UNIVERSIDADE DE SÃO PAULO

Generation of bivariate lifetime data assuming the Block & Basu exponential distribution

ROSELI A. LEANDRO

JORGE A. ACHCAR

Nº 19

NOTAS

Instituto de Ciências Matemáticas de São Carlos

Instituto de Ciências Matemáticas de São Carlos

ISSN - 0103-2577

Generation of bivariate lifetime data assuming the Block & Basu exponential distribution

ROSELI A. LEANDRO

JORGE A. ACHCAR

Nº 19

NOTAS DO ICMSC
Série Estatística

São Carlos
jul. / 95

Generation of Bivariate Lifetime Data assuming the BLOCK & BASU Exponential Distribution

Roseli A. Leandro *

Jorge A. Achcar †

Abstract

In this paper, we present a generation procedure for a very special bivariate lifetime distribution: the BLOCK & BASU exponential distribution. This model should have a special simulation procedure, since the marginal distributions are not exponentially distributed. We also present an illustrative example.

Keywords BLOCK & BASU exponential distribution, generation of bivariate lifetime data.

1 Introduction

In many applications of life testing, we usually have two lifetimes X and Y associated to each unit. Among the different existing bivariate lifetime models to be used in these applications, one family of models has been extensively explored in the literature: the bivariate exponential distribution (see for example, FREUND, 1961; MARSHALL & OLKIN, 1967; SARKAR, 1987; BLOCK & BASU, 1974).

*ESALQ-USP, 13418-900, Piracicaba,S.P.

†ICMSC-USP, 13560-970, São Carlos, S.P.

In this paper, we present a procedure to generate bivariate lifetime data assuming the BLOCK AND BASU (1974) exponential distribution.

The bivariate exponential distribution (BVED) of BLOCK & BASU with parameters λ_1, λ_2 and λ_3, for the lifetimes X and Y has a joint density function given by

$$f(x,y) = \begin{cases} f_1(x,y) = \frac{\lambda\lambda_1\lambda_{23}}{\lambda_{12}} exp\{-\lambda_1 x - \lambda_{23} y\} & \text{if } x < y \\ f_2(x,y) = \frac{\lambda\lambda_2\lambda_{13}}{\lambda_{12}} exp\{-\lambda_{13} x - \lambda_2 y\} & \text{if } x \geq y. \end{cases} \tag{1}$$

where $\lambda_{12} = \lambda_1 + \lambda_2$, $\lambda_{13} = \lambda_1 + \lambda_3$, $\lambda_{23} = \lambda_2 + \lambda_3$ and $\lambda = \lambda_1 + \lambda_2 + \lambda_3$.

The joint generating function for the BVED is given by

$$m(s,t) = E(e^{sx+ty}) = \frac{\lambda}{\lambda_{12}(\lambda - t - s)} \left\{ \frac{\lambda_1\lambda_{23}}{\lambda_{23} - t} + \frac{\lambda_2\lambda_{13}}{\lambda_{13} - s} \right\} \tag{2}$$

From (2), we get the moments of interest for X and Y; thus, the means and variances for X and Y are given by,

$$\begin{aligned} E(X) &= \frac{1}{\lambda_{13}} + \frac{\lambda_2\lambda_3}{\lambda\lambda_{12}\lambda_{13}} \\ E(Y) &= \frac{1}{\lambda_{23}} + \frac{\lambda_1\lambda_3}{\lambda\lambda_{12}\lambda_{23}} \\ \sigma_X^2 &= var(X) = \frac{1}{\lambda_{13}^2} + \frac{\lambda_2\lambda_3(2\lambda_1\lambda+\lambda_2\lambda_3)}{\lambda^2\lambda_{12}^2\lambda_{13}^2} \\ \sigma_Y^2 &= var(Y) = \frac{1}{\lambda_{23}^2} + \frac{\lambda_1\lambda_3(2\lambda_2\lambda+\lambda_1\lambda_3)}{\lambda^2\lambda_{12}^2\lambda_{23}^2} \end{aligned} \tag{3}$$

The correlation coefficient for X and Y is given by

$$\rho_{XY} = \frac{\lambda_3[(\lambda_1^2 + \lambda_2^2)\lambda + \lambda_1\lambda_2\lambda_3]}{\phi_1\phi_2} \tag{4}$$

where

$\phi_1 = [\lambda_{12}^2\lambda_{13}^2 + \lambda_2(\lambda_2 + 2\lambda_1)\lambda^2]^{1/2}$

and

$\phi_2 = [\lambda_{12}^2\lambda_{23}^2 + \lambda_1(\lambda_1 + 2\lambda_2)\lambda^2]^{1/2}$

Observe that $0 \leq \rho_{XY} \leq 1$ and $\rho_{XY} = 0$ only for the trivial cases $\lambda_3 = 0$ or $\lambda_1 = \lambda_2 = 0$.

Some properties of the BVED are given by,

1. $W = min(X,Y)$ has an exponential distribution with mean $1/\lambda$, $\lambda = \lambda_1 + \lambda_2 + \lambda_3$.

2. $Z = X - Y$ has distribution function,

$$F(z) = \begin{cases} \frac{\lambda_1}{\lambda_{12}} e^{\lambda_{23} z} & \text{if } z \leq 0 \\ 1 - \frac{\lambda_2}{\lambda_{12}} e^{-\lambda_{13} z} & \text{if } z > 0 \end{cases} \tag{5}$$

3. $W = min(X,Y)$ is independent of $Z = X - Y$ (and also of $|X - Y|$).

4. If (X,Y) has a BVED, then,

$$\begin{aligned} P(X > x) &= \frac{\lambda}{\lambda_{12}} e^{-\lambda_{13} x} - \frac{\lambda_3}{\lambda_{12}} e^{-\lambda x}, \quad x > 0 \\ P(Y > y) &= \frac{\lambda}{\lambda_{12}} e^{-\lambda_{23} y} - \frac{\lambda_3}{\lambda_{12}} e^{-\lambda y}, \quad y > 0 \end{aligned} \tag{6}$$

Observe that if (X,Y) has a BVED with density (1), the random variables X and Y are not exponentially dsitributed. The probability density functions for X and Y are given by

$$\begin{aligned} f_X(x) &= \frac{\lambda}{\lambda_{12}} \left\{ \lambda_{13} e^{-\lambda_{13} x} - \lambda_3 e^{-\lambda x} \right\}, \quad x > 0 \\ f_Y(y) &= \frac{\lambda}{\lambda_{12}} \left\{ \lambda_{23} e^{-\lambda_{23} y} - \lambda_3 e^{-\lambda y} \right\}, \quad y > 0 \end{aligned} \tag{7}$$

2 Generation of the BVED lifetime data

To generate a random sample of size n of bivariate lifetime observations X and Y with a BVED (1), follow the steps:

1. Generate a large number N ($N > n$) of observations W, where $W = min(X,Y)$ has an exponential distribution with density $f(w) = \lambda e^{-\lambda w}$, where

$$\lambda = \lambda_1 + \lambda_2 + \lambda_3$$

2. Associated to each value of W, consider an indicator variable δ, where $\delta = 1$ if $X < Y$ and $\delta = 0$ if $X \geq Y$. Thus, generate N observations from a Bernoulli distribution with parameter p given (from (1)) by

$$p = P(X < Y) = \int_0^\infty \int_x^\infty \frac{\lambda\lambda_1\lambda_{23}}{\lambda_{12}} e^{-\lambda_1 x - \lambda_{23} y} dy dx \tag{8}$$

That is,

$$p = \frac{\lambda_1}{\lambda_{12}} \tag{9}$$

3. From steps (i) and (ii), we have a large sample of size N of pairs (W_i, δ_i), $i = 1, 2, 3, ..., N$, where $\delta_i = 1$ or 0. If $\delta_i = 1$, that is, $X_i < Y_i$, we have $W_i = X_i$ and we should find a prediction value for Y_i to have the bivariate lifetime data (X_i, Y_i). If $\delta_i = 0$, that is, $X_i \geq Y_i$, we have $W_i = Y_i$ and we should find a prediction value for X_i to have the bivariate lifetime pair (X_i, Y_i).

4. With the obtained generated data (X_i, Y_i), $i = 1, 2, ..., N$ following steps (i), (ii) and (iii), construct an observed frequency distribution in $K_1 \times K_2$ classes of the form $(a_{1j} < X < b_{1j}, a_{2l} < Y < b_{2l})$, $j = 1, 2, ..., K_1$ and $l = 1, 2, ..., K_2$.

5. To get a random sample of size n ($n < N$), randomly choose np_{jl} observations for each frequency cell $(a_{1j} < X < b_{1j}, a_{2l} < Y < b_{2l})$, where,

$$p_{jl} = P\{a_{1j} < X < b_{1j}, a_{2l} < Y < b_{2l}\} \tag{10}$$

and this probability is obtained by using the BVED (1).

That is,

$$p_{jl} = \begin{cases} \frac{\lambda}{\lambda_{12}}(e^{-\lambda_1 b_{1j}} - e^{-\lambda_1 a_{1j}})(e^{-\lambda_{23} b_{2l}} - e^{-\lambda_{23} a_{2l}}) & \text{if } b_{1j} \leq a_{2l} \\ \frac{\lambda}{\lambda_{12}}(e^{-\lambda_2 a_{2l}} - e^{-\lambda_2 b_{2l}})(e^{-\lambda_{13} a_{1j}} - e^{-\lambda_{13} b_{1j}}) & \text{if } b_{2l} < a_{1j} \\ \frac{\lambda}{\lambda_{12}} e^{-\lambda_{23} b_{2l}}(e^{-\lambda_1 b_{1j}} - e^{-\lambda_1 a_{1j}}) + & \\ \frac{\lambda}{\lambda_{12}} e^{-\lambda_2 a_{2l}}(e^{-\lambda_{13} a_{1j}} - e^{-\lambda_{13} b_{1j}}) + & \\ \frac{\lambda_3}{\lambda_{12}}(e^{-\lambda b_{1j}} - e^{-\lambda a_{1j}}) & \text{if } a_{1j} = a_{2l} \text{ and } b_{1j} = b_{2l} \end{cases} \tag{11}$$

3 Prediction for Y

When $\delta = 1$, that is, $W = X$, a prediction value for Y (see for example, Bickel and Doksum, 1977) is given by

$$E(Y|Y > x) = \int_x^\infty \frac{y f_Y(y) dy}{P(Y > x)} \tag{12}$$

where $f_Y(y)$ is the marginal probability density function for Y (see (7)). That is,

$$E(Y|Y > x) = \frac{\lambda}{\lambda_{12} P(Y > x)} \left\{ x e^{-\lambda_{23} x} + \frac{1}{\lambda_{23}} e^{-\lambda_{23} x} - \frac{\lambda_3}{\lambda} x e^{-\lambda x} - \frac{\lambda_3}{\lambda^2} e^{-\lambda x} \right\} \tag{13}$$

where

$$P(Y > x) = \frac{\lambda}{\lambda_{12}} e^{-\lambda_{23} x} - \frac{\lambda_3}{\lambda_{12}} e^{-\lambda x}$$

4 Prediction for X

When $\delta = 0$, that is, $W = Y$, a prediction value for X is given by

$$E(X|X > y) = \int_y^\infty \frac{x f_X(x) dx}{P(X > y)} \tag{14}$$

where $f_X(x)$ is the marginal probability density function for X (see (7)).

$$E(X|X > y) = \frac{\lambda}{\lambda_{12} P(X > y)} \left\{ y e^{-\lambda_{13} y} + \frac{1}{\lambda_{13}} e^{-\lambda_{13} y} - \frac{\lambda_3}{\lambda} y e^{-\lambda y} - \frac{\lambda_3}{\lambda^2} e^{-\lambda y} \right\} \tag{15}$$

where

$$P(X > y) = \frac{\lambda}{\lambda_{12}} e^{-\lambda_{13} y} - \frac{\lambda_3}{\lambda_{12}} e^{-\lambda y}$$

5 An Example

Let us consider the generation of $n = 100$ bivariate lifetimes (X_i, Y_i), $i = 1, 2, ..., 100$ assuming the BVED (1) with $\lambda_1 = 0.08$, $\lambda_2 = 0.06$ and $\lambda_3 = 0.06$. Thus, $\lambda_{12} = 0.14$, $\lambda_{13} = 0.14$, $\lambda_{23} = 0.12$, $\lambda = 0.2$, $E(X) = 8.0612$, $E(Y) = 9.7619$, $\sigma_X = 7.7046$, $\sigma_Y = 9.0382$ and $\rho_{XY} = 0.1497$ (see (3) and (4)).

From steps (i) e (ii) of section (2), we first generate $N = 1000$ observations $W_i = min(X_i, Y_i)$ from an exponential density with parameter $\lambda = 0.20$ and also $N = 1000$ indicator observations δ_i from a Bernoulli distribution with parameter $p = \frac{\lambda_1}{\lambda_{12}} = 0.5714$ (see(9)). Following step (iii), we get the bivariate lifetime data (X_i, Y_i), using prediction values for X_i or Y_i, depending on the obtained value for δ_i, $i = 1, 2, ..., 1000$. That is, if $\delta_i = 1$, we have $X_i < Y_i$ or $W_i = X_i$ and we find a prediction value for Y_i using (13). If $\delta_i = 0$, we have $X_i \geq Y_i$, that is, $W_i = Y_i$ and we find from (15), a prediction value for X_i.

Following steps (iv) and (v), we choose $n = 100$ observations (X_i, Y_i) such that the observed frequencies of observations in constructed classes are in close agreement with the expected frequencies np_{jl} following BVED (1) (see (11)).

This generated data set is given in table 1. In table 2, we observe good agreement between the observed and the expected frequencies of the generated lifetime data (also observe figures 1 and 2).

Using Newton-Raphson iterative procedure, we find maximum likelihood estimators para λ_1, λ_2 and λ_3. Considering a random sample of size n of the BVED (1), the logarithm of the likelihood function is given by

$$l(\lambda_1, \lambda_2, \lambda_3) = n_1 ln\lambda_1 + n_2 ln\lambda_2 + n ln\lambda + n_1 ln(\lambda_2 + \lambda_3) + n_2 ln(\lambda_1 + \lambda_3) - n ln(\lambda_1 + \lambda_2) - n\overline{x}\lambda_1 - n\overline{y}\lambda_2 - R\lambda_3 \qquad (16)$$

where $n_1 = \sum_{i=1}^{n} \delta_i$, $n_2 = \sum_{i=1}^{n}(1 - \delta_i)$, $\delta_i = 1$ if $x_i < y_i$, $\delta_i = 0$ if $x_i \geq y_i$, $n\overline{x} = \sum_{i=1}^{n} x_i$, $n\overline{y} = \sum_{i=1}^{n} y_i$ and $R = \sum_{i=1}^{n}[y_i\delta_i + x_i(1 - \delta_i)]$.

Table 1: Generated Bivariate Data (n=100)

X	*Y*	*X*	*Y*	*X*	*Y*
0.01590	9.7752	0.92239	10.5434	17.4887	9.9211
0.03607	9.7921	0.95066	10.5677	1.4520	11.0002
0.06805	9.8189	1.00951	10.6182	1.6665	11.1866
0.08325	9.8316	1.03304	10.6384	1.8994	11.3900
0.11342	9.8570	1.07366	10.6733	2.2443	11.6929
0.13787	9.8775	1.10142	10.6972	2.6338	12.0372
0.14587	9.8842	1.11178	10.7061	3.1124	12.4633
0.20908	9.9374	1.14278	10.7328	3.4978	12.8087
0.23578	9.9599	1.17866	10.7637	3.8483	13.1245
0.28811	10.0040	1.20725	10.7884	4.3102	13.5430
0.32763	10.0373	1.25401	10.8288	4.7532	13.9467
0.35206	10.0580	8.08814	0.0292	5.3150	14.4614
0.37796	10.0798	8.2762	0.2331	5.6941	14.8105
0.39686	10.0958	8.5279	0.5053	6.1598	15.2411
0.40914	10.1062	8.7821	0.7797	6.7099	15.7522
0.41374	10.1101	9.0515	1.0697	7.5385	16.5264
0.42769	10.1219	9.2952	1.3316	8.0414	16.9987
0.44100	10.1332	9.5615	1.6169	8.6759	17.5969
0.46664	10.1549	9.8797	1.9573	9.4783	18.3567
0.47383	10.1610	10.1831	2.2810	10.3666	19.2020
0.50343	10.1861	10.4243	2.5378	11.1450	19.9458
0.51873	10.1991	10.8824	3.0244	12.1213	20.8825
0.53567	10.2135	11.3216	3.4897	18.6037	11.0695
0.56358	10.2372	11.7383	3.9299	19.3300	11.8160
0.61135	10.2778	12.2368	4.4552	20.1483	12.6555
0.64623	10.3075	12.7206	4.9637	22.5853	15.1477
0.68866	10.3436	13.2373	5.5054	13.3829	22.0983
0.70947	10.3614	14.0623	6.3677	14.4032	23.0855
0.74952	10.3955	15.0128	7.3577	15.1537	23.8137
0.81160	10.4486	15.3570	7.7153	23.0677	31.5669
0.85319	10.4842	15.4718	7.8345	31.8155	24.5140
0.87853	10.5059	15.8906	8.2688	25.5306	33.9989
0.91137	10.5340	16.5681	8.9702	32.1644	40.5758
		17.1588	9.5806		

Table 2: Frequency distribution for n=100 pairs (X_i, Y_i) in 6 × 5 classes (o=observed frequency; e=expected frequency)

X \ Y	(0, 11]	(11, 22]	(22, 33]	(33, 44]	(44, 55]
(0.00, 7.500]	e11=42.35 o11=44.00	e12=12.62 o12=14.00	e13=3.37 o13=0.00	e14=0.90 o14=0.00	e15=0:24 015=0.00
(7.50, 15.00]	e21=15.78 o21=17.00	e22= 6.76 o22= 7.00	e23=1.84 o23=2.00	e24=0.49 o24=0.00	e25=0.13 o25=0.00
(15.00, 22.50]	e31= 5.49 o31= 7.00	e32= 2.83 o32= 3.00	e33=1.01 o33=1.00	e34=0.27 o34=0.00	e35=0.07 o35=0.00
(22.50, 30.00]	e41= 1.92 o41= 0.00	e42=0.99 o42=1.00	e43=0.49 o43=1.00	e44=0.14 o44=1.00	e45=0.04 o45=0.00
(30.00, 37.50]	e51= 0.67 o51= 0.00	e52=0.35 o52=0.00	e53=0.18 o53=1.00	e54=0.08 o54=1.00	e55=0.02 o55=1.00
(37.50, 45.00]	e61=0.23 o61=0.00	e62=0.12 o62=0.00	e63=0.06 o63=0.00	e64=0.03 o64=0.00	e65=0.01 o65=0.00

Figure 1: Observed frequencies for the $n = 100$ generated data of table 1

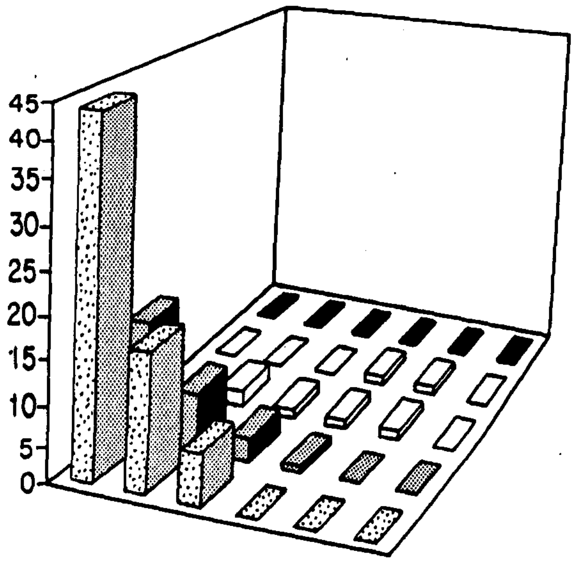
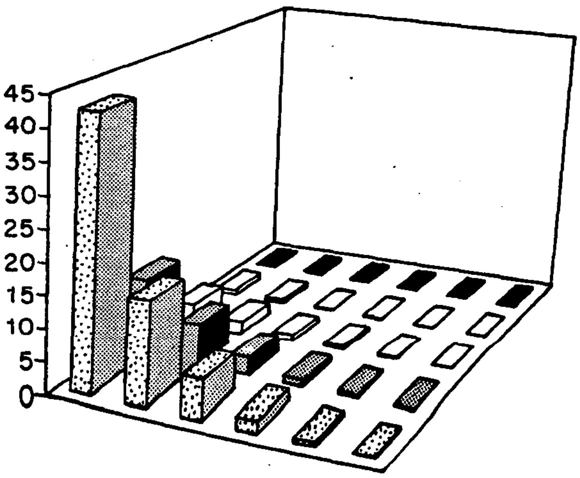


Figure 2: Expected frequencies for $n = 100$ observations using BVED (1) with parameters $\lambda_1 = 0.08$, $\lambda_2 = 0.06$ e $\lambda_3 = 0.06$.



With the generated data of table 2, and using the initial values $\lambda_1^0 = 0.04$, $\lambda_2^0 = 0.03$, $\lambda_3^0 = 0.03$ in the iterative procedure, we find after 9 iterations, the maximum likelihood estimates $\hat{\lambda}_1 = 0.10782$, $\hat{\lambda}_2 = 0.04896$, $\hat{\lambda}_3 = 0.05396$. We also find maximum likelihood estimators for functions of $\lambda_1, \lambda_2, \lambda_3$. Thus, $\hat{E}(X) = 6.6756$, $\hat{E}(Y) = 11.4275$, $\hat{\sigma}_X = 6.5311$, $\hat{\sigma}_Y = 10.3788$, and $\hat{\rho}_{XY} = 0.1241$ (see (3) and (4)).

Remark: For the generation of the bivariate lifetime data of table 1 following BVED (1), we developed a computer program using the S-PLUS software and the SAS FOR WINDOWS software (copies available under request).

6 References

BICKEL, P. J.; DOKSUM, K. A. *Mathematical Statistics - Basic Ideas and selected topics*, San Francisco: Holden Day, 1977.

BLOCK, H. W.; BASU, A. P. A. Continuous bivariate exponential, *Journal of the American Statistical Association*, Washington, **69, 348**, p. 1031-1037, 1974.

NOTAS DO ICMSC

SÉRIE ESTATÍSTICA

018/95 ACHCAR, J.A.; - Use of approximate bayesian inference for software reliability

017/95 ACHCAR, J.A.; FAVORETTI, A.C. - Accurate inferences for the Michaelis-Menten model

016/95 RODRIGUES, J.; CHILTA, S.P. - Bayesian Analysis in M/M/1 queues using Sampling Methods

015/95 ACHCAR, J.A.; PEGORINM M.A. - Laplace's approximations for posterior expectations when the mode is not in the parameter space.

014/94 ACHCAR, J. A.; FOGO, J.C. - Accurate inferences for the reliability function considering accelerated life tests.

013/94 RODRIGUES, J. - Bayesian Solutions to a lass of selections problems using weighted loss functions.

012/94 ACHCAR, J.A.; DAMASCENO, V.L. - Extreme value models: an useful reparametrization for the survival function

011/94 ACHCAR, J.A.- Approximate bayesian analysis for non-normal hierarchical classification model.

010/94 RODRIGUES, J.; RODRIGUES, E.F. - Bayesian estimationin the study of tampered Random variables.

009/94 ACHCAR, J.A.; FOGO, J.C. - An useful reparametrization for the reliability in the Weibull case.

008/94 RODRIGUES, J.; LEITE, J.G. - A note on Bayesian analysis in M/M/1 queues from confidence intervals.

007/94 ACHCAR, J.A.; DAMACENO, V.L. - An useful reparametrization for the survival function considering an exponential regression model.

006/94 RODRIGUES, J. - Bayesian estimation of a normal mean parameter using linex loss function and robustness considerations.