Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v11n1p108

Basu-Dhar's bivariate geometric distribution in presence of censored data and covariates: some computational aspects

By de Oliveira, Achcar

Published: 26 April 2018

Electronic Journal of Applied Statistical Analysis
Vol. 11, Issue 01, April 2018, 108-136
DOI: 10.1285/i20705948v11n1p108

Basu-Dhar's bivariate geometric distribution in presence of censored data and covariates: some computational aspects

Ricardo Puziol de Oliveira*and Jorge Alberto Achcar

University of São Paulo, Department of Social Medicine, Medical School
Ribeirão Preto, S.P., Brazil

Published: 26 April 2018

Some computational aspects to obtain classical and Bayesian inferences for the Basu and Dhar (1995) bivariate geometric distribution in presence of censored data and covariates are discussed in this paper. The posterior summaries of interest are obtained using standard existing MCMC (Markov Chain Monte Carlo) simulation methods available in popular free softwares as the OpenBugs software and the R software. Numerical illustrations are introduced considering simulated and real datasets showing that the use of discrete bivariate distributions may be a good alternative to the use of continuous bivariate distributions, in many areas of application.

keywords: Basu-Dhar distribution, Bayesian estimates, censored data, covariates, maximum likelihood estimates.

*Corresponding author: rpuziol.oliveira@gmail.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

1 Introduction

Many bivariate distributions for continuous random variables are introduced in the literature to be used in data analysis, especially in applications of survival data in the presence of censored data and covariates (see, for example, Block and Basu, 1974; Marshall and Olkin, 1967a,b; Downton, 1970; Freund, 1961; Sarkar, 1987; Arnold and Strauss, 1988; Gumbel, 1960; Hanagal, 2006; Hanagal and Ahmadi, 2008; Hawkes, 1972; Hougaard, 1986). Alternatively, it can be observed in the literature that it is not very common the use of bivariate distributions for survival data assuming discrete data. Some discrete bivariate distributions have been introduced in the literature as the bivariate geometric distribution of Basu and Dhar (1995) or the bivariate geometric distribution of Arnold (1975), but these discrete distributions are still not very popular in the analysis of bivariate lifetime data, especially in the presence of censored data and covariates (see also, Arnold and Strauss, 1988; Davarzani et al., 2015; Basu and Dhar, 1995; Dhar, 2003; Dhar and Balaji, 2006; Krishna and Pundir, 2009; Muraleedharan Nair and Unnikrishnan Nair, 1988; Sun and Basu, 1995).

It is important to observe that many data sets assumed as continuous data are in fact, discrete data. For example, in analysis of bivariate survival data such as reinfection in the kidneys of the same patient, times of loss of vision in eyes, among others there are bivariate responses measured in days, weeks or months which characterize discrete data. In engineering applications it is common data related to the number of cycles to failures, again characterizing discrete data.

In this paper it is explored a very promising bivariate distribution for discrete data: the geometric Basu-Dhar distribution denoted as BD distribution (Basu and Dhar, 1995; Dhar, 1998; Dhar, 2003). Assuming complete data, censored data and presence of covariates some computational aspects are presented for the attainment of classical inferences based on maximum likelihood estimation method and Bayesian inferences using MCMC (Markov Chain Monte Carlo) simulation methods.

The use of Bayesian methods is very popular for bivariate continuous or discrete random variables in presence of censored data and covariates (see for example, Achcar and Leandro (1998); dos Santos and Achcar (2011)) given the difficulties in the use of standard asymptotically maximum likelihood estimates (see for example, Lawless (1982); Klein and Moeschberger (2005)) especially using MCMC (Markov Chain Monte Carlo) methods (see for example, Chib and Greenberg, 1995; Gelfand and Smith, 1990).

An important goal of this paper is the discussion of useful computational issues related to the use of the bivariate Basu-Dhar geometric distribution in applications with complete and censored datasets and the introduction of computer codes using free R software (R Core Team, 2016) for programs to obtain the classical and Bayesian inferences of interest. Some applications with real data and simulated data are introduced showing some aspects of accuracy and computational costs considering different sample sizes to illustrate the proposed methodology.

2 The Basu-Dhar Bivariate Geometric Distribution

An useful bivariate geometric distribution introduced by Basu and Dhar (1995) could be a good alternative in applications with lifetime bivariate data in presence of censored data and covariates. Inferences for this distribution under a Bayesian approach is introduced by Achcar et al. (2016a). The joint survival function for the Basu-Dhar distribution is given by,

$$P(X_1 > x_1, X_2 > x_2) = p_1^{x_1} p_2^{x_2} p_{12}^{\max(x_1,x_2)} \tag{1}$$

where $x_1, x_2 \in \mathbb{N}^*$, $0 < p_1 < 1$, $0 < p_2 < 1$ and $0 < p_{12} \leq 1$. It is seen that the survival function satisfies the loss of memory property without any additional parameter restrictions, i.e.,

$$P(X_1 > x_1 + t, X_2 > x_2 + t \mid X_1 > x_1, X_2 > x_2) = (p_1 p_2 p_{12})^t \tag{2}$$

The bivariate probability mass function can easily be obtained from the joint survival function. In fact, from (1), we have that the bivariate probability mass function of the Basu-Dhar bivariate geometric distribution, for two discrete random variables X_1 and X_2, is given by,

$$\begin{aligned} P(X_1 = x_1, X_2 = x_2) &= P(X_1 > x_1 - 1, X_2 > x_2 - 1) - P(X_1 > x_1, X_2 > x_2 - 1) \\ &- P(X_1 > x_1 - 1, X_2 > x_2) + P(X_1 > x_1, X_2 > x_2) \end{aligned} \tag{3}$$

From (3) the bivariate probability mass function (pmf) for the Basu-Dhar distribution is given by,

$$P(X_1 = x_1, X_2 = x_2) = \begin{cases} p_1^{x_1-1}(p_2 p_{12})^{x_2-1} q_1 (1 - p_2 p_{12}) & \text{if } X_1 < X_2 \\ (p_1 p_2 p_{12})^{x_1-1}(1 - p_1 p_{12} - p_2 p_{12} + p_1 p_2 p_{12}) & \text{if } X_1 = X_2 \\ p_2^{x_2-1}(p_1 p_{12})^{x_1-1} q_2 (1 - p_1 p_{12}) & \text{if } X_1 > X_2 \end{cases} \tag{4}$$

where $X_1, X_2 \in \mathbb{N}^*$, $0 < p_1 < 1$, $0 < p_2 < 1$, $0 < p_{12} \leq 1$, $q_1 = 1 - p_1$ and $q_2 = 1 - p_2$. For convenience, we denote the Basu-Dhar distribution for a bivariate vector of discrete random variables (X_1, X_2) as BD distribution.

Remark: *An interpretation of the BD distribution could be given in terms of a system of two components where the failure of the system is given when the component 1 fails, the component 2 fails and both components fail simultaneously. Assuming that the probability of failure of the component 1 is equal to $1 - p_1$, observe that event $X_1 > x_1$ occurs if and only if there were no faults until $X_1 = x_1$ (a binomial distribution $b(x_1, 1 - p_1)$, that is, $P(X_1 > x_1) = p_1^{x_1}$). In the same way, $P(X_2 > x_2) = p_2^{x_2}$ and $P(\max(X_1, X_2) > \max(x_1, x_2)) = p_{12}^{\max(x_1,x_2)}$. Thus the probability of the system is working is given by,*

$$P(X_1 > x_1, X_2 > x_2) = p_1^{x_1} p_2^{x_2} p_{12}^{\max(x_1,x_2)} \tag{5}$$

3 Some Mathematical Properties for the BD Distribution

In this section we derive some mathematical properties of the BD distribution. The derived properties include the marginal probability distributions, conditional distribution (Section 3.1), distribution of $\min(X_1, X_2)$ (Section 3.2), conditional distributions (Section 3.3), moment properties (Section 3.4) and covariance and correlation (Section 3.5).

3.1 Marginal Probability Distributions for X_1 and X_2

Let (X_1, X_2) be a bivariate discrete random vector with a BD distribution. Clearly, the marginal distributions of X_1 and X_2 are geometric distributions with the corresponding parameters p_1p_{12} and p_2p_{12}. In fact, from Equation (4), we have, for X_1, that the marginal survival function can be expressed by:

$$P(X_1 > x_1) = P(X_1 > x_1, X_2 > 0) = p_1^{x_1} p_{12}^{x_1}, \quad x_1 \in \mathbb{N}^* \tag{6}$$

from that, the marginal distribution of X_1 is given by:

$$\begin{aligned} P(X_1 = x_1) &= P(X_1 > x_1 - 1) - P(X_1 > x_1) \\ &= (p_1p_{12})^{x_1-1}(1 - p_1p_{12}) \end{aligned} \tag{7}$$

where $\mathbb{E}(X_1) = \frac{1}{(1-p_1p_{12})}$ and $\mathrm{Var}(X_1) = \frac{p_1p_{12}}{(1-p_1p_{12})^2}$. In the same way for the random variable X_2, the marginal survival function can be expressed by:

$$P(X_2 > x_2) = P(X_1 > 0, X_2 > x_2) = p_2^{x_2} p_{12}^{x_2}, \quad x_2 \in \mathbb{N}^* \tag{8}$$

that is,the marginal distribution of X_2 is given by:

$$\begin{aligned} P(X_2 = x_2) &= P(X_2 > x_2 - 1) - P(X_2 > x_2) \\ &= (p_2p_{12})^{x_2-1}(1 - p_2p_{12}) \end{aligned} \tag{9}$$

where $\mathbb{E}(X_2) = \frac{1}{(1-p_2p_{12})}$ and $\mathrm{Var}(X_2) = \frac{p_2p_{12}}{(1-p_2p_{12})^2}$. Thus, we conclude that the marginal distributions for X_1 and X_2 of BD are respectively geometric distributions, that is, $X_1 \sim Geo(p_1p_{12})$ and $X_2 \sim Geo(p_2p_{12})$.

The probability generating function of X_1 and X_2 are given, respectively, by:

$$G(t_1) = \frac{(1 - p_1p_{12})t_1}{1 - p_1p_{12}t_1} \tag{10}$$

and,

$$G(t_2) = \frac{(1 - p_2p_{12})t_2}{1 - p_2p_{12}t_2} \tag{11}$$

3.2 Probability Distribution for $\min(X_1, X_2)$

Let us assume the transformation of the random variables X_1 and X_2 given by $W = \min(X_1, X_2)$. In the same way as it was considered for the marginal distributions, it is easy to see that the distribution of $\min(X_1, X_2)$ is a geometric distribution with corresponding parameter $p_1p_2p_{12}$. In fact, the cumulative distribution function of W is given by:

$$P(W \leq w) = 1 - P(W > w) = 1 - p_1^w p_2^w p_{12}^{\max(w,w)} = 1 - (p_1p_2p_{12})^w \tag{12}$$

That is, $P(W > w) = (p_1p_2p_{12})^w$ that implies $W \sim Geo(p_1p_2p_{12})$ with $\mathbb{E}(W) = \frac{1}{(1-p_1p_2p_{12})}$ and $\mathrm{Var}(W) = \frac{p_1p_2p_{12}}{(1-p_1p_2p_{12})^2}$.

The probability generating function for W is given by

$$G(t) = \frac{(1 - p_1p_2p_{12})t}{1 - p_1p_2p_{12}t}. \tag{13}$$

3.3 Conditional Distributions

From BD distribution pmf and marginal distributions, the conditional distribution of X_2 given X_1, presented by Li and Dhar (2013), is given by:

$$P(X_2 = x_2 \mid X_1 = x_1) = \begin{cases} p_2^{x_2-1} q_2 & \text{if} \quad x_2 < x_1 \\ \dfrac{p_2^{x_2-1}(1 - p_1p_{12} - p_2p_{12} + p_1p_2p_{12})}{1 - p_1p_{12}} & \text{if} \quad x_2 = x_1 \\ \dfrac{p_2^{x_2-1} p_{12}^{x_2-x_1} q_1 (1 - p_2p_{12})}{1 - p_1p_{12}} & \text{if} \quad x_2 > x_1 \end{cases} \tag{14}$$

and the conditional distribution of X_1 given X_2 is given by:

$$P(X_1 = x_1 \mid X_2 = x_2) = \begin{cases} p_1^{x_1-1} q_1 & \text{if} \quad x_2 < x_1 \\ \dfrac{p_1^{x_1-1}(1 - p_1p_{12} - p_2p_{12} + p_1p_2p_{12})}{1 - p_2p_{12}} & \text{if} \quad x_2 = x_1 \\ \dfrac{p_1^{x_1-1} p_{12}^{x_1-x_2} q_2 (1 - p_1p_{12})}{1 - p_2p_{12}} & \text{if} \quad x_2 > x_1 \end{cases} \tag{15}$$

3.4 Moment Properties

Let (X_1, X_2) be a bivariate discrete random vector with a BD(p_1, p_2, p_{12}) distribution. Following Li and Dhar (2013), the probability generating function of (X_1, X_2) for the bivariate Basu-Dhar geometric distribution can be expressed as:

$$G(t_1,t_2) = \frac{t_1t_2}{1-t_1t_2\gamma}\left[\frac{t_2q_1p_2p_{12}(1-p_2p_{12})}{1-t_2p_2p_{12}} + \frac{t_1q_2p_1p_{12}(1-p_1p_{12})}{1-t_1p_1p_{12}} + (1-p_1p_{12}-p_2p_{12}+\gamma)\right]. \tag{16}$$

where $0 < p_1, p_2, p_{12} < 1$, $|t_1| < 1/p_1$, $|t_2| < 1/p_2$, $|t_1t_2| < 1/p_1p_2p_{12}$ and $\gamma = p_1p_2p_{12}$. The same probability generating function also could be obtained from a generalization of the geometric distribution proposed by Hawkes (1972).

Using the relationship presented in Kocherlakota and Kocherlakota (1992), between the probability generating function and the moment generating function, $M(t_1,t_2) = G(\mathrm{e}^{t_1},\mathrm{e}^{t_2})$, the corresponding moment generating function is:

$$M(t_1,t_2) = \frac{\mathrm{e}^{t_1}\mathrm{e}^{t_2}}{1-\mathrm{e}^{t_1}\mathrm{e}^{t_2}\gamma}\left[\frac{\mathrm{e}^{t_2}q_1p_2p_{12}(1-p_2p_{12})}{1-\mathrm{e}^{t_2}p_2p_{12}} + \frac{\mathrm{e}^{t_1}q_2p_1p_{12}(1-p_1p_{12})}{1-\mathrm{e}^{t_1}p_1p_{12}} + (1-p_1p_{12}-p_2p_{12}+\gamma)\right]. \tag{17}$$

In the same way, using the relationship presented in Kocherlakota and Kocherlakota (1992), $N(t_1,t_2) = G(t_1+1,t_2+1)$, the corresponding factorial moment generating function of (X_1,X_2) is defined by:

$$N(t_1,t_2) = \frac{(t_1+1)(t_2+1)}{1-(t_1+1)(t_2+1)\gamma}\left[\frac{(t_2+1)q_1p_2p_{12}(1-p_2p_{12})}{1-p_2p_{12}(t_2+1)} + \frac{(t_1+1)q_2p_1p_{12}(1-p_1p_{12})}{1-p_1p_{12}(t_1+1)} + (1-p_1p_{12}-p_2p_{12}+\gamma)\right]. \tag{18}$$

3.5 Covariance and Correlation

Let (X_1,X_2) be a bivariate discrete random vector with a BD(p_1,p_2,p_{12}) distribution. Following Li and Dhar (2013) and from Equation (18), the product moment of X_1 and X_2 is derived as follows:

$$\mathbb{E}[X_1X_2] = \frac{1-p_1p_2p_{12}^2}{(1-p_1p_{12})(1-p_2p_{12})(1-p_1p_2p_{12})} \tag{19}$$

From (19), the covariance and correlation coefficient for the BD(p_1,p_2,p_{12}) distribution are respectively given by,

$$\mathrm{Cov}[X_1,X_2] = \frac{p_1p_2p_{12}(1-p_{12})}{(1-p_1p_{12})(1-p_2p_{12})(1-p_1p_2p_{12})} \tag{20}$$

and,

$$\rho[X_1, X_2] = \frac{(1 - p_{12})(p_1 p_2)^{1/2}}{1 - p_1 p_2 p_{12}}. \tag{21}$$

4 Estimation

4.1 Method of Moments

The method of moments could be used (see Dhar, 1998) to estimate the parameters of the BD(p_1, p_2, p_{12}) distribution. In this subsection, it is presented this estimation method. In order to apply method of moments, recall the population moments presented in Sections 3.1 and 3.2 and replace the population moments by their sample equivalents as follows:

$$\mathbb{E}(X_1) = \frac{1}{(1{-}p_1 p_{12})} = \overline{X_1},\ \ \mathbb{E}(X_2) = \frac{1}{(1{-}p_2 p_{12})} = \overline{X_2} \tag{22}$$

and,

$$\mathbb{E}(W) = \frac{1}{(1{-}p_1 p_2 p_{12})} = \overline{W} \tag{23}$$

where $W = \min(X_1, X_2)$ and $\overline{X_1} = \sum_{i=1}^{n} X_{1_i}/n, \overline{X_2} = \sum_{i=1}^{n} X_{2i}/n, \overline{W} = \sum_{i=1}^{n} \min(X_{1i}, X_{2i})/n$ denote the sample moments. These equations are solved to yield the estimates for p_1, p_2 and p_{12} obtained by the method of moments:

$$\tilde{p_1} = \frac{\overline{X_2}(1 - \overline{W})}{\overline{W}(1 - \overline{X_2})},\ \ \tilde{p_2} = \frac{\overline{X_1}(\overline{W} - 1)}{\overline{W}(\overline{X_1} - 1)} \tag{24}$$

and,

$$\tilde{p_{12}} = \frac{\overline{W}(\overline{X_1} - 1)(\overline{X_2} - 1)}{(\overline{W} - 1)\overline{X_2 X_2}} \tag{25}$$

4.2 Method of Maximum Likelihood

In this section, it is introduced the maximum likelihood estimation (MLE) method in two situations: the situation assuming censored data and the situation with complete data. In both cases, the maximum likelihood estimator do not have a closed form and it is needed the use of numerical methods like Newton-Rapshon, Nelder-Mead and others to get the MLE for each parameter.

4.2.1 Complete Data

Let $(X_{11}, X_{21}), (X_{12}, X_{22}), \dots, (X_{1n}, X_{2n})$ be a random sample of size n from a BD distribution and define two indicator variables given by $v_{1i} = 1$ if $X_{1i} < X_{2i}$ else 0;

$v_{2i} = 1$ if $X_{1i} > X_{2i}$ else 0 for $i = 1, 2, \ldots, n$ two indicator variables. Then, we have three possible situations considering these indicator variables:

$$\begin{aligned} (v_{1i}, v_{2i}) &= (1,0) \quad \text{if} \quad X_{1i} < X_{2i} \\ (v_{1i}, v_{2i}) &= (0,1) \quad \text{if} \quad X_{1i} > X_{2i} \\ (v_{1i}, v_{2i}) &= (0,0) \quad \text{if} \quad X_{1i} = X_{2i} \end{aligned} \tag{26}$$

Thus, from Equations (4) and (26), the likelihood function for the parameters p_1, p_2 e p_{12} is given by,

$$L(p_1, p_2, p_{12}) = \prod_{i=1}^{n} f_1^{v_{1i}(1-v_{2i})}(x_{1i}, x_{2i}) f_2^{(1-v_{1i})(1-v_{2i})}(x_{1i}, x_{2i}) f_3^{v_{2i}(1-v_{1i})}(x_{1i}, x_{2i}) \tag{27}$$

where

$$\begin{aligned} f_1(x_{1i}, x_{2i}) &= p_1^{x_1-1}(p_2p_{12})^{x_2-1} q_1 (1 - p_2p_{12}) \\ f_2(x_{1i}, x_{2i}) &= (p_1p_2p_{12})^{x_1-1}(1 - p_1p_{12} - p_2p_{12} + p_1p_2p_{12}) \\ f_3(x_{1i}, x_{2i}) &= p_2^{x_2-1}(p_1p_{12})^{x_1-1} q_2 (1 - p_1p_{12}) \end{aligned} \tag{28}$$

Defining $n_1 = \sum_{i=1}^{n} v_{1i}, n_2 = \sum_{i=1}^{n} v_{2i}, n_3 = \sum_{i=1}^{n}(1 - v_{1i})(1 - v_{2i}), n = n_1 + n_2 + n_3$ and $\sum_{i=1}^{n} v_{1i}v_{2i} = 0$, the likelihood function (28) can be rewritten as follows:

$$\begin{aligned} L(p_1, p_2, p_{12}) &= p_1^{T_1-n_1}(p_2p_{12})^{T_2-n_1} q_1^{n_1}(1 - p_2p_{12})^{n_1}(p_1p_2p_{12})^{T_3-n_3} \\ &\times \; p_2^{T_6-n_2}(p_1p_{12})^{T_5-n_2} q_2^{n_2}(1 - p_1p_{12})^{n_2} \\ &\times \; (1 - p_1p_{12} - p_2p_{12} + p_1p_2p_12)^{n_3} \\ &= p_1^{T_1+T_3+T_5-n} p_2^{T_2+T_3+T_6-n} p_{12}^{T_2+T_3+T_5-n} q_1^{n_1} q_2^{n_2} (1 - p_2p_{12})^{n_1}(1 - p_1p_{12})^{n_2} \\ &\times \; (1 - p_1p_{12} - p_2p_{12} + p_1p_2p_{12})^{n_3} \end{aligned} \tag{29}$$

where $T_1 = \sum_{i=1}^{n} v_{1i}(1 - v_{2i})x_{1i}, T_2 = \sum_{i=1}^{n} v_{1i}(1 - v_{2i})x_{2i}, T_3 = T_4 = \sum_{i=1}^{n}(1 - v_{1i})(1 - v_{2i})x_{1i}, T_5 = \sum_{i=1}^{n} v_{2i}(1 - v_{1i})x_{2i}$ and $T_6 = \sum_{i=1}^{n} v_{2i}(1 - v_{1i})x_{2i}$. From (29), the log-likelihood function for the parameters p_1, p_2 and p_{12} is given by:

$$\begin{aligned} \ell(p_1, p_2, p_{12}) &= (T_1 + T_3 + T_5 - n)\log(p_1) + (T_2 + T_3 + T_6 - n)\log(p_2) \\ &+ \; (T_1 + T_3 + T_5 - n)\log(p_{12}) + n_1 \log(q_1) + n_2 \log(q_2) + \\ &+ \; n_1 \log(1 - p_2p_{12}) + n_2 \log(1 - p_1p_{12}) \\ &+ \; n_3 \log(1 - p_1p_{12} - p_2p_{12} + p_1p_2p_{12}) \end{aligned} \tag{30}$$

The maximum likelihood estimators (MLEs) for p_1, p_2 and p_{12} are solutions of the equations $\dfrac{\partial \ell}{\partial p_1} = 0, \dfrac{\partial \ell}{\partial p_2} = 0$ e $\dfrac{\partial \ell}{\partial p_{12}} = 0$. From the log-likelihood (Equation 30), the first derivatives of $\ell(p_1, p_2, p_{12})$ with respect to p_1, p_2 and p_{12} are given respectively, by,

$$
\begin{aligned}
\frac{\partial \ell}{\partial p_1} &= \frac{T_1 + T_3 + T_5 - n}{p_1} - \frac{n_1}{1-p_1} - \frac{n_2 p_{12}}{1 - p_1 p_{12}} - \frac{n_3(p_{12} - p_2 p_{12})}{1 - p_1 p_{12} - p_2 p_{12} + p_1 p_2 p_{12}} \\
\frac{\partial \ell}{\partial p_2} &= \frac{T_2 + T_3 + T_6 - n}{p_2} - \frac{n_2}{1-p_2} - \frac{n_1 p_{12}}{1 - p_2 p_{12}} - \frac{n_3(p_{12} - p_2 p_{12})}{1 - p_1 p_{12} - p_2 p_{12} + p_1 p_2 p_{12}} \\
\frac{\partial \ell}{\partial p_{12}} &= \frac{T_2 + T_3 + T_5 - n}{p_{12}} - \frac{n_1 p_2}{1 - p_2 p_{12}} - \frac{n_2 p_1}{1 - p_1 p_{12}} - \frac{n_3(p_1 + p_2 - p_1 p_{12})}{1 - p_1 p_{12} - p_2 p_{12} + p_1 p_2 p_{12}}
\end{aligned}
\tag{31}
$$

Defining $a = T_1 + T_3 + T_5, b = T_2 + T_3 + T_6, c = T_2 + T_3 + T_5$, then the MLEs, denoted by $\widehat{p_1}, \widehat{p_2}$ and $\widehat{p_{12}}$, are the solutions of the following system of equations:

$$
\begin{aligned}
&\frac{a-n}{\widehat{p_1}} - \frac{n_1}{1-\widehat{p_1}} - \frac{n_2 \widehat{p_{12}}}{1 - \widehat{p_1}\widehat{p_{12}}} - \frac{n_3(\widehat{p_{12}} - \widehat{p_2}\widehat{p_{12}})}{1 - \widehat{p_1}\widehat{p_{12}} - \widehat{p_2}\widehat{p_{12}} + \widehat{p_1}\widehat{p_2}\widehat{p_{12}}} = 0 \\
&\frac{b-n}{\widehat{p_2}} - \frac{n_2}{1-\widehat{p_2}} - \frac{n_1 \widehat{p_{12}}}{1 - \widehat{p_2}\widehat{p_{12}}} - \frac{n_3(\widehat{p_{12}} - \widehat{p_2}\widehat{p_{12}})}{1 - \widehat{p_1}\widehat{p_{12}} - \widehat{p_2}\widehat{p_{12}} + \widehat{p_1}\widehat{p_2}\widehat{p_{12}}} = 0 \\
&\frac{c-n}{\widehat{p_{12}}} - \frac{n_1 \widehat{p_2}}{1 - \widehat{p_2}\widehat{p_{12}}} - \frac{n_2 p_1}{1 - \widehat{p_1}\widehat{p_{12}}} - \frac{n_3(\widehat{p_1} + \widehat{p_2} - \widehat{p_1}\widehat{p_{12}})}{1 - \widehat{p_1}\widehat{p_{12}} - \widehat{p_2}\widehat{p_{12}} + \widehat{p_1}\widehat{p_2}\widehat{p_{12}}} = 0
\end{aligned}
\tag{32}
$$

It is important to point out that the MLEs of p_1, p_2 and p_{12} have no closed form. However, using (32), we can estimate $\widehat{p_1}, \widehat{p_2}$ and $\widehat{p_{12}}$ using standard numeric optimization algorithms such the Newton-Raphson or the Nelder-Mead methods. Under standard asymptotic maximum likelihood theory, the observed information Fisher's matrix is obtained from the second derivatives of the log-likelihood function with respect to p_1, p_2 and p_{12} locally at the obtained MLE's, that is, the Fisher's observed information matrix, $I_0 = \left(\dfrac{\partial^2 \ell}{\partial^2 p_k p_j} \right), k, j = 1, 2, 12$, is given by,

$$
I_0(\widehat{p_1}, \widehat{p_2}, \widehat{p_{12}}) = \begin{pmatrix}
-\dfrac{\partial^2 \ell}{\partial^2 p_1^2} & -\dfrac{\partial^2 \ell}{\partial^2 p_1 p_2} & -\dfrac{\partial^2 \ell}{\partial^2 p_1 p_{12}} \\[2ex]
-\dfrac{\partial^2 \ell}{\partial^2 p_2 p_1} & -\dfrac{\partial^2 \ell}{\partial^2 p_2^2} & -\dfrac{\partial^2 \ell}{\partial^2 p_2 p_{12}} \\[2ex]
-\dfrac{\partial^2 \ell}{\partial^2 p_{12} p_1} & -\dfrac{\partial^2 \ell}{\partial^2 p_{12} p_2} & -\dfrac{\partial^2 \ell}{\partial^2 p_{12}^2}
\end{pmatrix}_{\widehat{p_1}, \widehat{p_2}, \widehat{p_{12}}}
\tag{33}
$$

where,

$$\begin{aligned}
\frac{\partial^2 \ell}{\partial^2 p_1 p_1} &= -\frac{a-n}{p_1^2} - \frac{n_1}{q_1^2} - \frac{n_2 p_{12}^2}{(1-p_1p_{12})^2} + \frac{n_3 p_{12} q_2 (p_2 p_{12} - p_{12})}{(1-\gamma)^2} \\
\frac{\partial^2 \ell}{\partial^2 p_1 p_2} &= \frac{\partial^2 \ell}{\partial^2 p_2 p_1} = \frac{n_3 p_{12}}{1-\gamma} + \frac{n_3 p_{12} q_2 (p_1 p_{12} - p_{12})}{(1-\gamma)^2} \\
\frac{\partial^2 \ell}{\partial^2 p_1 p_{12}} &= \frac{\partial^2 \ell}{\partial^2 p_{12} p_1} = -\frac{n_2}{1-p_1p_{12}} - \frac{n_2 p_1 p_{12}}{(1-p_1p_{12})^2} - \frac{n_3 q_2}{1-\gamma} + \frac{n_3 p_{12} q_2 (p_1 p_{12} - p_1 - p_2)}{(1-\gamma)^2} \\
\frac{\partial^2 \ell}{\partial^2 p_2 p_2} &= -\frac{b-n}{p_2^2} - \frac{n_2}{q_2^2} - \frac{n_1 p_{12}^2}{(1-p_2p_{12})^2} + \frac{n_3 p_{12} q_1 (p_1 p_{12} - p_{12})}{(1-\gamma)^2} \\
\frac{\partial^2 \ell}{\partial^2 p_2 p_{12}} &= \frac{\partial^2 \ell}{\partial^2 p_{12} p_2} = -\frac{n_1}{1-p_2p_{12}} - \frac{n_1 p_2 p_{12}}{(1-p_2p_{12})^2} - \frac{n_3 q_1}{1-\gamma} + \frac{n_3 p_{12} q_1 (p_1 p_{12} - p_1 - p_2)}{(1-\gamma)^2} \\
\frac{\partial^2 \ell}{\partial^2 p_{12} p_{12}} &= -\frac{c-n}{p_{12}^2} - \frac{n_1 p_2^2}{(1-p_2p_{12})^2} - \frac{n_1 p_1^2}{(1-p_1p_{12})^2} + \frac{n_3 (p_1 + p_2 - p_1 p_2)(p_1 p_2 - p_1 - p_2)}{(1-\gamma)^2}
\end{aligned}$$

and $q_1 = 1 - p_1, q_2 = 1 - p_2, \gamma = p_1 p_{12} + p_2 p_{12} - p_1 p_2 p_{12}$.

Hypotheses tests and confidence intervals for p_1, p_2 and p_{12} can be obtained by using the asymptotically normality of the MLEs $\widehat{p_1}, \widehat{p_2}$ and $\widehat{p_{12}}$, i. e.,

$$(\widehat{p_1}, \widehat{p_2}, \widehat{p_{12}}) \sim N((\widehat{p_1}, \widehat{p_2}, \widehat{p_{12}}), I_0^{-1})$$

where I_0 is the Fisher's observed information matrix described previously.

Remark: *Another way to write the likelihood function for p_1, p_2 and p_{12} is given directly by the pmf based on survival function expressed in Equation (3), that is,*

$$P(X_1 = x_1, X_2 = x_2) = p_1^{x_1-1} p_2^{x_2-1} p_{12}^{z_1} - p_1^{x_1} p_2^{x_2-1} p_{12}^{z_2} - p_1^{x_1-1} p_2^{x_2} p_{12}^{z_3} + p_1^{x_1} p_2^{x_2} p_{12}^{z_4} \quad (34)$$

where $z_1 = \max(x_1 - 1, x_2 - 1), z_2 = \max(x_1, x_2 - 1), z_3 = \max(x_1 - 1, x_2)$ e $z_4 = \max(x_1, x_2)$. *Thus, the likelihood function is rewritten as follows:*

$$L(p_1, p_2, p_{12}) = \prod_{i=1}^{n} \left[p_1^{x_{1i}-1} p_2^{x_{2i}-1} p_{12}^{z_{1i}} - p_1^{x_{1i}} p_2^{x_{2i}-1} p_{12}^{z_{2i}} - p_1^{x_{1i}-1} p_2^{x_{2i}} p_{12}^{z_{3i}} + p_1^{x_{1i}} p_2^{x_{2i}} p_{12}^{z_{4i}} \right] \quad (35)$$

4.2.2 Censored Data

Often in applications with lifetimes it is common the presence of censored data, that could be right, left or interval censoring. In this study, we assume the presence of right censoring data. Let $(X_{11}, X_{21}), (X_{12}, X_{22}), \ldots, (X_{1n}, X_{2n})$ be a random sample of size n from a BD(p_1, p_2, p_{12}) distribution. In presence of right random data,we define the following indicator variables:

$$\begin{cases} \delta_{1i} = 1 & \text{if} \quad X_{1i} < C_{1i} \quad \text{and 0, for the other part.} \\ \delta_{2i} = 1 & \text{if} \quad X_{2i} < C_{2i} \quad \text{and 0, for the other part.} \end{cases} \tag{36}$$

where $i = 1, 2, ..., n; (C_{1i}, C_{2i})$ are the right censoring times. In this case, we have four possible situations:

1. Both, X_{1i} and X_{2i}, are complete observations ($\delta_{1i} = 1, \delta_{2i} = 1$),
2. X_{1i} are complete and X_{2i} are censored ($\delta_{1i} = 1, \delta_{2i} = 0$),
3. X_{1i} are censored and X_{2i} are complete ($\delta_{1i} = 0, \delta_{2i} = 1$),
4. Both, X_{1i} and X_{2i}, are censored observations ($\delta_{1i} = 0, \delta_{2i} = 0$).

In all cases, the observed data is given by the expressions $t_{1i} = \min(X_{1i}, C_{1i}), i = 1, 2, ..., n$ and $t_{2i} = min(X_{2i}, C_{2i}), i = 1, 2, ..., n$. Then, the contributions for the likelihood function on the ith-observation are given by:

1. X_{1i} and X_{2i} are complete observations: $[P(X_{1i} = x_{1i}, X_{2i} = x_{2i})]^{\delta_{1i}\delta_{2i}}$,
2. X_{1i} is a complete observation and X_{2i} censored observation: $[P(X_{1i} = x_{1i}, X_{2i} > c_{2i})]^{\delta_{1i}(1-\delta_{2i})}$,
3. X_{1i} is a censored observation and X_{2i} is a complete observation: $[P(X_{1i} > c_{1i}, X_{2i} = x_{2i})]^{(1-\delta_{1i})\delta_{2i}}$,
4. X_{1i} and X_{2i} are censored observations: $[P(X_{1i} > c_{1i}, X_{2i} > c_{2i})]^{(1-\delta_{1i})(1-\delta_{2i})}$.

Thus, the likelihood function is is given by,

$$\begin{aligned} L(p_1, p_2, p_{12}) &= \prod_{i=1}^{n} [P(X_{1i} = x_{1i}, X_{2i} = x_{2i})]^{\delta_{1i}\delta_{2i}} [P(X_{1i} = x_{1i}, X_{2i} > c_{2i})]^{\delta_{1i}(1-\delta_{2i})} \\ &\times [P(X_{1i} > c_{1i}, X_{2i} = x_{2i})]^{(1-\delta_{1i})\delta_{2i}} [P(X_{1i} > c_{1i}, X_{2i} > c_{2i})]^{(1-\delta_{1i})(1-\delta_{2i})} \end{aligned} \tag{37}$$

Remarks: *For $\delta_{1i} = \delta_{2i} = 1$, the first term of likelihood function is:*

$$P(T_{1i} = t_{1i}, T_{2i} = t_{2i}) = p_1^{t_{1i}-1} p_2^{t_{2i}-1} p_{12}^{z_{1i}} - p_1^{t_{1i}} p_2^{t_{2i}-1} p_{12}^{z_{2i}} - p_1^{t_{1i}-1} p_2^{t_{2i}} p_{12}^{z_{3i}} + p_1^{t_{1i}} p_2^{t_{2i}} p_{12}^{z_{4i}} \tag{38}$$

for the case of $\delta_{1i} = 1, \delta_{2i} = 0$, the second term of likelihood function is given by:

$$P(T_{1i} = t_{1i}, T_{2i} > t_{2i}) = p_1^{t_{1i}-1} p_2^{t_{2i}} p_{12}^{z_{3i}} - p_1^{t_{1i}} p_2^{t_{2i}} p_{12}^{z_{4i}}; \tag{39}$$

for the case of $\delta_{1i} = 0, \delta_{2i} = 1$, the third term of likelihood function is given by:

$$P(T_{1i} > t_{1i}, T_{2i} = t_{2i}) = p_1^{t_{1i}} p_2^{t_{2i}-1} p_{12}^{z_{2i}} - p_1^{t_{1i}} p_2^{t_{2i}} p_{12}^{z_{4i}}; \tag{40}$$

and for the case of $\delta_{1i} = \delta_{2i} = 0$*, the last term of likelihood function is given by:*

$$P(T_{1i} > t_{1i}, T_{2i} > t_{2i}) = p_1^{t_{1i}} p_2^{t_{2i}} p_{12}^{z_{4i}}, \tag{41}$$

where $z_{1i} = \max(T_{1i} - 1, T_{2i} - 1), z_{2i} = \max(T_{1i}, T_{2i} - 1), z_{3i} = \max(T_{1i} - 1, T_{2i})$ *and* $z_{1i} = \max(T_{1i}, T_{2i})$.

Finally, the log-likelihood of bivariate Basu-Dhar geometric distribution considering right censoring data is expressed as follows

$$\begin{aligned}
\ell(p_1, p_2, p_{12}) &= \sum_{i=1}^{n} \delta_{1i}\delta_{2i} \log[p_1^{t_{1i}-1} p_2^{t_{2i}-1} p_{12}^{z_{1i}} - p_1^{t_{1i}} p_2^{t_{2i}-1} p_{12}^{z_{2i}} - p_1^{t_{1i}-1} p_2^{t_{2i}} p_{12}^{z_{3i}} + p_1^{t_{1i}} p_2^{t_{2i}} p_{12}^{z_{4i}}] \\
&+ \sum_{i=1}^{n} \delta_{1i}(1-\delta_{2i}) \log[p_1^{t_{1i}-1} p_2^{t_{2i}} p_{12}^{z_{3i}} - p_1^{t_{1i}} p_2^{t_{2i}} p_{12}^{z_{4i}}] \\
&+ \sum_{i=1}^{n} (1-\delta_{1i})(1-\delta_{2i}) \log[p_1^{t_{1i}} p_2^{t_{2i}} p_{12}^{z_{4i}}] \\
&+ \sum_{i=1}^{n} \delta_{2i}(1-\delta_{1i}) \log[p_1^{t_{1i}} p_2^{t_{2i}-1} p_{12}^{z_{2i}} - p_1^{t_{1i}} p_2^{t_{2i}} p_{12}^{z_{4i}}]
\end{aligned} \tag{42}$$

where $z_{1i} = \max(T_{1i}-1, T_{2i}-1), z_{2i} = \max(T_{1i}, T_{2i}-1), z_{3i} = \max(T_{1i}-1, T_{2i})$ and $z_{1i} = \max(T_{1i}, T_{2i})$. Defining $a = p_1^{t_{1i}-1} p_2^{t_{2i}-1}, b = p_1^{t_{1i}} p_2^{t_{2i}-1}, c = p_1^{t_{1i}-1} p_2^{t_{2i}}$ and $d = p_1^{t_{1i}} p_2^{t_{2i}}$, them, from the log-likelihood (Equation 42), the first derivatives in censored case for p_1, p_2 and p_{12} are, respectively, given by:

$$\begin{aligned}
\frac{\partial \ell}{\partial p_1} &= \sum_{i=1}^{n} \delta_{1i}\delta_{2i} \left[\frac{(t_{1i}-1)ap_1^{-1}p_{12}^{z_{1i}} - at_{1i}p_{12}^{z_{2i}}}{ap_{12}^{z_{1i}} - bp_{12}^{z_{2i}} - cp_{12}^{z_{3i}} + dp_{12}^{z_{4i}}} - \frac{(t_{1i}-1)cp_1^{-1}p_{12}^{z_{3i}} - ct_{1i}p_{12}^{z_{4i}}}{ap_{12}^{z_{1i}} - bp_{12}^{z_{2i}} - cp_{12}^{z_{3i}} + dp_{12}^{z_{4i}}} \right] \\
&+ \sum_{i=1}^{n} \frac{(1-\delta_{1i})(1-\delta_{2i})t_{1i}}{p_1} + \sum_{i=1}^{n} \delta_{1i}(1-\delta_{2i}) \left[\frac{(t_{1i}-1)cp^{-1}p_{12}^{z_{3i}} - ct_{1i}p_{12}^{z_{4i}}}{cp_{12}^{z_{3i}} - dp_{12}^{z_{4i}}} \right] \\
&+ \sum_{i=1}^{n} \delta_{2i}(1-\delta_{1i}) \left[\frac{at_{1i}p_{12}^{z_{2i}} - ct_{1i}p_{12}^{z_{4i}}}{bp_{12}^{z_{2i}} - dp_{12}^{z_{4i}}} \right] \\
\frac{\partial \ell}{\partial p_2} &= \sum_{i=1}^{n} \delta_{1i}\delta_{2i} \left[\frac{(t_{2i}-1)ap_2^{-1}p_{12}^{z_{1i}} - at_{2i}p_{12}^{z_{3i}}}{ap_{12}^{z_{1i}} - bp_{12}^{z_{2i}} - cp_{12}^{z_{3i}} + dp_{12}^{z_{4i}}} - \frac{(t_{2i}-1)bp_2^{-1}p_{12}^{z_{2i}} - bt_{2i}p_{12}^{z_{4i}}}{ap_{12}^{z_{1i}} - bp_{12}^{z_{2i}} - cp_{12}^{z_{3i}} + dp_{12}^{z_{4i}}} \right] \\
&+ \sum_{i=1}^{n} \frac{(1-\delta_{1i})(1-\delta_{2i})t_{2i}}{p_2} + \sum_{i=1}^{n} \delta_{1i}(1-\delta_{2i}) \left[\frac{at_{2i}p_{12}^{z_{3i}} - bt_{2i}p_{12}^{z_{4i}}}{cp_{12}^{z_{3i}} - dp_{12}^{z_{4i}}} \right] \\
&+ \sum_{i=1}^{n} \delta_{2i}(1-\delta_{1i}) \left[\frac{(t_{2i}-1)bp_2^{-1}p_{12}^{z_{2i}} - bt_{2i}p_{12}^{z_{4i}}}{bp_{12}^{z_{2i}} - dp_{12}^{z_{4i}}} \right]
\end{aligned} \tag{43}$$

$$\begin{aligned}\frac{\partial \ell}{\partial p_{12}} &= \sum_{i=1}^{n} \delta_{1i}\delta_{2i}\left[\frac{z_{1i}ap_{12}^{z_{1i}-1} - z_{2i}bp_{12}^{z_{2i}-1} - z_{3i}cp_{12}^{z_{3i}-1} + z_{4i}dp_{12}^{z_{4i}-1}}{ap_{12}^{z_{1i}} - bp_{12}^{z_{2i}} - cp_{12}^{z_{3i}} + dp_{12}^{z_{4i}}}\right] \\ &+ \sum_{i=1}^{n} \frac{(1-\delta_{1i})(1-\delta_{2i})z_{4i}}{p_{12}} + \sum_{i=1}^{n} \delta_{1i}(1-\delta_{2i})\left[\frac{z_{3i}cp_{12}^{z_{3i}-1} - z_{4i}dp_{12}^{z_{4i}-1}}{cp_{12}^{z_{3i}} - dp_{12}^{z_{4i}}}\right] \\ &+ \sum_{i=1}^{n} \delta_{2i}(1-\delta_{1i})\left[\frac{z_{2i}bp_{12}^{z_{2i}-1} - z_{4i}dp_{12}^{z_{4i}-1}}{bp_{12}^{z_{2i}} - dp_{12}^{z_{4i}}}\right]\end{aligned}$$

4.3 A Bayesian approach and the presence of Covariates

For a Bayesian analysis, we assume uniform prior distribution $U(0,1)$ for the three parameters (p_1, p_2, and p_{12}) not considering the presence of covariates. It is also assumed mean square error with Euclidean norm as the risk function. Notice that, it is not possible to get closed forms for the Bayes estimators of p_1, p_2 and p_{12} due to the complex form of the joint posterior distribution even using those selected priors. In this case, in terms of computational aspects, the joint posterior distribution is similar to the likelihood function in the MLEs computation.

In presence of a vector of k covariates, denoted by $S_i = (S_{1i}, S_{2i}, \ldots, S_{ki})$, we assume logistic regression models for the parameters p_1 and p_2, that is,

$$\begin{cases}\text{logit}(p_{1i}) = \beta_{10} + \beta_{11}S_{1i} + \ldots + \beta_{1k}S_{ki} \\ \text{logit}(p_{2i}) = \beta_{20} + \beta_{21}S_{1i} + \ldots + \beta_{2k}S_{ki}\end{cases} \tag{44}$$

where $\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$. For a Bayesian analysis of this regression model it is assumed normal prior distributions for the regression parameters with mean equals to zero and large variance (non-informative priors) and an uniform prior distribution, $U(0,1)$, for the parameter p_{12}.

5 Simulation Study

In this section, using Monte Carlo simulations, we present some simulation results for the Basu-Dhar bivariate geometric distribution to evaluate the performance of the moments estimators (MOM), the maximum likelihood estimators (MLEs) and the Bayes estimators. For the maximum likelihood estimation for the parameters of the model, we considered the two forms of log-likelihood: the first one, named MLE I, is obtained from Equation (35), and the second, named MLE II, is obtained from Equation (30). For the Bayesian estimation approach, we considered uniform $U(0,1)$ independent priors for the parameters p_1, p_2 and p_{12} for both forms of log-likelihood function.

To simulate observations from this model, the marginal distribution of X and the conditional distribution of Y given X are used. It is important to point out that an observation from the marginal distribution of the random variable X is generated using the inverse-transformation method. Using this realization of x, a value of Y is generated using the inverse-transformation method again based on the conditional distribution of Y given $X = x$ as given by (Li and Dhar, 2013):

$$P(X_2 = x_2 \mid X_1 = x_1) = \begin{cases} p_2^{x_2-1} q_2 & \text{if} \quad x_2 < x_1 \\ \dfrac{p_2^{x_2-1}(1 - p_1 p_{12} - p_2 p_{12} + p_1 p_2 p_{12})}{1 - p_1 p_{12}} & \text{if} \quad x_2 = x_1 \\ \dfrac{p_2^{x_2-1} p_{12}^{x_2-x_1} q_1 (1 - p_2 p_{12})}{1 - p_1 p_{12}} & \text{if} \quad x_2 > x_1 \end{cases} \tag{45}$$

The simulation was performed using the R software version 3.3.0 (R Core Team, 2016). For this study, we have taken sample sizes $n = 40, 60, 120, 160, 200, 250, 300$ and fixed parameter values $(p_1, p_2, p_{12}) = (0.90, 0.95, 0.97)$. For each combination (n, p_1, p_2, p_{12}), we generated 500 simulations from Basu-Dhar bivariate geometric distribution using the following algorithm (Li and Dhar, 2013):

1. Generate $X_i \sim Geo(1 - p_1 p_{12})$;
2. Generate $U \sim U(0, 1)$;
3. Set $w = \dfrac{1 - p_1}{1 - p_1 p_{12}}$;
4. If $x_i = 1$ and $u_i < 1 - p_2 p_{12} w$, then set $y_i = 1$; else
5. If $x_i = 1$, $u_i \geq 1 - w(p_2 p_{12})^{s-1}$ and $u_i < 1 - w(p_2 p_{12})^s$, then set $y_i = s$, for $s = 2, \ldots 500$; else
6. If $u_i \geq 1 - p_2^j$ and $u_i < 1 - p_2^{j+1}$, then set $y_i = j + 1$, for $j = 1, \ldots x_i - 2$; else
7. If $u_i \geq 1 - p_2^{x_i-1}$ and $u_i < 1 - p_2^{x_i} p_{12} w$, then set $y_i = x_i$; else
8. If $u_i \geq 1 - p_2^{x_i} p_{12} w$ and $u_1 < 1 - p_2^{x_i+1} p_{12}^2 w$, then set $y_i = x_i + 1$; else
9. If $u_i \geq 1 - p_{12}^{1-x_i} (p_2 p_{12})^k w$ and $u_i < 1 - p_{12}^{1-x_i} (p_2 p_{12})^{k+1} w$, then set $y_i = k + 1$, for $k = 1 + x_i, \ldots, 500$.

The estimates for the parameters p_1, p_2 and p_{12} were obtained using the method of moments (MOM) using the equations (24), the MLE estimates were obtained by the Nelder-Mead method using **maxLik** function and **optim.method = "NM"** from *maxLik* package (Henningsen and Toomet, 2011). The Bayes estimates were obtained by **MCMCmetrop1R** function from *MCMCpack* package (Martin et al., 2011) with arguments **burn = 1000, mcmc = 10000, thin = 1**, **optim.method = "Nelder-Mead"**. To assess the performance of the methods, we calculated the bias and the MSE (mean-squared-error) for the simulated estimates of p_1, p_2 and p_{12}. Also, we calculated the execution time of simulation for each method. The results are presented in Tables 1 and 2 and the R codes used for simulation study are given in Appendix II at the end of this manuscript.

From these simulated observations, the mean of the 500 estimated vectors of the parameter $p = (p_1, p_2, p_{12})$, the biases and the MSE based on these 500 vectors are computed. The performances of the five estimation methods are assessed based on the estimated expected value of the estimator vector and their estimated biases and estimated MSE for different sample sizes. The biases and MSE were calculated as:

$$\text{BIAS}(\widehat{p}) = \frac{1}{B}\sum_{i=1}^{B}(\widehat{p}_i - p_i) \quad \text{and} \quad \text{MSE}(\widehat{p}) = \frac{1}{B}\sum_{i=1}^{B}(\widehat{p}_i - p_i)^2$$

where B is the number of simulations and $p = (p_1, p_2, p_{12})$.

Table 1 shows, respectively, the mean of the estimates (for MOM, MLE I, MLE II, Bayes I and Bayes II), the biases and the MSE of the obtained estimates for p_1, p_2 and p_{12} for each simulated sample. The results are also illustrated in the Figures 1 and 2. From the results of Table 1, it is observed that:

1. When n is small, the biases and MSEs for MOM, MLE I, MLE I and Bayes I estimates are greater when compared to the case when $n \to \infty$. In addition, the Bayes I estimates have negative biases for all sample sizes. Except for two cases, the same happens to MLE I;

2. The Bayes II and MLE II have the same problems to estimate the parameter p_{12} in the simulation study which leads to poor estimates for biases and MSEs. This fact may be related to the MCMC iterations or the correlation structure between X_1 and X_2.

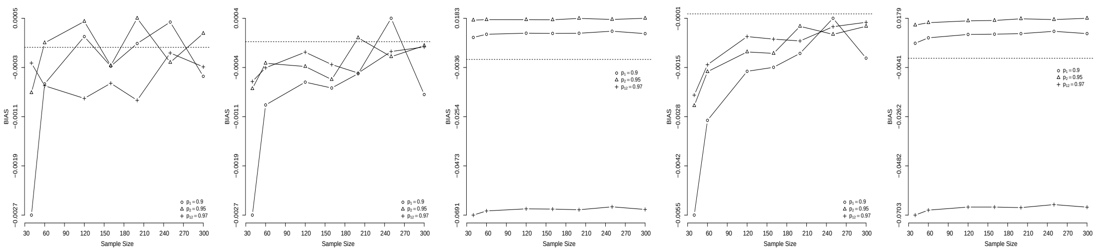

Figure 1: Biases for the BD model parameters considering all estimation methods (MOM $\to$ Bayes II) in the simulation study under a Bayesian approach.

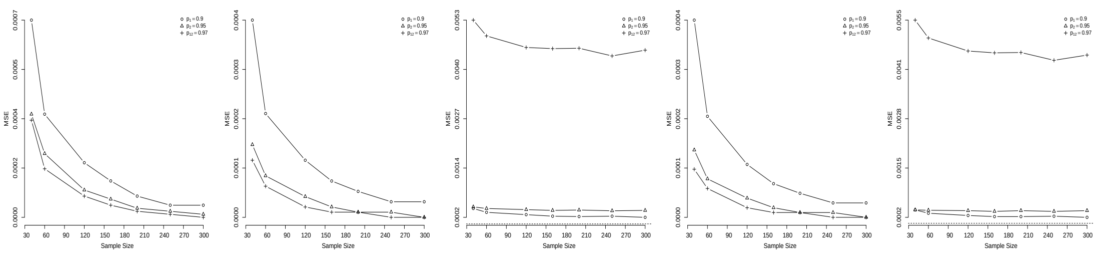

Figure 2: MSEs for the BD model parameters considering all estimation methods (MOM $\to$ Bayes II) in the simulation study under a Bayesian approach.

Table 1: Simulation results.

n	True Par.	MOM (Mean)	BIAS (MOM)	MSE (MOM)	MLE I (Mean)	BIAS (MLE I)	MSE (MLE I)	MLE II (Mean)	BIAS (MLE II)	MSE (MLE II)
40	$p_1 = 0.90$	0.89734	-0.00266	0.00069	0.89734	-0.00266	0.00040	0.90976	0.00976	0.00040
	$p_2 = 0.95$	0.94928	-0.00072	0.00038	0.94928	-0.00072	0.00016	0.96745	0.01745	0.00044
	$p_{12} = 0.97$	0.96975	-0.00025	0.00036	0.96939	-0.00061	0.00013	0.90087	-0.06913	0.00529
60	$p_1 = 0.90$	0.89942	-0.00058	0.00038	0.89903	-0.00097	0.00022	0.91117	0.01117	0.00030
	$p_2 = 0.95$	0.95007	0.00007	0.00025	0.94967	-0.00033	0.00010	0.96766	0.01766	0.00040
	$p_{12} = 0.97$	0.96939	-0.00061	0.00020	0.96960	-0.00040	0.00008	0.90272	-0.06728	0.00488
120	$p_1 = 0.90$	0.90017	0.00017	0.00022	0.89938	-0.00062	0.00013	0.91165	0.01165	0.00024
	$p_2 = 0.95$	0.95041	0.00041	0.00013	0.94962	-0.00038	0.00006	0.96763	0.01763	0.00037
	$p_{12} = 0.97$	0.96919	-0.00081	0.00011	0.96984	-0.00016	0.00004	0.90365	-0.06635	0.00458
160	$p_1 = 0.90$	0.89970	-0.00030	0.00016	0.89929	-0.00071	0.00009	0.91155	0.01155	0.00020
	$p_2 = 0.95$	0.94971	-0.00029	0.00010	0.94942	-0.00058	0.00004	0.96760	0.01760	0.00035
	$p_{12} = 0.97$	0.96943	-0.00057	0.00008	0.96965	-0.00035	0.00003	0.90354	-0.06646	0.00455
200	$p_1 = 0.90$	0.90006	0.00006	0.00011	0.89951	-0.00049	0.00007	0.91162	0.01162	0.00019
	$p_2 = 0.95$	0.95046	0.00046	0.00007	0.95006	0.00006	0.00003	0.96822	0.01822	0.00036
	$p_{12} = 0.97$	0.96916	-0.00084	0.00006	0.96952	-0.00048	0.00003	0.90324	-0.06676	0.00456
250	$p_1 = 0.90$	0.90040	0.00040	0.00008	0.90036	0.00036	0.00005	0.91254	0.01254	0.00020
	$p_2 = 0.95$	0.94976	-0.00024	0.00006	0.94977	-0.00023	0.00003	0.96775	0.01775	0.00034
	$p_{12} = 0.97$	0.96991	-0.00009	0.00005	0.96985	-0.00015	0.00002	0.90455	-0.06545	0.00436
300	$p_1 = 0.90$	0.89954	-0.00046	0.00008	0.89919	-0.00081	0.00005	0.91145	0.01145	0.00017
	$p_2 = 0.95$	0.95022	0.00022	0.00005	0.94994	-0.00006	0.00002	0.96827	0.01827	0.00035
	$p_{12} = 0.97$	0.96969	-0.00031	0.00004	0.96992	-0.00008	0.00002	0.90337	-0.06663	0.00451

n	True Par.	MOM (Mean)	BIAS (MOM)	MSE (MOM)	Bayes I (Mean)	BIAS (Bayes I)	MSE (Bayes I)	Bayes II (Mean)	BIAS (Bayes II)	MSE (Bayes II)
40	$p_1 = 0.90$	0.89734	-0.00266	0.00069	0.89450	-0.00550	0.00043	0.90670	0.00670	0.00036
	$p_2 = 0.95$	0.94928	-0.00072	0.00038	0.94749	-0.00251	0.00016	0.96485	0.01485	0.00036
	$p_{12} = 0.97$	0.96975	-0.00025	0.00036	0.96778	-0.00222	0.00012	0.89971	-0.07029	0.00546
60	$p_1 = 0.90$	0.89942	-0.00058	0.00038	0.89709	-0.00291	0.00023	0.90914	0.00914	0.00027
	$p_2 = 0.95$	0.95007	0.00007	0.00025	0.94842	-0.00158	0.00010	0.96593	0.01593	0.00035
	$p_{12} = 0.97$	0.96939	-0.00061	0.00020	0.96861	-0.00139	0.00008	0.90194	-0.06806	0.00498
120	$p_1 = 0.90$	0.90017	0.00017	0.00022	0.89843	-0.00157	0.00013	0.91066	0.01066	0.00021
	$p_2 = 0.95$	0.95041	0.00041	0.00013	0.94896	-0.00104	0.00006	0.96675	0.01675	0.00034
	$p_{12} = 0.97$	0.96919	-0.00081	0.00011	0.96938	-0.00062	0.00004	0.90331	-0.06669	0.00463
160	$p_1 = 0.90$	0.89970	-0.00030	0.00016	0.89854	-0.00146	0.00009	0.91078	0.01078	0.00018
	$p_2 = 0.95$	0.94971	-0.00029	0.00010	0.94892	-0.00108	0.00004	0.96692	0.01692	0.00032
	$p_{12} = 0.97$	0.96943	-0.00057	0.00008	0.96931	-0.00069	0.00003	0.90331	-0.06669	0.00458
200	$p_1 = 0.90$	0.90006	0.00006	0.00011	0.89892	-0.00108	0.00007	0.91102	0.01102	0.00018
	$p_2 = 0.95$	0.95046	0.00046	0.00007	0.94966	-0.00034	0.00003	0.96767	0.01767	0.00034
	$p_{12} = 0.97$	0.96916	-0.00084	0.00006	0.96965	-0.00074	0.00003	0.90306	-0.06694	0.00459
250	$p_1 = 0.90$	0.90040	0.00040	0.00008	0.89988	-0.00012	0.00005	0.91206	0.01206	0.00019
	$p_2 = 0.95$	0.94976	-0.00024	0.00006	0.94944	-0.00056	0.00003	0.96732	0.01732	0.00032
	$p_{12} = 0.97$	0.96991	-0.00009	0.00005	0.96965	-0.00035	0.00002	0.90442	-0.06558	0.00438
300	$p_1 = 0.90$	0.89954	-0.00046	0.00008	0.89879	-0.00121	0.00005	0.91103	0.01103	0.00016
	$p_2 = 0.95$	0.95022	0.00022	0.00005	0.94966	-0.00034	0.00002	0.96790	0.01790	0.00034
	$p_{12} = 0.97$	0.96969	-0.00031	0.00004	0.96977	-0.00023	0.00002	0.90326	-0.06674	0.00452

From Figures 1 and 2, it is possible to conclude that the MLE II and Bayes II estimators do not have nice properties showing inconsistency for the parameter p_{12}. This fact could be justified by the existing correlation structure among the parameters and the complexity of the log-likelihood function presented by equation (30) when compared to the log-likelihood presented in equation (35). For the others estimators, MOM, MLE I and Bayes I, the BD model parameters have good properties when $n \to \infty$.

Table 2 shows, respectively, the CPU time for user and system in the simulation study. It is important to point out that the simulation study was performed in a Core i3-3240 (3.40 Ghz) machine with 8 GB DDR3 RAM and Windows 10 Pro (version 1703) as operating system.

Table 2: Execution time (in seconds) of simulation.

Method	MOM	MLE I	MLE II	Bayes I	Bayes II
User	6014.95	6142.58	6180.63	12905.80	7722.79
System	1.07	0.78	2.19	1.27	2.15
Total	6081.79	6180.60	6266.66	12939.05	7792.48

From the results of Table 2, it is possible to see that the smaller system time is obtained for MLE I estimation method and higher in MLE II estimation method. Between the two Bayes estimators, the Bayes I have almost the double of execution time than Bayes II, but, by Table 1, Bayes I have better estimates, bias and MSE when compared to the Bayes II estimation. For the two MLEs estimation methods, it is observed that both have close execution times, but the MLE I has more consistent estimates when compared to the MLE II estimates.

6 Applications

6.1 An example with no censored observations

Let us assume a real data set from Dhar (2003). This data set consists of scores given by seven sport judges in an international diver competition from seven different countries in the form of a video recording. The score given by each judge is a discrete random variable taking positive integer values and also the midpoints of consecutive integers between zero and ten (dataset in the Appendix I at the end of the manuscript).

For the analysis, we assumed a BD with pmf given by Equation (4). For the moments estimation method, it is calculated from the sample the sample means $\overline{X} = 13.947$, $\overline{Y} = 14.368$ and $\overline{W} = 13.789$, from which it is obtained the estimates, $\tilde{p}_1 = 0.9968$, $\tilde{p}_2 = 0.9991$ and $\tilde{p}_{12} = 0.9312$. It is important to point out that the estimates based on the moments method were used as initial values for MLEs and Bayes estimates computation.

The model, using all methods, was estimated using the R software. The MLE estimates and confidence intervals for p_1, p_2 and p_{12} are presented in Table 3. For a Bayesian analysis, we considered uniform $U(0, 1)$ priors for p_1, p_2 and p_{12}, and it was used MCMC (Markov Chain Monte Carlo) simulation methods to get the Bayesian estimates.

Table 3: MLE and Bayes posterior summaries for the application 1.

Par.	MLE I	S.E.	95% Conf. Int.	MLE II	S.E.	95% Conf. Int.
p_1	0.9616	0.0124	(0.9373, 0.9858)	0.9618	0.0123	(0.9376, 0.9859)
p_2	0.9854	0.0098	(0.9663, 0.9999)	0.9859	0.0095	(0.9673, 0.9999)
p_{12}	0.9401	0.0158	(0.9093, 0.9711)	0.9371	0.0164	(0.9051, 0.9692)
Par.	Bayes I	S.D.	95% Cred. Int.	Bayes II	S.D.	95% Cred. Int.
p_1	0.9575	0.0129	(0.9290, 0.9791)	0.9577	0.0132	(0.9286, 0.9798)
p_2	0.9796	0.0108	(0.9529, 0.9957)	0.9799	0.0109	(0.9542, 0.9959)
p_{12}	0.9391	0.0157	(0.9048, 0.9670)	0.9358	0.0168	(0.9013, 0.9652)

From the results of Table 3, it is concluded that the four used methods are quite similar to estimate the model parameters and do not change considerably moving from one method to another. Similar results also are obtained for the Bayesian credibility intervals and confidence intervals. Also it is observed that for all methods, the standard errors for the estimates are very low that implies in good length of confidence and credibility intervals, an indication that BD has a good performance for this dataset.

6.2 An example with censored observations and covariates

As another application of the proposed methodology, let us now to consider a survival dataset related to kidney infection (McGilchrist and Aisbett, 1991) where the recurrence of infection of 38 kidney patients, using portable dialysis machines, are recorded. Infections may occur at the location of insertion of the catheter. The time recorded, called infection time, is either the survival times (in complete weeks) of the patient until an infection occurred and the catheter had to be removed, or the censored time, where the catheter was removed for others reasons. The catheter is reinserted after some time and the second infection time is again observed or censored (dataset in the Appendix I at the end of the manuscript).

For the analysis not considering the presence of covariates, we assumed a BD with pmf given by Equation (4). It is important to say that the MLE estimates and confidence intervals for p_1, p_2 and p_{12} were obtained by log-likelihood with censoring and estimated on R software (R Core Team, 2016) and the results are presented in Table 4. For a Bayesian analysis, we considered as prior an uniform prior distribution, $U(0, 1)$, for p_1, p_2 and p_{12}, using the MCMC (Markov Chain Monte Carlo) simulation method to get the Bayes estimates.

Table 4: MLE and Bayes posterior summaries for the model not considering the covariate gender for the application 2.

Par.	MLE	S.E.	95% Conf. Int.	Bayes	S.D.	95% Cred. Int.
p_1	0.9512	0.0122	(0.9273, 0.9751)	0.9495	0.0095	(0.9290, 0.9660)
p_2	0.9485	0.0110	(0.9269, 0.9700)	0.9487	0.0105	(0.9278, 0.9682)
p_{12}	0.9950	0.0041	(0.9870, 1.0000)	0.9947	0.0042	(0.9844, 0.9998)

From the results of Table 4, it is concluded that the MLE and Bayesian estimates are quite similar and do not change considerably moving from one method to another. Similar results also are observed for the confidence and credibility intervals. Also, it is observed that the standard errors of the estimates are very low in both methods. That implies in good length of confidence and credibility intervals which suggests the good performance of the BD distribution for the dataset in presence of censored data.

In the analysis in presence of the covariate gender, we assumed logit regression models for the parameters p_1 and p_2 given by Equation (44). Only the Bayesian estimates are considered in this case. For the Bayesian analysis, we assumed normal priors for the regression parameters ($\beta_{10}, \beta_{20}, \beta_{11}, \beta_{21} \sim N(0, 100)$) and an uniform prior distribution, $U(0, 1)$, for the parameter p_{12}. Posterior summaries for the model were obtained using the OpenBugs software (Spiegelhalter et al., 2007) (see Table 5).

Table 5: Posterior summaries in presence of the gender covariate for application 2.

Par.	Mean	S.D.	95% Conf. Int.
β_{10}	-0.4816	0.6892	(-1.9504, 0.7237)
β_{11}	1.8721	0.3945	(1.2143, 2.6514)
β_{20}	2.8243	0.8804	(1.2073, 4.3396)
β_{21}	0.1189	0.4903	(-0.8292, 1.0592)
p_{12}	0.9944	0.0042	(0.9855, 0.9998)

From the results of Table 5, it is concluded that similar to the other examples, the BD distribution is a good alternative for the data analysis in presence of censored data and covariates.

7 Concluding remarks

In this paper, we presented some computational aspects of bivariate Basu-Dhar distribution introduced by Basu and Dhar (1995). The results obtained in a simulation study leads to the conclusion that BD have good computational aspects in terms of system time. Also, the biases and the MSEs are lower among all estimation methods considered in the simulation study.

Another point to interest, in mathematical terms, the log-likelihood (with censored data and complete data) are quite simple and do not have terms depending on exponential function as it is common with other discrete distributions leading to good computational aspects avoiding instability in parameter estimation. It is not illustrated here, but the convergence of all Markov chains were observed from trace plots of the simulated Gibbs samples.

Finally, as observed in the applications with real datasets, the BD distribution could be a good alternative to analyze lifetime data assuming discrete rather than continuous data with small computational costs to get the inferences of interest as compared to many existing bivariate parametric lifetime distributions or bivariate models derived from copula functions for continuous bivariate lifetime data (see for example, Achcar et al., 2016b).

References

Achcar, J., Davarzani, N., and Souza, R. (2016a). Basu–Dhar bivariate geometric distribution in the presence of covariates and censored data: a bayesian approach. *Journal of Applied Statistics*, 43(9):1636–1648.

Achcar, J. A. and Leandro, R. A. (1998). Use of markov chain monte carlo methods in a bayesian analysis of the block and basu bivariate exponential distribution. *Annals of the Institute of Statistical Mathematics*, 50(3):403–416.

Achcar, J. A., Martinez, E. Z., and Cuevas, J. R. T. (2016b). Bivariate lifetime modelling using copula functions in presence of mixture and non-mixture cure fraction models, censored data and covariates. *Model Assisted Statistics and Applications*, 11(4):261–276.

Arnold, B. C. (1975). A characterization of the exponential distribution by multivariate geometric compounding. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 164–173.

Arnold, B. C. and Strauss, D. (1988). Bivariate distributions with exponential conditionals. *Journal of the American Statistical Association*, 83(402):522–527.

Basu, A. P. and Dhar, S. (1995). Bivariate geometric distribution. *Journal of Applied Statistical Science*, 2(1):33–44.

Block, H. W. and Basu, A. (1974). A continuous, bivariate exponential extension. *Journal of the American Statistical Association*, 69(348):1031–1037.

Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335.

Davarzani, N., Achcar, J. A., Smirnov, E. N., and Peeters, R. (2015). Bivariate lifetime geometric distribution in presence of cure fractions. *Journal of Data Science*, 13(4):755–770.

Dhar, S. K. (1998). Data analysis with discrete analogue of freund's model. *Journal of Applied Statistical Science*, 7:169–183.

Dhar, S. K. (2003). Modeling with a bivariate geometric distribution. *Advances on Methodological and Applied Aspects of Probability and Statistics*, 1:101–109.

Dhar, S. K. and Balaji, S. (2006). On the characterization of a bivariate geometric distribution. *Communications in Statistics—Theory and Methods*, 35(5):759–765.

dos Santos, C. A. and Achcar, J. A. (2011). A bayesian analysis for the block and basu bivariate exponential distribution in the presence of covariates and censored data. *Journal of Applied Statistics*, 38(10):2213–2223.

Downton, F. (1970). Bivariate exponential distributions in reliability theory. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 408–417.

Freund, J. E. (1961). A bivariate extension of the exponential distribution. *Journal of the American Statistical Association*, 56(296):971–977.

Gelfand, A. E. and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409.

Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, 55(292):698–707.

Hanagal, D. and Ahmadi, K. (2008). Estimation of parameters by em algorithm in bivariate exponential distribution based on censored samples. *Econ. Quality Control*, 23(2):257–66.

Hanagal, D. D. (2006). Bivariate weibull regression model based on censored samples. *Statistical Papers*, 47(1):137–147.

Hawkes, A. G. (1972). A bivariate exponential distribution with applications to reliability. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 129–131.

Henningsen, A. and Toomet, O. (2011). maxlik: A package for maximum likelihood estimation in R. *Computational Statistics*, 26(3):443–458.

Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika*, pages 671–678.

Klein, J. P. and Moeschberger, M. L. (2005). *Survival analysis: techniques for censored and truncated data.* Springer Science & Business Media.

Kocherlakota, S. and Kocherlakota, K. (1992). *Bivariate discrete distributions.* Wiley Online Library.

Krishna, H. and Pundir, P. S. (2009). A bivariate geometric distribution with applications to reliability. *Communications in Statistics—Theory and Methods*, 38(7):1079–1093.

Lawless, J. F. (1982). *Statistical models and methods for lifetime data.* John Wiley & Sons.

Li, J. and Dhar, S. K. (2013). Modeling with bivariate geometric distributions. *Communications in Statistics-Theory and Methods*, 42(2):252–266.

Marshall, A. W. and Olkin, I. (1967a). A generalized bivariate exponential distribution. *Journal of Applied Probability*, 4(02):291–302.

Marshall, A. W. and Olkin, I. (1967b). A multivariate exponential distribution. *Journal of the American Statistical Association*, 62(317):30–44.

Martin, A. D., Quinn, K. M., and Park, J. H. (2011). MCMCpack: Markov chain monte carlo in R. *Journal of Statistical Software*, 42(9):22.

McGilchrist, C. and Aisbett, C. (1991). Regression with frailty in survival analysis. *Biometrics*, pages 461–466.

Muraleedharan Nair, K. and Unnikrishnan Nair, N. (1988). On characterizing the bivariate exponential and geometric distributions. *Annals of the institute of Statistical Mathematics*, 40(2):267–271.

R Core Team (2016). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Sarkar, S. K. (1987). A continuous bivariate exponential distribution. *Journal of the American Statistical Association*, 82(398):667–675.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2007). Openbugs user manual, version 3.0. 2. *MRC Biostatistics Unit, Cambridge.*

Sun, K. and Basu, A. P. (1995). A characterization of a bivariate geometric distribution. *Statistics & probability letters*, 23(4):307–311.

Appendix I - Datasets

Table 6: Sports data: scores given by different judges.

Item	Diver	X:max score, Asian judge	Y:max score, West judge
1	Sun Shuwei, China	19	19
2	David Pichler, USA	15	15
3	Jan Hempel, Germany	13	14
4	Roman Volodkuv, Ukraine	11	12
5	Sergei Kudrevich, Belarus	14	14
6	Patrick Jeffrey, USA	15	14
7	Valdimir Timoshinin, Russia	13	16
8	Dimitry Sautin, Russia	7	5
9	Xiao Hailiang, China	13	13
10	Sun Shuwei, China	15	16
11	David Pichler, USA	15	15
12	Jan Hempel, Germany	17	18
13	Roman Volodkuv, Ukraine	16	16
14	Sergei Kudrevich, Belarus	12	13
15	Patrick Jeffrey, USA	14	14
16	Valdimir Timoshinin, Russia	12	13
17	Dimitry Sautin, Russia	17	18
18	Xiao Hailiang, China	9	10
19	Sun Shuwei, China	18	18

Table 7: Recurrence times of infections for 38 kidney patients.

Patient	First Time	Second Time	Status: First Time	Status: Second Time	Sex
1	1	2	1	1	1
2	3	2	1	0	2
3	3	4	1	1	1
4	64	45	1	1	2
5	4	2	1	1	1
6	3	35	1	1	2
7	1	1	1	1	1
8	73	4	1	1	2
9	8	28	1	1	2
10	2	22	1	1	1
11	1	48	1	1	2
12	20	1	1	0	2
13	14	5	1	1	2
14	21	10	0	0	2
15	77	4	1	0	2
16	2	1	1	0	1
17	26	17	1	1	2
18	42	16	1	1	2
19	3	23	0	0	2
20	2	15	1	0	2
21	22	80	1	1	1
22	57	3	1	0	2
23	2	9	1	1	2
24	6	7	1	0	2
25	2	6	1	1	1
26	16	29	0	1	2
27	19	22	1	1	2
28	5	4	1	1	2
29	0	4	1	1	1
30	19	4	1	1	2
31	4	8	1	1	2
32	1	6	0	1	2
33	22	4	1	1	2
34	27	1	1	0	2
35	17	1	1	1	2
36	8	2	0	0	2
37	1	11	0	1	2
38	9	1	1	0	1

Status: Censoring(0), infection(1);Sex: male(1), female(2)

Appendix II - R/OpenBugs Codes

1. Simulation

```
# Seed
set.seed(1212)

# True Parameters
p1              <- 0.90
p2              <- 0.95
p12             <- 0.97

# Simulation Structure
B               <- 500                                  # Number of Simulations
n               <- c(40, 60, 120, 160, 200, 250, 300)   # Samples
a               <- 1
out             <- array(0, dim = c(B, 3))
means           <- array(0, dim = c(length(n), 3))
bias            <- array(0, dim = c(length(n), 3))
mse             <- array(0, dim = c(length(n), 3))

for(m in n)
{
for(l in 1:B)
{
# Generating Xi
x               <- rgeom(m, 1 - p1 * p3) + 1
x               <- array(x, dim = c(m, 1))

# Generating Yi
u               <- runif(m)
u               <- array(u, dim = c(m, 1))
y               <- array(c(0), dim = c(m, 1))
for (i in 1:m)
{
        if(x[i] == 1 && u[i] < 1 - p2 * p3 * (1 - p1)/(1 - p1 * p3)) {y[i] = 1}
        else
        if(x[i] == 1)
                {for(s in 2:1000) if(u[i] >= 1 - (1 - p1) * (p2 * p3)^(s - 1)/(1 - p1 * p3) &&
                u[i] < 1 - (1 - p1) * (p2 * p3)^s/(1 - p1 * p3)) {y[i] = s}}
        else
                for (j in 1:x[i] - 2) if(u[i] >= 1 - p2^j && u[i] < 1 - p2^(j + 1)) {y[i] = j + 1}
        else
        if(u[i] >= 1 - p2^(x[i] - 1) && u[i] < 1 - (p2^x[i]) * p3 * (1 - p1)/(1 - p1 * p3))
                {y[i] = x[i]}
        else
        if(u[i] >= 1 - (p2^x[i]) * p3 * (1 - p1)/(1 - p1 * p3) &&
                u[i] < 1 - (1 - p1) * (p2^(x[i] + 1)) * p3^2/(1 - p1 * p3)) {y[i] = x[i] + 1}
        else
        if(u[i] >= 1 - (1 - p1) * (p2^(x[i] + 1)) * p3^2/(1 - p1 * p3))
                {for(k in 1 + x[i]:1000) if(u[i] >= 1 - (1 - p1) * p3^(-x[i] + 1) *
                (p2 * p3)^k/(1 - p1 * p3) && u[i] < 1 - (1 - p1) * p3^(-x[i] + 1) *
                (p2 * p3)^(k+1)/(1 - p1 * p3)) {y[i] = k + 1}}
}

# Generating min(Xi,Yi)
min_xy          <- array(c(0), dim = c(m, 1))
for(i in 1:m)
{
  min_xy[i]     <- min(x[i],y[i])
}


#  Moments estimator
xbar            <-  mean(x)
ybar            <-  mean(y)
zbar            <-  mean(min_xy)
estimative_p1   <-  (ybar - ybar * zbar)/(zbar - ybar * zbar)
estimative_p2   <-  (xbar - xbar * zbar)/(zbar - xbar * zbar)
estimative_p12  <-  (zbar - xbar * zbar - ybar * zbar + xbar * ybar * zbar)/(xbar * ybar * zbar - xbar * ybar)

## Out
out[l, 1]       <-  estimative_p1
out[l, 2]       <-  estimative_p2
out[l, 3]       <-  estimative_p12
}
means[a,1]      <-  mean(out[,1])
means[a,2]      <-  mean(out[,2])
means[a,3]      <-  mean(out[,3])
bias[a, 1]      <-  mean(out[,1] - p1, na.rm = TRUE)
bias[a, 2]      <-  mean(out[,2] - p2, na.rm = TRUE)
bias[a, 3]      <-  mean(out[,3] - p12, na.rm = TRUE)
```

```
mse[a, 1]        <- mean((out[,1] - p1)^2, na.rm = TRUE)
mse[a, 2]        <- mean((out[,2] - p2)^2, na.rm = TRUE)
mse[a, 3]        <- mean((out[,3] - p12)^2, na.rm = TRUE)
a <- a + 1
}

## For MLE I - Change moments estimator for MLE I estimator:

logvero          <- function(param)
{
        p1       <- param[1]
        p2       <- param[2]
        p12      <- param[3]
        if(param[1]<=0)return(-Inf)
        if(param[2]<=0)return(-Inf)
        if(param[3]<=0 && param[3] <=1)return(-Inf)
        z1       <- pmax(x - 1, y - 1)
        z2       <- pmax(x, y - 1)
        z3       <- pmax(x - 1, y)
        z4       <- pmax(x, y)
        L        <- log(p1^(x - 1) * p2^(y - 1) * p12^(z1) - p1^(x) * p2^(y - 1) * p12^(z2)
        - p1^(x - 1) * p2^(y) * p12^(z3) + p1^(x) * p2^(y) * p12^(z4))
        logL     <- sum(L)
        return(logL)
}

## For MLE II - Change moments estimator for MLE II estimator:

logvero          <- function(param)
{
        p1       <- param[1]
        p2       <- param[2]
        p12      <- param[3]
        if(param[1]<=0)return(-Inf)
        if(param[2]<=0)return(-Inf)
        if(param[3]<=0 && param[3] <=1)return(-Inf)
        v1       <- ifelse(x < y, 1, 0)
        v2       <- ifelse(x > y, 1, 0)
        n1       <- sum(v1)
        n2       <- sum(v2)
        n3       <- sum((1 - v1) * (1 - v2))
        n        <- n1 + n2 + n3
        T1       <- sum(v1 * (1 - v2) * x)
        T2       <- sum(v1 * (1 - v2) * y)
        T3       <- sum((1 - v1) * (1 - v2) * x)
        T4       <- sum((1 - v1) * (1 - v2) * y)
        T5       <- sum(v2 * (1 - v1) * x)
        T6       <- sum(v2 * (1 - v1) * y)
        L1       <- (T1 + T3 + T5 - n) * log(p1)
        L2       <- (T2 + T3 + T6 - n) * log(p2)
        L3       <- (T1 + T3 + T5 - n) * log(p12)
        L4       <- n1 * log(1 - p1) + n2 * log(1 - p2) + n3 * log(1 - p1 * p12 - p2 * p12
        + p1 * p2 * p12)
        L5       <- n1 * log(1 - p2 * p12) + n2 * log(1 - p1 * p12)
        L        <- L1 + L2 + L3 + L4 + L5
        logL     <- sum(L)
        return(logL)
}

## For Bayes I - Change moments estimator for Bayes I estimator:

log.post         <- function(t1, t2, theta)
{
        p1       <- theta[1]
        p2       <- theta[2]
        p12      <- theta[3]
        if(theta[1]<=0)return(Inf)
        if(theta[2]<=0)return(Inf)
        if(theta[3]<=0)return(Inf)
        z1       <- pmax(t1 - 1, t2 - 1)
        z2       <- pmax(t1, t2 - 1)
        z3       <- pmax(t1 - 1, t2)
        z4       <- pmax(t1, t2)
        l1       <- p1^(t1 - 1) * p2^(t2 - 1) * p12^(z1)
        l2       <- p1^(t1) * p2^(t2 - 1) * p12^(z2)
        l3       <- p1^(t1 - 1) * p2^(t2) * p12^(z3)
        l4       <- p1^(t1) * p2^(t2) * p12^(z4)
        l5       <- l1 - l2 - l3 + l4
        log.like <- log(l5)
        priori   <- dunif(p1, 0, 1) * dunif(p2, 0, 1) * dunif(p12, 0, 1)
        log.p    <- log(priori)
        L        <- sum(log.like+log.p)
        if(is.na(L)==TRUE){return(-Inf)}else{return(L)}
}
```

```
## For Bayes II - Change moments estimator for Bayes II estimator:

log.post        <- function(t1, t2, theta)
{
        p1      <- theta[1]
        p2      <- theta[2]
        p12     <- theta[3]
        if(theta[1]<=0)return(Inf)
        if(theta[2]<=0)return(Inf)
        if(theta[3]<=0)return(Inf)
        v1      <- ifelse(x < y, 1, 0)
        v2      <- ifelse(x > y, 1, 0)
        n1      <- sum(v1)
        n2      <- sum(v2)
        n3      <- sum((1 - v1) * (1 - v2))
        n       <- n1 + n2 + n3
        T1      <- sum(v1 * (1 - v2) * x)
        T2      <- sum(v1 * (1 - v2) * y)
        T3      <- sum((1 - v1) * (1 - v2) * x)
        T4      <- sum((1 - v1) * (1 - v2) * y)
        T5      <- sum(v2 * (1 - v1) * x)
        T6      <- sum(v2 * (1 - v1) * y)
        L1      <- (T1 + T3 + T5 - n) * log(p1)
        L2      <- (T2 + T3 + T6 - n) * log(p2)
        L3      <- (T1 + T3 + T5 - n) * log(p12)
        L4      <- n1 * log(1 - p1) + n2 * log(1 - p2) + n3 * log(1 - p1 * p12 - p2 * p12
        + p1 * p2 * p12)
        L5      <- n1 * log(1 - p2 * p12) + n2 * log(1 - p1 * p12)
        log.like<- L1 + L2 + L3 + L4 + L5
        priori  <- dunif(p1, 0, 1) * dunif(p2, 0, 1) * dunif(p12, 0, 1)
        log.p   <- log(priori)
        L       <- sum(log.like+log.p)
        if(is.na(L)==TRUE){return(-Inf)}else{return(L)}
}
```

2. Application 7.1 (No censored data)

Use the same codes provided by simulation codes for MLE I, MLE II, Bayes I, Bayes II.

3. Application 7.2 (Censored data)

```
## Maximum Likelihood Estimators

logvero2        <- function(param)
{
        # Parameters
        p1      <- theta[1]
        p2      <- theta[2]
        p3      <- theta[3]
        if(theta[1]<=0)return(Inf)
        if(theta[2]<=0)return(Inf)
        if(theta[3]<=0)return(Inf)

        # Max's
        z1      <- pmax(t1 - 1, t2 - 1)
        z2      <- pmax(t1, t2 - 1)
        z3      <- pmax(t1 - 1, t2)
        z4      <- pmax(t1, t2)

        # Likelihood parts
        L1      <- c1 * c2 * log(p1^(t1 - 1) * p2^(t2 - 1) * p3^(z1)
                   - p1^(t1) * p2^(t2 - 1) * p3^(z2)
                   - p1^(t1 - 1) * p2^(t2) * p3^(z3)
                   + p1^(t1) * p2^(t2) * p3^(z4))
        L2      <- c1 * (1 - c2) * log(p1^(t1 - 1) * p2^(t2) * p3^(z3)
                   + p1^(t1) * p2^(t2) * p3^(z4))
        L3      <- c2 * (1 - c1) * log(p1^(t1) * p2^(t2 - 1) * p3^(z2)
                   - p1^(t1) * p2^(t2) * p3^(z4))
        L4      <- (1 - c1) * (1 - c2) * log(p1^(t1) * p2^(t2) * p3^(z4))

        # Final log-likelihood
        logL    <- sum(L1) + sum(L2) + sum(L3) + sum(L4)
        return(logL)
}
```

```
## Bayesian Estimators
log.post        <- function(t1, t2, theta)
{
        # Parameters
        p1      <- theta[1]
        p2      <- theta[2]
        p3      <- theta[3]
        if(theta[1]<=0)return(Inf)
        if(theta[2]<=0)return(Inf)
        if(theta[3]<=0)return(Inf)

        # Max's
        z1      <- pmax(t1 - 1, t2 - 1)
        z2      <- pmax(t1, t2 - 1)
        z3      <- pmax(t1 - 1, t2)
        z4      <- pmax(t1, t2)

        # Log-likelihood parts
        L1      <- c1 * c2 * log(p1^(t1 - 1) * p2^(t2 - 1) * p3^(z1)
                   - p1^(t1) * p2^(t2 - 1) * p3^(z2)
                   - p1^(t1 - 1) * p2^(t2) * p3^(z3)
                   + p1^(t1) * p2^(t2) * p3^(z4))
        L2      <-      c1 * (1 - c2) * log(p1^(t1 - 1) * p2^(t2) * p3^(z3)
                   + p1^(t1) * p2^(t2) * p3^(z4))
        L3      <-      c2 * (1 - c1) * log(p1^(t1) * p2^(t2 - 1) * p3^(z2)
                   - p1^(t1) * p2^(t2) * p3^(z4))
        L4      <- (1 - c1) * (1 - c2) * log(p1^(t1) * p2^(t2) * p3^(z4))

        # Final log-likelihood
        log.ll  <- sum(L1) + sum(L2) + sum(L3) + sum(L4)

        # Priors
        priori  <- dunif(p1, 0, 1) * dunif(p2, 0, 1) * dunif(p3, 0, 1)
        log.p   <- log(priori)

        L       <- sum(log.ll + log.p)
        if(is.na(L)==TRUE){return(-Inf)}else{return(L)}
}
```

4. Application 7.2 (Censored data and covariates)

```
model
{
        for (i in 1:N)
        {
                zeros[i] <-     0
                phi[i]   <-  -log(L[i])
                zeros[i]~dpois(phi[i])

                # Regression model
                logit(p1[i]) <- beta10 + beta11 * sex[i]
                logit(p2[i]) <- beta20 + beta21 * sex[i]

                # Max's
                z1[i]   <- max(t1[i]-1,t2[i]-1)
                z2[i]   <- max(t1[i],t2[i]-1)
                z3[i]   <- max(t1[i]-1,t2[i])
                z4[i]   <- max(t1[i],t2[i])

                # Likelihood parts
                A[i]    <- pow(p1[i], t1[i] - 1) * pow(p2[i], t2[i] - 1) * pow(p12, z1[i])
                B[i]    <- pow(p1[i], t1[i]) * pow(p2[i], t2[i] - 1) * pow(p12, z2[i])
                C[i]    <- pow(p1[i], t1[i] - 1) * pow(p2[i], t2[i]) * pow(p12, z3[i])
                D[i]    <- pow(p1[i], t1[i]) * pow(p2[i], t2[i]) * pow(p12, z4[i])
                E[i]    <- c1[i] * c2[i] * log(A[i] - B[i] - C[i] + D[i])
                F[i]    <- c1[i] * (1 - c2[i]) * log(C[i] - D[i])
                G[i]    <- c2[i] * (1 - c1[i]) * log(B[i] - D[i])
                H[i]    <- (1 - c1[i]) * (1 - c2[i]) * log(D[i])

                # Final likelihood
                L[i]    <- exp(E[i] + F[i] + G[i] + H[i])
        }
beta10~dnorm(0,0.01)
beta11~dnorm(0,0.01)
beta20~dnorm(0,0.01)
beta21~dnorm(0,0.01)
p12~dunif(0,1)
}
```