ISSN 0974-3235

ProbStat Forum, Volume 02, July 2009, Pages 68-77

Estimation of residual entropy function for exponential distribution from censored samples

E.S.Jeevanand
Department of Mathematics and Statistics
Union Christian College, Aluva-683 102, India.
e-mail: *radhajeevanand@yahoo.co.in*

E.I.Abdul-Sathar
Department of Statistics
University of Kerala, Thiruvananthapuram - 695 581, India.
e-mail: *sathare@gmail.com*

Paper received on 14 October 2008; revised, 24 April 2009; accepted, 01 May 2009.

Abstract. In this paper, we obtain estimators for the residual entropy function of exponential distribution under two censoring schemes namely progressive and double censoring respectively using Bayes and Maximum likelihood estimation procedures. We compared the performance of the estimators using a simulation study.

Key Words: Progressive censoring, Double censoring, Bayes estimator, Maximum likelihood estimator, Squared error loss function, LINEX loss function, Residual entropy.

1 Introduction

In the recent past, many researchers have taken 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 non-negative random variable having an absolutely continuous distribution function F with probability density function f, then the Shannon's entropy associated with the random variable X is defined as

$$H(X) = H(f) = -\int_0^\infty f(x)\log f(x)dx. \tag{1}$$

The entropy measures the uniformity of a distribution. As $H(f)$ increases, $f(x)$ approaches to uniform. Consequently, the concentration of probabilities decreases and it becomes more difficult to predict an outcome of a draw from $f(x)$. In fact, a very sharply peaked distribution has a very low entropy, whereas if the probability is spread out the entropy is much higher. In this sense $H(X)$ is a measure of uncertainty associated with f.

Since this entropy is not applicable to a system that has survived for some units of time, the concept of residual entropy has been developed in the literature. Ebrahimi and Pellerey (1995) and Ebrahimi (1996) introduced the concept of residual entropy in terms of conditional Shannon's 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 $X - t$, truncated at $t \geq 0$, and is defined as

$$\begin{aligned} H(f,t) &= -\int_t^{\infty} \frac{f(x)}{\overline{F}(t)} \log \frac{f(x)}{\overline{F}(t)} dx \\ &= 1 - \frac{1}{\overline{F}(t)} \int_t^{\infty} f(x) \log h(x) dx, \overline{F}(t) > 0 \end{aligned} \tag{2}$$

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 , respectively. Navaro et al (2004) has 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 about its remaining life. That is, $H(f,t)$ measures concentration of conditional probabilities. 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 clear that for practical purposes, we need to develop some inference techniques about this measure. This research area has been initiated by Ebrahimi(1997), with the proposal of a test for exponentiality against DURL (IURL) alternatives. Belzunce et.al. (2001) proposed nonparametric estimators of the residual entropy using smoothing techniques.

The aim of this paper is to study estimation of residual entropy $H(f,t)$ of exponential distribution under two censoring schemes namely progressive and double censoring respectively using Bayes and Maximum likelihood estimation procedures. Bayes estimators have been developed under squared error loss function as well as under LINEX loss functions. In the next Section, we consider the estimation under progressively type-II sample. First, we discuss the Bayes estimation. Then the maximum likelihood estimator is obtained. In Section 3, the Bayes estimator and maximum likelihood estimator under double censoring scheme are discussed. In Section 4, the performance of the estimators are compared using a simulation study.

2 Estimation of H under Progressively Type-II sample

In this section we obtain the Bayes estimate and MLE of $H(f,t)$ under progressively type II censored scheme in which n units are placed on a test at time zero, with m failures to be observed. When the first failure is observed, r_1 of the surviving units are randomly selected and removed. At the second observed failure, r_2 of the surviving units are randomly selected and removed. This experiment stops at the time when the m^{th} failure is observed and the remaining $r_m = n - r_1 - r_2 - \; - r_{m-1} - m$ surviving units all are removed. Let $X_1, ..., X_m$ be a progressively type-II censored sample from the exponential distribution with density function

$$f(x, \lambda) = \lambda e^{-\lambda x}, x > 0, \lambda > 0. \tag{3}$$

For the model (3), residual entropy simplifies to

$$H(f, t) = 1 - \log \lambda, \lambda > 0. \tag{4}$$

The likelihood function corresponding to this set-up can be written as

$$l(x|\theta) = k \prod_{i=1}^{m} f(x_{i:m:n}) [1 - F(x_{i:m:n})]^{r_i}$$

where

$$k = n(n-1-r_1)(n-2-r_1-r_2)...(n-m+1-r_1-...-r_{m-1}).$$

Substituting (3) in the above equation, the likelihood function gives

$$l(x|\theta) = k\lambda^m \exp\left\{-\lambda \left(\sum_{i=1}^{m} x_i (1 + r_i)\right)\right\}. \tag{5}$$

2.1 Bayes estimator

In the Bayesian approach to statistical inference the posterior distribution summarizes the information about a parameter. This distribution depends on a probability model and a prior distribution and is conditional on the observed data.
The likelihood function provides a conjugate prior for λ namely

$$g(\lambda) = C_1 \lambda^{p-1} \exp(-\tau\lambda), p, \tau, \lambda > 0. \tag{6}$$

The symbol C with various suffixes stands for the normalizing constants. Combining (5) and (6) and using Bayes theorem, the posterior density turns out to be

$$f(\lambda|\underline{x}) = C_2 \lambda^{N-1} \exp(-\lambda T), \lambda \geq 0 \tag{7}$$

where

$$N = m + p, T = \tau + \sum_{i=1}^{m} x_i\,(1 + r_i).$$

Replacing λ in terms of H by that of Eq.(4), we obtain the posterior density function of H as

$$f\,(H|\underline{x}) = [C_3(0)]^{-1} \exp\left[(1 - H)\,N - T \exp\left(1 - H\right)\right], -\infty < H < \infty \tag{8}$$

where

$$C_3(d) = \int_{-\infty}^{\infty} H^d \exp\left[(1 - H)\,N - T \exp\left(1 - H\right)\right] dH. \tag{9}$$

From a decision-theoretic viewpoint, in order to select a single value as representing our best estimator of H, a loss function must be specified. Under squared-error loss, the Bayes estimator of residual entropy is the mean of the posterior density given by

$$\overline{\mathrm{H}}_{bsp} = 1 + \log T - \mathit{Polygamma}\,[0, \mathrm{N}] \tag{10}$$

and the posterior risk (minimum posterior expected loss) of $\overline{\mathrm{H}}_{bsp}$ is the posterior variance given by

$$\mathrm{Var}\left(\overline{\mathrm{H}}_{bsp}\right) = \mathit{Polygamma}\,[1, \mathrm{N}]\,. \tag{11}$$

$\mathit{Polygmma}\,[\mathrm{n}, \mathrm{x}]$ gives the n^{th} derivative of the digamma function

$$\Psi\,(x) = \frac{d \log \Gamma\,(x)}{dx}, x > 0.$$

Another loss function in popular use is the LINEX loss function(LLF) introduced by Varian (1975). The LINEX loss function may be expressed as

$$L(\delta) \propto e^{c\delta} - c\delta - 1, c \neq 0, \tag{12}$$

where $\delta = \hat{\beta} - \beta$. The sign and magnitude of c reflects the direction and degree of asymmetry, respectively. The Bayes estimator for β relative to LINEX loss function, denoted by $\hat{\beta}_L$, is given by

$$\hat{\beta}_L = -\frac{1}{c} \ln E_\beta \left(e^{-c\beta}\right), \tag{13}$$

provided that $E_\beta\left(e^{-c\beta}\right)$ exists and is finite, where E_β denotes the expected value. Under LINEX loss function, the Bayes estimate of H using (13) is

$$\hat{H}_{bxp} = \frac{1}{a} \ln G_1 \tag{14}$$

where

$$G_1 = [C_3(0)]^{-1} \int_{-\infty}^{\infty} \exp\left[aH + (1 - H)\,N - T \exp\left(1 - H\right)\right] dH. \tag{15}$$

To evaluate (9) and (15) we seek numerical integration.

2.2 Maximum likelihood estimator

The likelihood function of H in this set-up can be expressed as

$$l\left(x|H\right) = C_4 \exp\left(m\left(1-H\right) - \sum_{i=1}^{m} x_i\left(1+r_i\right)\exp\left(1-H\right)\right).$$

We derive the MLE, using the usual procedure and is given by

$$\overline{H}_{mlp} = 1 - \log\left(\frac{m}{\sum\limits_{i=1}^{m} x_i\left(1+r_i\right)}\right). \tag{16}$$

It may be noted that $\overline{H}_{mlp}$ is consistent and asymptotically normal with

$$Var\left(\overline{H}_{mlp}\right) = \frac{1}{m}. \tag{17}$$

3 Estimation of H under double censored sample

In this section, we obtain the Bayes estimate and MLE of $H(f,t)$ under double censored scheme in which a certain number (or a proportion) of observations are censored on the left or right or both. For example, in early childhood learning centres, interest often focuses upon testing children to determine when a child learns to accomplish certain specified tasks. The age at which a child learns the task would be considered the time-to-event. Often, some children can already perform the task when they start their study. Such event times are considered left censored. Some children undergoing testing, may not learn the task during the entire study period, in which case such event times would be right censored. Thus, the sample would also be doubly censored.

Consider a doubly censored sample $y_{r+1}, y_{r+2}, ..., y_{n-s}$ with r observations censored on the left and s observations censored on the right, where $r = [nq_1]+1$, and $s = [nq_2]+1$, from an Exponential distribution with density given by (3). The likelihood of the sample is given by

$$l\left(x|\theta\right) = \left(F(Y_{r+1})\right)^r\left(1-F(Y_{n-s})\right)^s \prod_{i=r+1}^{n-s} f\left(Y_i\right).$$

Substituting (3) in the above equation, the likelihood function gives

$$l\left(x|\lambda\right) = \left(1-\exp\left(-\lambda y_{r+1}\right)\right)^r \exp\left(-\lambda\left(\sum_{i=r+1}^{n-s} y_i + s y_{n-s}\right)\right)\lambda^{n-r-s}. \tag{18}$$

3.1 Bayes estimator

With a gamma prior as in (6) and using (18), the posterior density of λ can be obtained as

$$f(\lambda|\underline{x}) = C_5 \left(1 - \exp\left(-\lambda y_{r+1}\right)\right)^r \exp\left(-\lambda \mathrm{Z}\right) \lambda^{N-1}, \lambda \geq 0 \tag{19}$$

where

$$N = p + n - r - s, Z = \tau + \sum_{i=r+1}^{n-s} y_i + s y_{n-s}.$$

The posterior distribution of H is derived as

$$f(H|\underline{x}) = [C_6(0)]^{-1} \left(1 - e^{-y_{r+1}\left(e^{1-H}\right)}\right)^r e^{\left(N(1-H) - \mathrm{Z}\left(\mathrm{e}^{1-\mathrm{H}}\right)\right)}, -\infty < H < -\infty \tag{20}$$

where

$$C_6(d) = \int_{-\infty}^{\infty} H^d \left(1 - e^{-y_{r+1}\left(e^{1-H}\right)}\right)^r e^{\left(N(1-H) - \mathrm{Z}\left(\mathrm{e}^{1-\mathrm{H}}\right)\right)} dH. \tag{21}$$

Under squared-error loss, the Bayes estimator of H is the mean of the posterior density given by

$$\overline{H}_{bsd} = \frac{C_6(1)}{C_6(0)} \tag{22}$$

with Bayes risk

$$Var\left(\overline{H}_{bsd}\right) = \frac{C_6(2)}{C_6(0)} - \left(\overline{H}_{bsd}\right)^2. \tag{23}$$

Under LINEX loss function, the Bayes estimate of H is

$$\hat{H}_{bxd} = \frac{1}{a} \ln G_2 \tag{24}$$

where

$$G_2 = [C_6(0)]^{-1} \int_{-\infty}^{\infty} \left(1 - e^{-y_{r+1}\left(e^{1-H}\right)}\right)^r e^{\left(aH + N(1-H) - \mathrm{Z}\left(\mathrm{e}^{1-\mathrm{H}}\right)\right)} dH. \tag{25}$$

To evaluate (21) and (25) we seek numerical integration.

3.2 Maximum likelihood estimator

In this set-up, we derive the MLE estimate of λ, denoted by λ_{mld}, by solving the following equation

$$\frac{r y_{r+1} e^{-\lambda y_{r+1}}}{1 - e^{-\lambda y_{r+1}}} + \frac{n - (r+s)}{\lambda} - \mathrm{Z} = 0. \tag{26}$$

The equations can be solved numerically using an iterative procedure and an estimate of the residual entropy function H is

$$\overline{H}_{mld} = 1 - \log\left(\lambda_{mld}\right). \tag{27}$$

4 Simulation results

In this section, we present the results of a simulation study in order to compare the performance of these estimators. In order to assess the performance of the estimators of the residual entropy, we perform a simulation study of 2000 samples of sizes $n = 25, 50, 100$ and 500 generated from (3) for values of $\lambda = 0.2, 0.4, 0.75$ and 1.25. We present the simulation results concerning the mean and mean square errors of all these estimators.

Table 1: Means and MSEs (in parentheses) of the estimates of H under Progressive censoring

		$\lambda = 0.2$	$\lambda = 0.4$	$\lambda = 0.75$	$\lambda = 1.25$
True H		2.60944	1.91629	1.28768	0.77686
$n = 20$	$\overline{H}_{bsp}$	2.60871	1.90913	1.30937	0.76578
		(0.0364)	(0.03916)	(0.03422)	(0.04337)
$m = \tau = 0$	$\overline{H}_{bxp}$	2.62731	1.92772	1.32796	0.78437
		(0.03672)	(0.03924)	(0.03537)	(0.04330)
	$\overline{H}_{mlp}$	2.58858	1.92137	1.27581	0.79015
		(0.03684)	(0.04593)	(0.04392)	(0.04745)
$n = 30$	$\overline{H}_{bsp}$	2.61809	1.91542	1.2807	0.76594
		(0.03304)	(0.02383)	(0.03065)	(0.03114)
$m = \tau = 1$	$\overline{H}_{bxp}$	2.63087	1.92821	1.29348	0.77872
		(0.03343)	(0.02397)	(0.03064)	(0.03102)
	$\overline{H}_{mlp}$	2.58524	1.88519	1.27265	0.77792
		(0.02538)	(0.03309)	(0.03415)	(0.02605)
$n = 50$	$\overline{H}_{bsp}$	2.59493	1.90332	1.27308	0.77202
		(0.01691)	(0.02086)	(0.0181474)	(0.01891)
$m = \tau = 2$	$\overline{H}_{bxp}$	2.60293	1.91133	1.28108	0.78003
		(0.01674)	(0.02071)	(0.01798)	(0.01889)
	$\overline{H}_{mlp}$	2.58565	1.90772	1.28565	0.77046
		(0.01912)	(0.02635)	(0.01775)	(0.01565)
$n = 100$	$\overline{H}_{bsp}$	2.60193	1.9190	1.28447	0.77886
		(0.00796)	(0.00855)	(0.00740)	(0.00689)
$m = 1, \tau = 2$	$\overline{H}_{bxp}$	2.60621	1.92328	1.28875	0.78314
		(0.00791)	(0.00858)	(0.00739)	(0.00693)
	$\overline{H}_{mlp}$	2.61813	1.91714	1.28219	0.75960
		(0.00988)	(0.00966)	(0.00857)	(0.00753)

Table 2: Means and MSEs (in parentheses) of the estimates of H under Double censoring

		$\lambda = 0.2$	$\lambda = 0.4$	$\lambda = 0.75$	$\lambda = 1.25$
True H		2.60944	1.91629	1.28768	0.77686
$n = 20$	$\overline{H}_{bsd}$	2.59913	1.90242	1.27694	0.73742
		(0.14884)	(0.14489)	(0.16365)	(0.13656)
$m = \tau = 0$	$\overline{H}_{bxd}$	2.67164	1.97493	1.34945	0.80992
		(0.15260)	(0.14815)	(0.16735)	(0.13610)
	$\overline{H}_{mld}$	2.55712	1.8389	1.20382	0.73565
		(0.13545)	(0.17133)	(0.17056)	(0.17063)
$n = 30$	$\overline{H}_{bsd}$	2.59566	1.88033	1.24015	0.80022
		(0.13189)	(0.17158)	(0.1057)	(0.12226)
$m = \tau = 1$	$\overline{H}_{bxd}$	2.66817	1.94278	1.3026	0.86268
		(0.13515)	(0.17099)	(0.10366)	(0.12908)
	$\overline{H}_{mld}$	2.51622	1.86637	1.17234	0.71207
		(0.14945)	(0.14637)	(0.16772)	(0.13581)
$n = 50$	$\overline{H}_{bsd}$	2.44047	1.80808	1.2728	0.79488
		(0.18229)	(0.12531)	(0.11485)	(0.06354)
$m = \tau = 2$	$\overline{H}_{bxd}$	2.49532	1.86292	1.32765	0.84973
		(0.16676)	(0.11645)	(0.11622)	(0.06852)
	$\overline{H}_{mld}$	2.51241	1.84054	1.22827	0.74857
		(0.173415)	(0.165477)	(0.13743)	(0.12130)
$n = 100$	$\overline{H}_{bsd}$	2.60122	1.90008	1.2798	0.775083
		(0.14525)	(0.15772)	(0.14603)	(0.147799)
$m = 1, \tau = 2$	$\overline{H}_{bxd}$	2.67373	1.97259	1.35231	0.84759
		(0.14932)	(0.16063)	(0.15015)	(0.15279)
	$\overline{H}_{mld}$	2.53183	1.84679	1.22106	0.69714
		(0.16961)	(0.15466)	(0.15845)	(0.15329)

In all the simulation results presented here, the bias of an estimator can be determined as the average value of the estimate report in the table - True value. The variance of an estimator was determined as the sample variance obtained from all the simulations carried out. Finally, the mean square error of estimator is (variance of the estimator + $(Bias)^2$). The means and mean squared errors (in parentheses) of the estimators are presented in Tables 1 and 2.

The Bayes estimator was evaluated for the prior hyper-parameters $m, \tau = 0, 1$ and 2. It is revealed that the Bayes estimator does not seem very sensitive

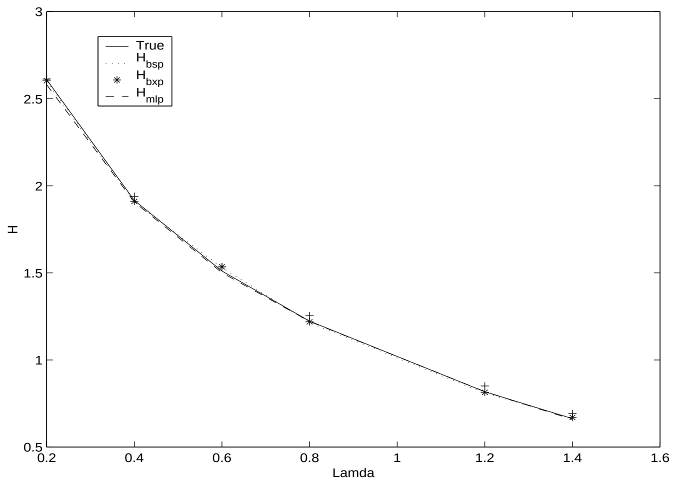

Figure 1: Estimates of Residual entropy (simulated data)

with variation of the prior parameters m and τ. It is to be noted that the bias and MSE of the Bayes estimator become smaller as the sample size increases. In most of the cases, the performance of Bayes estimator is better in terms of MSE compared to MLE estimators.

In estimating the residual entropy of exponential distribution under the two censoring schemes discussed above, we advocate the use of MLE for small samples and Bayes estimator for large samples. It is to be noted that the performance of the estimators under both loss functions are more or less similar. In Figure 1, we plot these estimates of residual entropy, for various values of λ. Figure 1 shows the performance of the estimators using different methods are similar.

Acknowledgment

The authors would like to express their gratitude to the Referee for his valuable suggestions which have considerably improved this paper.

References

[1] Asadi, M.,Ebrahimi, N., (2000) Residual entropy and its characterizations in terms of hazard function and mean residual life function. *Statist.Prob.Lett.* **49**, 263-269.

[2] Belzunce F., Guillamon, A., Navarro, J., Ruiz, J.M. (2001) Kernel estimation of residual entropty. *Commun.Statist. - Theory Meth.*, **30(7)**, 1243-1255.

[3] Ebrahimi, N. (1997) Testing whether lifetime distribution is decreasing uncertainty. *Journal of Statistical Planning and Inference*, **64**, 919.

[4] Ebrahimi, N., (1996) How to measure uncertainty in the residual life distributions. *Sankhya Ser.A* **58**, 48-56.

[5] Ebrahimi, N.,Pellerey, F.,(1995) New partial ordering of survival functions based on the notion of uncertainty. *J.Appl.Probab.***32**, 202-211.

[6] Ebrahimi, N.,Kirmani, S.N.U.A.,(1996) Some results on ordering of survival functions through uncertainty. *Statist.Prob.Lett.***11**,167-176.

[7] Nair, K.R.M., Rajesh, G., (1998) Characterisation of Probability idstributions using the reidual entropy function. J.Indian Statist.Assoc.36,157-166.

[8] Navarro, J., Belzunce, F. and Ruiz, J.M. (2004) Some results on residual entropy function. *Metrica*, **59**, 147-161.

[9] Shannon, C.E. (1948) A mathematical theory of communication. Bell System Tech.J.27 379-423 and 623-656.

[10] Varian, H.R. (1975) A Bayesian approach to Real Estate Assessment. In Studies in Bayesian Econometrics and Statistics in honour of L.J.Savage, Eds. S.E.Feinberg and A. Zellner, Amsterdam, North Holland, 195-208.

ProbStat Forum is an e-journal. For details please visit; *www.probstat.org.in.*