Estimation of P(X

Nasser Davarzani, Firoozeh Haghighi, Ahmad Parsian

Estimation of $P(X \leq Y)$ for a Bivariate Weibull Distribution

Nasser Davarzani[a], Firoozeh Haghighi[b] and Ahmad Parsian[b]

a Department of Statistics,Faculty of Science, Payammenoor University, Tehran, IRAN.
b School of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, IRAN.

Résumé

Nous étudions ici l'estimation de $R = P(X \leq Y)$, quand X et Y sont des variables aléatoires qui suivent la loi bivariate Weibull et X est censurée à Y. On obtient la loi marginale pour les données observées et on en tire MLE, UMVUE et MME de R. Ainsi qu'on obtient les estimateurs de Bayse sur la fonction SEL. On a effectué une simulation de Monte-Carlo pour comparer ces estimateurs.

Abstract

In this paper we deal with estimation of $R = P(X \leq Y)$, when X and Y are random variables from a bivariate weibull distribution and X is censored at Y. we obtain the marginal distribution for observed data and drive MLE, UMVUE and MME of R. Also we obtain Bayes estimators of R under squared error loss (SEL) function. Monte Carlo simulations carried out to compare these estimators.

1 Introduction

In the context of reliability the stress-strength model describes the life of a component which has a random strength Y and is subjected to random stress X. The component fails at the instant that the stress applied to it exceeds the strength and the component will function satisfactorily whenever $X \leq Y$. Thus $R = P(X \leq Y)$ is a measure of component reliability. Estimation of $R = P(X \leq Y)$, when X and Y are random variables following a specified distributions has been discussed extensively in the literature in both distribution free and parametric frame work. The problem of estimating R when the stress and strength are dependent and follow bivariate normal has been discussed by Enis and Geisser (1971) and Mukherjee and Saran (1985). Jana (1994) and Hanagal (1995) discussed the estimation of R when (X,Y) follow bivariate exponential of Marshall-Olkin(1967). Hanagal (1997) and Jeevanand (1997) discussed the estimation of R when

(X,Y) has a bivariate Pareto distribution. Hanagal (1997) obtained an estimator of R when the stress is censored at strength and (X,Y) has Marshall-Olkin's Bivariate exponential (MOBE) distribution. In this paper we consider estimation of R when the stress is censored at strength and (X,Y) follows Marshall-Olkin's bivariate weibull (MOBW) distribution. Hanagal (1996) proposed a family of k-variate weibull distributions, which includes the multivariate exponential distribution of Marshall-Olkin(1967) as an especial case. When k=2, the bivariate weibull distribution has the following probability density function:

$$f(x,y) = \begin{cases} f_1(x,y) = \theta_2(\theta_1+\theta_3)\sigma^2\ (xy)^{(\sigma-1)}\exp(-x^\sigma(\theta_1+\theta_3)-\theta_2 y^\sigma), & x>y \\ f_2(x,y) = \theta_1(\theta_2+\theta_3)\sigma^2\ (xy)^{(\sigma-1)}\exp(-y^\sigma(\theta_2+\theta_3)-\theta_1 x^\sigma), & x<y \\ f_\circ(x,x) = \theta_3\exp(-\theta x^\sigma), \qquad \theta=\theta_1+\theta_2+\theta_3, & x=y \end{cases} \tag{1}$$

where $\theta_1, \theta_2, \theta_3, \sigma > 0$. Notice that, the marginal distributions of X and Y are Weibull with scale parameters $\theta_1^* = \theta_1 + \theta_3$ and $\theta_2^* = \theta_2 + \theta_3$, respectively and σ is the common shape parameter. The parameter θ_3 quantifies the dependence between the two variables (X,Y) and $\theta_3 = 0$ implies that X and Y are independent . Also, notice that the model introduced in (1) is not absolutely continuous and contains a singular component in which $P(X=Y) > 0$. When (X,Y) follows MOBW distribution with pdf as given in (1), it is easy to verify that R as a measure of component reliability is given by

$$R = \frac{\theta_1^*}{\theta_1^* + \theta_2} \tag{2}$$

In section 2, we consider the dependent right censoring and suppose that X is censored at Y and (X,Y) follows the MOBWdistribution. We drive the marginal distribution for observed data. In section 3, we obtain MLE, MME and UMVUE of R . In section 4, we mainly consider the Bayesian inference on R under SEL function when the prior distributions of θ_1^* and θ_2 are independent and σ is known. It is observed that the Bayes estimators cannot be expressed in an explicit form in both cases. We use Monte Carlo Markov Chain (MCMC) method and Lindley's approximation to obtain the desired Bayes estimates and their estimated risks under SEL function. In section 5 we consider the numerical methods for a comparison purpose of different methods.

2 Preliminaries

Let $(X_1\,,\,Y_1)\,,\,(X_2\,,\,Y_2)\,,\,...,(X_n\,,\,Y_n)$ be a random sample (rs) of size n from MOBW distribution. If X_i is right censored at Y_i, the observed data are $Z_i = \min(X_i, Y_i)$ and $V_i = u(Y_i - X_i)\,,\ \ i = 1, 2, ..., n$, where

$$u(t) = \begin{cases} 1 & if\ \ t \geq 0 \\ 0 & if\ \ t < 0\,, \end{cases}$$

i.e., the potential data are $(X_1\,,\,Y_1)\,,\,...,(X_n\,,\,Y_n)$, but the actual observed data are $\mathbf{W} = (\mathbf{W_1}, ..., \mathbf{W_n})$ where $W_i = (Z_i, V_i)\,, i = 1, 2, ..., n$. Notice that:

- $Z_1, ..., Z_n$ are independent
- $V_1, ..., V_n$ are independent.

we obtain pdf of $W_i, i = 1, ..., n$, as

$$\begin{aligned} h\left(w_i\right) &= \left[h_1\left(w_i\right)\right]^{1-v_i} \left[h_2\left(w_i\right)\right]^{v_i} \\ &= \sigma\ z_i^{\sigma-1} e^{-\theta z_i^{\sigma}} \theta_1^{*v_i} \theta_2^{1-v_i}. \end{aligned} \tag{3}$$

Notice that $V_i, i = 1, ..., n$ has Bernouli distribution with parameter R=$\frac{\theta_1^*}{\theta}$ and $Z_i^{\sigma}, i = 1, ..., n$ has an exponential distribution with failure rate θ, and Z_i^{σ} and V_i are independent. Let $V = \sum\limits_{i=1}^{n} V_i$ and $Z_{\sigma} = \sum\limits_{i=1}^{n} Z_i^{\sigma}$.

3 MLs, MMs and UMVUs Estimators of θ_1^* and θ_2

Let $(X_1\,,\,Y_1)\,,\,(X_2\,,\,Y_2)\,,\,...,(X_n\,,\,Y_n)$ be a random sample of size n from MOBW distribution with the pdf defined in (1). Then the likelihood function based on actual observed data $\mathbf{w}$ is given by

$$L\left(\sigma, \theta_1^*, \theta_2\right) = \sigma^n \left(\prod_{i=1}^{n-1} z_i\right)^{\sigma} e^{-\theta z_{\sigma}} \theta_1^{*v} \theta_2^{n-v}, \tag{4}$$

and the log-likelihood function is

$$l\left(\sigma, \theta_1^*, \theta_2\right) = n \ln \sigma + (\sigma - 1) \sum_{i=1}^{n} \ln z_i - \theta z_{\sigma} + v \ln {\theta_1}^* + (n - v) \ln \theta_2. \tag{5}$$

The MLEs of σ, θ_1^* and θ_2, say $\hat{\sigma}$, $\hat{\theta}_1^*$ and $\hat{\theta}_2$ respectively, can be obtained as

$$\hat{\theta}_1^*(\sigma) = \frac{V}{Z_{\sigma}} \qquad , \quad \hat{\theta}_2(\sigma) = \frac{n - V}{Z_{\sigma}} \qquad , \quad \hat{\theta}(\sigma) = \frac{n}{Z_{\sigma}}. \tag{6}$$

It is observed that the MLE of σ cannot be expressed in an explicit form and it can be obtained by numerical methods. So the MMEs of θ_1^* and θ_2 can be obtained from (6) as $\hat{\theta}_1^*(\hat{\sigma})$ and $\hat{\theta}_2(\hat{\sigma})$, respectively. Hence, using the invariance property of MLE's the MLE of R can be obtained from (2)

$$R = \frac{\hat{\theta}_1^*(\hat{\sigma})}{\hat{\theta}_1^*(\hat{\sigma}) + \hat{\theta}_2(\hat{\sigma})} = \frac{V}{n}. \tag{7}$$

When σ is known, the MMEs of θ_1^*,θ_2, θ and R, say $\tilde{\theta_1^*}$, $\tilde{\theta_2}$, $\tilde{\theta}$ and $\tilde{R}$, can be obtained from the following equations

$$\begin{cases} \frac{\Gamma(1+\frac{1}{\sigma})}{\theta^{\frac{1}{\sigma}}} = \frac{1}{n}\sum\limits_{i=1}^{n} Z_i \\ \frac{\theta_1^*}{\theta} = R = \frac{V}{n}. \end{cases}$$

Therefore,

$$\tilde{\theta}_1^* = \left(\frac{\Gamma(1+\frac{1}{\sigma})}{\frac{1}{n}\sum\limits_{i=1}^{n} Z_i}\right)^{\sigma} \frac{V}{n} \quad , \quad \tilde{\theta}_2 = \left(\frac{\Gamma(1+\frac{1}{\sigma})}{\frac{1}{n}\sum\limits_{i=1}^{n} Z_i}\right)^{\sigma} \frac{n-V}{n} \quad , \quad \tilde{\theta} = \left(\frac{\Gamma(1+\frac{1}{\sigma})}{\frac{1}{n}\sum\limits_{i=1}^{n} Z_i}\right)^{\sigma} , \quad \tilde{R} = \frac{V}{n}. \tag{8}$$

We know that $V \sim Bin(n, R)$ so it is interesting to observe that MLE, MME, and UMVUE of R are the same in both parametric and non-parametric cases and is equal to $\frac{V}{n}$.

4 Bayesian Estimation of R

In this section, we deal with Bayes estimation of R based on actual observed data $\mathbf{w}$, under SEL function when σ is known or unknown. We see that in both cases the Bayes estimators cannot be obtained in a closed form, so we apply an MCMC methods and Lindley's approximation, for a numerical comparison purpose.

4.1 Bayesian Estimation of R when σ is known

Suppose θ_1^* and θ_2 have independent gamma prior distributions, with the following densities.

$$\pi_1\left(\theta_1^*\right) = \frac{\theta_1^{*\alpha_1-1} e^{-\theta_1^*\eta_1} \eta_1^{\alpha_1}}{\Gamma\left(\alpha_1\right)}, \tag{9}$$

$$\pi_2\left(\theta_2\right) = \frac{\theta_2^{\alpha_2-1} e^{-\theta_2\eta_2} \eta_2^{\alpha_2}}{\Gamma\left(\alpha_2\right)}, \tag{10}$$

where $\alpha_i, \eta_i > 0$, $i = 1, 2$, and σ are known. The joint posterior density function of θ_1^* and θ_2 is

$$\pi\left(\theta_2, \theta_1^* \left|\mathbf{w}\right.\right) \propto \sigma^n \left(\prod_{i=1}^{n-1} z_i\right)^{\sigma} \theta_1^{*v+\alpha_1-1} e^{-\theta_1^*(z_\sigma+\eta_1)} \theta_2^{n-v+\alpha_2-1} e^{-\theta_2\,(z_\sigma+\eta_2)}. \tag{11}$$

From (11), it is easy to verify that

$$\theta_1^* \left|\mathbf{w}\right. \sim \Gamma\left(v+\alpha_1,\, \eta_1+z_\sigma\right) \quad and \quad \theta_2 \left|\mathbf{w}\right. \sim \Gamma\left(\alpha_2+n-v, \eta_2+z_\sigma\right). \tag{12}$$

Notice that, R is a function of θ_1^* and θ_2, hence, using the transformation methods one can easily obtain the posterior density function of R based on the actual observed data $\mathbf{w}$ in the form

$$f\left(r|\,\mathbf{w}\right) = \mathbf{C}\frac{\mathbf{r^{v+\alpha_1-1}(1-r)^{n-v+\alpha_2-1}}}{\mathbf{[r(z_\sigma+\eta_1)+(1-r)(z_\sigma+\eta_2)]^{n+\alpha_1+\alpha_2}}}, \tag{13}$$

where

$$C = \frac{\Gamma(n+\alpha_1+\alpha_2)}{\Gamma(n-v+\alpha_2)\Gamma(v+\alpha_1)}(z_\sigma+\eta_1)^{v+\alpha_1}(z_\sigma+\eta_2)^{n-v+\alpha_2}.$$

Hence, the Bayes estimator of R under SEL function is given by

$$\delta\left(\mathbf{w}\right) = \int_\circ^\infty \frac{r^{v+\alpha_1}(1-r)^{n-v+\alpha_2-1}}{[r(z_\sigma+\eta_1)+(1-r)(z_\sigma+\eta_2)]^{n+\alpha_1+\alpha_2}}dr. \tag{14}$$

Notice that, when $\eta_1 = \eta_2 = (\eta)$, (14) reduces to

$$\delta\left(\mathbf{w}\right) = \frac{v+\alpha_1}{n+\alpha_1+\alpha_2}$$

Since, the obtained estimator is unique Bayes estimator hence, it is admissible.
We will employ an MCMC Method to obtain estimated risk and bias of R. Also, using the Lindley's approximation (1980) with the approach of Ahmad and et al.(1997), it can be easily seen that the approximate Bayes estimate of R under SEL function is

$$\hat{R}_B = \hat{\hat{R}}[1+\frac{\hat{\hat{\theta}}_1\hat{\hat{R}}}{\hat{\hat{\theta}}_1^{*\,2}(n-v+\alpha_2+1)(v+\alpha_1+1)}(\hat{\hat{\theta}}_2(n-v+\alpha_2-1)-\hat{\hat{\theta}}^*(v+\alpha_1-1))], \tag{15}$$

where

$$\hat{\hat{\theta}}_1^* = \frac{v+\alpha_1-1}{z_\sigma+\eta_1}, \quad \hat{\hat{\theta}}_2 = \frac{n-v+\alpha_2-1}{z_\sigma+\eta_2}, \quad \hat{\hat{R}} = \frac{\hat{\hat{\theta}}_2}{\hat{\hat{\theta}}_1^*+\hat{\hat{\theta}}_2}$$

5 Numerical Results

As we recognized, in both cases of known and unknown σ, the obtained estimators lead to some computational complexities. An analytic calculations of estimators and their risks for comparison is not possible. Obviously, one sample dose not tell us too much. However, it is not difficult to carry out an empirical comparison. For this purpose a simulation study was conducted to generate a sequences of independent observations using SAS9 package. The desired Bayes estimates when σ is known are calculated by Metropolis-Hasting algorithm (MCMC method) and Lindley's approximation method.

When σ is unknown we obtain Bayes estimate using Gibbs sampling technique. We obtained estimates $N = 10^4$ times and calculate Estimated Risks (ER) given by

$$ER\left(\hat{R}\right) = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{R}_i - R\right)^2,$$

Where $\hat{R}_i$ is an estimate of R.
We take $\theta_1 = 4, \theta_2 = 6, \theta_3 = 2$, $\sigma = 2$, and follow the following steps for calculating ER. The results are tabulated in Table 1
From Table 1 some of the points are quite clear. Even for small sample size, the performance of the estimates are quite satisfactory in terms of biases and estimated MSEs. It observed that when n increases, the estimated MSEs for MLEs decrease. This verifies the consistency property of MLEs. The performance of the Bayes estimates in both methods are also quite satisfactory and as the sample size increases, their estimated MSEs decrease.

Table 1. Biases and MSEs of MLEs and both Lindley's approximation and MCMC method to Bayes estimates of R, when σ is known and $\theta_1 = 4, \theta_2 = 6, \theta_3 = 2$ and $\sigma = 2$ for different values of $\alpha_1, \alpha_2, \eta_1$ and η_2.*

		BAYES ESTIMATES				LINDLEYS's APPROXIMATION			
		$\alpha_1 = 0.6, \eta_1 = 0.1$		$\alpha_1 = 1, \eta_1 = 1/6$		$\alpha_1 = 0.6, \eta_1 = 0.1$		$\alpha_1 = 1, \eta_1 = 1/6$	
n	MLE	$\alpha_2 = 0.6$ $\eta_2 = 0.1$	$\alpha_2 = 1$ $\eta_2 = 1/6$	$\alpha_2 = 0.6$ $\eta_2 = 0.1$	$\alpha_2 = 1$ $\eta_2 = 1/6$	$\alpha_2 = 0.6$ $\eta_2 = 0.1$	$\alpha_2 = 1$ $\eta_2 = 1/6$	$\alpha_2 = 0.6$ $\eta_2 = 0.1$	$\alpha_2 = 1$ $\eta_2 = 1/6$
10	0.04378	0.03490 (-0.1293)	0.03586 (-0.1383)	0.02761 (-0.1172)	0.02967 (-0.1175)	0.03168 (-0.1232)	0.02408 (-0.0874)	0.03176 (-0.1578)	0.02830 (-0.1139)
20	0.03143	0.02797 (-0.1329)	0.02986 (-0.1406)	0.02480 (−0.1229)	0.02610 (-0.1284)	0.02737 (-0.1314)	0.02246 (-0.1093)	0.0319 (-0.1493)	0.02559 (-0.1271)
30	0.2809	0.02597 (-0.1371)	0.02708 (-0.1419)	0.02392 (-0.1305)	0.02484 (-0.1340)	0.02573 (-0.1365)	0.02141 (-0.1190)	0.02958 (-0.1519)	0.02462 (-0.1334)
40	0.02590	0.02441 (-0.1377)	0.02564 (-0.1421)	0.02268 (-0.1317)	0.02378 (-0.1361)	0.02428 (-0.1374)	0.02101 (-0.1240)	0.02728 (-0.1489)	0.02367 (-0.1358)
50	0.02516	0.02399 (-0.1401)	0.02450 (-0.14202)	0.02261 (-0.1352)	0.02295 (-0.1367)	0.02391 (-0.1399)	0.02062 (-0.1270)	0.02654 (-0.1496)	0.02288 (-0.1364)

*In each box the first row presents estimated MSE of the estimates and the second row is its biases reported within parentheses.

Bibliographie

[1] Enis, P. and Geisser, S. (1971), Estimation of the probability that $P(Y < X)$. J. Amer. Statist. Assn., 66, 162-186.

[2] Hanagal, D. D. (1995), Testing reliability in a bivariate exponential stress-strength model. Journal of the Indian Statistical Association, 33, 41-45.
[3] Hanagal, D. D. (1996), A multivariate Weibull distribution. Economic Quality Control 11, 193-200.
[5] Hanagal, D. D. (1997), Estimation of reliability when stress is censored at strength. Communicatin in statistics, Theory and Methods, vol. 26(4), 911-919.
[6] Jana, P. K. (1994), Estimation of $P(Y < X)$ in the bivariate exponential case due to Marshall-Olkin. Journal of the Indian Statistical Association, vol. 32, 35-37.
[7] Lindley, D. V. (1980), Approximate Bayesian method", Trabajos de Estadistica, vol.31, 223 - 237.
[8] Marshall, A. W. and Olkin, I. A. (1967a), A multivariate exponential distribution, J. Amer. Statist. Assoc., vol. 62, 30-44.
[9] Mukherjee, S. P. and Saran, L. K. (1985), Estimation of failure probability from a bivariate normal stress-strength distribution. Microelectonics and Reliability, 25, 692-702.