Bayes estimation of entropy of exponential distribution based on multiply Type II censored competing risks data[†]

Kyeongjun Lee[1] · Youngseuk Cho[2]

[12]Department of Statistics, Pusan National University

Abstract

In lifetime data analysis, it is generally known that the lifetimes of test items may not be recorded exactly. There are also situations wherein the withdrawal of items prior to failure is prearranged in order to decrease the time or cost associated with experience. Moreover, it is generally known that more than one cause or risk factor may be present at the same time. Therefore, analysis of censored competing risks data are needed. In this article, we derive the Bayes estimators for the entropy function under the exponential distribution with an unknown scale parameter based on multiply Type II censored competing risks data. The Bayes estimators of entropy function for the exponential distribution with multiply Type II censored competing risks data under the squared error loss function (SELF), precautionary loss function (PLF) and DeGroot loss function (DLF) are provided. Lindley's approximate method is used to compute these estimators. We compare the proposed Bayes estimators in the sense of the mean squared error (MSE) for various multiply Type II censored competing risks data. Finally, a real data set has been analyzed for illustrative purposes.

Keywords: Bayes estimate, competing risks, exponential distribution, multipy Type II censoring.

1. Introduction

Entropy, which is one of the important terms in statistical mechanics, was originally defined in physics especially in the second law of thermodynamics. Shannon (1948) re-defined it in an information theorem using the concepts of probability and statistics. The differential entropy $H(X)$ of the random variable X is defined by Cover and Thomas (2005) to be

$$H(X) = -\int_{-\infty}^{\infty} f(x)\log f(x)dx,$$

where $f(x)$ denote a probability density function (PDF) of random variable X. Estimation of entropy of various distributions has been studied by many researchers. Baratpour *et al.*

---

[1] Post doc. fellow, Department of Statistics, Pusan National University, Busan 609-735, Korea.
[2] Corresponding author: Professor, Department of Statistics, Pusan National University, Busan 609-735, Korea. E-mail: choys@pusan.ac.kr

(2007) considered lower/upper bounds for entropy of upper record values. They provideed the entropy of upper record values. Kang *et al.* (2012) considered the estimators of entropy of a double exponential distribution based on multiply Type II censoring. Cho *et al.* (2014) provided the estimators of entropy of a Rayleigh distriution based on doubly-generalized Type II hybrid censoring scheme. Cho *et al.* (2015) provided the estimators of entropy of a Weibull distribution based on generalized progressive hybrid censoring scheme.

Let us consider a life-testing experiment where n items is kept under observation until failure. These items could be some system, components, or computer chips in reliability study experiments, or they could be patients put under certain drug or clinical conditions. However, it is generally known that the lifetimes of test items may not be recorded exactly. There are also situations wherein the withdrawal of items prior to failure is prearranged in order to decrease the time or cost associated with experience. Moreover, in lifetime data analysis, it is generally known that more than one cause or risk factor may be present at the same time. Following Cox (1959), we will refer to this model as competing risks model. In competing risks model, it is generally supposed that the among risk factors are statistically independent. It is also assumed that competing risks data consists of a observed failure time and an indicator denoting the risk factor of failure. Lately, researchers are interested with one special cause in the presence of other risk factors.

In this paper, therefore, we derive the Bayes estimators for the entropy function of the exponential distribution with an unknown scale parameter based on multiply Type II censored competing risks data. We also compare the proposed Bayes estimators in the sense of the mean squared error (MSE) for various multiply Type II censored competing risks data.

The rest of this paper is organized as follows. In Section 2, we describe the computation of the entropy function with maximum likelihood estimators (MLEs). In Section 3, Bayes estimators of entropy function under squared error loss function (SELF), precautionary loss function (PLF) and DeGroot loss function (DLF) are also obtained under the assumption of the gamma prior distribution on the unknown parameters. Lindley's approximate method is used to compute these Bayes estimators.To know the performance of proposed Bayes estimators of entropy function based on exponential distribution with multiply Type II censored competing risks data, a numerical study is conducted in section 4. A real life data analysis has been provided in section 5. Finally, section 6 concludes.

2. Maximum likelihood estimation

Suppose n randomly selected items with competing risks data for exponential distribution were placed on a life test. We also suppose that the lifetimes $X_1, X_2, \cdots, X_n$ are independent and identically distributed (iid) with an exponential distribution. Here, $X_i = \min\{X_{1i}, X_{2i}, \cdots, X_{ji}\}$, X_{ji} denotes the lifetime of the ith item under the jth failure risk factor with cumulative distribution function (CDF) and PDF such as $G_j(x) = 1 - \exp(-\lambda_j x)$ and $g_j(x) = \lambda_j \exp(-\lambda_j x)$. Lately, researchers are interested with one special cause in the presence of other risk factors. Therefore, we assume that there are two risk factors for the failure of items. Then, it is easy to obtain the CDF and PDF of lifetime as

$$F(x) = 1 - \prod_{i=1}^{2}(1 - G_i(x)) = 1 - \exp\left[-\left(\lambda_1 + \lambda_2\right)x\right]$$

and

$$f(x) = (\lambda_1 + \lambda_2) \exp\left[-(\lambda_1 + \lambda_2) x\right].$$

The differential entropy of exponential distribution with competing risks data simplifies to

$$H(f) = 1 - \ln(\lambda_1 + \lambda_2). \tag{2.1}$$

Let $X_{a_1:n}, X_{a_2:n}, \cdots, X_{a_s:n}$ denote ordered multiply Type II censored lifetime data of s items, and $D = (\delta_{a_1}, \delta_{a_2}, \cdots, \delta_{a_s})$ denote the indicator of risk cause corresponding to the ordered multiply Type II censored lifetimes. Here $\delta_{a_j} = 1$, $j = 1, 2, \cdots, s$, denotes the failure of the a_jth item caused by first risk factor. On the other hands, $\delta_{a_j} = 0$ denotes that second risk factor is responsible for the a_jth failure. From multiply Type II censoring scheme, therefore, we have the following observations;

$$(X_{a_1:n}, \delta_{a_1}), (X_{a_2:n}, \delta_{a_2}), \cdots, (X_{a_s:n}, \delta_{a_s}).$$

Based on above assumption, the joint PDF of lifetime and corresponding factor (X, D) is given by

$$f_{X,D}(x, i) = \lambda_i \exp\left[-(\lambda_1 + \lambda_2) x\right], \ i = 1, 2.$$

Based on the multiply Type II censored competing risks data, the likelihood equation (Eq.) of s failure items is given by

$$L = \frac{n!}{\prod_{i=1}^{s} (a_i - a_{i-1} - 1)!} \left[F(x_{a_1:n})\right]^{a_1 - 1} \left[1 - F(x_{a_s:n})\right]^{n - a_s}$$

$$\times \prod_{i=1}^{s} f_{X,D}(x_{a_i:n}, 1)^{\delta_i} f_{X,D}(x_{a_i:n}, 2)^{1 - \delta_i} \prod_{i=2}^{s} \left[F(x_{a_i:n}) - F(x_{a_{i-1}:n})\right]^{a_i - a_{i-1} - 1}. \tag{2.2}$$

From likelihood Eq. (2.2), the natural logarithm of the likelihood Eq. (2.2) is given by

$$\ln L = \ln \frac{n!}{\prod_{i=1}^{s} (a_i - a_{i-1} - 1)!} + s_1 \ln \lambda_1 + s_2 \ln \lambda_2$$

$$- (\lambda_1 + \lambda_2) \left[\sum_{i=1}^{s} x_{a_i:n} + (n - a_s) x_{a_s:n}\right] + (a_1 - 1) \ln \left[1 - e^{-(\lambda_1 + \lambda_2) x_{a_1:n}}\right]$$

$$+ \sum_{i=2}^{s} (a_i - a_{i-1} - 1) \ln \left[e^{-(\lambda_1 + \lambda_2) x_{a_{i-1}:n}} - e^{-(\lambda_1 + \lambda_2) x_{a_i:n}}\right],$$

where s_i denote the total number of observed failure of items due to the risk factor i, $i = 1, 2$. Then, we have

$$\frac{\partial \ln L}{\partial \lambda_1} = \frac{s_1}{\lambda_1} - \left[\sum_{i=1}^{s} x_{a_i:n} + (n - a_s) x_{a_s:n}\right] + (a_1 - 1) \frac{x_{a_1:n} e^{-(\lambda_1 + \lambda_2) x_{a_1:n}}}{1 - e^{-(\lambda_1 + \lambda_2) x_{a_1:n}}}$$

$$- \sum_{i=2}^{s} (a_i - a_{i-1} - 1) \frac{x_{a_{i-1}:n} e^{-(\lambda_1 + \lambda_2) x_{a_{i-1}:n}} - x_{a_i:n} e^{-(\lambda_1 + \lambda_2) x_{a_i:n}}}{e^{-(\lambda_1 + \lambda_2) x_{a_{i-1}:n}} - e^{-(\lambda_1 + \lambda_2) x_{a_i:n}}} = 0 \tag{2.3}$$

and

$$\frac{\partial \ln L}{\partial \lambda_2} = \frac{s_2}{\lambda_2} - \left[\sum_{i=1}^{s} x_{a_i:n} + (n - a_s) x_{a_s:n}\right] + (a_1 - 1) \frac{x_{a_1:n} e^{-(\lambda_1 + \lambda_2) x_{a_1:n}}}{1 - e^{-(\lambda_1 + \lambda_2) x_{a_1:n}}}$$

$$- \sum_{i=2}^{s} (a_i - a_{i-1} - 1) \frac{x_{a_{i-1}:n} e^{-(\lambda_1 + \lambda_2) x_{a_{i-1}:n}} - x_{a_i:n} e^{-(\lambda_1 + \lambda_2) x_{a_i:n}}}{e^{-(\lambda_1 + \lambda_2) x_{a_{i-1}:n}} - e^{-(\lambda_1 + \lambda_2) x_{a_i:n}}} = 0. \tag{2.4}$$

The MLEs of λ_1 and λ_2, say $\hat{\lambda}_1$ and $\hat{\lambda}_2$, respectively, are the solution of Eq.s (2.3) and (2.4). These Eq.s are in implicit form, so it may be subsequently solved with a numerical method such as Newton-Raphson method. See for example the work of Cho *et al.* (2013), Kwon *et al.* (2014), Lee *et al.* (2014) and Shin *et al.* (2014). The MLE of entropy function may be obtained by replacing (λ_1, λ_2) by $(\hat{\lambda}_1, \hat{\lambda}_2)$ in Eq. (2.1).

3. Bayes estimation

3.1. Loss function

In this Section,we derive the Bayes estimates of the entropy function of the exponential distribution with multiply Type II censored competing risks data under symmetric, as well asymmetric loss functions. A very well-known symmetric loss function is the SELF, which is defined as $L_1(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$. The Bayes estimats under SELF is the posterior mean. The SELF is widely employed in the Bayes inference due to its computational simplicity. It is a symmetric loss function that gives equal weight to overestimation as well as underestimation. However,this is not a good criterion from a practical point of view. In estimating reliability and failure rate functions, an overestimation causes more damage than underestimation. To resolve such situation, asymmetrical loss functions are more appropriate.

Norstrom (1996) introduced the PLF, which is defined as $L_2(\hat{\theta}, \theta) = (\theta - \hat{\theta})^2/\hat{\theta}$, and claims that it is useful to derive conservative estimators since it approaches infinity near the origin and prevents underestimation. This loss function is suitable for situations where underestimation may lead to serious consequences. The Bayes estimator under PLF is given by $\hat{\theta}_P = \sqrt{E(\theta^2)}$, provided that $E(\theta^2)$ exists and finite.

DeGroot (2005) introduced the DLF, which is defined as $L_3(\hat{\theta}, \theta) = [(\theta - \hat{\theta})/\hat{\theta}]^2$. The Bayes estimator under DLF is given by $\hat{\theta}_D = E(\theta^2)/E(\theta)$, provided that $E(\theta^2)$ and $E(\theta)$ exist and finite.

3.2. Prior and posterior distributions

The Bayes estimate requires the choice of appropriate priors for the unknown parameters in addition to the experimental data. The model under consideration has two competing risks parameters; and continuous conjugate priors for these parameters do not exist. Here, it is assumed that the parameters follow the gamma (α_1, β_1) and gamma (α_2, β_2) prior distributions with $\alpha_1 > 0$, $\alpha_2 > 0$, $\beta_1 > 0$ and $\beta_2 > 0$. Also, we assumed that priors of λ_1 and λ_2 are independent. So, the joint prior distribution of λ_1 and λ_2 is obtained as

$$\pi(\lambda_1, \lambda_2) \propto \lambda_1^{\alpha_1-1}\lambda_2^{\alpha_2-1}e^{-\beta_1\lambda_1-\beta_2\lambda_2}, \lambda_1 > 0, \lambda_2 > 0.$$

Note that, when $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$, the prior distributions are the non-informative priors of λ_1 and λ_2. Then, the joint density of the parameters and data is obtained as follows.

$$\pi(\lambda_1, \lambda_2, \mathbf{X}) \propto \lambda_1^{s_1+\alpha_1-1}\lambda_2^{s_2+\alpha_2-1}e^{-\beta_1\lambda_1-\beta_2\lambda_2}e^{-(\lambda_1+\lambda_2)\left[\sum_{i=1}^{s}x_{a_i:n}+(n-a_s)x_{a_s:n}\right]}$$
$$\left[1 - e^{-(\lambda_1+\lambda_2)x_{a_1:n}}\right]^{a_1-1}\prod_{i=2}^{s}\left[e^{-(\lambda_1+\lambda_2)x_{a_{i-1}:n}} - e^{-(\lambda_1+\lambda_2)x_{a_i:n}}\right]^{a_i-a_{i-1}-1},$$

where $\mathbf{X} = (X_{a_1:n}, X_{a_2:n}, \cdots, X_{a_s:n})$.

Therefore, the posterior distribution of λ_1 and λ_2, given data is obtained as

$$\pi(\lambda_1, \lambda_2 | \mathbf{X}) \propto \frac{\pi(\lambda_1, \lambda_2, \mathbf{X})}{\int_0^\infty \int_0^\infty \pi(\lambda_1, \lambda_2, \mathbf{X}) \, d\lambda_1 d\lambda_2} \qquad (3.1)$$

Now, we can obtain the Bayes estimator of entropy under SELF, PLF and DLF. It is obtained as

$$\tilde{H}_S = \frac{\int_0^\infty [1 - \ln(\lambda_1 + \lambda_2)] \pi(\lambda_1, \lambda_2, \mathbf{X}) \, d\lambda_1 d\lambda_2}{\int_0^\infty \int_0^\infty \pi(\lambda_1, \lambda_2, \mathbf{X}) \, d\lambda_1 d\lambda_2}. \qquad (3.2)$$

$$\tilde{H}_P = \left[\frac{\int_0^\infty [1 - \ln(\lambda_1 + \lambda_2)]^2 \pi(\lambda_1, \lambda_2, \mathbf{X}) \, d\lambda_1 d\lambda_2}{\int_0^\infty \int_0^\infty \pi(\lambda_1, \lambda_2, \mathbf{X}) \, d\lambda_1 d\lambda_2} \right]^{1/2}. \qquad (3.3)$$

and

$$\tilde{H}_D = \frac{\int_0^\infty [1 - \ln(\lambda_1 + \lambda_2)]^2 \pi(\lambda_1, \lambda_2, \mathbf{X}) \, d\lambda_1 d\lambda_2}{\int_0^\infty [1 - \ln(\lambda_1 + \lambda_2)] \pi(\lambda_1, \lambda_2, \mathbf{X}) \, d\lambda_1 d\lambda_2}. \qquad (3.4)$$

3.3. Lindley's approximation

In above subsection, the Bayes estimators are in the Eq. (3.2), (3.3), and (3.4) for which closed forms are not available. Therefore, we apply Lindley's approximation method (Lindley (1980)) to obtain the Bayes estimator of entropy. For the (λ_1, λ_2), the Lindley's approximate Bayes estimator of entropy can be obtained as

$$\hat{g} = g\left(\hat{\lambda}_1, \hat{\lambda}_2\right) + \frac{1}{2}(A + L_{30}B_{12} + L_{03}B_{21} + L_{21}C_{12} + L_{12}C_{21}) + p_1 A_{12} + p_2 A_{21}, \qquad (3.5)$$

where

$$A = \sum_{i=1}^{2} \sum_{j=1}^{2} w_{ij}\tau_{ij}, \;\; w_{ij} = \frac{\partial^2 g}{\partial \lambda_i \lambda_j}, \;\; L_{ij} = \frac{\partial^{i+j} \ln L(\lambda_1, \lambda_2)}{\partial \lambda_1^i \partial \lambda_2^j}, \;\; B_{ij} = (w_i \tau_{ii} + w_j \tau_{ij})\tau_{ii},$$

$$C_{ij} = 3w_i \tau_{ii}\tau_{ij} + w_j(\tau_{ii}\tau_{jj} + 2\tau_{ij}^2), \;\; w_i = \frac{\partial g}{\partial \lambda_i}, \;\; p = \ln\pi(\lambda_1, \lambda_2), \;\; p_i = \frac{\partial p}{\partial \lambda_i}, \;\; A_{ij} = w_i \tau_{ii} + w_j \tau_{ji}.$$

Here, τ_{ij} denotes the (i, j)th element of the matrix $[-\partial^2 \ln L / \partial \lambda_1^i \partial \lambda_2^j]^{-1}$. For our problem, we have

$$L_{30} = \frac{\partial^3 \ln L}{\partial \lambda_1^3} = \frac{2s_1}{\lambda_1^3} + (a_1 - 1)\left[C_{2,3} + C_{2,1}^2(3x_{a_1:n} + 2C_{2,1})\right]$$

$$+ \sum_{i=2}^{s}(a_i - a_{i-1} - 1)\left[-C_{1,i,3} + C_{1,i,2}C_{1,i,1} + 2C_{1,i,1}\left(C_{1,i,2} - C_{1,i,1}^2\right)\right],$$

$$L_{03} = \frac{\partial^3 \ln L}{\partial \lambda_2^3} = \frac{2s_2}{\lambda_2^3} + (a_1 - 1)\left[C_{2,3} + C_{2,1}^2(3x_{a_1:n} + 2C_{2,1})\right]$$

$$+ \sum_{i=2}^{s}(a_i - a_{i-1} - 1)\left[-C_{1,i,3} + C_{1,i,2}C_{1,i,1} + 2C_{1,i,1}\left(C_{1,i,2} - C_{1,i,1}^2\right)\right]$$

and

$$L_{12} = \frac{\partial^3 \ln L}{\partial \lambda_1 \partial \lambda_2^2} = (a_1 - 1) \left[C_{2,3} + C_{2,1}^2 (3x_{a_1:n} + 2C_{2,1}) \right]$$

$$+ \sum_{i=2}^{s} (a_i - a_{i-1} - 1) \left[-C_{1,i,3} + C_{1,i,2} C_{1,i,1} + 2C_{1,i,1} \left(C_{1,i,2} - C_{1,i,1}^2 \right) \right]$$

$$= L_{21} = \frac{\partial^3 \ln L}{\partial \lambda_1^2 \partial \lambda_2},$$

where

$$C_{1,i,j} = \frac{x_{a_{i-1}:n}^j e^{-(\lambda_1 + \lambda_2) x_{a_{i-1}:n}} - x_{a_i:n}^j e^{-(\lambda_1 + \lambda_2) x_{a_i:n}}}{e^{-(\lambda_1 + \lambda_2) x_{a_{i-1}:n}} - e^{-(\lambda_1 + \lambda_2) x_{a_i:n}}} \quad \text{and} \quad C_{2,j} = \frac{x_{a_1:n}^j e^{-(\lambda_1 + \lambda_2) x_{a_1:n}}}{1 - e^{-(\lambda_1 + \lambda_2) x_{a_1:n}}}.$$

Now, we derive the Bayes estimator of entropy using Lindley's approximation. First of all, we compute the Bayes estimator of entropy under SELF. In this case, we observe that

$$g(\lambda_1, \lambda_2) = 1 - \ln(\lambda_1 + \lambda_2) \quad, \quad w_1 = w_2 = -\frac{1}{\lambda_1 + \lambda_2}$$

and

$$w_{11} = w_{22} = w_{12} = w_{21} = \frac{1}{(\lambda_1 + \lambda_2)^2}.$$

Using Eq. (3.5), the Bayes estimator of entropy under SELF is obtained as

$$\hat{H}_{LS} = \left[1 - \ln\left(\hat{\lambda}_1 + \hat{\lambda}_2\right) \right] + \frac{1}{2\left(\hat{\lambda}_1 + \hat{\lambda}_2\right)} \left[\frac{\tau_{11} + 2\tau_{12} + \tau_{22}}{\hat{\lambda}_1 + \hat{\lambda}_2} - L_{30}\tau_{11} (\tau_{11} + \tau_{12}) \right.$$

$$- L_{03}\tau_{22} (\tau_{22} + \tau_{21}) - L_{21} \left\{ 3\tau_{12}(\tau_{11} + \tau_{22}) + 2\tau_{11}\tau_{22} + 4\tau_{12}^2 \right\}$$

$$\left. - 2 \left\{ p_1 (\tau_{11} + \tau_{21}) + p_2 (\tau_{22} + \tau_{12}) \right\} \right].$$

Also, we compute the Bayes estimator of entropy under PLF. Here, we observe that

$$g(\lambda_1, \lambda_2) = [1 - \ln(\lambda_1 + \lambda_2)]^2 \quad, \quad w_1 = w_2 = -\frac{2[1 - \ln(\lambda_1 + \lambda_2)]}{\lambda_1 + \lambda_2}$$

and

$$w_{11} = w_{22} = w_{12} = w_{21} = \frac{2[2 - \ln(\lambda_1 + \lambda_2)]}{(\lambda_1 + \lambda_2)^2}.$$

Using Eq. (3.5), the Bayes estimator of entropy under PLF is obtained as

$$\hat{H}_{LP} = \left[\left[1 - \ln\left(\hat{\lambda}_1 + \hat{\lambda}_2\right) \right]^2 + \frac{1}{2\left(\hat{\lambda}_1 + \hat{\lambda}_2\right)} \left\{ \frac{\tau_{11} + 2\tau_{12} + \tau_{22}}{\hat{\lambda}_1 + \hat{\lambda}_2} - L_{30}\tau_{11} (\tau_{11} + \tau_{12}) \right. \right.$$

$$- L_{03}\tau_{22} (\tau_{22} + \tau_{21}) - L_{21} \left\{ 3\tau_{12}(\tau_{11} + \tau_{22}) + 2\tau_{11}\tau_{22} + 4\tau_{12}^2 \right\}$$

$$\left. \left. - 2 \left\{ p_1 (\tau_{11} + \tau_{21}) + p_2 (\tau_{22} + \tau_{12}) \right\} \right\} \right]^{1/2}.$$

Using Eq. (3.5), finally, the Bayes estimator of entropy under DLF is obtained as

$$\hat{H}_{LD} = \frac{E_H\left[H^2|\mathbf{X}\right]}{E_H\left[H|\mathbf{X}\right]},$$

where

$$
\begin{aligned}
E_H\left[H^2|\mathbf{X}\right] = {} & \left[1 - \ln\left(\hat{\lambda}_1 + \hat{\lambda}_2\right)\right]^2 + \frac{1}{2\left(\hat{\lambda}_1 + \hat{\lambda}_2\right)}\left[\frac{\tau_{11} + 2\tau_{12} + \tau_{22}}{\hat{\lambda}_1 + \hat{\lambda}_2} - L_{30}\tau_{11}\left(\tau_{11} + \tau_{12}\right)\right. \\
& - L_{03}\tau_{22}\left(\tau_{22} + \tau_{21}\right) - L_{21}\left\{3\tau_{12}(\tau_{11} + \tau_{22}) + 2\tau_{11}\tau_{22} + 4\tau_{12}^2\right\} \\
& \left. - 2\left\{p_1\left(\tau_{11} + \tau_{21}\right) + p_2\left(\tau_{22} + \tau_{12}\right)\right\}\right].
\end{aligned}
$$

and

$$
\begin{aligned}
E_H\left[H|\mathbf{X}\right] = {} & \left[1 - \ln\left(\hat{\lambda}_1 + \hat{\lambda}_2\right)\right] + \frac{1}{2\left(\hat{\lambda}_1 + \hat{\lambda}_2\right)}\left[\frac{\tau_{11} + 2\tau_{12} + \tau_{22}}{\hat{\lambda}_1 + \hat{\lambda}_2} - L_{30}\tau_{11}\left(\tau_{11} + \tau_{12}\right)\right. \\
& - L_{03}\tau_{22}\left(\tau_{22} + \tau_{21}\right) - L_{21}\left\{3\tau_{12}(\tau_{11} + \tau_{22}) + 2\tau_{11}\tau_{22} + 4\tau_{12}^2\right\} \\
& \left. - 2\left\{p_1\left(\tau_{11} + \tau_{21}\right) + p_2\left(\tau_{22} + \tau_{12}\right)\right\}\right].
\end{aligned}
$$

4. Simulation study

Since the performance of the different methods can not compared theoretically, we use Monte Carlo simulations to compare different methods for different sampling schemes.

Multiply Type II censored competing risks data can be easily generated as follows. For a given n and a_j, $j = 1, 2, \cdots, s$, we generate $X_{a:n}, \cdots, X_{a_s:n}$. Also, we generate new random variable U_i, for $i = 1, 2, \cdots, s$. Now if $U_i < \lambda_1/(\lambda_1 + \lambda_2)$, then assign $\delta_i = 1$ otherwise $\delta_i = 0$. Then, the corresponding multiply Type II censored competing risks data is $\{(X_{a_1:n}, \delta_1), (X_{a_2:n}, \delta_2), \cdots, (X_{a_s:n}, \delta_{a_s})\}$. Without loss of generality we take $\lambda_1 = 0.6$ and $\lambda_2 = 0.4$ in each case. We replicate the process 10,000 times in each case. To make the comparison meaningful, it is assumed that the priors are non-informative, and they are $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$. Note that in this case the priors are non-proper also. Press (2001) suggested to use very small non-negative values of the hyper parameters in this case, and it will make the priors proper. We have tried $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0.0001$. Bayes estimates of entropy are derived with respect to three different loss functions, SELF, PLF and DLF. The Lindley's approximation method has been then used to derive approximate explicit expressions for these estimates. Finally, different schemes have been taken into consideration to compute MSE and bias values of all estimates, and these values are tabulated in Table 4.1. The corresponding biases are reported within brackets.

From Table 4.1, the following general observations can be made. The MSEs and biases decrease as sample size n increases. For fixed sample size n, the MSEs and biases decrease generally as the number of multiply Type II censored competing risks data s decreases.

Furthermore, we observed that Bayes estimators of entropy are superior to the respective MLE in terms of MSEs and biases. For estimating the entropy, the choice DLF seems to be a reasonable choice for Bayes estimation.

Kyeongjun Lee · Youngseuk Cho

Table 4.1 The relative MSEs and biases for the proposed estimators of entropy

n	s	a_j	MSE (Bias)			
			$\hat{H}$	$\hat{H}_{LS}$	$\hat{H}_{LP}$	$\hat{H}_{LD}$
20	18	$1 \sim 5, 8 \sim 20$	0.0654 (-0.1147)	0.0604 (-0.0901)	0.0600 (-0.0997)	0.0598 (-0.1091)
		$1 \sim 4, 6 \sim 10, 12 \sim 20$	0.0653 (-0.1145)	0.0603 (-0.0899)	0.0599 (-0.0995)	0.0596 (-0.1089)
		$1 \sim 4, 6 \sim 19$	0.0692 (-0.1215)	0.0638 (-0.0964)	0.0633 (-0.1060)	0.0630 (-0.1154)
		$2 \sim 5, 7 \sim 20$	0.0654 (-0.1146)	0.0604 (-0.0900)	0.0599 (-0.0996)	0.0597 (-0.1090)
	16	$1 \sim 5, 10 \sim 20$	0.0933 (-0.2020)	0.0845 (-0.1797)	0.0842 (-0.1863)	0.0833 (-0.1919)
		$1 \sim 4, 7 \sim 10, 13 \sim 20$	0.0929 (-0.2014)	0.0841 (-0.1790)	0.0838 (-0.1827)	0.0835 (-0.1921)
		$1 \sim 4, 7 \sim 18$	0.1076 (-0.2232)	0.0972 (-0.1987)	0.0964 (-0.2054)	0.0953 (-0.2117)
		$3 \sim 5, 8 \sim 20$	0.0930 (-0.2017)	0.0842 (-0.1793)	0.0839 (-0.1860)	0.0835 (-0.1923)
	14	$1 \sim 5, 12 \sim 20$	0.1323 (-0.2818)	0.1213 (-0.2632)	0.1192 (-0.2660)	0.1228 (-0.2658)
		$1 \sim 4, 8 \sim 10, 14 \sim 20$	0.1313 (-0.2805)	0.1202 (-0.2615)	0.1186 (-0.2646)	0.1222 (-0.2691)
		$1 \sim 4, 8 \sim 17$	0.1664 (-0.3251)	0.1515 (-0.0314)	0.1472 (-0.3037)	0.1524 (-0.3084)
		$4 \sim 5, 9 \sim 20$	0.1311 (-0.2809)	0.1199 (-0.2618)	0.1183 (-0.2650)	0.1185 (-0.2683)
	12	$1 \sim 5, 14 \sim 20$	0.2440 (-0.4041)	0.2310 (-0.3907)	0.2191 (-0.3850)	0.1984 (-0.3726)
		$1 \sim 4, 9 \sim 10, 15 \sim 20$	0.1705 (-0.3444)	0.1589 (-0.3299)	0.1560 (-0.3299)	0.1512 (-0.3284)
		$1 \sim 4, 9 \sim 16$	0.2522 (-0.4315)	0.2330 (-0.4088)	0.2183 (-0.4032)	0.2329 (-0.4080)
		$5, 10 \sim 20$	0.1695 (-0.3424)	0.1577 (-0.3275)	0.1550 (-0.3277)	0.1507 (-0.3266)
30	28	$1 \sim 5, 8 \sim 30$	0.0416 (-0.0842)	0.0391 (-0.0675)	0.0394 (-0.0747)	0.0397 (-0.0819)
		$1 \sim 4, 6 \sim 10, 12 \sim 30$	0.0416 (-0.0843)	0.0391 (-0.0676)	0.0393 (-0.0748)	0.0397 (-0.0819)
		$1 \sim 4, 6 \sim 29$	0.0432 (-0.0858)	0.0406 (-0.0690)	0.0408 (-0.0762)	0.0411 (-0.0833)
		$2 \sim 5, 7 \sim 30$	0.0416 (-0.0843)	0.0391 (-0.0676)	0.0394 (-0.0748)	0.0397 (-0.0819)
	26	$1 \sim 5, 10 \sim 30$	0.0557 (-0.1450)	0.0512 (-0.1290)	0.0520 (-0.1352)	0.0528 (-0.1413)
		$1 \sim 4, 7 \sim 10, 13 \sim 30$	0.0555 (-0.1450)	0.0511 (-0.1289)	0.0518 (-0.1351)	0.0527 (-0.1413)
		$1 \sim 4, 7 \sim 28$	0.0609 (-0.1541)	0.0560 (-0.1375)	0.0567 (-0.1437)	0.0575 (-0.1499)
		$3 \sim 5, 8 \sim 30$	0.0555 (-0.1448)	0.0511 (-0.1287)	0.0518 (-0.1349)	0.0527 (-0.1411)
	24	$1 \sim 5, 12 \sim 30$	0.0756 (-0.2023)	0.0697 (-0.1873)	0.0706 (-0.1923)	0.0715 (-0.1972)
		$1 \sim 4, 8 \sim 10, 14 \sim 30$	0.0754 (-0.2022)	0.0694 (-0.1872)	0.0703 (-0.1921)	0.0713 (-0.1971)
		$1 \sim 4, 8 \sim 27$	0.0879 (-0.2215)	0.0809 (-0.2051)	0.0817 (-0.2102)	0.0825 (-0.2152)
		$4 \sim 5, 9 \sim 30$	0.0754 (-0.2021)	0.0695 (-0.1871)	0.0704 (-0.1921)	0.0714 (-0.1970)
	22	$1 \sim 5, 14 \sim 30$	0.1007 (-0.2572)	0.0937 (-0.2437)	0.0944 (-0.2473)	0.0952 (-0.2508)
		$1 \sim 4, 9 \sim 10, 15 \sim 30$	0.1007 (-0.2570)	0.0936 (-0.2433)	0.0944 (-0.2470)	0.0951 (-0.2506)
		$1 \sim 4, 9 \sim 26$	0.1247 (-0.2898)	0.1157 (-0.2737)	0.1159 (-0.2773)	0.1160 (-0.2808)
		$5, 10 \sim 30$	0.1003 (-0.2563)	0.0932 (-0.2427)	0.0940 (-0.2463)	0.0947 (-0.2499)
40	38	$1 \sim 5, 8 \sim 40$	0.0314 (-0.0709)	0.0298 (-0.0583)	0.0300 (-0.0639)	0.0303 (-0.0695)
		$1 \sim 4, 6 \sim 10, 12 \sim 40$	0.0314 (-0.0709)	0.0297 (-0.0583)	0.0300 (-0.0639)	0.0303 (-0.0695)
		$1 \sim 4, 6 \sim 39$	0.0321 (-0.0718)	0.0304 (-0.0592)	0.0307 (-0.0648)	0.0310 (-0.0703)
		$2 \sim 5, 7 \sim 40$	0.0314 (-0.0708)	0.0297 (-0.0583)	0.0300 (-0.0639)	0.0303 (-0.0695)
	36	$1 \sim 5, 10 \sim 40$	0.0401 (-0.1174)	0.0374 (-0.1051)	0.0380 (-0.1102)	0.0387 (-0.1153)
		$1 \sim 4, 7 \sim 10, 13 \sim 40$	0.0401 (-0.1173)	0.0374 (-0.1050)	0.0380 (-0.1101)	0.0387 (-0.1152)
		$1 \sim 4, 7 \sim 38$	0.0425 (-0.1214)	0.0396 (-0.1089)	0.0402 (-0.1140)	0.0409 (-0.1192)
		$3 \sim 5, 8 \sim 40$	0.0401 (-0.1174)	0.0374 (-0.1050)	0.0380 (-0.1102)	0.0387 (-0.1152)
	34	$1 \sim 5, 12 \sim 40$	0.0526 (-0.1620)	0.0489 (-0.1501)	0.0497 (-0.1546)	0.0506 (-0.1591)
		$1 \sim 4, 8 \sim 10, 14 \sim 40$	0.0525 (-0.1618)	0.0488 (-0.1499)	0.0497 (-0.1544)	0.0506 (-0.1589)
		$1 \sim 4, 8 \sim 37$	0.0587 (-0.1727)	0.0545 (-0.1603)	0.0554 (-0.1649)	0.0563 (-0.1694)
		$4 \sim 5, 9 \sim 40$	0.0525 (-0.1618)	0.0488 (-0.1499)	0.0496 (-0.1544)	0.0505 (-0.1589)
	32	$1 \sim 5, 14 \sim 40$	0.0683 (-0.2047)	0.0638 (-0.1934)	0.0647 (-0.1973)	0.0657 (-0.2011)
		$1 \sim 4, 9 \sim 10, 15 \sim 40$	0.0681 (-0.2044)	0.0636 (-0.1931)	0.0646 (-0.1969)	0.0655 (-0.2008)
		$1 \sim 4, 9 \sim 36$	0.0785 (-0.2223)	0.0732 (-0.2100)	0.0741 (-0.2140)	0.0751 (-0.2179)
		$5, 10 \sim 40$	0.0681 (-0.2044)	0.0636 (-0.1931)	0.0645 (-0.1969)	0.0655(-0.2007)

5. Data analysis

In order to analyze the real life data set, in this section, we use the proposed estimators in the above section. The real life data set were from some small electronic appliances exposed to the automatic test machine (Lawless (2011)). This dataset was analyzed by Mao *et al.* (2014). There were 18 failure risk factors of the small electronic appliances. Among the 18 failure risk factors, the ninth failure risk factor appeared the most times. Therefore, it was desirable to consider inference of ninth failure risk factor in the presence of other failure risk factors. Let us express the jth failure appliance due to ninth failure risk factor with $\delta_j = 1$, and $\delta_j = 0$ denotes failure caused by other failure risk factors. The ordered data are displayed in Table 5.1.

Table 5.1 Some small electronic appliances exposed to the automatic test machine

x_i	11	35	49	170	329	381	708	958	1062	1167	1594	1925
δ_i	0	0	0	0	0	0	0	0	0	1	0	1
x_i	1990	2223	2327	2400	2451	2471	2551	2565	2568	2694	2702	2761
δ_i	1	1	0	1	0	1	1	0	1	1	0	0
x_i	2831	3034	3059	3112	3214	3478	3504	4329	6367	6976	7846	13403
δ_i	0	1	0	1	1	1	1	1	0	1	1	0

From the above data, we take multiply Type II censoring scheme such as $a_j = 1 \sim 5$, $10 \sim 14$, $20 \sim 30$. Then, we had $s = 21$, $s_1 = 11$ and $s_2 = 10$. For Bayes estimators, we use non-informative prior distributions with all the hyperparameters equal to zero ($\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$).

Using the proposed estimators described in above sections, we obtain the MLE and Bayes estimates of entropy, to be $\hat{H} = 8.095153$, $\hat{H}_{LS} = 8.094907$, $\hat{H}_{LP} = 8.095137$ and $\hat{H}_{LD} = 8.095368$. We observed that Bayes estimates of entropy using Lindley's approximation method is marginally smaller than the corresponding MLE of entropy while opposite is true for the Bayes estimate of entropy using Lindley's approximation method obtained under the DLF.

6. Conclusions

In lifetime data analysis, it is generally known that more than one cause or risk factor may be present at the same time. That is, a failure of test item is often resulted by one of the several risk factors. Also, it is generally known that the lifetimes of test items may not be recorded exactly.

In this paper, therefore, we consider the estimators of entropy of exponential distribution with multiply Type II censored competing risks data. We obtain the MLE of entropy of exponential distribution with multiply Type II censored competing risks data. We also obtain the Bayes estimators of entropy for the exponential distribution with multiply Type II censored competing risks data under the SELF, PLF and DLF. Lindley's approximate method is used to compute these Bayes estimators. The Bayes estimators of entropy are superior to the respective MLE in terms of MSEs and biases. The choice of DLF seems to be a reasonable choice for Bayes estimation of entropy.

Although we focused on the entropy estimate of the exponential distribution with multiply Type II censored competing risk data, estimation of the entropy from other distributions with multiply Type II censored competing risk data is of potential interest in future research.

References

Baratpour, S., Ahmadi, J. and Arghami, N. R. (2007). Entropy properties of record statistics. *Statistical Papers*, **48**, 197-213.

Cho, Y., Lee, C. and Shin, H. (2013). Estimation for the generalized exponential distribution under progressive type I interval censoring. *Journal of the Korean Data & Information Science Society*, **24**, 1309-1317.

Cho, Y., Sun, H. and Lee, K. (2014). An estimation of the entropy for a Rayleigh distribution based on doubly-generalized Type-II hybrid censored samples. *Entropy*, **16**, 3655-3669.

Cho, Y., Sun, H. and Lee, K. (2015). Estimating the entropy of a Weibull distribution under generalized progressive hybrid censoring. *Entropy*, **17**, 101-122.

Cover, T. M. and Thomas, J. A. (2005). *Elements of information theory*, Wiley, Hoboken, NJ, USA.

Cox, D. (1959). The analysis of exponentially distributed lifetimes with two types of failure. *Journal of the Royal Statistical Society B (Methodological)*, **21**, 411-421.

DeGroot, M. H. (2005). *Optimal statistical decision*, John Wiley & Sons Inc., New York.

Kang, S., Cho, Y., Han, J. and Kim, J. (2012). An estimation of the entropy for a double exponential distribution based on multiply Type-II censored samples. *Entropy*, **14**, 161-173.

Kwon, B., Lee, K. and Cho, Y. (2014). Estimation for the Rayleigh distribution based on Type I hybrid censored sample. *Journal of the Korean Data & Information Science Society*, **25**, 431-438.

Lawless, J. F. (2011). *Statistical models and methods for lifetime data*, Wiley, Hoboken.

Lee, K., Sun, H. and Cho, Y. (2014). Estimation of the exponential distribution based on multiply Type I hybrid censored sample. *Journal of the Korean Data & Information Science Society*, **25**, 633-641.

Lindley, D. (1980). Approximate Bayesian methods. *Trabajos Estadistica*, **31**, 223-237.

Mao, S. and Sun, Y. (2014). Exact inference for competing risks model with generalized type I hybrid censored exponential data. *Journal of Statistical Computation and Simulation*, **84**, 2506-2521.

Norstrom, J. (1996). The use of precautionary loss functions in risk analysis. *Reliability, IEEE Transactions on*, **45**, 400-403.

Press, S. J. (2001). *The subjectivity of scientists and the Bayesian approach*, Wiley, New York.

Shin, H., Kim, J. and Lee, C. (2014). Estimation of the half-logistic distribution based on multiply Type I hybrid censored sample. *Journal of the Korean Data & Information Science Society*, **25**, 1581-1589.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379-423.