Jeevanand ES. Bayesian Analysis of residual Entropy function of finite Range distribution under different priors, Journal of Information Storage and Processing Systems 2021b;20(1):141-153, ISSN1099-8047
3. Asadi M, Ebrahimi N. Residual entropy and its characteriza-tions in terms of hazard function and mean residual life function. Sta- tist. Prob. Lett. 2000;49:263-269.
4. Belzunce F, Navarro J, Ruiz JM, del Aguila Y. Some results on residual entropy function. Metrika 2004;59(2):147-161.
5. Shannon CE, Weaver W. *The Mathematical Theory of Communication* (University of Illinois Press, Urbana) 1948.
6. Ebrahimi N, Kirmani. Some results on ordering of survival functions through uncertainty, Statistics and Probability Letters 1995;29(2):167-176.
7. Ebrahimi N. How to measure uncertainty in the residual life distributions. Sankhya Ser. 1996A;58:48-56.
8. Abdul-Sathar EI, Jeevanand ES, Rajesh G. Bayes Estimation of residual entropy function from classical Pareto model containing spurious observations, IAPQR Transaction 2010;35(1):15-23.
9. Jeevanand ES, Abdul-Sathar EI. Estimation of residual entropy function for exponential distribution from a censored sample, ProbStat Froum 2009;2:68-77.
10. Mathachan Pathiyil, Jeevanand ES. Estimation of the Residual Entropy Function of the Geometric Distribution Using Record Values, in the Proceedings of the National Workshop on Bayesian Statistics and MCMC Methods using BUGS and R, St. Thomas College, Pala, 2009a, 26-30.
11. Mathachan Pathiyil, Jeevanand ES. Estimation of the Residual Entropy Function of the Geometric Distribution when Outliers are Present in the Sample, Proceedings of the UGC Sponsored Seminar on National Seminar on Modern Applied Statistical Methods in Statistical Softwares, Department of Statistics Nirmala College, Muvattupuzha 2009b, 37-48.