Inference on $P(X < Y)$ for Exponentiated Family of Distributions

Sudhansu S, Maiti
Department of Statistics, Visva-Bharati University
Santiniketan, India
dssm1@rediffmail.com

Sudhir Murmu
Department of Statistics, Visva-Bharati University
Santiniketan, India
sudhir.murmu@yahoo.com

Abstract

Inference on $R = P(X < Y)$ has been considered when X and Y belong to independent exponentiated family of distributions. Maximum Likelihood Estimator (MLE), Uniformly Minimum Variance Unbiased Estimator (UMVUE) and Bayes Estimator of R has been derived and compared through simulation study. Exact and approximate confidence intervals and Bayesian credible intervals have also been derived.

Keywords: Bayes Estimator, Confidence Interval, Credible Interval, Delta Method, Markov Chain Monte Carlo, Maximum Likelihood Estimator, Uniformly Minimum Variance Unbiased Estimator.

1. Introduction

$R = P(X < Y)$ is used in various applications e.g. stress-strength reliability, statistical tolerancing, measuring demand-supply system performance, measuring heritability of a genetic trait, bio-equivalence study etc. Some examples are as follows.

i) If X represents the maximum chamber pressure generated by ignition of a solid propellant and Y represents the strength of the rocket chamber, then R is the probability of successful firing of the engine.

ii) If X represents the diameter of a shaft and Y represents the diameter of a bearing that is to be mounted on the shaft, then R is the probability that the bearing fits without interference.

iii) If X represents a patient's remaining years of life if treated with drug A and Y represents a patient's remaining years of life if treated with drug B, inference about R represents a comparison of the effectiveness of the two drugs.

iv) If X and Y represent lifetimes of two electronic devices, then R is the probability that one fails before the other.

v) A certain unit - be it a receptor in a human eye or ear or any other organ - operates only if it is stimulated by the source of random magnitude Y and the stimulus exceeds a lower threshold X specific for that unit. In this case, R is the probability that the unit functions.

The statistical formulation of R appears to be given first by Birnbaum (1956). The problem he considered was to find both the point estimate and an interval estimate of R on the basis of m independent observations $X_1, X_2, ..., X_m$ on X and n independent observations $Y_1, Y_2, ..., Y_n$ on Y. Birnbaum used Mann-Whitney statistic to estimate R and found the confidence interval of R in nonparametric set up following the Hodges-Lehmann approach. This paper opened up the flood gates and was followed by a deluge of papers [Birnbaum and McCarthy (1958), Owen et al. (1964), Govindarajulu (1967, 1968), Church and Harris (1970), Majumdar (1970), Enis and Geisser (1971), Bhattacharyya and Johnson (1974), Tong (1974, 1975), Kelley et al. (1976), Beg (1980), Sathe and Shah (1981), Shah and Sathe (1982), Iwase (1987), Guttman et al. (1988), McCool (1991), Weerhandi and Johnson (1992), Reiser and Farragi (1994), Ivshin (1996), Cramé r and Kamps (1997), Sinha and Zielinski (1997), Surles and Padgett (2001), Banerjee and Biswas (2003), Ali et al. (2004), Nadarajah (2004), Pal et al. (2005), Mokhlis (2005), Kundu and Gupta (2005), etc.], more or less on the same theme. For excellent reviews we refer to Johnson (1988) and Kotz et al. (2003).

Let $F_X(x,\theta) = F_0^{\alpha}(x,\theta)$ and $G_Y(y) = F_0^{\beta}(y,\theta)$, where $F_0(.,\theta)$ is the continuous baseline distribution and θ may be vector valued, and α and β are positive shape parameters. Then, X and Y are said to be belonged to the exponentiated family of distributions (abbreviated as EFD) or the proportional reversed hazard family. If X and Y are independent, then $R = P(X < Y) = \frac{\beta}{\alpha + \beta}$. In particular, we take $F_0(x,\theta) = F_0\left(\frac{x-\mu}{\sigma}\right)$ i.e. the location-scale family. If $\mu = 0$, then $F_0(x,\theta)$ belongs to the scale family and this case was studied in detail by Kakade et al. (2007).

If $\overline{F}(x,\theta) = \overline{F}_0^{\alpha}(x,\theta)$ and $\overline{G}(y,\theta) = \overline{F}_0^{\beta}(y,\theta)$ i.e. they belong to the proportional hazard family, then $R = 1 - \frac{\beta}{\alpha + \beta}$.

Kundu and Gupta (2005) and Kakade et al. (2008) considered inferential aspect of R assuming $F_0(x,\theta)$ as exponential and Gumbel distributions respectively. Surles and Padgett (2001) and Raqab and Kundu (2005) considered the same problem for scaled Burr type X distribution that eventually belongs to the exponentiated family of distributions. Awad and Gharraf (1986), Mokhlis (2005) and Rezaei et al. (2010) considered inferential aspect of R for Burr type XII, Burr type III and generalized Pareto distributions respectively which are nothing but the exponentiated family of distributions with some baseline distributions.

Our objective in this article is to draw inference, parametric as well as Bayesian, about R when X and Y belong to the exponentiated family of distributions. We look into the problem in more general set up under any known baseline distribution not necessarily restricted to the location-scale family. The problem is also studied for the baseline distribution unknown through parameter(s), in particular for the folded Crammer distribution. An outline is also given for inference about R in general case i.e. when the baseline distributions for X and Y are different through parameter.

The paper is organized as follows. In section 2, we derive the expression of R for parallel system. Section 3 discusses inference about R when the baseline distribution is completely known. In this section MLE, UMVUE and Bayes estimate of R have been derived in a general set up. Also Confidence Interval, approximate as well as exact and Bayesian Credible Intervals have been derived. In section 4, Inference about R for unknown baseline distribution through parameters has been considered. In particular, the folded Crammer distribution has been attempted and Confidence limits have been found out using bootstrap methods in section 5. In section 6, Bayes estimate of R have been calculated adopting Markov Chain Monte Carlo (MCMC) approach. An outline is given in section 7 for estimation of R in general case. Simulation results have been discussed in Section 8. Section 9 concludes.

2. Expression of R for parallel system

A system consisting of n units is said to be parallel if at least one of the units must succeed for the system to succeed. If $X_1, X_2, ..., X_n$ are the life lengths of the units, then the life length of the system is $X_{(n)} = max\,(X_1, X_2, ..., X_n)$. The following theorem holds for parallel system when the life length of each unit belongs to the exponentiated family of distribution.

Theorem 2.1 *If the X_i are independent and belong to the exponentiated family of distribution $EFD(x, \alpha_i)$, for $i = 1, 2, ..., n,$ then $X_{(n)} = max\,(X_1, X_2, ..., X_n)$ is distributed as the $EFD(x, \alpha = \sum_{i=1}^{n} \alpha_i)$.*

Remark 2.1 *If the baseline distribution is normal, then Gupta and Gupta (2008) called it as power normal and their theorem 3.1 is particular case of the above theorem.*

If any one or both of X and Y is realized as resultant of a parallel system, then with the help of theorem 2.1, one can find out the expression for R.

3. Inference about R when the baseline distribution is completely known

Without loss of generality, we assume that $\mu = 0$ and $\sigma = 1$. If we transform the random variables $U = -\ln F_0(X)$ (i.e. $U = \xi(X)$) and $V = -\ln F_0(Y)$ (i.e. $V = \xi(Y)$), then U and V follow independent exponential distributions with parameters α and β respectively. Therefore, all the results of R for independent exponential distributions will follow. Moreover, $R = P(X < Y) = P(U < V)$. We summarize inferential results in sequel.

3.1 Maximum Likelihood Estimator of R

To compute the MLE of R, we will use the following theorem in Kotz et al. (2003)[p.40].

Theorem 3.1 *Let* $\hat{\Psi}(\underline{U}, \underline{V})$ *be the MLE of* R *based on observations* $\underline{U} = (U_1, U_2, ..., U_m)$ *and* $\underline{V} = (V_1, V_2, ..., V_n)$, *where* $m = n$ *whenever* U *and* V *are dependent. Then the MLE* $\hat{R}$ *of* R *based on* $\underline{X}$ *and* $\underline{Y}$ *is given by*

$$\hat{R} = \hat{\Psi}(\underline{U}, \underline{V}) = \hat{\Psi}(\xi(\underline{X}), \xi(\underline{Y})) \tag{1}$$

where $\xi(\underline{X}) = (\xi(X_1), \xi(X_2), ..., \xi(X_m)), \xi(\underline{Y}) = (\xi(Y_1), \xi(Y_2), ..., \xi(Y_n))$, and $\underline{X}$ and $\underline{Y}$ are observation vectors.

Using the theorem 3.1 and writing $W_1 = \sum_{i=1}^{m} U_i = -\sum_{i=1}^{m} \ln F_0(X_i)$ and $W_2 = \sum_{i=1}^{n} V_i = -\sum_{i=1}^{n} \ln F_0(Y_i)$, we obtain the MLE of R is

$$\hat{R}_1 = \frac{\frac{n}{W_2}}{\frac{m}{W_1} + \frac{n}{W_2}}. \tag{2}$$

Here $\underline{X} = (X_1, X_2, ..., X_m)$ is a random sample from $EFD(\alpha)$ and $\underline{Y} = (Y_1, Y_2, ..., Y_n)$ is a random sample from $EFD(\beta)$ and the MLE of α is $\hat{\alpha} = -\frac{m}{W_1}$ and that of β is $\hat{\beta} = -\frac{n}{W_2}$.

Theorem 3.2 *If* $\hat{R}_1$ *is the MLE of* R, *then*

$$Var(\hat{R}_1) \cong R^2(1-R)^2\left[\frac{m^2}{(m-1)^2(m-2)} + \frac{n^2}{(n-1)^2(n-2)}\right] \cong R^2(1-R)^2\left[\frac{1}{m} + \frac{1}{n}\right]$$

Proof: $Var(\hat{R}_1) \cong \left[\frac{\partial \hat{R}_1}{\partial \hat{\alpha}}\right]^2_{\hat{\alpha}=\alpha, \hat{\beta}=\beta} Var(\hat{\alpha}) + \left[\frac{\partial \hat{R}_1}{\partial \hat{\beta}}\right]^2_{\hat{\alpha}=\alpha, \hat{\beta}=\beta} Var(\hat{\beta})$.

Now, $Var(\hat{\alpha}) = \frac{m^2\alpha^2}{(m-1)^2(m-2)}$, $Var(\hat{\beta}) = \frac{n^2\beta^2}{(n-1)^2(n-2)}$, $\left[\frac{\partial \hat{R}_1}{\partial \hat{\alpha}}\right]^2_{\hat{\alpha}=\alpha, \hat{\beta}=\beta} = \frac{\beta^2}{(\alpha+\beta)^4}$ and

$$\left[\frac{\partial \hat{R}_1}{\partial \hat{\beta}}\right]^2_{\hat{\alpha}=\alpha, \hat{\beta}=\beta} = \frac{\alpha^2}{(\alpha+\beta)^4}.$$

Hence, $Var(\hat{R}_1) \cong R^2(1-R)^2\left[\frac{m^2}{(m-1)^2(m-2)}+\frac{n^2}{(n-1)^2(n-2)}\right] \cong R^2(1-R)^2\left[\frac{1}{m}+\frac{1}{n}\right],$

since $R=\frac{\beta}{\alpha+\beta}$.

3.2 Uniformly Minimum Variance Unbiased Estimator of R

The upcoming theorem in Kotz et al. (2003)[p.40] will be used to obtain the UMVUE of R.

Theorem 3.3 *Let $T_{U,V}$ be a sufficient statistic for τ based on $(\underline{U},\underline{V})$ and let there exist an UMVUE $\widetilde{\Psi}(T_{U,V})$ of R based on observations $(\underline{U},\underline{V})$. Then, $T_{X,Y}=T_{U,V}(\xi(\underline{X}),\xi(\underline{Y}))$ is a sufficient statistic for θ based on the sample $(\xi(\underline{X}),\xi(\underline{Y}))$ and the UMVUE $\widetilde{R}$ of R based on $\underline{X}$ and $\underline{Y}$ is given by*

$$\widetilde{R}=\widetilde{\Psi}(T_{X,Y}), \tag{3}$$

where the scalar or vector-valued parameter τ is connected to θ by the one-to-one transformation ρ with the inverse $\nu:\theta=\nu(\tau)\Leftrightarrow\tau=\rho(\theta)$.

Since (W_1,W_2) is a complete sufficient statistic for (α,β), using theorem 3.3, the UMVUE of R, say $\hat{R}_2$, can be obtained [see also the result of Tong (1974, 1975)] as

$$\hat{R}_2=\sum_{s=0}^{n-1}(-1)^s\frac{(m-1)!(n-1)!}{(m+s-1)!(n-s-1)!}\left(\frac{W_2}{W_1}\right)^s \quad if\ W_2<W_1$$

$$=1-\sum_{s=0}^{m-1}(-1)^s\frac{(m-1)!(n-1)!}{(m+s-1)!(n-s-1)!}\left(\frac{W_1}{W_2}\right)^s \quad if\ W_1<W_2.$$

This can also be expressed in the following form

$$\hat{R}_2=F\left(1,-(m-1);n,\frac{W_2}{W_1}\right) \quad if\ W_2<W_1$$

$$=1-F\left(1,-(n-1);m,\frac{W_1}{W_2}\right) \quad if\ W_1<W_2,$$

where $F(\alpha,\beta;\gamma,z)$ is the Gauss hypergeometric function given by

$$F(\alpha,\beta;\gamma,z)=1+\frac{\alpha.\beta}{\gamma.1}z+\frac{\alpha(\alpha+1).\beta(\beta+1)}{\gamma(\gamma+1).1.2}z^2+....,$$

see Grandshteyn and Ryzhik (2000) (formula 9.100).

Now, we are interested to find out the variance of $\hat{R}_2$. From Blight and Rao (1974) and Ghosh and Sathe (1987), the Bhattacharya bound converges to the variance of UMVUE for the family of exponential distributions. Hence,

$$Var(\hat{R}_2) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}{}_{i+j\geq 1} \frac{\tau_{ij}^2}{A_i^2 B_j^2},$$

where $\tau_{ij}^2 = (-1)^{i+j} \frac{(i\beta - j\alpha)}{(\alpha+\beta)^{i+j+1}}$, $A_i^2 = \frac{(m+i-1)!\,i!}{(m-1)!}\alpha^{2i}$, $B_j^2 = \frac{(n+j-1)!\,j!}{(n-1)!}\beta^{2j}$.

3.3 Bayes Estimator of R

We state the following theorem in Kotz et al. (2003)[p.41] that will be used to obtain a Bayes estimator of R.

Theorem 3.4 *Let $\hat{\Psi}(\underline{U},\underline{V})$ be a Bayes estimator of R based on observations $\underline{U} = (U_1, U_2, ..., U_m)$ and $\underline{V} = (V_1, V_2, ..., V_n)$ and the prior pdf $\pi(\tau)$. Then a Bayes estimator $\hat{R}$ of R based on $\underline{X}$ and $\underline{Y}$ is given by*

$$\hat{R} = \hat{\Psi}(\xi(\underline{X}), \xi(\underline{Y})) \tag{4}$$

where $\xi(\underline{X}) = (\xi(X_1), \xi(X_2), ..., \xi(X_m)), \xi(\underline{Y}) = (\xi(Y_1), \xi(Y_2), ..., \xi(Y_n))$ and the prior pdf $\pi(\rho(\theta))\,|J_\rho(\theta)|$. Here $|J_\rho(\theta)|$ is the Jacobian of the transformation $\rho(\theta)$.

3.3.1 Conjugate Prior Distributions

We obtain the Bayes estimator of R under the assumption that the shape parameters α and β are random variables for both the populations. It is assumed that α and β have independent gamma prior with pdfs:

$$\pi(\alpha) = \frac{b_1^{a_1}}{\Gamma(a_1)}\alpha^{a_1-1}e^{-b_1\alpha}; \alpha > 0,$$

and

$$\pi(\beta) = \frac{b_2^{a_2}}{\Gamma(a_2)}\beta^{a_2-1}e^{-b_2\beta}; \beta > 0,$$

$a_1, b_1, a_2, b_2 > 0$ respectively. The prior pdfs of α and β are as follows:

$$\alpha/W_1 : Gamma(a_1 + m, b_1 + W_1),$$

$$\beta/W_2 : Gamma(a_2 + n, b_2 + W_2).$$

Since apriori α and β are independent, the posterior pdf of R becomes

$$f_R(r) = c.\frac{r^{a_1+m-1}(1-r)^{a_2+n-1}}{\{(b_1+W_1)r + (b_2+W_2)(1-r)\}^{m+n+a_1+a_2}} \quad for \;\; 0 < r < 1,$$
$$= 0 \quad otherwise,$$
$$where \;\; c = \frac{1}{B(a_1+m, a_2+n)}(b_1+W_1)^{a_1+m}(b_2+W_2)^{a_2+n}.$$

Here, the Bayes estimator of R with respect to the squared error loss function is

$$\hat{R}_3 = E\left[R/(W_1,W_2)\right]$$

$$= \left(\frac{\lambda_1}{\lambda_2}\right)^{\delta_1}\left(\frac{\delta_2}{\delta_1+\delta_2}\right)F\left(\delta_1+\delta_2,\delta_1;\delta_1+\delta_2+1,1-\frac{\lambda_1}{\lambda_2}\right) \quad if\ \lambda_1 \le \lambda_2$$

$$= \left(\frac{\lambda_2}{\lambda_1}\right)^{\delta_2}\left(\frac{\delta_2}{\delta_1+\delta_2}\right)F\left(\delta_1+\delta_2,\delta_2+1;\delta_1+\delta_2+1,1-\frac{\lambda_2}{\lambda_1}\right) \quad if\ \lambda_2 < \lambda_1,$$

where $\delta_1 = a_1+m, \lambda_1 = b_1+W_1, \delta_2 = a_2+n$ and $\lambda_2 = b_2+W_2$.

It is to be noted that the Bayes estimator $\hat{R}_3$ depends on the parameters of the prior distributions of α and β. These parameters could be estimated by means of an empirical Bayes procedure, see Lindley (1969) and Awad and Gharraf (1986). Given the random samples $(X_1,X_2,...,X_m)$ and $(Y_1,Y_2,...,Y_n)$, the likelihood functions of α and β are gamma densities with parameters $(m+1,W_1)$ and $(n+1,W_2)$ respectively. Hence it is proposed to estimate the prior parameters a_1 and b_1 from the samples by $m+1$ and W_1. Similarly, a_2 and b_2 could be estimated from the samples by $n+1$ and W_2. Therefore, the Bayes estimator of R with respect to the squared error loss function could be given as

$$\hat{R}_4 = \left(\frac{W_1}{W_2}\right)^{2n+1}\left(\frac{2n+1}{2m+2n+2}\right)F\left(2m+2n+2,2m+1;2m+2n+3,1-\frac{W_1}{W_2}\right) if\ W_1 \le W_2$$

$$= \left(\frac{W_2}{W_1}\right)^{2n+1}\left(\frac{2n+1}{2m+2n+2}\right)F\left(2m+2n+2,2n+2;2m+2n+3,1-\frac{W_2}{W_1}\right) \quad if\ W_2 < W_1.$$

3.3.2 Non Informative Prior Distributions:

In this subsection we obtain the Bayes estimator of R under the assumption that the shape parameters α and β are random variables having independent noninformative priors $\pi_1(\alpha) \propto \frac{1}{\alpha}$ and $\pi_2(\beta) \propto \frac{1}{\beta}$ respectively.

Hence, the Bayes estimator with respect to the squared error loss function will be

$$\hat{R}_5 = \left(\frac{W_1}{W_2}\right)^{m}\left(\frac{m}{m+n}\right)F\left(m+n,m+1;m+n+1,1-\frac{W_1}{W_2}\right) \quad if\ W_1 < W_2$$

$$= \left(\frac{W_2}{W_1}\right)^{n}\left(\frac{m}{m+n}\right)F\left(m+n,n;m+n+1,1-\frac{W_2}{W_1}\right) \quad if\ W_2 \le W_1.$$

3.4 Interval Estimation of R

3.4.1 Approximate Confidence Interval

It is to be noted that the MLE $\hat{R}_1$ is asymptotically normal with mean R and variance $\sigma^2_{\hat{R}_1} \cong R^2(1-R)^2\left[\frac{m^2}{(m-1)^2(m-2)}+\frac{n^2}{(n-1)^2(n-2)}\right] \cong R^2(1-R)^2\left[\frac{1}{m}+\frac{1}{n}\right].$

Hence an approximate $100(1-\gamma)\%$ confidence interval for R would be (L_1, U_1), where

$$L_1 = \hat{R}_1 - \tau_{\gamma/2}\sqrt{\left(\frac{1}{m}+\frac{1}{n}\right)\hat{R}_1(1-\hat{R}_1)},$$

and

$$U_1 = \hat{R}_1 + \tau_{\gamma/2}\sqrt{\left(\frac{1}{m}+\frac{1}{n}\right)\hat{R}_1(1-\hat{R}_1)},$$

with $\tau_{\gamma/2}$ being the upper $\gamma/2$ point of the standard normal distribution.

3.4.2 Exact Confidence Interval

Before obtaining a confidence interval for R, we state the following theorem in Kotz et al. (2003)[p.42].

Theorem 3.5 *Let $L_2(\underline{U},\underline{V}), U_2(\underline{U},\underline{V})$ be a confidence interval for R with the confidence coefficient $(1-\gamma)$. Then $[L_2(\xi(\underline{X}),\xi(\underline{Y})), U_2(\xi(\underline{X}),\xi(\underline{Y}))]$ is the confidence interval for R based on $(\underline{X},\underline{Y})$ with the same coverage probability.*

Notice that $2\alpha W_1$ and $2\beta W_2$ are two independent chi-square random variables kwith $2m$ and $2n$ degrees of freedom. Now, $\hat{R}_1$ can be rewritten as

$$\hat{R}_1 = \left(1+\frac{\hat{\alpha}}{\hat{\beta}}\right)^{-1}$$
$$= \left(1+\frac{m}{n}\frac{\alpha}{\beta}F_1\right)^{-1},$$

where $F_1 = \frac{\beta W_2}{\alpha W_1}$ is an F distributed random variable with $(2n,2m)$ degrees of freedom. We see that $F_1 = \frac{W_2}{W_1}\frac{R}{1-R}$. Using F_1 as a pivotal quantity, we obtain a $100(1-\gamma)\%$ confidence interval for R as (L_2, U_2), where

$$L_2 = F_{(1-\frac{\gamma}{2})}(2n,2m)\left[F_{(1-\frac{\gamma}{2})}(2n,2m)+\frac{W_2}{W_1}\right]^{-1},$$

and

$$U_2 = F_{\frac{\gamma}{2}}(2n,2m)\left[F_{\frac{\gamma}{2}}(2n,2m)+\frac{W_2}{W_1}\right]^{-1}.$$

3.5 Bayesian Credible Intervals

3.5.1 Conjugate Prior Distributions:

Assuming α and β are independent, we have seen in subsection 3.3.1 that the posterior distributions of α and β corresponding to gamma priors are gamma with parameters $(2m+1,2W_1)$ and $(2n+1,2W_2)$, respectively. Thus $4\alpha W_1$ and $4\beta W_2$ are independent chi-square random variables with $2(2m+1)$ and $2(2n+1)$ degrees of freedom. Thus

$$F_2 = \frac{4\beta W_2}{4\alpha W_1} = \frac{W_2}{W_1}\frac{R}{1-R}$$

is an $F-$distributed random variable with $[2(2n+1),2(2m+1)]$ degrees of freedom. Using F_2 as a pivotal quantity, we obtain a $100(1-\gamma)\%$ Bayes credible interval for R as (L_3,U_3), where

$$L_3 = F_{(1-\frac{\gamma}{2})}(2(2n+1),2(2m+1))\left[F_{(1-\frac{\gamma}{2})}(2(2n+1),2(2m+1))+\frac{W_2}{W_1}\right]^{-1},$$

and

$$U_3 = F_{\frac{\gamma}{2}}(2(2n+1),2(2m+1))\left[F_{\frac{\gamma}{2}}(2(2n+1),2(2m+1))+\frac{W_2}{W_1}\right]^{-1}.$$

3.5.2 Non-informative Prior Distributions

We have seen in subsection 3.3.2 that assuming independence and non-informative prior distributions for α and β, the posterior distributions of α and β are gamma with parameters (m,W_1) and (n,W_2), respectively. Therefore, $2\alpha W_1$ and $2\beta W_2$ are independent chi-square random variables with $2m$ and $2n$ degrees of freedom. Thus

$$F_3 = \frac{2\beta W_2}{2\alpha W_1} = \frac{W_2}{W_1}\frac{R}{1-R}$$

is an $F-$distributed random variable with $(2n,2m)$ degrees of freedom. Using F_3 as a pivotal quantity, we obtain a $100(1-\gamma)\%$ Bayes credible interval for R with lower and upper bounds exactly the same as those given in subsection 3.4.2.

4. Inference about R when the baseline distribution is unknown through parameter

4.1 Maximum Likelihood Estimation on R

To compute the MLE of R, we have to obtain the MLEs of α and β. Suppose $(X_1, X_2, ..., X_m)$ is a random sample from $f_X(\alpha,\theta)$ and $(Y_1, Y_2, ..., Y_n)$ is a random sample from $g_Y(\beta,\theta)$. Hence, the underlying log-likelihood function is

$$l(\alpha,\beta,\theta) = m\ln\alpha + n\ln\beta + \sum_{i=1}^{m}\{\ln f_0(x_i,\theta) + (\alpha-1)\ln F_0(x_i,\theta)\}$$
$$+\sum_{j=1}^{n}\{\ln f_0(y_j,\theta) + (\beta-1)\ln F_0(y_j,\theta)\}$$

Then the MLE of α is to be obtained from the relation

$$\hat{\alpha}(\theta) = \frac{m}{-\sum_{i=1}^{m}\ln F_0(x_i,\theta)}$$

and that of β is from

$$\hat{\beta}(\theta) = \frac{n}{-\sum_{j=1}^{n}\ln F_0(y_j,\theta)}$$

and the MLE of components of θ are to be obtained by solving the equations

$$\frac{\partial l(\alpha,\beta,\theta)}{\partial\theta_t} = 0;\, t = 1, 2, ..., k.$$

An estimate $\hat{R}$ of R is to be obtained from expression replacing α and β by $\hat{\alpha}(\hat{\theta})$ and $\hat{\beta}(\hat{\theta})$ respectively. Here we will use delta method to obtain approximate confidence intervals of R.

Let us write

$$\mathbf{W} = \begin{pmatrix} a_{\alpha\alpha} & a_{\alpha\beta} & a_{\alpha 1} & a_{\alpha 2} & \cdots & a_{\alpha k} \\ a_{\alpha\beta} & a_{\beta\beta} & a_{\beta 1} & a_{\beta 2} & \cdots & a_{\beta k} \\ a_{\alpha 1} & a_{\beta 1} & a_{11} & a_{12} & \cdots & a_{1k} \\ a_{\alpha 2} & a_{\beta 2} & a_{12} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{\alpha k} & a_{\beta k} & a_{1k} & a_{2k} & \cdots & a_{kk} \end{pmatrix}$$

$$= \begin{pmatrix} W_{11} & W_{12} \\ W_{12}' & W_{22} \end{pmatrix}$$

where $-a_{\alpha\alpha}=E\left(\frac{\partial^2 l(\alpha,\beta,\theta)}{\partial\alpha^2}\right)$, $-a_{\beta\beta}=E\left(\frac{\partial^2 l(\alpha,\beta,\theta)}{\partial\beta^2}\right)$,

$-a_{\alpha\beta}=E\left(\frac{\partial^2 l(\alpha,\beta,\theta)}{\partial\alpha\partial\beta}\right)$, $-a_{\alpha j}=E\left(\frac{\partial^2 l(\alpha,\beta,\theta)}{\partial\alpha\partial\theta_j}\right)$, $-a_{\beta j}=E\left(\frac{\partial^2 l(\alpha,\beta,\theta)}{\partial\beta\partial\theta_j}\right)$,

and $-a_{ij}=E\left(\frac{\partial^2 l(\alpha,\beta,\theta)}{\partial\theta_i\partial\theta_j}\right); i,j=1,2,...,k.$

Now, the asymptotic variance-covariance matrix of $(\hat{\alpha},\hat{\beta},\hat{\theta})$ is given by

$$V=W^{-1}=W^{11}W^{12}W^{12'}W^{22}.$$

Let $G=(G_1,G_2)'$, with $G_1=\left(\frac{\partial R}{\partial\alpha},\frac{\partial R}{\partial\beta}\right)$, $G_2=\left(\frac{\partial R}{\partial\theta_1},\frac{\partial R}{\partial\theta_2},...,\frac{\partial R}{\partial\theta_k}\right)=(0,0,...,0)$ yield the asymptotic variance of $\hat{R}$ as $S_\Delta^2(\hat{R})=G'VG=G_1'W^{11}G_1$. Here $\frac{\partial R}{\partial\alpha}=-\frac{\beta}{(\alpha+\beta)^2}$ and $\frac{\partial R}{\partial\beta}=-\frac{\alpha}{(\alpha+\beta)^2}$. Assuming that $\frac{R-\hat{R}}{S_\Delta(\hat{R})}$ as a standard normal variate, confidence intervals to R can be constructed.

5. Inference on R for Exponentiated Folded Crammer Distribution

The Folded Crammer distribution has the density function

$$f_X(x;\sigma)=\frac{\sigma}{(\sigma+x)^2};\ x,\sigma>0$$

and the distribution function

$$F_X(x;\sigma)=\frac{x}{\sigma+x}.$$

Hence the density function of Exponentiated Folded Crammer (EFC) distribution is given by

$$f(x;\sigma,\alpha)=\alpha\left(\frac{x}{\sigma+x}\right)^{\alpha-1}.\frac{\sigma}{(\sigma+x)^2};\ \ x,\sigma,\alpha>0.$$

For convenience, we re-parametrized this distribution by defining $\frac{1}{\sigma}=\lambda$.

Therefore,

$$f(x;\lambda,\alpha)=\alpha\lambda(\lambda x)^{(\alpha-1)}(1+\lambda x)^{-(\alpha+1)};\ x>0,\lambda>0$$

5.1 Maximum Likelihood Estimation of R

Let $X : EFC(\alpha,\lambda)$ and $Y : EFC(\beta,\lambda)$, where X and Y are independent random variables. To compute the MLE of R, first we obtain the MLEs of α and β. Suppose $(X_1, X_2,.....X_m)$ is random sample from $EFC(\alpha,\lambda)$ and $(Y_1, Y_2,.....Y_n)$ is random sample from $EFC(\beta,\lambda)$. Therefore, the log-likelihood function of the observed samples is

$$L(\alpha,\beta,\lambda) = (m+n)\ln\lambda + m\ln\alpha + n\ln\beta + (\alpha-1)\sum_{i=1}^{m}\ln(\lambda x_i) + (\beta-1)\sum_{j=1}^{n}\ln(\lambda y_j)$$
$$-(\alpha+1)\sum_{i=1}^{m}\ln(1+\lambda x_i) - (\beta+1)\sum_{j=1}^{n}\ln(1+\lambda y_j)$$

The MLE's of α, β and λ say $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\lambda}$ respectively, can be obtained as the solutions of $\frac{\partial L}{\partial \alpha}=0, \frac{\partial L}{\partial \beta}=0$ and $\frac{\partial L}{\partial \lambda}=0.$ After calculation, we obtain

$$\hat{\alpha} = -\frac{m}{\sum_{i=1}^{m}\ln\frac{\lambda x_i}{1+\lambda x_i}} \tag{5}$$

$$\hat{\beta} = -\frac{n}{\sum_{j=1}^{n}\ln\frac{\lambda y_j}{1+\lambda y_j}} \tag{6}$$

and $\hat{\lambda}$ can be obtained as the solution of the non-linear equation

$$g(\lambda) = -\frac{\left(\frac{m^2}{\sum_{i=1}^{m}\ln\frac{\lambda x_i}{1+\lambda x_i}} + \frac{n^2}{\sum_{j=1}^{n}\ln\frac{\lambda y_j}{1+\lambda y_j}}\right)}{\lambda} - \left(1-\frac{m}{\sum_{i=1}^{m}\ln\frac{\lambda x_i}{1+\lambda x_i}}\right)\times\sum_{i=1}^{m}\frac{x_i}{1+\lambda x_i}$$
$$-\left(1-\frac{n}{\sum_{j=1}^{n}\ln\frac{\lambda y_j}{1+\lambda y_j}}\right)\times\sum_{j=1}^{n}\frac{y_j}{1+\lambda y_j} = 0 \tag{7}$$

Therefore, $\hat{\lambda}$ can be obtained as a solution of the non-linear equation of the form

$$h(\lambda) = \lambda \tag{8}$$

where

$$h(\lambda)=-\left(\frac{m^2}{\sum_{i=1}^{m}\ln\frac{\lambda x_i}{1+\lambda x_i}}+\frac{n^2}{\sum_{j=1}^{n}\ln\frac{\lambda y_j}{1+\lambda y_j}}\right)\times$$

$$\left[\left(1-\frac{m}{\sum_{i=1}^{m}\ln\frac{\lambda x_i}{1+\lambda x_i}}\right)\times\sum_{i=1}^{m}\frac{x_i}{1+\lambda x_i}+\left(1-\frac{n}{\sum_{j=1}^{n}\ln\frac{\lambda y_j}{1+\lambda y_j}}\right)\times\sum_{j=1}^{n}\frac{y_j}{1+\lambda y_j}\right]^{-1}.$$

It can be obtained by using a simple iterative scheme as follows

$$h(\lambda_{(j)})=\lambda_{(j+1)} \tag{9}$$

where $\lambda_{(j)}$ is the j^{th} iterate of $\hat{\lambda}$. The iteration procedure should be stopped when $\left|\lambda_{(j)}-\lambda_{(j+1)}\right|$ is sufficiently small. Once we obtain $\hat{\lambda}$, $\hat{\alpha}$ and $\hat{\beta}$ can be obtained from (5.5) and (5.6) respectively. Therefore, the MLE of R become

$$\hat{R}=\frac{\hat{\beta}}{\hat{\alpha}+\hat{\beta}} \tag{10}$$

5.2 Asymptotic distribution and confidence intervals

In this section, the asymptotic distribution of $\hat{\theta}=(\hat{\alpha},\hat{\beta},\hat{\lambda})$ and the asymptotic distribution of $\hat{R}$ are obtained. Based on the asymptotic distribution of $\hat{R}$, the asymptotic confidence interval of R is derived. Let us denote the Fisher information matrix of $\theta=(\alpha,\beta,\lambda)$ as $I(\theta)=\left(I_{ij},(\theta)\right); i,j=1,2,3$. Therefore,

$$I(\theta)=E\left(\frac{\partial^2 L}{\partial\alpha^2}\right)E\left(\frac{\partial^2 L}{\partial\alpha\partial\beta}\right)E\left(\frac{\partial^2 L}{\partial\alpha\partial\lambda}\right)-E\left(\frac{\partial^2 L}{\partial\beta\partial\alpha}\right)E\left(\frac{\partial^2 L}{\partial^2\beta}\right)E\left(\frac{\partial^2 L}{\partial\beta\partial\lambda}\right)E\left(\frac{\partial^2 L}{\partial\lambda\partial\alpha}\right)E\left(\frac{\partial^2 L}{\partial\lambda\partial\beta}\right)E\left(\frac{\partial^2 L}{\partial\lambda^2}\right)$$
$$=I_{11}I_{12}I_{13}I_{21}I_{22}I_{23}I_{31}I_{32}I_{33}(say).$$

Using the integrals of the form

$$\int_0^\infty x^{r-1}(1+\lambda x)^{-v}dx=\lambda^{-r}B(r,v-r)$$

for $0<r<v$, where $B(x,y)$ is the beta function, we have

$$E\left(\frac{\partial^2 L}{\partial\alpha^2}\right)=-\frac{m}{\alpha^2},$$

$$E\left(\frac{\partial^2 L}{\partial\beta^2}\right)=-\frac{n}{\beta^2},$$

$$E\left(\frac{\partial^2 L}{\partial\alpha\partial\beta}\right)=E\left(\frac{\partial^2 L}{\partial\beta\partial\alpha}\right)=0,$$

$$E\left(\frac{\partial^2 L}{\partial\alpha\partial\lambda}\right)=E\left(\frac{\partial^2 L}{\partial\lambda\partial\alpha}\right)=\frac{m}{\lambda}-\frac{m\alpha}{\lambda}B(\alpha+1,1),$$
$$E\left(\frac{\partial^2 L}{\partial\beta\partial\lambda}\right)=E\left(\frac{\partial^2 L}{\partial\lambda\partial\beta}\right)=\frac{n}{\lambda}-\frac{n\alpha}{\lambda}B(\beta+1,1),$$
$$E\left(\frac{\partial^2 L}{\partial\lambda^2}\right)=-\frac{m\alpha+n\beta}{\lambda^2}+\frac{m\alpha(\alpha+1)}{\lambda^2}B(\alpha+2,1)+\frac{n\beta(\beta+1)}{\lambda^2}B(\beta+2,1).$$

Theorem 5.1 *As* $m\to\infty$ *and* $n\to\infty$ *and* $\frac{m}{n}\to p$ *then*

$$\left[\sqrt{m}(\hat{\alpha}-\alpha),\sqrt{n}(\hat{\beta}-\beta),\sqrt{m}(\hat{\lambda}-\lambda)\right]\to N_3\left(0,U^{-1}(\alpha,\beta,\lambda)\right)$$

where

$$U(\alpha,\beta,\lambda)=u_{11}0u_{13}0u_{22}u_{23}u_{31}u_{32}u_{33}$$

and

$$u_{11}=-\frac{1}{m}I_{11}=\frac{1}{\alpha^2}$$
$$u_{13}=u_{31}=-\frac{1}{m}I_{13}=-\frac{1}{\lambda}+\frac{\alpha}{\lambda}B(\alpha+1,1)$$
$$u_{22}=-\frac{1}{n}I_{22}=\frac{1}{\beta^2}$$
$$u_{23}=u_{32}=-\frac{\sqrt{p}}{n}I_{23}=-\frac{1}{\lambda\sqrt{p}}+\frac{\alpha}{\lambda\sqrt{p}}B(\beta+1,1)$$
$$u_{33}=-\frac{1}{m}I_{33}=\frac{\alpha+p\beta}{p\lambda^2}-\frac{\alpha(\alpha+1)}{\lambda^2}B(\alpha+2,1)-\frac{\beta(\beta+1)}{p\lambda^2}B(\beta+2,1).$$

Proof: The proof follows from the asymptotic normality of MLE.

Theorem 5.2 *As* $m\to\infty$ *and* $n\to\infty$ *and* $\frac{m}{n}\to p$ *then*

$$\sqrt{m}(\hat{R}-R)\to N(0,B),$$

where

$$B=\frac{1}{k(\alpha+\beta)^4}\left[\beta^2(u_{22}u_{33}-u_{23}^2)-2\alpha\beta\sqrt{p}u_{23}u_{31}+\alpha^2 p(u_{11}u_{33}-u_{13}^2)\right]$$

$$k=u_{11}u_{22}u_{33}-u_{11}u_{23}u_{32}-u_{13}u_{22}u_{31}.$$

Proof. It is clear that

$$Var[\sqrt{m}(\hat{R}-R)]=E[\sqrt{m}(\hat{R}-R)]^2$$
$$=E\left[\frac{\alpha\sqrt{n}(\hat{\beta}-\beta)-\beta\sqrt{m}(\hat{\alpha}-\alpha)}{(\alpha+\beta)(\hat{\alpha}+\hat{\beta})}\right]^2$$

$$=E\left[\frac{\frac{m}{n}\alpha^2[\sqrt{n}(\hat\beta-\beta)]^2+\beta^2[\sqrt{m}(\hat\alpha-\alpha)]^2-\sqrt{\frac{m}{n}}2\alpha\beta[\sqrt{m}(\hat\alpha-\alpha)\sqrt{n}(\hat\beta-\beta)]}{(\alpha+\beta)^2(\hat\alpha+\hat\beta)^2}\right]$$

Using Theorem 5.1, the consistency and asymptotic normality of MLE, the proof is complete.

Note that Theorem 5.2 can be used to construct asymptotic confidence intervals. To compute the confidence interval of R, the variance B needs to be estimated. To estimate it, the empirical Fisher information matrix and the MLEs of α, β and λ are used, as follows;

$$\hat u_{11}=-\frac{1}{m}I_{11}=\frac{1}{\hat\alpha^2}$$

$$\hat u_{13}=\hat u_{31}=-\frac{1}{m}I_{13}=-\frac{1}{\hat\lambda}+\frac{\hat\alpha}{\hat\lambda}B(\hat\alpha+1,1)$$

$$\hat u_{22}=-\frac{1}{n}I_{22}=\frac{1}{\hat\beta^2}$$

$$\hat u_{23}=\hat u_{32}=-\frac{\sqrt{p}}{n}I_{23}=-\frac{1}{\hat\lambda\sqrt{p}}+\frac{\hat\alpha}{\hat\lambda\sqrt{p}}B(\hat\beta+1,1)$$

$$\hat u_{33}=-\frac{1}{m}I_{33}=\frac{\hat\alpha+p\hat\beta}{p\hat\lambda^2}-\frac{\hat\alpha(\hat\alpha+1)}{\hat\lambda^2}B(\hat\alpha+2,1)-\frac{\hat\beta(\hat\beta+1)}{p\hat\lambda^2}B(\hat\beta+2,1).$$

5.3 Bootstrap Confidence Limits

In this subsection, we propose to use two confidence limits based on the parametric bootstrap methods; (i) percentile bootstrap method (we call it from now on as Boot-p) based on the idea of Efron (1982), (ii) bootstrap-t method (we refer it as Boot-t from now on) based on the idea of Hall (1988). We illustrate briefly how to estimate confidence limits of R using both methods.

Boot-p Methods

Step 1: From the sample $\{x_1,.......,x_m\}$ and $\{y_1,.......,y_n\}$, compute $\hat\alpha$, $\hat\beta$ and $\hat\lambda$.

Step 2: Using $\hat\alpha$ and $\hat\lambda$ generate a bootstrap sample $\{x_1^*,.......,x_m^*\}$ and similarly using $\hat\beta$ and $\hat\lambda$ generate a bootstrap sample $\{y_1^*,.......,y_n^*\}$. Based on $\{x_1^*,.......,x_m^*\}$ and $\{y_1^*,.......,y_n^*\}$ compute the bootstrap estimate of R using (5.10), say $\hat R^*$.

Step 3: Repeat step 2, **N** times.

Step 4: Let $G(x)=P(\hat R^*\le x)$, be the cumulative distribution function of $\hat R^*$.

Define $\hat{R}_{Boot-p}(x) = G^{-1}(x)$ for a given x. The approximate $100(1-\gamma)\%$ confidence ineterval of R is given by

$$\left[\hat{R}_{Boot-p}(\frac{\gamma}{2}), \hat{R}_{Boot-p}(1-\frac{\gamma}{2})\right]$$

Bootstrap-t Confidence Limits

Step 1: From the samples $\{x_1,......., x_m\}$ and $\{y_1,......., y_n\}$, compute $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\lambda}$.

Step 2: Using $\hat{\alpha}$ and $\hat{\lambda}$ generate a bootstrap sample $\{x_1^*,......., x_m^*\}$ and similarly using $\hat{\beta}$ and $\hat{\lambda}$ generate a bootstrap sample $\{y_1^*,......., y_n^*\}$. Based on $\{x_1^*,......., x_m^*\}$ and $\{y_1^*,......., y_n^*\}$ compute the bootstrap estimate of R using (5.10), say $\hat{R}^*$ and the following statistic:

$$T^* = \frac{\sqrt{m}(\hat{R}^* - \hat{R})}{\sqrt{V(\hat{R}^*)}},$$

where $V(\hat{R}^*)$ is obtained using the expected Fisher information matrix.

Step 3: Repeat step 2, **N** times.

Step 4: From the T^* values obtained, determine the lower and the upper bound of the $100(1-\gamma)\%$ confidence limits of R as follows: Let $H(x) = P(T^* \le x)$ be the cumulative distribution function of T^*. For a given x, define

$$\hat{R}_{Boot-t} = \hat{R} + m^{-\frac{1}{2}}\sqrt{V(\hat{R})}H^{-1}x .$$

Here also, $V(\hat{R})$ can be computed similarly as for the $V(\hat{R}^*)$. The approximate $100(1-\gamma)\%$ confidence interval of R is given by

$$\left[\hat{R}_{Boot-t}(\frac{\gamma}{2}), \hat{R}_{Boot-t}(1-\frac{\gamma}{2})\right].$$

6. Bayes estimation of R

In this section, we obtain the Bayes estimation of R under assumption that the shape parameters α, β and λ are random variables. We mainly obtain the Bayes estimate of R under the squared error loss using the Gibbs sampling technique. It is assumed that α, β and λ have independent gamma priors with the parameters $(a_1,b_1),(a_2,b_2)$ and (a_3,b_3) respectively. Based on the above assumptions, we have the likelihood function of the observed data as

$$L(data \mid \alpha,\beta,\lambda) = (\alpha\lambda)^m \prod_{i=1}^{m}(\lambda x_i)^{\alpha-1}(1+\lambda x_i)^{-(\alpha+1)}.(\beta\lambda)^n \prod_{j=1}^{n}(\lambda y_j)^{\beta-1}(1+\lambda y_j)^{-(\beta+1)}$$

Therefore, the joint density of the data, α, β and λ can be obtained as

$$L(data,\alpha,\beta,\lambda) = L(data \mid \alpha,\beta,\lambda)\pi(\alpha)\pi(\beta)\pi(\lambda)$$

where $\pi(.)$ is the prior distribution. Therefore, the joint posterior density of α, β and λ given the data is

$$L(\alpha,\beta,\lambda \mid data) = \frac{L(data,\alpha,\beta,\lambda)}{\int_0^\infty\int_0^\infty\int_0^\infty L(data,\alpha,\beta,\lambda)\,d\alpha\,d\beta\,d\lambda}$$

We adopt the Gibbs sampling technique to compute the Bayes estimate of R. The posterior pdfs of α, β and λ are as follows:

$$\alpha \mid \beta,\lambda,data : Gamma\left(a_1+m, b_1 - \sum_{i=1}^{m}\ln\frac{\lambda x_i}{1+\lambda x_i}\right)$$

$$\beta \mid \alpha,\lambda,data : Gamma\left(a_2+n, b_2 - \sum_{j=1}^{n}\ln\frac{\lambda y_j}{1+\lambda y_j}\right)$$

and

$$f_\lambda(\lambda \mid \alpha,\beta,data) \propto \lambda^{a_3+m\alpha+n\beta-1} e^{-\lambda b_3-(\alpha+1)\sum_{i=1}^{m}\ln(1+\lambda x_i)-(\beta+1)\sum_{j=1}^{n}\ln(1+\lambda y_j)}$$

The posterior pdfs of λ are not known, but the plots of them show that they are similar to normal distribution. So to generate random numbers from these distributions, we use the Metropolis-Hastings method with normal proposal distribution. Therefore, the algorithm of Gibbs sampling is as follows:

Step 1: Start with an initial guess $(\alpha^0,\beta^0,\lambda^0)$.

Step 2: Set $t=1$.

Step 3: Using the Metropolis-Hastings, generate $\alpha^{(t)}$ from

$$Gamma\left(a_1+m, b_1 - \sum_{i=1}^{m}\ln\frac{\lambda^{(t-1)}x_i}{1+\lambda^{(t-1)}x_i}\right).$$

Step 4: Using the Metropolis-Hastings, generate $\beta^{(t)}$ from

$$Gamma\left(a_2+n, b_2 - \sum_{j=1}^{n}\ln\frac{\lambda^{(t-1)}y_j}{1+\lambda^{(t-1)}y_j}\right).$$

Step 5: Using the Metropolis-Hastings, generate $\lambda^{(t)}$ from f_λ with the $N(\lambda^{(t-1)},1)$ proposal distribution.

Step 6: Compute R^t from (5.13)

Step 7: Set $t=t+1$.

Step 8: Repeat step 3-7, T times.

Note that in steps 5, we use the Metropolis-Hastings algorithm with $q(\lambda^{(t-1)}, \sigma^2)$ proposal distribution as follows:

1. Let $x = \lambda^{(t-1)}$.
2. Generate y from the proposal distribution q.
3. Let $p(x, y) = min\{1, f_\lambda(y)/f_\lambda(x).q(x)/q(y)\}$.
4. Accept y with the probability $p(x, y)$ or accept x with the probability $1 - p(x, y)$.

Now the approximate posterior mean, and posterior variance of R become

$$\hat{E}(R \mid data) = \frac{1}{T}\sum_{t=1}^{T} R^t$$

and

$$MSE(R \mid data) = \frac{1}{T}\sum_{t=1}^{T}\left(R^t - R\right)^2$$

respectively.

7. Estimation of R in general case

Computing the R when the parameter λ is different for X and Y, is considered in this section. Surles and Padgett (1998, 2001) considered this case also. In Surles and Padgett (2001), there is no exact expression for R, but they presented a bound for it.

7.1 Maximum likelihood estimator of R

Let $X : EFC(\alpha, \lambda_1)$ and $Y : EFC(\beta, \lambda_2)$, where X and Y are independent random variables. Therefore,

$$\begin{aligned} R &= \int_0^\infty P(X < Y \mid Y = y)P(Y = y)dy \\ &= \beta\int_0^\infty t^{\alpha+\beta-1}(1+t)^{-(\beta+1)}\left(\frac{\lambda_2}{\lambda_1} + t\right)^{-\alpha} dt \\ &= \frac{\beta}{\alpha+\beta}\left(\frac{\lambda_2}{\lambda_1}\right)^{\beta} F\left(\beta+1, \alpha+\beta, \alpha+\beta+1, 1-\frac{\lambda_2}{\lambda_1}\right) \end{aligned} \tag{11}$$

where $F(.)$ is the Gauss hypergeometric function, see Grandshteyn and Ryzhik (2000) (formula 9.100).

To compute the MLE of R, Suppose $(X_1, X_2, ...X_m)$ is random sample from $EFC(\alpha, \lambda_1)$ and $(Y_1, Y_2, ...Y_n)$ is random sample from $EFC(\beta, \lambda_2)$. Therefore, the log-likelihood function of the observed samples is

$$\begin{aligned} L(\alpha, \beta, \lambda_1, \lambda_2) &= m\ln\alpha + m\ln\lambda_1 + (\alpha-1)\sum_{i=1}^{m}\ln(\lambda_1 x_i) - (\alpha+1)\sum_{i=1}^{m}\ln(1+\lambda_1 x_i) \\ &+ n\ln\beta + n\ln\lambda_2 + (\beta-1)\sum_{j=1}^{n}\ln(\lambda_2 y_j) - (\beta+1)\sum_{j=1}^{n}\ln(1+\lambda_2 y_j) \end{aligned}$$

The MLE's of α, β, λ_1 and λ_2 say $\hat{\alpha}$, $\hat{\beta}$, $\hat{\lambda}_1$ and $\hat{\lambda}_2$ respectively, can be obtained as the solutions of

$$\frac{\partial L}{\partial \alpha}=0, \frac{\partial L}{\partial \beta}=0, \frac{\partial L}{\partial \lambda_1}=0, \frac{\partial L}{\partial \lambda_2}=0$$

After calculation, we obtain

$$\hat{\alpha}=-\frac{m}{\sum_{i=1}^{m}\ln\frac{\lambda_1 x_i}{1+\lambda_1 x_i}} \tag{12}$$

$$\hat{\beta}=-\frac{n}{\sum_{j=1}^{n}\ln\frac{\lambda_2 y_j}{1+\lambda_2 y_j}} \tag{13}$$

and $\hat{\lambda}_1$ and $\hat{\lambda}_2$ can be obtained as the solution of the non-linear equation

$$g(\lambda_1)=\frac{m}{\lambda_1}-\left(1+\frac{m}{\sum_{i=1}^{m}\ln\frac{\lambda_1 x_i}{1+\lambda_1 x_i}}\right)\times\frac{m}{\lambda_1}-\left(1-\frac{m}{\sum_{i=1}^{m}\ln\frac{\lambda_1 x_i}{1+\lambda_1 x_i}}\right)\times\sum_{i=1}^{m}\frac{x_i}{1+\lambda_1 x_i}=0 \tag{14}$$

and

$$g(\lambda_2)=\frac{n}{\lambda_2}-\left(1+\frac{n}{\sum_{j=1}^{n}\ln\frac{\lambda_2 y_j}{1+\lambda_2 y_j}}\right)\times\frac{n}{\lambda_2}-\left(1-\frac{n}{\sum_{j=1}^{n}\ln\frac{\lambda_2 y_j}{1+\lambda_2 y_j}}\right)\times\sum_{j=1}^{n}\frac{y_j}{1+\lambda_2 y_j}=0 \tag{15}$$

respectively.

By invariance property of the ML estimators, the MLE of R becomes

$$\hat{R}=\frac{\hat{\beta}}{\hat{\alpha}+\hat{\beta}}\left(\frac{\hat{\lambda}_2}{\hat{\lambda}_1}\right)^{\hat{\beta}} F\left(\hat{\beta}+1, \hat{\alpha}+\hat{\beta}, \hat{\alpha}+\hat{\beta}+1, 1-\frac{\hat{\lambda}_2}{\hat{\lambda}_1}\right) \tag{16}$$

7.2 Asymptotic distribution

The asymptotic distribution of $\hat{\theta}=(\hat{\alpha}, \hat{\beta}, \hat{\lambda}_1, \hat{\lambda}_2)$ is to be obtained using the approach of Theorems 5.1 and hence the asymptotic distribution of $\hat{R}$ could be obtained using the approach of 5.2. We denote the expected Fisher information matrix of $\theta=(\alpha, \beta, \lambda_1, \lambda_2)$ as $I(\theta)=(I_{ij},(\theta))$; $i, j=1,2,3,4$. Therefore

$$I(\theta)=-E\left(\frac{\partial^2 L}{\partial \alpha^2}\right)E\left(\frac{\partial^2 L}{\partial \alpha \partial \beta}\right)E\left(\frac{\partial^2 L}{\partial \alpha \partial \lambda_1}\right)E\left(\frac{\partial^2 L}{\partial \alpha \partial \lambda_2}\right)E\left(\frac{\partial^2 L}{\partial \beta \partial \alpha}\right)E\left(\frac{\partial^2 L}{\partial^2 \beta}\right)$$
$$E\left(\frac{\partial^2 L}{\partial \beta \partial \lambda_1}\right)E\left(\frac{\partial^2 L}{\partial \beta \partial \lambda_2}\right)E\left(\frac{\partial^2 L}{\partial \lambda_1 \partial \alpha}\right)E\left(\frac{\partial^2 L}{\partial \lambda_1 \partial \beta}\right)E\left(\frac{\partial^2 L}{\partial \lambda_1^2}\right)E\left(\frac{\partial^2 L}{\partial \lambda_1 \lambda_2}\right)$$
$$E\left(\frac{\partial^2 L}{\partial \lambda_2 \partial \alpha}\right)E\left(\frac{\partial^2 L}{\partial \lambda_2 \partial \beta}\right)E\left(\frac{\partial^2 L}{\partial \lambda_2 \lambda_1}\right)E\left(\frac{\partial^2 L}{\partial \lambda_2^2}\right).$$

It is easy to see that

$$E\left(\frac{\partial^2 L}{\partial\alpha^2}\right)=-\frac{m}{\alpha^2},\ E\left(\frac{\partial^2 L}{\partial\alpha\partial\lambda_1}\right)=\frac{m}{\lambda_1}-\frac{m\alpha}{\lambda_1}B(\alpha+1,1),\ E\left(\frac{\partial^2 L}{\partial\alpha\partial\beta}\right)=E\left(\frac{\partial^2 L}{\partial\alpha\partial\lambda_2}\right)=0,$$

$$E\left(\frac{\partial^2 L}{\partial\beta^2}\right)=-\frac{n}{\beta^2},\ E\left(\frac{\partial^2 L}{\partial\alpha\partial\beta}\right)=E\left(\frac{\partial^2 L}{\partial\beta\partial\lambda_1}\right)=0,\ E\left(\frac{\partial^2 L}{\partial\alpha\partial\lambda_2}\right)=\frac{n}{\lambda_2}-\frac{n\beta}{\lambda_2}B(\beta+1,1),$$

$$E\left(\frac{\partial^2 L}{\partial\lambda_1^2}\right)=-\frac{m}{\lambda_1^2}+(\alpha-1)\frac{m}{\lambda_1^2}-\frac{m\alpha(\alpha+1)}{\lambda_1^2}B(\alpha+2,1),$$

$$E\left(\frac{\partial^2 L}{\partial\lambda_1\partial\alpha}\right)=\frac{m}{\lambda_1}-\frac{m\alpha}{\lambda_1}B(\alpha+1,1),$$

$$E\left(\frac{\partial^2 L}{\partial\lambda_1\partial\beta}\right)=E\left(\frac{\partial^2 L}{\partial\lambda_1\partial\lambda_2}\right)=0,$$

$$E\left(\frac{\partial^2 L}{\partial\lambda_2\partial\beta}\right)=\frac{n}{\lambda_2}-\frac{n\beta}{\lambda_2}B(\beta+1,1),\ E\left(\frac{\partial^2 L}{\partial\lambda_2\partial\beta}\right)=E\left(\frac{\partial^2 L}{\partial\lambda_2\partial\lambda_1}\right)=0.$$

Based on the above Fisher information matrix, it is possible to present confidence intervals of R based on the percentile bootstrap and bootstrap-t method. They are very similar to those mentioned in Section 5.3. The Bayes estimate of R could be found out using the Metropolis-Hastings algorithm assuming two independent gamma priors for λ_1 and λ_2 following the same procedure as in section 6. For saving space, we omit them.

8. Simulation and discussion

In this section we present some results based on the Monte Carlo simulations to compare the performance of different methods. All computations were performed using R-software and these are available on request from the corresponding author. We consider to draw inference on R when the baseline distribution of exponentiated distribution is (a) known and (b) unknown through parameters. In our study we take sample sizes $(m,n)=(15,15),(20,25),(25,25),(50,50)$ and take $(\alpha,\beta)=(3,0.4),(0.8,0.4),(1,1)\ (0.4,0.8),(0.4,3)$ respectively. For the unknown case, we take $\lambda=0.5,1.5,1,3,2$. All the results are based on 1000 replications.

We have used the initial estimate to be 1 and the iterative process stops when the difference between the two consecutive iterates are less than 10^{-4} for both α and β using the iterative equations. We choose the initial estimate to be 1, since for that value exponentiated distribution reduces to the baseline distribution. We obtain the MLE of R substituting $\hat{\alpha}$ and $\hat{\beta}$ in the expression.

First we consider the case when the baseline distribution is completely known. We report the estimates of R, $\hat{R}_1$, $\hat{R}_2$ and $\hat{R}_4$ using the MLE, UMVUE and empirical Bayes procedure assuming conjugate priors (in each cell first, second and third row respectively), and the average biases and mean squared errors (MSEs) of R in tables 1-4 over 1000 replications. We also compute the 95% confidence limits of R, both approximate [(L_1, U_1)] and exact [(L_2, U_2)], and Bayesian credible intervals [(L_3, U_3)], and hence report average confidence lengths and coverage proportions (cp) based on 1000 replications in table 5.

Some of the points are quite clear from this experiment. The performance of the MLEs are quite satisfactory with respect to the UMVUEs in terms of biases and MSEs. Though differences are marginal, the MLEs have computational ease. Since for the MLE, the exact distribution is known therefore it can be used to construct confidence intervals. As expected, with the help of prior information, the Bayes estimates of R perform better than the MLEs and UMVUEs. For all the methods, when m and n increase, the average biases and MSEs decrease. The Bayesian interval (with conjugate priors), (L_3, U_3) has the shortest average length for all values of R and (m, n). The average lengths of all intervals decrease as m, n increase. The interval (L_2, U_2) has the largest average probability coverage which is approximately the anticipated 95%. The interval (L_3, U_3) has the smallest average probability coverage and it is far from 0.95. The average probability coverage of (L_1, U_1) is approximately 0.95 for large m, n.

For the second case, we have assumed that the common scale parameter λ of the folded Crammer distribution is unknown. From the sample, we compute the estimate of λ using the iterative algorithm (5.9). Once we estimate λ, we obtain the MLE of R using (5.10). We report the average biases and mean squared errors (MSEs) in table 6, and report 95% confidence intervals based on the delta, Boot-p and Boot-t methods in table 7 using 1000 bootstrap replications in both cases. The performance of the MLEs are quite satisfactory in terms of biases and MSEs. It is observed that when m, n increase, the MSEs decrease. It verifies the consistency property of the MLE of R. The confidence intervals based on the delta method work quite well, as it offers much narrower intervals, unless the sample size is very small, say (15, 15). For very small samples, the Boot-t confidence intervals perform well.

We do not have any prior information on R, therefore, we prefer to use the non-informative prior to compute different Bayes estimates. Since the non-informative prior, i.e. $a_1 = a_2 = b_1 = b_2 = 0$ provides prior distributions which are not proper, we adopt the suggestion of Congdon (2001, p.20) and Kundu and Gupta (2005), i.e. choose $a_1 = a_2 = b_1 = b_2 = 0.0001$, which are almost like Jeffreys prior, but they are proper. Under the same prior distributions, we compute the Bayes estimate of α and β and have approximate the Bayes estimates of R under the squared error loss function. To generate random observations from the posterior distributions of α, β and λ, we use the Metropolis-Hastings method. The algorithms of Gibbs sampling is described in section 6. The burn in sample in each case is taken

5000. The results are reported in table 8. It is observed that as expected when m, n increase then the average biases and the MSEs decrease.

The calculations for general case will be in similar way as have been done in second case with some modifications. That is why we omit this portion here.

9. Concluding Remark

In this article, we have discussed inference problem of $R = P(X < Y)$ for exponentiated family of distributions. This family is obtained by adding a parameter to the exponent of a distribution function (called a baseline distribution function) to make resulting distribution richer and more flexible for modeling data. We have considered the cases when the baseline distribution is known or unknown through parameter(s). At first we look into inference of R in more general set up under any known baseline distribution not necessarily restricted to the location-scale family. Based on the simulation results, we recommend to use the MLE for R from the frequentist view point. From the Bayesian view point, the Bayes estimate of R is also recommended with conjugate priors. The confidence interval (L_2, U_2) based on the exact distribution of the MLE is recommended for its largest average probability coverage, even though the credible interval (L_3, U_3) has the shortest average length. When the baseline distribution is unknown through parameter(s), in particular for the folded Crammer distribution, it is observed that the MLE works quite well. The confidence intervals based on the delta method is recommended to use. For very small samples, the Boot-t confidence intervals perform well and it is recommended to use.

Acknowledgment

Sudhir Murmu likes to thank University Grants Commission for financial support in the form of Rajiv Gandhi National Fellowship (Sanction No. 14-2(ST)/2008(SA-III)).

References

1. Ali, M.M., Woo, J. and Pal, M. (2004). Inference on reliability $P(Y < X)$ in a two-parameter exponential distribution. *International Journal of Statistical Sciences*, 3, 119-125.
2. Awad, A. M. and Gharraf, M. K.(1986). Estimation of $P(Y < X)$ in the Burr case: A comparative study. *Comm. Statist. Simul. Computat.*, 15, 189-203.
3. Banerjee, T. and Biswas, A. (2003). A new formulation of stress-strength reliability in a regression setup. *Journal of Statistical Planning and Inference*, 112, 147-157.
4. Beg, M.A. (1980). On the Estimation of $Pr\{Y < X\}$ for the Two-Parameter Exponential Distribution. *Metrika*, 27, 29-34.

5. Bhattacharyya, G.K. and Johnson, R.A. (1974). Estimation of reliability in a multicomponent stress-strength model. *Journal of the American Statistical Association*, 69, 966-970.
6. Birnbaum, Z.W. (1956). On a use of the Mann-Whitney statistic. *Proceedings of the Third Berkeley Symposium on Mathematics, Statistics and Probability*, 1, 13-17.
7. Birnbaum, Z.W. and McCarthy, R.C. (1958). A distribution free upper confidence bound for $Pr(Y < X)$ based on independent samples of X and Y. *Annals of Mathematical Statistics*, 29, 558-562.
8. Blight, B. J. N. and Rao, P. V. (1974). The convergence of Bhattacharya bounds. *Biometrika*, 61, 137-142.
9. Church, J.D. and Harris, B. (1970). The estimation of reliability from stress-strength relationship. *Technometrics*, 12, 49-54.
10. Congdon, P. (2001): Bayesian Statistical Modeling. John Wiley, New York.
11. Crame'r, E. and Kamps, U. (1997). A note on UMVUE of $Pr(Y < X)$ in the Exponential case. *Commun. Statist.-Theory and Methods*, 26, 1051-1055.
12. Efron, B.(1982): The jackknife, the bootstrap and other resampling plans. *In: CBMS-NSF Regional Conference Series in Applied Mathematics*, 38, SIAM, Phiadelphia, PA.
13. Enis, P. and Geisser, S. (1971). Estimation of the Probability that $Y < X$. *Journal of the American Statistical Association*, 66, 162-168.
14. Ghosh, J. K. and Sathe, Y. S. (1987). Convergence of Bhattacharya bounds-revisited. *Sankhya, Ser.A*, 49, 37-42.
15. Govindarajulu, Z. (1967). Two sided confidence limits for $P(Y < X)$ for normal samples of X and Y. *Sankhya B*, 29, 35-40.
16. Govindarajulu, Z. (1968). Distribution-free confidence bounds for $P(Y < X)$. *Ann. Inst. Statist. Math.*, 20, 229-238.
17. Grandshteyn, I. S. and Ryzhik, I. M. (2000). *Table of Integrals, Series, and Products.*, Academic Press, USA.
18. Gupta, R. D. and Gupta, R. C. (2008). Analyzing skewed data by power normal model. *Test*, 17, 197-210.
19. Guttman, I., Johnson, R.A., Bhattacharyya, G.K. and Reiser, B. (1988). Confidence limits for stress-strength models with explanatory variables. *Technometrics*, 30, 161-168.
20. Hall, P. (1988): Theoretical comparision of bootstrap confidence intervals. *Annals of Statistics*, 16, 927-953.
21. Ivshin, V.V. (1996). Unbiased Estimators of $P(X < Y)$ and their variances in the case of Uniform and Two-Parameter Exponential Distributions. *Journal of Mathematical Sciences*, 81, 2790-2793.
22. Iwase, K. (1987). On UMVUE Estimators of $Pr(Y < X)$ in the Two-Parameter Exponential case. *Mem. Fac. Hiroshima Univ.*, 9, 21-24.

23. Johnson, R.A. (1988). Stress-strength models for reliability. In: Krishnaiah, P.R. and Rao, C.R. (Eds.), *Handbook of Statistics*, 7, North-Holand, Amsterdam.

24. Kakade, C.S. and Shirke, D.T. (2007). Inference of $P(Y < X)$ for Exponentiated Scale Family of Distributions. *Journal of the Indian Statistical Association*, 45, 13-31.

25. Kakade, C.S. and Shirke, D.T. and Kundu, D. (2008). Inference for $P(Y < X)$ in Exponentiated Gumbel Distribution. *J. Stat. and Appl.*, 3, 121-133.

26. Kelley, G.D., Kelley, J.A. and Schucany, W. (1976). Efficient Estimation of $Pr(Y < X)$ in the Exponential case. *Technometrics*, 18, 359-360.

27. Kotz, S., Lumelskii, Y. and Pensky, M. (2003). The Stress-Strength Model and Its Generalizations: Theory and Applications. River Edge, NJ: World Scientific Publishing Co.

28. Kundu, D. and Gupta, R.D. (2005). Estimation of $P[Y < X]$ for generalized exponential distribution. *Metrika*, 61, 291-308.

29. Lindely, D. V. (1969). *Introduction to Probability and Statistics from a Bayesian Viewpoint*, 1, Cambridge: Cambridge University Press.

30. Mazumdar, M. (1970). Some estimates of reliability using interference theory. *Naval Research Logistic Quarterly*, 17, 159-165.

31. McCool, J.I. (1991). Inference on $P(X < Y)$ in the Weibull case. *Commun. Statist.-Simula. and Comp.*, 20, 129-148.

32. Mokhlis, N. A. (2005). Reliability of a stress-strength model with Burr type III distributions. *Comm. in Statist.-Theory and Methods*, 34, 1643-1647.

33. Nadarajah, S. (2004). Reliability for lifetime distributions. *Mathematical and Computer Modeling*, 37, 683-688.

34. Owen, D.B., Craswell, K.J. and Hanson, D.L. (1964). Nonparametric upper confidence bound for $P(Y < X)$ and confidence limits for $P(Y < X)$ when X and Y are normal. *Journal of the American Statistical Association*, 59, 906-924.

35. Pal, M., Ali, M.M. and Woo, J. (2005). Estimation and testing of $P(Y > X)$ in two-parameter exponential distributions. *Statistics*, 39, 415-428.

36. Raqab, M.Z. and Kundu, D. (2005). Comparison of different estimators of $P(Y < X)$ for a scaled Burr type X distribution. *Comm. in Staist.-Simula. and Comp.*, 34, 465-483.

37. Reiser, B. and Farragi, D. (1994). Confidence bounds for $P(a'x > b'y)$. *Statistics*, 25, 107-111.

38. Rezaei, S., Tahmasbi, R. and Mahmoodi, M. (2010). Estimation of $P[Y < X]$ for genralized Pareto distribution. *Journal of Statistical Planning and Inference*, 140, 480-494.

39. Sathe, Y.S. and Shah, S.P. (1981). On estimating of $Pr(X>Y)$ for the exponential distribution. *Communications in Statistics- Theory and Methods*, 10, 39-47.
40. Shah, S.P. and Sathe, Y.S. (1982). Erratum: On estimating of $Pr(X>Y)$ for the exponential distribution. *Communications in Statistics- Theory and Methods*, 11, 2357.
41. Sinha, B.K. and Zielinski, R. (1997). Estimating $Pr(X>Y)$ in exponential model revisited. *Statistics*, 29, 299-316.
42. Surles, J. G. and Padgett, W. J. (1998). Inference for $P(Y<X)$ in the Burr type X model. *Journal of Applied Statistical Sciences*, 7, 225-238.
43. Surles, J. G. and Padgett, W. J. (2001). Inference for reliability and stress-strength for a scaled Burr type X distribution. *Lifetime Data Analysis*, 7, 187-200.
44. Tong, H. (1974, 1975). A note on the estimation of $Pr(Y<X)$ in the exponential case. *Technometrics*, 16, 625, 17, Errata:, 395.
45. Weerahandi, S. and Johnson, R.A. (1992). Testing reliability in a stress-strength model when X and Y are normally distributed. *Technometrics*, 34, 83-91.

Table 1: Biases and Mean Squared Errors of estimates of R when baseline distributions are completely known and $m=n=15$

α, β	R	$\hat{R}$	Bias	MSE
		0.1241110	0.0064639	0.0018093
3, 0.4	0.1176471	0.1186844	0.0010373	0.0016983
		0.1266500	0.0090028	0.0018782
		0.3400351	0.0067017	0.0054612
0.8, 0.4	0.3333333	0.3351727	0.0018393	0.0056811
		0.3421717	0.0088383	0.0053780
		0.5028957	0.0028957	0.0081285
1, 1	0.5	0.5030025	0.0030025	0.0086792
		0.5028498	0.0028498	0.0078948
		0.6612367	-0.0054299	0.0069237
0.4, 0.8	0.6666667	0.6660255	-0.0006411	0.0072197
		0.6591277	-0.0075389	0.0068062
		0.8775772	-0.0047756	0.0015257
0.4, 3	0.882353	0.8829957	0.0006427	0.0014362
		0.8750404	-0.0073125	0.0015850

Table 2: Biases and Mean Squared Errors of estimates of R when baseline distributions are completely known and $m = 20, n = 25$

α, β	R	$\hat{R}$	Bias	MSE
		0.1216647	0.0040176	0.0010981
3, 0.4	0.1176471	0.1185643	0.0009172	0.0010577
		0.1231408	0.0054937	0.0011230
		0.3385388	0.0052054	0.0042621
0.8, 0.4	0.3333333	0.3363988	0.0030655	0.0043901
		0.3394884	0.0061550	0.0042053
		0.4941221	-0.0058779	0.0056908
1, 1	0.5	0.4952429	-0.0047570	0.0059368
		0.4935758	-0.0064241	0.0055822
		0.6612858	-0.0053809	0.0045092
0.4, 0.8	0.6666667	0.6656551	-0.0010115	0.0046101
		0.6592534	-0.0074132	0.0044756
		0.8786907	-0.0036622	0.0010396
0.4, 3	0.882353	0.8828578	0.0005048	0.0009864
		0.876689	-0.0056638	0.0010768

Table 3: Biases and Mean Squared Errors of estimates of R when baseline distributions are completely known and $m = n = 25$

α, β	R	$\hat{R}$	Bias	MSE
		0.1206998	0.0030527	0.0008488
3, 0.4	0.1176471	0.1174729	-0.0001741	0.0008160
		0.1222521	0.0046050	0.0008715
		0.3352984	0.0019650	0.0037183
0.8, 0.4	0.3333333	0.3323512	-0.0009821	0.0038204
		0.3366616	0.0033282	0.0036765
		0.4983102	-0.0016897	0.0049426
1, 1	0.5	0.4982778	-0.0017221	0.0051422
		0.4983251	-0.0016749	0.0048524
		0.661385	-0.0052816	0.0037909
0.4, 0.8	0.6666667	0.664281	-0.0023856	0.0038718
		0.6600459	-0.0066207	0.0037586
		0.879125	-0.0032279	0.0009735
0.4, 3	0.882353	0.882344	-8.9956×10^{-6}	0.0009370
		0.8775767	-0.0047762	0.0009977

Table 4: Biases and Mean Squared Errors of estimates of R when baseline distributions are completely known and $m=n=50$

α, β	R	$\hat{R}$	Bias	MSE
		0.1205489	0.0029018	0.0004714
3, 0.4	0.1176471	0.1189402	0.0012931	0.0004581
		0.1213379	0.0036908	0.0004798
		0.3361455	0.0028121	0.0019635
0.8, 0.4	0.3333333	0.3346827	0.0013493	0.0019840
		0.3368485	0.0035151	0.0019551
		0.500021	2.0961×10^{-5}	0.0025857
1, 1	0.5	0.5000206	2.0567×10^{-5}	0.0026374
		0.5000211	2.1138×10^{-5}	0.0025612
		0.6650935	-0.0015731	0.0019151
0.4, 0.8	0.6666667	0.6665655	-0.0001011	0.0019387
		0.664386	-0.0022806	0.0019052
		0.881857	-0.0004959	0.0004406
0.4, 3	0.882353	0.8834483	0.0010953	0.0004352
		0.8810763	-0.0012766	0.0004451

Table 5: Average length of the intervals and coverage probability, $1-\gamma=0.95$ when baseline distributions are completely known

R	$m=15, n=15$		$m=20, n=25$		$m=25, n=25$		$m=50, n=50$	
	Avg. length	cp	Avg. length	cp	Avg. length	cp	Avg. length	cp
	0.1530690	0.931	0.1243955	0.938	0.1167395	0.951	0.0827526	0.94
0.1176471	0.1610088	0.948	0.1107205	0.906	0.1211200	0.961	0.0849714	0.95
	0.1088123	0.823	0.0749692	0.719	0.0828934	0.846	0.0586488	0.839
	0.3134595	0.956	0.2583569	0.941	0.2429865	0.941	0.1734145	0.942
0.3333333	0.3114038	0.973	0.2401906	0.906	0.2426449	0.962	0.1739129	0.943
	0.2170539	0.868	0.16703	0.745	0.1696754	0.847	0.1218414	0.831
	0.3462040	0.932	0.2873023	0.93	0.2717007	0.933	0.1939692	0.936
0.5	0.3402838	0.95	0.2808955	0.874	0.2692232	0.95	0.1936650	0.941
	0.2385059	0.828	0.1969246	0.703	0.1890375	0.819	0.1360222	0.837
	0.3107594	0.908	0.2581351	0.935	0.2441317	0.944	0.1731287	0.946
0.6666667	0.308911	0.942	0.2712599	0.883	0.2436993	0.949	0.1736381	0.949
	0.2152449	0.812	0.1901887	0.712	0.1704457	0.843	0.1216441	0.829
	0.1516271	0.934	0.1241446	0.932	0.1167498	0.94	0.0813341	0.941
0.882353	0.1596477	0.954	0.1494543	0.895	0.1211098	0.939	0.0835561	0.957
	0.107836	0.833	0.1028255	0.721	0.0828937	0.832	0.0576519	0.822

In each cell first, second and third row represent for $A=(L_1,U_1)$, $B=(L_2,U_2)$ and $C=(L_3,U_3)$.

Table 6: Biases and Mean Squared Errors of estimates of R when baseline distribution is unknown through parameter.

m, n	α, β, λ	R	$\hat{R}$	Bias	MSE
	3, 0.4, 0.5	0.117647	0.110523	-0.007123	0.002425
	0.8, 0.4, 1.5	0.333333	0.328390	-0.004943	0.007279
15, 15	1, 1, 1	0.5	0.501738	0.001738	0.009832
	0.4, 0.8, 3	0.666667	0.676385	0.009718	0.008039
	0.4, 3, 2	0.882353	0.891497	0.009144	0.002609
	3, 0.4, 0.5	0.117647	0.113055	-0.004591	0.001633
	0.8, 0.4, 1.5	0.333333	0.328322	-0.0050107	0.004883
20, 25	1, 1, 1	0.5	0.497996	-0.0020038	0.006136
	0.4, 0.8, 3	0.666667	0.667912	0.001245	0.005218
	0.4, 3, 2	0.882353	0.887199	0.004846	0.001874
	3, 0.4, 0.5	0.117647	0.112801	-0.004845	0.001438
	0.8, 0.4, 1.5	0.333333	0.327344	-0.005988	0.004682
25, 25	1, 1, 1	0.5	0.496659	-0.0033405	0.005591
	0.4, 0.8, 3	0.666667	0.668385	0.001718	0.004698
	0.4, 3, 2	0.882353	0.886460	0.004107	0.001508
	3, 0.4, 0.5	0.117647	0.117314	-0.000332	0.000762
	0.8, 0.4, 1.5	0.333333	0.330085	-0.003248	0.002126
50, 50	1, 1, 1	0.5	0.499593	-0.000406	0.002577
	0.4, 0.8, 3	0.666667	0.665511	-0.001155	0.002120
	0.4, 3, 2	0.882353	0.886149	0.003796	0.000766

Table 7: Confidence Intervals of R when baseline distribution is unknown through parameter

m, n	R	CI_d	CI_{boot-p}	CI_{boot-t}
	0.117647	(0.033603, 0.187443)	(0.028184, 0.197734)	(0.082433, 0.193374)
	0.333333	(0.172729, 0.484051)	(0.117986, 0.444176)	(0.259907, 0.407828)
15, 15	0.5	(0.329508, 0.673968)	(0.293958, 0.687201)	(0.349942, 0.580655)
	0.666667	(0.524033, 0.828736)	(0.458024, 0.815290)	(0.445414, 0.834927)
	0.882353	(0.822886, 0.960108)	(0.665628, 0.934786)	(0.736418, 0.938587)
	0.117647	(0.058805, 0.167305)	(0.027543, 0.143901)	(0.074772, 0.324875)
	0.333333	(0.213413, 0.443231)	(0.174119, 0.444362)	(0.189632, 0.616123)
20, 25	0.5	(0.369590, 0.626402)	(0.382204, 0.682432)	(0.196976, 0.789454)
	0.666667	(0.552942, 0.782882)	(0.588276, 0.857327)	(0.394075, 0.752330)
	0.882353	(0.834471, 0.939928)	(0.820402, 0.962061)	(0.833712, 0.947217)
	0.117647	(0.051065, 0.174537)	(0.060964, 0.238650)	(0.024918, 0.565707)
	0.333333	(0.205569, 0.449119)	(0.215327, 0.491937)	(0.156457, 0.559833)
25, 25	0.5	(0.360983, 0.632335)	(0.369239, 0.657253)	(0.294208, 0.618601)
	0.666667	(0.547042, 0.787928)	(0.619468, 0.861645)	(0.554524, 0.936866)
	0.882353	(0.830453, 0.942468)	(0.812885, 0.956534)	(0.753327, 0.982590)
	0.117647	(0.071726, 0.162902)	(0.038928, 0.118739)	(0.068552, 0.279955)
	0.333333	(0.242745, 0.417425)	(0.246952, 0.428971)	(0.280111, 0.459289)
50, 50	0.5	(0.402536, 0.596650)	(0.404521, 0.610471)	(0.464106, 0.671751)
	0.666667	(0.578313, 0.752709)	(0.597613, 0.784652)	(0.544028, 0.780716)
	0.882353	(0.846121, 0.926178)	(0.762388, 0.896096)	(0.604966, 0.998084)

Table 8: Biases and Mean Squared Errors of Bayes estimates of R when baseline distribution is unknown through parameter.

m, n	α, β, λ	R	$\hat{R}$	Bias	MSE
	3, 0.4, 0.5	0.117647	0.107534	-0.010113	0.001164
	0.8, 0.4, 1.5	0.333333	0.303510	-0.029823	0.006342
15, 15	1, 1, 1	0.5	0.498338	-0.001661	0.007842
	0.4, 0.8, 3	0.666667	0.641738	-0.0249270	0.008540
	0.4, 3, 2	0.882353	0.775178	-0.107174	0.015074
	3, 0.4, 0.5	0.117647	0.147594	0.029947	0.002248
	0.8, 0.4, 1.5	0.333333	0.233583	-0.099749	0.013007
20, 25	1, 1, 1	0.5	0.498658	-0.001341	0.006335
	0.4, 0.8, 3	0.666667	0.642149	-0.024516	0.005845
	0.4, 3, 2	0.882353	0.892355	0.010024	0.000832
	3, 0.4, 0.5	0.117647	0.040825	-0.076821	0.006040
	0.8, 0.4, 1.5	0.333333	0.355480	0.022147	0.005181
25, 25	1, 1, 1	0.5	0.516483	0.016483	0.005083
	0.4, 0.8, 3	0.666667	0.525415	-0.014125	0.024947
	0.4, 3, 2	0.882353	0.776394	-0.105959	0.013335
	3, 0.4, 0.5	0.117647	0.064869	-0.052777	0.002915
	0.8, 0.4, 1.5	0.333333	0.357916	0.024582	0.002537
50, 50	1, 1, 1	0.5	0.480819	-0.019180	0.002637
	0.4, 0.8, 3	0.666667	0.630305	-0.036361	0.003143
	0.4, 3, 2	0.882353	0.860816	-0.021536	0.000987