

Estimation on dependent right censoring scheme in an ordinary bivariate geometric distribution

Citation for published version (APA):

Davarzani, N., Golpalvar, L., Parsian, A., & Peeters, R. (2017). Estimation on dependent right censoring scheme in an ordinary bivariate geometric distribution. *Journal of Applied Statistics*, 44(8), 1369-1384. <https://doi.org/10.1080/02664763.2016.1206064>

Document status and date:

Published: 01/01/2017

DOI:

[10.1080/02664763.2016.1206064](https://doi.org/10.1080/02664763.2016.1206064)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Estimation on dependent right censoring scheme in an ordinary bivariate geometric distribution

N. Davarzani, L. Golparvar, A. Parsian & R. Peeters

To cite this article: N. Davarzani, L. Golparvar, A. Parsian & R. Peeters (2017) Estimation on dependent right censoring scheme in an ordinary bivariate geometric distribution, Journal of Applied Statistics, 44:8, 1369-1384, DOI: [10.1080/02664763.2016.1206064](https://doi.org/10.1080/02664763.2016.1206064)

To link to this article: <https://doi.org/10.1080/02664763.2016.1206064>

 View supplementary material [↗](#)

 Published online: 13 Jul 2016.

 Submit your article to this journal [↗](#)

 Article views: 154

 View related articles [↗](#)

 View Crossmark data [↗](#)

 Citing articles: 1 View citing articles [↗](#)



Estimation on dependent right censoring scheme in an ordinary bivariate geometric distribution

N. Davarzani^a, L. Golparvar^b, A. Parsian^c and R. Peeters^a

^aDepartment of Data Science and Knowledge Engineering, Maastricht University, Maastricht, Netherlands;

^bDepartment of Mathematics and Statistics, McGill University, Montreal, Canada; ^cSchool of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, Iran

ABSTRACT

Discrete lifetime data are very common in engineering and medical researches. In many cases the lifetime is censored at a random or predetermined time and we do not know the complete survival time. There are many situations that the lifetime variable could be dependent on the time of censoring. In this paper we propose the dependent right censoring scheme in discrete setup when the lifetime and censoring variables have a bivariate geometric distribution. We obtain the maximum likelihood estimators of the unknown parameters with their risks in closed forms. The Bayes estimators as well as the constrained Bayes estimates of the unknown parameters under the squared error loss function are also obtained. We considered an extension to the case where covariates are present along with the data. Finally we provided a simulation study and an illustrative example with a real data.

ARTICLE HISTORY

Received 18 October 2015

Accepted 22 June 2016

KEYWORDS

Bayes estimation; bivariate geometric distribution; dependent right censoring; discrete lifetime

AMS SUBJECT CLASSIFICATION

62C10; 62F15; 62N01; 62N02

1. Introduction

In the context of reliability and lifetime data analysis, lifetime variables are usually assumed to be nonnegative continuous random variables. Very often, lifetime values are available to the investigator in discrete form because they are measured discretely, that is, there is rounding to some small measurement unit. For instance, when we deal with the incubation period of a disease such as AIDS, or when the outcome variable of interest is the remission time of cancer, the time to event could be counted in number of weeks, months or so forth, see [24]. Sometimes the lifetime variables are discrete in theory meaning that time periods to the event of interest are countable instead of continuous. For instance, consider a study in which women who stop using oral contraception are followed until pregnancy. The outcome is defined as the number of menstrual cycles until pregnancy, see [15] for more examples.

In survival analysis, censoring is a key analytical problem that arises in a variety of ways. The most common type of censoring in both medical and engineering experiments is the right censoring. Statistical inference based on right censorship has been discussed

CONTACT A. Parsian  ahmad_p@khayam.ut.ac.ir

 Supplemental data for this article can be accessed here. <http://dx.doi.org/10.1080/02664763.2016.1206064>

extensively in survival analysis literature. For more details see, [1,4,16,22] and references therein.

Despite the high frequency of discrete measuring of lifetime values, in both clinical trials and engineering experiments, just a few works addressed the censoring in discrete setup. Rezaei and Arghami [31], obtained the maximum likelihood estimators (MLEs) of the parameters while the lifetime variables has an Exponential-geometric distribution in the presence of right censoring, Davarzani and Parsian [10], proposed the Middle censoring in a discrete setup and Davarzani and Parsian [11], considered an independent right censoring when both the lifetime and censoring variables have geometric distributions. Jammalamadaka and Leong [17] considered analysis of discrete lifetime data under middle-censoring in the presence of covariates.

Usually, lifetime variables are assumed to be independent from the censoring variables. Whereas, there are many situations that the lifetime variable could be dependent on the time of censoring. For instance, in some clinical trials, potentially harmful therapies (e.g. radiation or chemotherapy) may have side effects on patients, depending on the toleration of patients and the harmful degree of the therapy. Then patients may need to be withdrawn from the clinic. In such cases the time of withdrawal depends on the risk of time-to-event (e.g. death). In some studies, clinicians are interested to see the time to death from a particular disease (main disease) but the patients died from another competing disease that could be correlated with the main disease. In this case the event of interest (time to death from the main disease) is censored at the time to death from the competing disease. In industrial experiments, it may occur that a component is taken away from the experiment (is censored) because it shows some sign of future failure.

Recently, dependent censoring has been discussed in a few works in the literature in continuous setup. In most of them the dependence between the censoring and lifetimes variable modeled by allocating a known bivariate distribution to the lifetime and censoring variables or by applying a Copula functions. Jiang *et al.* [18] considered the problem of right censoring in a semi-parametric model where the dependence between the censoring mechanism and lifetimes is modeled by a gamma frailty copula. Braekers and Veraverbeke [6], obtained a copula-graphic estimator for the conditional survival function. Davarzani and Parsian [9], considered dependent right censoring where the lifetime and censoring variables have a Marshall–Olkin bivariate exponential (MOBE) distribution. Davarzani *et al.* [12], dealt with a dependent middle censoring in a dependent setup using MOBE distribution.

In this paper we propose the dependent right censoring scheme in discrete setup using a discrete bivariate distribution. Several bivariate discrete models have been proposed in statistics to represent lifetime data, which can be applied in this regard. Basu and Dhar [5] proposed a bivariate geometric distribution. Arnold [2] introduced multivariate geometric distributions which showed that it leads in a natural way to the Marshall–Olkin multivariate exponential distribution [28]. Nair and Nair [29] studied the characterizations of the geometric distributions. Basu and Dhar [5] introduced a bivariate geometric model which is analog to the bivariate distribution of Marshall and Olkin [27]. In our proposed scheme, it is not possible that a response of an individual to be observed and censored at the same time. Hence, among all the proposed geometric distributions, we found the bivariate geometric distribution proposed by Arnold [2] more applicable and flexible in the presence of dependent censoring of such an scheme.

The rest of the paper is organized as: In Section 2, we briefly introduce the bivariate geometric distribution provided by Arnold [2]. In Section 3, we introduce the potential data under the right censoring scheme and derive the likelihood and obtain the MLEs of the unknown parameters with their risks in closed forms. In Section 4, we deal with Bayes estimates of the unknown parameters and survival function under the squared error loss (SEL) function. Using constrained Bayes methodology, we obtained the constrained Bayes estimates of the unknown parameters under the SEL function. In Section 5, we propose the Bayes prediction of the next observation based on the observed data. In Section 6, we consider estimation under dependent right censoring in the presence of covariates. We obtain MLEs of the unknown parameters via method of scoring. Finally in Section 7, we provide a simulation study and an illustrative example in connection with our obtained estimators using a simulation data and a real data.

2. Preliminaries

Consider a sequence $\{T_1, T_2, T_3, \dots\}$ of i.i.d. random variables with states S_1, S_2 and S_3 , that refer to event of type one, type two and type three, respectively. In the context of survival analysis S_1, S_2 and S_3 states, can be supposed as failure (as an event of interest), censoring and neither censoring nor failure, respectively. Define the corresponding probability of every state as follows

$$P(T_i \in S_1) = \theta_1, \quad P(T_i \in S_2) = \theta_2, \quad P(T_i \in S_3) = \theta_3, \quad i = 1, 2, \dots,$$

where, $\theta_1, \theta_2, \theta_3 > 0$ and $\theta_3 = 1 - \theta_1 - \theta_2$. Define X and Y as

$$X = \min\{x : T_x \in S_1\}, \quad Y = \min\{y : T_y \in S_2\},$$

then we obtain the joint probability mass function (pmf) of X and Y as follows:

$$P_{\theta}(X = x, Y = y) = \begin{cases} \theta_1 \theta_2 (1 - \theta_1 - \theta_2)^{x-1} (1 - \theta_2)^{y-x-1} & y > x \geq 1 \\ \theta_1 \theta_2 (1 - \theta_1 - \theta_2)^{y-1} (1 - \theta_1)^{x-y-1} & x > y \geq 1 \\ 0 & x = y, \end{cases} \quad (1)$$

see Omev and Minkova [30] for more details. We call this distribution ‘Ordinary Bivariate Geometric’ (OBG) with parameter $\theta = (\theta_1, \theta_2)^T$, and denote it by $OBG(\theta_1, \theta_2)$. The marginal distributions of X and Y are ordinary univariate geometric distributions with parameters θ_1 and θ_2 , respectively.

Lemma 2.1: *Let (X, Y) has an $OBG(\theta_1, \theta_2)$ distribution defined in Equation (1). Let $Z = \min(X, Y)$ and $V = u(Y - X)$ where $u(t) = 1$ if $t \geq 0$ and is 0 if $t < 0$. Then V has a Bernoulli distribution with parameter $\theta_1 / (\theta_1 + \theta_2)$, Z has a geometric distribution with mean $1 / (\theta_1 + \theta_2)$, and Z and V are independent.*

Proof: If $V = 0$, the joint pmf of Z and V is given by

$$\begin{aligned} P_{\theta}(Z = z, V = 0) &= P_{\theta}(Z = z, X > Y) \\ &= P_{\theta}(Y = z, X \geq z + 1) \\ &= \theta_2(1 - \theta_1 - \theta_2)^{z-1}, \quad z \geq 1. \end{aligned} \tag{2}$$

If $V = 1$, the pmf is given by

$$\begin{aligned} P_{\theta}(Z = z, V = 1) &= P_{\theta}(Z = z, X \leq Y) \\ &= P_{\theta}(X = z, Y \geq z) \\ &= \theta_1(1 - \theta_1 - \theta_2)^{z-1}, \quad z \geq 1. \end{aligned} \tag{3}$$

Now, combining Equations (2) and (3), we obtain the joint pmf of (Z, V) as follows:

$$\begin{aligned} P_{\theta}(Z = z, V = v) &= [\theta_1(1 - \theta_1 - \theta_2)^{z-1}]^v [\theta_2(1 - \theta_1 - \theta_2)^{z-1}]^{1-v} \\ &= \theta_1^v \theta_2^{1-v} (1 - \theta_1 - \theta_2)^{z-1}, \quad z \geq 1, \quad v = 0, 1. \end{aligned} \tag{4}$$

Now it is easy to verify that the pmf of (Z, V) in Equation (4) is a multiplication of marginal pmfs of Z and V , and this completes the proof. ■

3. MLEs of θ_1 and θ_2 in the presence of right censoring

Suppose that we have a random sample of n individuals and we only check them once in a season. Although, due to problems such as lack of follow up and competing risks, we do not always have opportunity of observing X_1, \dots, X_n . The corresponding potential censoring times are denoted by Y_1, \dots, Y_n . Hence the observed data are $Z_i = \min(X_i, Y_i)$ and $V_i = u(Y_i - X_i)$, $i = 1, \dots, n$, where $u(t) = 1$ if $t \geq 0$ and is 0 if $t < 0$. So the potential data reduces to the actual observed data $\mathbf{W} = (\mathbf{W}_1, \dots, \mathbf{W}_n)$, where $\mathbf{W}_i = (Z_i, V_i)$, $i = 1, \dots, n$ and has pmf obtained in Equation (4). This kind of censoring is called right censoring, see Davarzani and Parsian [9].

In this setup we assume the i th patient has lifetime variable X_i that depends on the right censoring variable, Y_i and (X_i, Y_i) , $i = 1, \dots, n$, have an OBG(θ_1, θ_2) distribution defined in Equation (1). Now using Equation (4), the likelihood function based on actual observed data \mathbf{w} is given by

$$L(\theta_1, \theta_2) = \theta_1^v \theta_2^{n-v} (1 - \theta_1 - \theta_2)^{z-n}, \tag{5}$$

where $v = \sum_{i=1}^n v_i$ and $z = \sum_{i=1}^n z_i$. Hence the log-likelihood function is

$$l(\theta_1, \theta_2) = v \ln \theta_1 + (n - v) \ln \theta_2 + (z - n) \ln(1 - \theta_1 - \theta_2).$$

Thus the MLEs of θ_1 and θ_2 , are the solutions of the following log-likelihood equations:

$$\begin{aligned} \frac{\partial}{\partial \theta_1} l(\theta_1, \theta_2) &= \frac{v}{\theta_1} - \frac{z - n}{1 - \theta_1 - \theta_2} = 0, \\ \frac{\partial}{\partial \theta_2} l(\theta_1, \theta_2) &= \frac{n - v}{\theta_2} - \frac{z - n}{1 - \theta_1 - \theta_2} = 0. \end{aligned}$$

So,

$$\hat{\theta}_1 = \frac{V}{Z}, \quad \hat{\theta}_2 = \frac{n - V}{Z}.$$

And it is also easy to verify that the survival function of X_i is given by

$$S(t) = (1 - \theta_1)^t,$$

hence the MLE of $S(t)$ is

$$\hat{S}(t) = \left(\frac{Z - V}{Z} \right)^t.$$

Notice that from Lemma 2.1, we know that $Z \sim \text{NB}(n, \theta_1 + \theta_2)$ and $V \sim \text{Bin}(n, \theta_1 / (\theta_1 + \theta_2))$ are independent. It is straightforward to show that V/n converges in distribution to a Normal distribution and Z/n converges in probability to $1/(\theta_1 + \theta_2)$. Following the well known asymptotic properties of MLEs, we calculated the Fisher information of θ and we discovered that $\sqrt{n}(\hat{\theta} - \theta)$ converges in distribution to Bivariate Normal with mean $\mathbf{0}$ and variance covariance matrix

$$\theta_1 \theta_2 (\theta_1 + \theta_2) \begin{pmatrix} \frac{1 - \theta_1}{\theta_2} & -1 \\ -1 & \frac{1 - \theta_2}{\theta_1} \end{pmatrix},$$

where $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)^T$. Now using Delta methods, we discovered that $\sqrt{n}(\hat{S}(t) - (1 - \theta_1)^t) \xrightarrow{d} N(0, t^2(1 - \theta_1)^{2t-1} \theta_1(\theta_1 + \theta_2))$. The following Lemma will help us to obtain the risk functions of MLEs in a closed form.

Lemma 3.1: *Let Z has a $\text{NB}(n, \theta_1 + \theta_2)$ distribution, then we have:*

(i)

$$E\left(\frac{1}{Z}\right) = -(\theta_1 + \theta_2)^n (1 - \theta_1 - \theta_2)^{-n} \times [(-1)^{n-1} \ln(\theta_1 + \theta_2) + \sum_{k=0}^{n-2} \binom{n-1}{k} (-1)^k \frac{(\theta_1 + \theta_2)^{k-n+1}}{k-n+1}],$$

(ii)

$$E\left(\frac{1}{Z^2}\right) = (\theta_1 + \theta_2)^n (1 - \theta_1 - \theta_2)^{-n} \left[(-1)^{n-1} \sum_{k=1}^{\infty} \frac{(1 - \theta_1 - \theta_2)^k}{k^2} + \sum_{\substack{r=0 \\ r \neq n-k-2}}^{\infty} \sum_{k=0}^{n-2} \binom{n-1}{k} \frac{(-1)^k (\theta_1 + \theta_2)^{k+r-n+2}}{(k-n+1)(k+r-n+2)} + \sum_{k=0}^{n-2} \binom{n-1}{k} \frac{(-1)^k}{k-n+1} \ln(\theta_1 + \theta_2) \right].$$

Proof: See the Appendix (see supplementary material). ■

Now, from Lemma 3.1, and using the facts that $E(V) = n\theta_1/(\theta_1 + \theta_2)$ and $E(V^2) = n\theta_1(\theta_2 + n\theta_1)/(\theta_1 + \theta_2)^2$, the risk functions of $\hat{\theta}_1$ and $\hat{\theta}_2$ under the SEL function, can be obtained as follows:

$$\begin{aligned}
 R(\boldsymbol{\theta}, \hat{\theta}_1) &= n\theta_1(\theta_1 + \theta_2)^{n-2}(1 - \theta_1 - \theta_2)^{-n} \left\{ (\theta_2 + n\theta_1) \left((-1)^{n-1} \sum_{k=1}^{\infty} \frac{(1 - \theta_1 - \theta_2)^k}{k^2} \right. \right. \\
 &+ \sum_{\substack{r=0 \\ r \neq n-k-2}}^{\infty} \sum_{k=0}^{n-2} \binom{n-1}{k} \frac{(-1)^k (\theta_1 + \theta_2)^{k+r-n+2}}{(k-n+1)(k+r-n+2)} \\
 &+ \left. \left. \sum_{k=0}^{n-2} \binom{n-1}{k} \frac{(-1)^k}{k-n+1} \ln(\theta_1 + \theta_2) \right) + 2\theta_1(\theta_1 + \theta_2) \right. \\
 &\times \left. \left((-1)^{n-1} \ln(\theta_1 + \theta_2) + \sum_{k=0}^{n-2} \binom{n-1}{k} (-1)^k \frac{(\theta_1 + \theta_2)^{k-n+1}}{k-n+1} \right) \right\} + \theta_1^2, \\
 R(\boldsymbol{\theta}, \hat{\theta}_2) &= n\theta_2(\theta_1 + \theta_2)^{n-2}(1 - \theta_1 - \theta_2)^{-n} \left\{ (\theta_1 + n\theta_2) \left((-1)^{n-1} \sum_{k=1}^{\infty} \frac{(1 - \theta_1 - \theta_2)^k}{k^2} \right. \right. \\
 &+ \sum_{\substack{r=0 \\ r \neq n-k-2}}^{\infty} \sum_{k=0}^{n-2} \binom{n-1}{k} \frac{(-1)^k (\theta_1 + \theta_2)^{k+r-n+2}}{(k-n+1)(k+r-n+2)} \\
 &+ \left. \left. \sum_{k=0}^{n-2} \binom{n-1}{k} \frac{(-1)^k}{k-n+1} \ln(\theta_1 + \theta_2) \right) + 2\theta_2(\theta_1 + \theta_2) \right. \\
 &\times \left. \left((-1)^{n-1} \ln(\theta_1 + \theta_2) + \sum_{k=0}^{n-2} \binom{n-1}{k} (-1)^k \frac{(\theta_1 + \theta_2)^{k-n+1}}{k-n+1} \right) \right\} + \theta_2^2.
 \end{aligned}$$

4. Bayesian analysis

It is well known, see [23], that estimating with uniformly minimum risk do not exist and there are two different ways in obtaining optimal estimates. One way is to restrict the estimators by requiring to satisfy some condition enforcing a certain degree of impartiality. Unbiasedness and invariance are two different impartiality conditions of this method. However, this approach of minimizing the risk uniformly in θ after restricting the estimators to be considered has limited applicability. Alternative and more general approach is to minimize an overall measure of the risk function associated with an estimator without

restricting the estimators to be considered. One global measure of the size of the risk associated with an estimator of θ leads to Bayes estimators. Although we derived the MLEs of the parameters and adopting an asymptotic theory to describe the variation of the derived estimations, in the censored data setting, one problem could be the scarcity of data. Therefore, the Bayesian approach provides a coherent framework. In this section, we consider the Bayes estimation and constrained Bayes estimation of unknown parameters of interest in Sections 4.1 and 4.2, respectively.

4.1. Bayes estimation

We believe, as stated in [3], that from a strictly Bayesian viewpoint there is clearly no way in which one can say that one prior is better than any other. Presumably one has one’s own subjective prior and must live with all of its lumps and bumps. It is more frequently the case that we elect to restrict attention to a given flexible family of priors and we choose one from that family that seems to best match our personal beliefs. If all members of a family of prior distribution possess undesirable features, such as insensitivity to the data, we should elect to use a different family of priors. Krishna and Pundir [21] obtained the Bayes estimate of parameters for some bivariate geometric distribution using Dirichlet distribution as prior. Li and Dhar [24] used a uniform prior distribution instead of using Dirichlet distribution. Due to the nature of the parameters, we choose a dependent joint conjugate prior for θ_1 and θ_2 that follows a Dirichlet distribution with parameters α_0, α_1 and α_2 with the following prior density:

$$\pi(\theta_1, \theta_2) \propto \theta_1^{\alpha_1-1} \theta_2^{\alpha_2-1} (1 - \theta_1 - \theta_2)^{\alpha_0-1}, \quad 0 < \theta_1 + \theta_2 < 1,$$

and we denote it by $\text{Dir}_2(\alpha_0, \alpha_1, \alpha_2)$. The joint posterior density function of θ_1 and θ_2 becomes as

$$\pi(\theta_1, \theta_2 | \mathbf{w}) \propto \theta_1^{v+\alpha_1-1} \theta_2^{n-v+\alpha_2-1} (1 - \theta_1 - \theta_2)^{z-n+\alpha_0-1},$$

that is, $\theta | \mathbf{W} = \mathbf{w} \sim \text{Dir}_2(z - n + \alpha_0, v + \alpha_1, n - v + \alpha_2)$.

Therefore, the Bayes estimators of θ_1, θ_2 and $S(t)$ under the SEL function are given, respectively, by

$$\begin{aligned} \hat{\theta}_1^B &= E(\theta_1 | \mathbf{w}) = \frac{v + \alpha_1}{z + \alpha_0 + \alpha_1 + \alpha_2}, \\ \hat{\theta}_2^B &= E(\theta_2 | \mathbf{w}) = \frac{n - v + \alpha_2}{z + \alpha_0 + \alpha_1 + \alpha_2}, \end{aligned}$$

and,

$$\hat{S}^B(t) = \sum_{i=0}^{\infty} \binom{t}{i} (-1)^i \frac{\Gamma(v + \alpha_1 + i) \Gamma(z + \alpha_0 + \alpha_1 + \alpha_2)}{\Gamma(z + i + \alpha_0 + \alpha_1 + \alpha_2) \Gamma(v + \alpha_1)},$$

where t is a real number and

$$\binom{t}{i} = \begin{cases} \frac{t(t-1) \cdots (t-i+1)}{i!} & i = 1, 2, \dots \\ 1 & i = 0. \end{cases}$$

When $\alpha_i \rightarrow 0, i = 0, 1, 2$, we obtain the limiting Bayes estimators (Generalized Bayes estimators w.r.t. the noninformative prior, $\pi(\theta_1, \theta_2) \propto 1/\theta_1\theta_2(1 - \theta_1 - \theta_2)$) of θ_1, θ_2 and $S(t)$

as

$$\hat{\theta}_1^{LB} = \hat{\theta}_1, \quad \hat{\theta}_2^{LB} = \hat{\theta}_2,$$

and,

$$\hat{S}^{LB}(t) = \sum_{i=0}^{\infty} \binom{t}{i} (-1)^i \frac{\Gamma(v+i)\Gamma(z)}{\Gamma(z+i)\Gamma(v)}.$$

4.2. Constrained Bayes estimation

It is well known that the Bayes estimation of θ under the SEL function is the vector of posterior means. As pointed out in [13,14,25], the empirical histogram of the posterior means is underdispersed as an estimate of the histogram of the corresponding population parameters. Thus adjustment of Bayes estimators is needed in order to meet the twin objective of simultaneous estimation and closeness of the histogram of the estimates with the posterior estimates of the parameter histogram. Hence, we consider constrained Bayes estimations of θ_1 and θ_2 under the SEL function by matching the first two empirical moments of the set of estimates with the corresponding posterior moments of the population parameters. In this subsection, we will derive the constrained Bayes (CB) estimate $\hat{\theta}^{CB} = (\hat{\theta}_1^{CB}, \hat{\theta}_2^{CB})$ of θ , where $\hat{\theta}^{CB}$ minimizes

$$E[(\theta_1 - t_1)^2 + (\theta_2 - t_2)^2 | \mathbf{w}]$$

within the class of all estimates $\mathbf{t} \equiv \mathbf{t}(\mathbf{w}) = (t_1(\mathbf{w}), t_2(\mathbf{w}))$ of θ that satisfy

$$E(\bar{\theta} | \mathbf{w}) = \bar{t}(\mathbf{w}) \quad \text{and} \quad E \left[\sum_{i=1}^2 (\theta_i - \bar{\theta})^2 | \mathbf{w} \right] = \sum_{i=1}^2 (t_i(\mathbf{w}) - \bar{t}(\mathbf{w}))^2.$$

Following [13], the $\hat{\theta}^{CB}$ is given by

$$\hat{\theta}_i^{CB} = a\hat{\theta}_i^B + (1 - a)\bar{\theta}^B, \quad i = 1, 2,$$

where $a \equiv a(\mathbf{w}) = [1 + H_1(\mathbf{w})/H_2(\mathbf{w})]^{1/2}$, $H_1(\mathbf{w}) = \text{tr}[V((\theta - \bar{\theta}\mathbf{1}_2) | \mathbf{w})] = \text{tr}[(\mathbf{I}_2 - \frac{1}{2}\mathbf{J}_2)V(\theta | \mathbf{w})]$ and $H_2(\mathbf{w}) = \sum_{i=1}^2 (\hat{\theta}_i^B - \bar{\theta}^B)^2 = [E(\theta | \mathbf{w})]^T (\mathbf{I}_2 - \frac{1}{2}\mathbf{J}_2)E(\theta | \mathbf{w})$, where $\mathbf{1}_2$ is a 2-component column vector with each element equal to one, \mathbf{I}_2 is the identity matrix of order 2 and $\mathbf{J}_2 = \mathbf{1}_2\mathbf{1}_2^T$. When θ has a $Dir_2(\alpha_0, \alpha_1, \alpha_2)$ -prior distribution, we have

$$\begin{aligned} E(\theta_1^2 | \mathbf{w}) &= \frac{(v + \alpha_1 + 1)(v + \alpha_1)}{(\alpha_0 + \alpha_1 + \alpha_2 + z + 1)(\alpha_0 + \alpha_1 + \alpha_2 + z)}, \\ E(\theta_2^2 | \mathbf{w}) &= \frac{(n - v + \alpha_2 + 1)(n - v + \alpha_2)}{(\alpha_0 + \alpha_1 + \alpha_2 + z + 1)(\alpha_0 + \alpha_1 + \alpha_2 + z)}, \\ V(\theta_1 | \mathbf{w}) &= \frac{(v + \alpha_1)(z - v + \alpha_0 + \alpha_2)}{(\alpha_0 + \alpha_1 + \alpha_2 + z)^2(\alpha_0 + \alpha_1 + \alpha_2 + z + 1)}, \end{aligned}$$

$$V(\theta_2 | \mathbf{w}) = \frac{(n - v + \alpha_2)(z - n + v + \alpha_0 + \alpha_1)}{(\alpha_0 + \alpha_1 + \alpha_2 + z)^2(\alpha_0 + \alpha_1 + \alpha_2 + z + 1)},$$

$$\text{Cov}(\theta_1, \theta_2 | \mathbf{w}) = -\frac{(v + \alpha_1)(n - v + \alpha_2)}{(\alpha_0 + \alpha_1 + \alpha_2 + z)^4(\alpha_0 + \alpha_1 + \alpha_2 + z + 1)^2}$$

$$\times \sqrt{(v + \alpha_1)(z - v + \alpha_0 + \alpha_2)(n - v + \alpha_2)(z - n + v + \alpha_0 + \alpha_1)},$$

and $H_1(\mathbf{w})$ and $H_2(\mathbf{w})$ reduce to

$$H_1(\mathbf{w}) = \frac{1}{2}V(\theta_1 | \mathbf{w}) + 2\text{Cov}(\theta_1, \theta_2 | \mathbf{w}) + \frac{1}{2}V(\theta_2 | \mathbf{w}),$$

$$H_2(\mathbf{w}) = \frac{1}{2}[E(\theta_1 | \mathbf{w})]^2 + 2\text{Cov}(\theta_1, \theta_2 | \mathbf{w}) + \frac{1}{2}[E(\theta_2 | \mathbf{w})]^2.$$

5. Bayes prediction

In this section we obtain the Bayes prediction of W_{n+1} based on the observed data $W = (W_1, \dots, W_n)$. The predictive distribution of $W_{n+1} | W = \mathbf{w}$ is given by

$$P(Z_{n+1} = z_{n+1}, V_{n+1} = v_{n+1} | W = \mathbf{w})$$

$$= \iint_{\{(\theta_1, \theta_2): 0 < \theta_1 + \theta_2 < 1\}} P(Z_{n+1} = z_{n+1}, V_{n+1} = v_{n+1} | \theta_1, \theta_2) \pi(\theta_1, \theta_2 | W = \mathbf{w}) d\theta_1 d\theta_2$$

$$= \frac{\Gamma(z + \alpha_0 + \alpha_1 + \alpha_2) \Gamma(v + v_{n+1} + \alpha_1) \Gamma(n + 1 - v - v_{n+1} + \alpha_2) \Gamma(z + z_{n+1} - n + \alpha_0 - 1)}{\Gamma(v + \alpha_1) \Gamma(n - v + \alpha_2) \Gamma(z - n + \alpha_0) \Gamma(z + z_{n+1} + \alpha_0 + \alpha_1 + \alpha_2)}$$

$$= (v + \alpha_1)^{v_{n+1}} (n - v + \alpha_2)^{1 - v_{n+1}} \frac{\Gamma(z + \alpha_0 + \alpha_1 + \alpha_2) \Gamma(z + z_{n+1} - n + \alpha_0 - 1)}{\Gamma(z - n + \alpha_0) \Gamma(z + z_{n+1} + \alpha_0 + \alpha_1 + \alpha_2)}.$$

We are interested to predict Z_{n+1} under the SEL function and V_{n+1} under zero-one loss function. These loss functions are chosen based on the nature of the variables that must be predicted. It is obvious that $E(Z_{n+1} | W = \mathbf{w})$ is the Bayes prediction of Z_{n+1} and reduces to

$$\tilde{Z}_{n+1} = (n + \alpha_1 + \alpha_2) \frac{\Gamma(z + \alpha_0 + \alpha_1 + \alpha_2)}{\Gamma(z - n + \alpha_0)} \sum_{k=1}^{\infty} k \frac{\Gamma(z + k - n + \alpha_0 - 1)}{\Gamma(z + k + \alpha_0 + \alpha_1 + \alpha_2)},$$

where $n > 2$.

On the other hand the Bayes prediction of V_{n+1} under zero-one loss function is given by the posterior mode which is called the maximum a posteriori (MAP) prediction. For this purpose the marginal predictive distribution of $V_{n+1} | W = \mathbf{w}$ is

$$P(V_{n+1} = v_{n+1} | W = \mathbf{w})$$

$$= (v + \alpha_1)^{v_{n+1}} (n - v + \alpha_2)^{1 - v_{n+1}} \frac{\Gamma(z + \alpha_0 + \alpha_1 + \alpha_2)}{\Gamma(z - n + \alpha_0)} \sum_{k=1}^{\infty} \frac{\Gamma(z + k - n + \alpha_0 - 1)}{\Gamma(z + k + \alpha_0 + \alpha_1 + \alpha_2)}.$$

It is easy to verify that the MAP predictor of V_{n+1} is given by

$$\tilde{V}_{n+1} = \begin{cases} 0 & n - 2v > \alpha_1 - \alpha_2, \\ 1 & n - 2v \leq \alpha_1 - \alpha_2. \end{cases}$$

6. The model in the presence of covariates

In this section, we consider a geometric lifetime with dependent right censoring in the presence of covariates. Let $U_i^T = (U_{i1}, U_{i2}, \dots, U_{im})$ be the covariate vector assigned to an individual with lifetime X_i , that is, each person has a survival time X_i and covariates U_i . We assume $\ln\{\theta_{1i}/(1 - \theta_{1i})\} = \beta^T U_i$, where θ_{1i} denotes the probability of failure for individual i , $i = 1, 2, \dots, n$. Further, assume $\beta = (\beta_0, \beta_1, \dots, \beta_m)$ is an unknown vector of coefficients. To set an appropriate model with substituting $\theta_{1i} = (\exp\{\beta^T U_i\}) / (1 + \exp\{\beta^T U_i\})$ in Equation (5), the likelihood function of β and θ_2 is obtained as follows:

$$L(\beta, \theta_2) \propto \prod_{i=1}^n \left(\frac{e^{\beta^T U_i}}{1 + e^{\beta^T U_i}} \right)^{v_i} \theta_2^{1-v_i} \left(\frac{1 - \theta_2(1 + e^{\beta^T U_i})}{1 + e^{\beta^T U_i}} \right)^{z_i - 1}.$$

Hence, the log-likelihood function is

$$l(\beta, \theta_2) \propto \sum_{i=1}^n v_i \beta^T U_i - \sum_{i=1}^n (v_i + z_i - 1) \ln(1 + e^{\beta^T U_i}) + (n - v) \ln \theta_2 + \sum_{i=1}^n (z_i - 1) \ln(1 - \theta_2(1 + e^{\beta^T U_i})),$$

and the MLEs of β and θ_2 can be obtained from the following equations:

$$\begin{aligned} \frac{\partial l(\beta, \theta_2)}{\partial \beta_j} &= \sum_{i=1}^n v_i U_{ij} - \sum_{i=1}^n (v_i + z_i - 1) \frac{U_{ij} e^{\beta^T U_i}}{1 + e^{\beta^T U_i}} \\ &\quad - \sum_{i=1}^n (z_i - 1) \frac{\theta_2 U_{ij} e^{\beta^T U_i}}{1 - \theta_2(1 + e^{\beta^T U_i})} = 0, \quad j = 1, \dots, m. \\ \frac{\partial l(\beta, \theta_2)}{\partial \theta_2} &= \frac{n - v}{\theta_2} - \sum_{i=1}^n (z_i - 1) \frac{1 + e^{\beta^T U_i}}{1 - \theta_2(1 + e^{\beta^T U_i})} = 0, \end{aligned}$$

From the above equations we can't find the MLEs of the parameters in closed forms. However, they can be solved numerically. The second derivatives of log-likelihood are given by

$$\begin{aligned} \frac{\partial^2 l(\beta, \theta_2)}{\partial \theta_2 \partial \beta_j} &= - \sum_{i=1}^n \frac{(z_i - 1) U_{ij} e^{\beta^T U_i}}{(1 - \theta_2(1 + e^{\beta^T U_i}))^2}, \\ \frac{\partial^2 l(\beta, \theta_2)}{\partial \theta_2^2} &= - \frac{n - v}{\theta_2^2} - \sum_{i=1}^n (z_i - 1) \frac{(1 + e^{\beta^T U_i})^2}{(1 - \theta_2(1 + e^{\beta^T U_i}))^2}, \\ \frac{\partial^2 l(\beta, \theta_2)}{\partial \beta_j^2} &= - \sum_{i=1}^n (v_i + z_i - 1) \frac{U_{ij}^2 e^{\beta^T U_i}}{(1 + e^{\beta^T U_i})^2} - \sum_{i=1}^n (z_i - 1) \frac{\theta_2(1 - \theta_2) U_{ij}^2 e^{\beta^T U_i}}{(1 - \theta_2(1 + e^{\beta^T U_i}))^2}, \\ \frac{\partial^2 l(\beta, \theta_2)}{\partial \beta_j \partial \beta_k} &= - \sum_{i=1}^n (v_i + z_i - 1) \frac{U_{ij} U_{ik} e^{\beta^T U_i}}{(1 + e^{\beta^T U_i})^2} - \sum_{i=1}^n (z_i - 1) \frac{\theta_2(1 - \theta_2) U_{ij} U_{ik} e^{\beta^T U_i}}{(1 - \theta_2(1 + e^{\beta^T U_i}))^2}. \end{aligned}$$

For convention let $\beta^{*T} = (\beta^T, \theta_2)$. Therefore, the Hessian of the log-likelihood matrix has the following form:

$$J(\beta^*) = - \begin{pmatrix} \frac{\partial^2 l}{\partial \theta_2^2} & \frac{\partial^2 l}{\partial \theta_2 \partial \beta_1} & \cdots & \frac{\partial^2 l}{\partial \theta_2 \partial \beta_m} \\ \frac{\partial^2 l}{\partial \beta_1 \partial \theta_2} & \frac{\partial^2 l}{\partial \beta_1^2} & \cdots & \frac{\partial^2 l}{\partial \beta_1 \partial \beta_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l}{\partial \beta_m \partial \theta_2} & \frac{\partial^2 l}{\partial \beta_m \partial \beta_1} & \cdots & \frac{\partial^2 l}{\partial \beta_m^2} \end{pmatrix}$$

Now using the method of scoring, at $(j + 1)$ th stage we reach at

$$\beta_{j+1}^* = \beta_j^* + J^{-1}(\beta_j^*) \left. \frac{\partial l(\beta^*)}{\partial \beta^*} \right|_{\beta^* = \beta_j^*},$$

where $\beta_j^{*T} = (\beta_j^T, \theta_{2j})$ and $J^{-1}(\beta^*)$ is the inverse of $J(\beta^*)$. Keep repeating this procedure till some convergence criterion for terminating the iteration obtain, that is, repeat the process until differences between successive estimates are sufficiently close to zero. The resulting estimates are denoted by $\hat{\beta}^* = (\hat{\beta}, \hat{\theta}_2)$. Under regularity conditions, $\hat{\beta}^*$ converges in distribution to multivariate normal distribution with mean equal to true parameter value and variance-covariance given by the inverse of the information matrix, that is, $\hat{\beta}^*$ has approximately $N_{m+1}(\beta^*, I^{-1}(\beta^*))$.

7. Numerical exploration

In this section we conduct a simulation study and analyze an illustrative example to demonstrate the application of the proposed approach.

7.1. Simulation study

In this subsection we conduct a Monte Carlo simulation study to investigate the finite sample properties of the proposed estimators. For this purpose, we will consider the following steps:

- (1) Generate a sequence of $S_1 = 1, S_2 = 2$ and $S_3 = 3$ with probabilities $\theta_1 = 1/2, \theta_2 = 1/3$ and $\theta_3 = 1/6$. Call this sequence as $\{T_1, T_2, T_3, \dots\}$,
- (2) Let

$$X = \min\{x : T_x \in \{1\}\}, \quad Y = \min\{y : T_y \in \{2\}\},$$

- (3) Repeat steps (1) and (2) n times to obtain a sample of size n that is, $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$.

Conducting the mentioned process for n times leads to independent $(X_i, Y_i), i = 1, \dots, n$ from OBG(1/2, 1/3), and obtain $z_i = \min(x_i, y_i)$ and $v_i = u(y_i - x_i)$. From the

Table 1. Estimates (SREMSE) of $\theta_1 = 1/2$ and $\theta_2 = 1/3$ with $\alpha_0 = 1/6, \alpha_1 = 1/2$ and $\alpha_2 = 1/3$.

n	Method	θ_1	θ_2
10	MLE	0.5136491(0.1454064)	0.3375778(0.1376167)
	Bayes	0.5120525(0.1338159)	0.3369063(0.1265844)
	Constrained Bayes	0.5147960(0.1369269)	0.3341628(0.1298954)
15	MLE	0.5082355(0.1196004)	0.3352056(0.1132365)
	Bayes	0.5075454(0.1131215)	0.3349424(0.1071312)
	Constrained Bayes	0.5096223(0.1150398)	0.3328656(0.1091918)
20	MLE	0.5079537(0.10111420)	0.3321635(0.09586175)
	Bayes	0.5074816(0.09702185)	0.3321068(0.09198160)
	Constrained Bayes	0.5092323(0.09844780)	0.3303560(0.09346534)
25	MLE	0.5017629(0.09146889)	0.3406905(0.08799243)
	Bayes	0.5016115(0.08847363)	0.3403801(0.08508828)
	Constrained Bayes	0.5029751(0.08958229)	0.3390165(0.08625248)
30	MLE	0.5008589(0.08592165)	0.3373053(0.08120470)
	Bayes	0.5007634(0.08357470)	0.3371500(0.07898357)
	Constrained Bayes	0.5019522(0.08447135)	0.3359612(0.07991623)
50	MLE	0.5014414(0.06359301)	0.3328116(0.05890157)
	Bayes	0.5013907(0.06254816)	0.3328042(0.05793552)
	Constrained Bayes	0.5021970(0.06296868)	0.3319979(0.05837054)
100	MLE	0.5023765(0.04454527)	0.3323350(0.04342938)
	Bayes	0.5023502(0.04417384)	0.3323375(0.04307006)
	Constrained Bayes	0.5027823(0.04432918)	0.3319054(0.04323500)
200	MLE	0.5001816(0.03239752)	0.3329480(0.02973611)
	Bayes	0.5001790(0.03226321)	0.3329485(0.02961254)
	Constrained Bayes	0.5003991(0.03232173)	0.3327284(0.02967325)

available generated (or simulated) data we obtain ML estimates and Bayes and constrained Bayes estimates of θ_1 and θ_2 using $(\alpha_0, \alpha_1, \alpha_2) = (1/6, 1/2, 1/3)$ and $(1/3, 1, 2/3)$ as informative and more informative prior choices of the hyper-parameters. Notice that the ML estimates are Bayes estimates w.r.t. our noninformative prior choices of hyper-parameters. We repeat this simulation for 1000 replicates and square root of the estimated mean squared error (SREMSE) of the estimates reported in parentheses in Tables 1 and 2. As one can see from Tables 1 and 2, the SREMSE of the Bayes and constrained Bayes estimates are smaller than MLEs and increasing the sample size will decrease the SREMSE. Tables 3 and 4 are summarized the Bayes prediction of Z_{n+1} and V_{n+1} and SREMSE of Z_{n+1} and mean absolute deviation (MAD) of V_{n+1} .

7.2. An illustrative example

In this section, we apply the proposed estimators and regression model on a data set about the survival times of patients with prostate cancer in stages 3 and 4 that have been followed-up during 1967–1969, see [7]. This data set has been analyzed in several studies, for example, [8,19,26]. In the original data set, patients could die from prostate cancer, cardiovascular disease (CVD), or other causes. In this study, we just include 161 patients (from 502 patients), who were died from prostate cancer or CVD. We are interested to analyze the time of death only from prostate cancer in the absence and in the presence of some covariates. These covariates are labeled by RX, AG, WT, PF, HX, HG, SZ, and SG, correspond, to treatment group (0, receiving low-dose of diethylstilbestrol (DES) and 1, receiving high-dose of DES), age group, weight index, performance rating, history of

Table 2. Estimates (SREMSE) of $\theta_1 = 1/2$ and $\theta_2 = 1/3$ with $\alpha_0 = 1/3, \alpha_1 = 1$ and $\alpha_2 = 2/3$.

n	Method	θ_1	θ_2
10	MLE	0.5102883(0.1440578)	0.3371433(0.1346620)
	Bayes	0.5078018(0.1229707)	0.3359783(0.1148933)
	Constrained Bayes	0.5104376(0.1257158)	0.3333425(0.1178023)
15	MLE	0.5041214(0.1157682)	0.3340964(0.1116663)
	Bayes	0.5032704(0.1040000)	0.3336826(0.1003649)
	Constrained Bayes	0.5053032(0.1058441)	0.3316498(0.1023140)
20	MLE	0.5060715(0.10396352)	0.3345641(0.09930793)
	Bayes	0.5055262(0.09259581)	0.3362295(0.08704348)
	Constrained Bayes	0.5071908(0.09387701)	0.3345649(0.08840277)
25	MLE	0.5001836(0.09225354)	0.3338332(0.08727419)
	Bayes	0.4999826(0.08646051)	0.3336751(0.08181079)
	Constrained Bayes	0.5013499(0.08748583)	0.3323078(0.08287403)
30	MLE	0.5035190(0.08343606)	0.3374599(0.08131070)
	Bayes	0.5032047(0.07896754)	0.3371343(0.07692521)
	Constrained Bayes	0.5043943(0.07978070)	0.3359447(0.07779294)
50	MLE	0.4975903(0.06276495)	0.3373152(0.06146491)
	Bayes	0.4976243(0.06071911)	0.3371487(0.05946371)
	Constrained Bayes	0.4983865(0.06114046)	0.3363865(0.05990958)
100	MLE	0.4989304(0.04553555)	0.3332674(0.04353102)
	Bayes	0.4989357(0.04478902)	0.3332597(0.04281664)
	Constrained Bayes	0.4993551(0.04495443)	0.3328403(0.04298878)
200	MLE	0.4989728(0.03261344)	0.3344915(0.03002419)
	Bayes	0.4989780(0.03234353)	0.3344803(0.02977579)
	Constrained Bayes	0.4991939(0.03240419)	0.3342644(0.02983845)

Table 3. Bayes prediction of Z_{n+1} and V_{n+1} and their SREMSE and MAD with $\alpha_0 = 1/6, \alpha_1 = 1/2$ and $\alpha_2 = 1/3$.

n	Z_{n+1}	V_{n+1}
10	1.223593(0.1601316)	1(0.276)
15	1.211933(0.1256167)	1(0.235)
20	1.21679(0.1090521)	1(0.12)
25	1.214456(0.1012657)	1(0.165)
30	1.202447(0.09057832)	1(0.107)
50	1.205398(0.06993069)	1(0.067)
100	1.202384(0.04861567)	1(0.016)

Table 4. Bayes prediction of Z_{n+1} and V_{n+1} and their SREMSE and MAD with $\alpha_0 = 1/3, \alpha_1 = 1$ and $\alpha_2 = 2/3$.

n	Z_{n+1}	V_{n+1}
10	1.212375(0.1382153)	1(0.274)
15	1.216787(0.1202065)	1(0.235)
20	1.210452(0.1081182)	1(0.133)
25	1.208571(0.09727655)	1(0.161)
30	1.20613(0.08424493)	1(0.097)
50	1.204757(0.06793364)	1(0.065)
100	1.203179(0.04909137)	1(0.014)

CVD, serum hemoglobin, size of primary lesion, and Gleason stage/grade category. We had one missing data in the SZ column and four missing data in the SG column, which we imputed them to the nearest integer of their column mean, respectively. In this study, time to death from prostate cancer has been taken as the lifetime of interest that can be censored

at time to death from CVD (as a competing risk). In this case we can never fully determine whether time to death from prostate cancer and time to death from CVD (as the censoring cause) are independent, see chapter 9 of [20].

Using the above information, it is easy to find the values of $z=4306$ and $\nu=130$, that depicts 80.75% of patients died from prostate cancer and 19.25% died from CVD and censored. Let $(\alpha_0, \alpha_1, \alpha_2) = (1/6, 1/2, 1/3)$, then the MLE, Bayes and Constrained Bayes Estimates of (θ_1, θ_2) are $(0.03019043, 0.007199257)$, $(0.03029951, 0.007274979)$ and $(0.03034977, 0.007224719)$, respectively.

Figure 1 displays the ML estimated survival curves for patients in different groups of treatment that indicates, the estimated survival curve for the group with low-dose of DES lies below that for the group with high-dose of DES. For testing differences in survival curves, that is, $H_0 : S_1(t) = S_2(t)$ vs $H_1 : S_1(t) \neq S_2(t)$ we split the data into two groups namely, the group with low-dose of DES with n_L patients and parameters θ_{1L}, θ_{2L} and the group with high-dose of DES with n_H patients and parameters θ_{1H}, θ_{2H} . Now using the asymptotic distributions of the ML estimates of θ_1 and θ_2 in these two group, that is, $\hat{\theta}_{1L}$ and $\hat{\theta}_{1H}$ under H_0 , $T = (\hat{\theta}_{1L} - \hat{\theta}_{1H}) / (\sqrt{\hat{\theta}_{1L}\hat{\theta}_{2L} + \hat{\theta}_{1H}\hat{\theta}_{2H}})$ has approximately t -distribution with $df = \lfloor (n_L - 1)(n_H - 1) / ((n_H - 1)c^2 + (n_L - 1)(1 - c^2)) \rfloor$, where $c^2 = \hat{\theta}_{1L}\hat{\theta}_{2L} / (\hat{\theta}_{1L}\hat{\theta}_{2L} + \hat{\theta}_{1H}\hat{\theta}_{2H})$ and $\lfloor x \rfloor$ is the integer part of x . In Byar sub-data, $T = 1.236394$, $c^2 = 0.5793856$, $df = 79$ and approximate p -value = 0.215 which guarantees acceptance of the null hypothesis $H_0 : S_1(t) = S_2(t)$ at 10% significance level.

To analyze the effect of covariates in the Byar sub-data, we entered eight covariates to the model. The outcome of running the data with eight covariates show that estimates of the regression coefficients are very close to zero except for RX, AG, WT, HX and HG with estimated coefficients (standard errors of coefficients) $-0.30786745(0.188251745)$, $-0.01311510(0.007513652)$, $-0.00866909(0.006332463)$, $-0.29698126(0.190059084)$ and $-0.110739786(0.042973908)$. Calculating $|\hat{\beta}/s.e.(\hat{\beta})|$ leads to 1.635403, 1.745503, 1.368992, 1.562573 and 2.576908 that guarantees the significance of AG at 10% significance

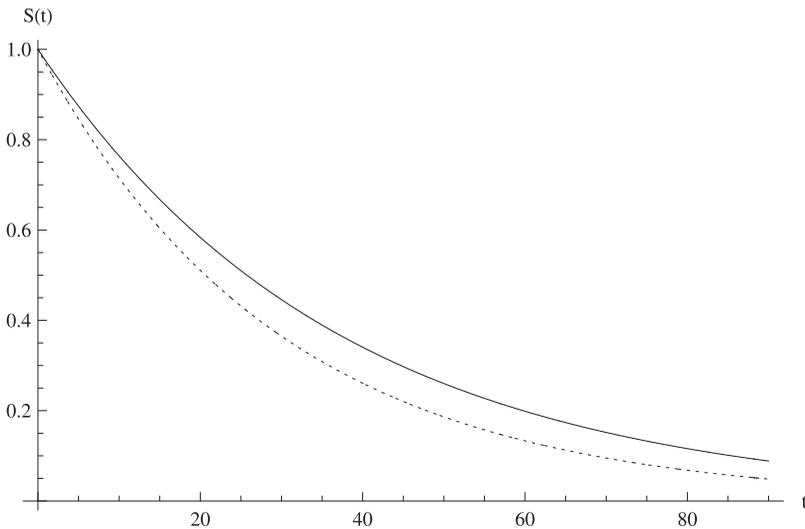


Figure 1. Maximum likelihood estimated survival curves for treatment group, High-dose DES(——) and Low-dose DES(.....).

level and HG at the 5% significance level. Kay [19] at his model (2) dealt with the effects of the explanatory variables on the causes separately. He investigated that RX, HG, SZ and SG affect risk of cancer death with higher values being associated with higher risk. Since our subpopulation is different from Kay [19], we discovered that at the 5% significance level HG is significant.

8. Discussion

Measuring lifetime values discretely or rounding them to small measurement units is very common clinical investigations. Depending on the follow up period of each individual, these values may be either real lifetimes or censored lifetimes. Usually, lifetime variables are assumed to be independent from the censoring variables, while there are many situations that the lifetime variable could be dependent on the time of censoring. In this paper we considered dependent right censoring while the lifetimes are in a discrete form. We proposed Maximum Likelihood, Bayes and Constrained Bayes estimators of the survival function and unknown parameters in closed forms. The simulation study demonstrated that all three methods perform well, even for small sample sizes, with respect to the SREMSEs and biases of estimations. However, the Bayes estimators outperform with respect to the SREMSE.

A multivariate regression model provided in this setup and applied to analyzing a real data. To the best knowledge of the authors there exists no key regression model dealing with dependent discrete censored data for assessing the validity of the results.

Authors observed that there are some discrepancies between the results obtained from the proposed model and results obtained from Cox regression model in terms of obtained p -values; however, they are not (very) inconsistent especially in terms of magnitudes and signs of estimated parameters. This inconsistency might be a consequence of ignoring the correlation between the lifetime and censoring variables in using Cox regression models.

Although, it will be difficult, but it might be interesting to consider a bivariate discrete distribution with nonconstant hazard rate which might be more fitted to clinical data. We believe, more work is needed along these directions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Ahmad Parsian's research was supported by the University of Tehran.

References

- [1] J. Ahmadi, M. Doostparast, and A. Parsian, *Estimation with left-truncated and right censored data: A comparison study*, *Statist. Probab. Lett.* 82 (2012), pp. 1391–1400.
- [2] B.C. Arnold, *A characterization of the exponential distribution by multivariate geometric compounding*, *Sankhya* 37 (1975), pp. 164–173.
- [3] B.C. Arnold and S.J. Press, *Bayesian inference for pareto populations*, *J. Econometrics* 21 (1983), pp. 287–306.
- [4] N. Balakrishnan and D. Mitra, *Left truncated and right censored Weibull data and likelihood inference with an illustration*, *Comput. Statist. Data Anal.* 56 (2012), pp. 4011–4025.

- [5] A.P. Basu and S.K. Dhar, *Bivariate geometric distribution*, J. Appl. Statist. Sci. 2 (1995), pp. 33–44.
- [6] R. Braekers and N. Veraverbeke, *A copula-graphic estimator for the conditional survival function under dependent censoring*, Canad. J. Statist. 33 (2005), pp. 429–447.
- [7] D.P. Byar and D.K. Corle, *Selecting optimal treatment in clinical trials using covariate information*, J. Chronic Dis. 30 (1977), pp. 445–459.
- [8] D.P. Byar and S.B. Green, *Prognostic variables for survival in a randomized comparison of treatments for prostate cancers*, Bulletin Cancer 67 (1980), pp. 477–490.
- [9] N. Davarzani and A. Parsian, *Bayesian inference in dependent right censoring*, Comm. Statist. Theory Methods 39 (2010), pp. 1270–1288.
- [10] N. Davarzani and A. Parsian, *Statistical inference for discrete middle-censored data*, J. Statist. Plann. Inference 411 (2011), pp. 1455–1462.
- [11] N. Davarzani and A. Parsian, *Inference under right censoring in a discrete setup*, Comm. Statist. Theory Methods 42 (2013), pp. 2362–2375.
- [12] N. Davarzani, A. Parsian, and R. Peeters, *Statistical inference on middle-censored data in a dependent setup*, J. Statist. Theory Pract. 9 (2014), pp. 646–657.
- [13] M. Ghosh, *Constrained Bayes estimation with applications*, J. Amer. Statist. Assoc. 87 (1992), pp. 533–540.
- [14] M. Ghosh, M.J. Kim, and D.H. Kim, *Constrained Bayes and empirical Bayes estimation under random effects normal ANOVA model with balanced loss function*, J. Statist. Plann. Inference 138 (2008), pp. 2017–2028.
- [15] S.D. Grimshaw, J. McDonald, G.R. McQueen, and S. Thorley, *Estimating hazard functions for discrete lifetimes*, Comm. Statist. Simulation Comput. 34 (2005), pp. 451–463.
- [16] Y.-T. Hwang and C.-C. Wang, *A goodness of fit test for left-truncated and right-censored data*, Statist. Probab. Lett. 78 (2008), pp. 2420–2425.
- [17] S.R. Jammalamadaka and E. Leong, *Analysis of discrete lifetime data under middle-censoring and in the presence of covariates*, J. Appl. Statist. 42 (2015), pp. 905–913.
- [18] H. Jiang, J.P. Fine, M.R. Kosorok, and R. Chappell, *Pseudo self-consistent estimation of a copula model with informative censoring*, Scand. J. Statist. 32 (2005), pp. 1–20.
- [19] R. Kay, *Treatment effects in competing-risks analysis of prostate cancer data*, Biometrics 42 (1986), pp. 203–211.
- [20] D.G. Kleinbaum and M. Klein, *Survival Analysis A Self-Learning Text*, 3rd ed., Springer, New York, 2013.
- [21] H. Krishna and P. S. Pundir, *A bivariate geometric distribution with applications to reliability*, Comm. Statist. Theory Methods 38 (2009), pp. 1079–1093.
- [22] J.F. Lawless, *Statistical Models and Methods for Lifetime Data*, 2nd ed., John Wiley & Sons, New York, 2004.
- [23] E.L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed., Springer, New York, 1998.
- [24] J. Li and S.K. Dhar, *Modeling with bivariate geometric distributions*, Comm. Statist. Theory Methods 42 (2013), pp. 252–266.
- [25] T.A. Louis, *Estimating a population of parameter values using Bayes and empirical Bayes methods*, J. Amer. Statist. Assoc. 79 (1984), pp. 393–398.
- [26] M. Lunn and D. McNeil, *Applying Cox regression to competing risks*, Biometrics 51 (1995), pp. 524–532.
- [27] A.W. Marshall and I. Olkin, *A multivariate exponential distribution*, J. Amer. Statist. Assoc. 62 (1967), pp. 30–44.
- [28] A.W. Marshall and I. Olkin, *A family of bivariate distributions generated by the bivariate bernoulli distribution*, J. Amer. Statist. Assoc. 80 (1985), pp. 332–338.
- [29] K.R.M. Nair and N.U. Nair, *On characterizing the bivariate exponential and geometric distributions*, Ann. Inst. Statist. Math. 40 (1988), pp. 267–271.
- [30] E. Omey and L.D. Minkova, *Bivariate geometric distributions*, Compt. Rendue Bulg. Acad. Sci. 67 (2014), pp. 1201–1210.
- [31] A.H. Rezaei and N.R. Arghami, *Right censoring in a discrete life model*, Metrika 55 (2002), pp. 151–160.