

The Library of the Department of Statistics  
North Carolina State University

**BIVARIATE EXPONENTIALS, DUAL COUNTING  
PROCESSES AND THEIR INTENSITIES**

by

A. C. Pedroso de Lima and Pranab K. Sen

Department of Biostatistics  
University of North Carolina

Institute of Statistics  
Mimeo Series No. 2166

August 1996

# BIVARIATE EXPONENTIALS, DUAL COUNTING PROCESSES AND THEIR INTENSITIES

A. C. Pedroso de Lima and Pranab K. Sen

Universidade de São Paulo, Brazil and University of North Carolina at Chapel Hill

July 29, 1996

Some popular absolutely continuous bivariate exponential densities are considered in the light of intensities of their dual (matrix-valued) counting processes that crop up in multi-event survival analysis. Statistical properties of such parameterized intensities provide meaningful hind-sights in many applications.

**1. Introduction.** Univariate counting processes have their genesis in the dual interarrival time distributions, and the simple negative exponential density plays basic role in this context. In survival analysis, parametric models often relate to exponential, Weibull, Pareto and Gamma family of densities, while semiparametric ones allow more flexibility for such underlying densities by allowing the dual counting processes to have more general intensity or hazard functions. Nevertheless, if a specific parametric model holds but we take recourse to a semiparametric one, its intensity function can be expressed in a parametric form in terms of the hazard function that is a functional of the underlying parametric survival function.

The situation is a bit more complex in multivariate problems. For example, even in a simpler bivariate model, stochastic interdependence (or association) of the two endpoints has to be incorporated in the formulation of their joint density or hazard functions. In this setup, hazard as well as conditional hazard functions interplay in the intensity processes associated with each component of the vector of dual counting processes, so that the interaction between the two components needs to be incorporated in this formulation as well. In a parametric mold, there is an hierarchy of bivariate exponential distributions; while some of these are not absolutely continuous, the others that are so remain amenable to multiple endpoints survival

---

AMS (1991) Subject classifications: 62M99

*Key words and phrases:* Association; Block-Basu model; conditional hazard; Gumbel model; hazard function; intensities; interactions; matrix-valued counting process; Sarkar model; semiparametrics; survival analysis for multiple endpoints.

analysis. In semiparametrics, the usual multivariate counting processes, arising in univariate models, have been extended to more general matrix-valued counting processes whose intensities can accommodate such stochastic interactions (or associations) in a meaningful way; we refer to [7] where other pertinent references have also been cited. In order to interpret statistically such semiparametric formulations, we need to establish their compatibility for parametric models as well. The treatise in [7] touched on this aspect only through a specific example, and we find it of genuine interest to provide, in the current study, an unified compatibility picture for general families of absolutely continuous bivariate exponential distributions.

Bivariate matrix-valued counting processes are introduced in Section 2. Section 3 deals with the three notable bivariate exponential distributions proposed by Gumbel [5], Sarkar [8] and Block and Basu [2], while the intensities of their dual counting processes are studied in Section 4. In this context the role of the corresponding association parameters is appraised properly. The concluding section is devoted to statistical interpretations of these three parametric forms of intensities, and stresses their effective roles in multiple endpoints survival analysis.

**2. Bivariate matrix-valued counting process.** Let  $(T_1, T_2)$  be a nonnegative random vector defined in a probability space  $(\Omega, \mathcal{F}, \Pr)$ , with joint survival function given by  $S_{12}(t_1, t_2)$ . The marginal survival functions are represented by  $S_1(t_1)$  and  $S_2(t_2)$ .

We define the counting processes

$$(1) \quad N_k(t) = \mathbf{I}\{T_k \leq t\}, \quad t \geq 0, \quad k = 1, 2,$$

representing a right-continuous function that assumes value zero, jumping to one when the particular event associated to  $T_k$  occurs. Since the quantities in (1) are defined on dependent random variables, it makes sense to consider also the random vector

$$(2) \quad \mathbf{N}(t) = (N_1(t), N_2(t))'.$$

The self-exciting filtration  $\{\mathcal{N}_t, t \geq 0\} = \sigma\{\mathbf{N}(s), 0 \leq s \leq t\}$  is considered in this model in the characterization of the counting processes above. Hence, we define the  $\mathcal{N}_t$ -predictable processes

$$(3) \quad Y_k(t) = \mathbf{I}\{T_k \geq t\}, \quad t \geq 0, \quad k = 1, 2.$$

Such a process is assumed to have its value at instant  $t$  known just before  $t$ . Then, the intensity process associated with  $N(t)$  and represented by the vector  $\lambda$ , can be written as

$$(4) \quad \lambda(t) = \begin{pmatrix} \lambda_1(t) \\ \lambda_2(t) \end{pmatrix},$$

and fully specifies the counting process defined in (2). We write

$$(5) \quad \lambda_k(t) = E\{ dN_k(t) \mid \mathcal{N}_{t-} \},$$

i.e., the average of jumps for component  $k$  given the information available just before  $t$ . We may note that in this case  $\mathcal{N}_{t-}$  contains information whether or not one (or both) component(s) have failed just before  $t$ . Since the processes  $Y_k$  are predictable, this means we know the value of  $Y_k$  at the instant  $t$ . If the component  $k$  have failed before  $t$ , then expression (5) equals zero. In other words, we need to consider the situations (i) no component has failed before time  $t$ , i.e.,  $Y_1(t) = Y_2(t) = 1$ ; (ii) the first component has failed before  $t$  but the second has not, i.e.,  $Y_1(t) = 0$  and  $Y_2(t) = 1$ ; (iii) only second component has failed before  $t$ , that is,  $Y_1(t) = 1$  and  $Y_2(t) = 0$ ; and (iv) both components failed before  $t$ , in which case  $Y_1(t) = Y_2(t) = 0$ . If we want to consider the intensity process for the first component, then we only consider cases where  $Y_1(t) = 1$ . This together with expression (5) allow us to write

$$(6) \quad \lambda_1(t) = E\{ dN_1(t) \mid \mathcal{N}_{t-} \} = p_1^{(1)}(t)Y_1(t)[1 - Y_2(t)] + p_2^{(1)}(t)Y_1(t)Y_2(t)$$

where  $p_1^{(1)}(t) = \lim_{\Delta t \rightarrow 0} (\Delta t)^{-1} \Pr\{T_1 \in [t, t + \Delta t) \mid T_1 \geq t; T_2 < t\}$  and  $p_2^{(1)}(t) = \lim_{\Delta t \rightarrow 0} (\Delta t)^{-1} \Pr\{T_1 \in [t, t + \Delta t) \mid T_1 \geq t; T_2 \geq t\}$  may be interpreted as conditional hazard functions, given the outcome for the other component. Similarly, for component 2,

$$(7) \quad \lambda_2(t) = E\{ dN_2(t) \mid \mathcal{N}_{t-} \} = p_2^{(2)}(t)Y_2(t)[1 - Y_1(t)] + p_1^{(2)}(t)Y_1(t)Y_2(t)$$

for  $p_2^{(2)}(t) = \lim_{\Delta t \rightarrow 0} (\Delta t)^{-1} \Pr\{T_2 \in [t, t + \Delta t) \mid T_1 < t; T_2 \geq t\}$  and  $p_1^{(2)}(t) = \lim_{\Delta t \rightarrow 0} (\Delta t)^{-1} \Pr\{T_2 \in [t, t + \Delta t) \mid T_1 \geq t; T_2 \geq t\}$ .

Based on (6) and (7) we can represent the intensity process by the product of matrices

$$(8) \quad \lambda(t) = \begin{pmatrix} \lambda_1(t) \\ \lambda_2(t) \end{pmatrix} = \begin{pmatrix} Y_1(t) & 0 \\ 0 & Y_2(t) \end{pmatrix} \begin{pmatrix} \alpha_{11}(t) & \alpha_{12}(t) \\ \alpha_{21}(t) & \alpha_{22}(t) \end{pmatrix} \begin{pmatrix} Y_1(t) \\ Y_2(t) \end{pmatrix} \\ = \text{Diag}(\mathbf{Y}(t))\boldsymbol{\alpha}(t)\mathbf{Y}(t)$$

where the elements of  $\boldsymbol{\alpha}$  are given by  $\alpha_{11}(t) = p_1^{(1)}(t)$  and  $\alpha_{12}(t) = p_2^{(1)}(t) - p_1^{(1)}(t)$  for the first component, and  $\alpha_{21}(t) = p_1^{(2)}(t) - p_2^{(2)}(t)$  and  $\alpha_{22}(t) = p_2^{(2)}(t)$  for the second.

The *matrix-valued counting process model* is expressed in terms of the vectors  $\mathbf{N}_1, \dots, \mathbf{N}_n$ , that is,  $n$  copies of the process  $\mathbf{N}$  defined on (2) and associated with a sample of  $n$  individuals. Then the matrix-valued counting process is given by

$$(9) \quad \mathbf{N}(t) = (\mathbf{N}_1(t), \dots, \mathbf{N}_n(t)),$$

with an associated intensity process given by (8). Note that the columns of  $\mathbf{N}$  are independent and in each column, the two may be stochastically dependent. We refer to [7] for further details, including the model with covariates.

**3. Bivariate exponential distributions.** The model developed in Section 2 is defined in terms of conditional and marginal hazard functions. Therefore, in order to assess the behavior of such a model in a more applied setup, we consider the densities of bivariate exponential distributions and compute the corresponding intensities.

It is well known that for dimensions greater than one, different approaches can lead us to different multivariate exponential distributions. Therefore, we consider here three absolutely continuous bivariate exponential distributions.

*3.1. Gumbel bivariate exponential distribution (GBED).* The GBED is a particular case of the Gumbel representation of multivariate distributions studied in details in [4], where a bivariate distribution function can be expressed as

$$F(x, y) = F_X(x)F_Y(y)[1 + \alpha\{1 - F_X(x)\}\{1 - F_Y(y)\}],$$

where  $F_X$  and  $F_Y$  are two given probability distribution functions and  $-1 \leq \alpha \leq 1$  is a dependence parameter, in the sense that if  $\alpha = 0$ ,  $X$  and  $Y$  are independent.

It follows immediately that whenever the densities  $f_X(\cdot)$  and  $f_Y(\cdot)$  exist, the joint bivariate density function is given by

$$(10) \quad f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = f_X(x)f_Y(y)[1 + \alpha(1 - 2F_X(x))(1 - 2F_Y(y))]$$

Given the expression above, the Gumbel bivariate exponential model is obtained setting  $f(x) = \theta_1 e^{-\theta_1 x}$  and  $f(y) = \theta_2 e^{-\theta_2 y}$ , such that the bivariate density is given by

$$(11) \quad f(x, y) = \theta_1 \theta_2 e^{-(\theta_1 x + \theta_2 y)} \{1 + \alpha(1 - 2e^{-\theta_1 x})(1 - 2e^{-\theta_2 y})\}, \quad x, y \geq 0,$$

with  $\theta_1 > 0$ ,  $\theta_2 > 0$ , and, as before,  $\alpha \in [-1.1]$ . For this model it can be shown that the correlation coefficient  $\rho_{GBED}$  is (see, e.g., [5])

$$\rho_{GBED} = \alpha/4,$$

and hence, for the GBED the correlation between  $X$  and  $Y$  varies in the interval  $[-0.25, 0.25]$ .

**3.2. Sarkar bivariate exponential distribution (SBED).** Another absolutely continuous bivariate exponential distribution has been defined by Sarkar in [8]. It is characterized by the joint survival function

$$(12) \quad \Pr\{T_1 \geq x; T_2 \geq y\} = \begin{cases} e^{-(\beta_2 + \beta_{12})y} \{1 - [A(\beta_1 y)]^{-\gamma} [A(\beta_1 x)]^{1+\gamma}\}, & 0 < x \leq y \\ e^{-(\beta_1 + \beta_{12})x} \{1 - [A(\beta_2 x)]^{-\gamma} [A(\beta_2 y)]^{1+\gamma}\}, & 0 < y \leq x, \end{cases}$$

where  $\beta_1 > 0$ ,  $\beta_2 > 0$ ,  $\beta_{12} \geq 0$ ,  $\gamma = \beta_{12}/(\beta_1 + \beta_2)$  and  $A(z) = 1 - e^{-z}$ ,  $z > 0$ . This model is based on modifications on the characterization property for bivariate exponential distributions, stating that : (i)  $X$  and  $Y$  are marginally exponential, (ii)  $\min(X, Y)$  is exponential and (iii)  $\min(X, Y)$  and  $X - Y$  are independent. Here whenever  $\beta_{12} = 0$ , the joint distribution can be expressed as the product of two exponential distributions and hence,  $X$  and  $Y$  are independent.

The corresponding density function can be expressed as

$$(13) \quad f(x, y) = \begin{cases} \beta_1(1 + \gamma)e^{-\beta_1 x - (\beta_2 + \beta_{12})y} \{\beta_1 \gamma + \beta_2(1 + \gamma)A(\beta_1 y)\} \\ \quad \times [A(\beta_1 x)]^{-\gamma} [A(\beta_1 y)]^{-(1+\gamma)}, & 0 < x \leq y, \\ \beta_2(1 + \gamma)e^{-\beta_2 y - (\beta_1 + \beta_{12})x} \{\beta_2 \gamma + \beta_1(1 + \gamma)A(\beta_2 x)\} \\ \quad \times [A(\beta_2 y)]^{-\gamma} [A(\beta_2 x)]^{-(1+\gamma)}, & 0 < y \leq x, \end{cases}$$

where  $\beta = \beta_1 + \beta_2 + \beta_{12}$ . For the SBED, it can be shown that the coefficient of correlation is given by

$$(14) \quad \rho_{SBED} = \frac{\beta_{12}}{\beta} + \frac{\beta_{12}}{\beta} \sum_{k=1}^{\infty} \frac{j!}{\beta + (\beta_1 + \beta_2)k} \left\{ \beta_1^k \prod_{l=1}^k \frac{1}{\beta + l\beta_1} + \beta_2^k \prod_{l=1}^k \frac{1}{\beta + l\beta_2} \right\},$$

and hence, it follows that the SBED allows only for positive association between  $X$  and  $Y$ , since by (14) we immediately conclude that  $0 \leq \rho_{SBED} \leq 1$ . Also note that when  $\beta_{12} = 0$ , not only the correlation coefficient equals zero, but also  $X$  and  $Y$  are stochastically independent.

**3.3. Block-Basu exponential distribution (BBBED).** The BBBED represents the absolutely continuous part of the bivariate exponential distribution of Marshal and Olkin [6] and is considered in [2]. For  $(\beta_1, \beta_2, \beta_{12}) \in (0, \infty) \times (0, \infty) \times [0, \infty)$ , its distribution function is given by

$$F(x, y) = \frac{\beta}{\beta_1 + \beta_2} \exp[-\beta_1 x - \beta_2 y - \beta_{12} \max(x, y)] - \frac{\beta_{12}}{(\beta_1 + \beta_2)} \exp[-\beta \max(x, y)] I_{(0, \infty)}(x) I_{(0, \infty)}(y),$$

where  $\beta = \beta_1 + \beta_2 + \beta_{12}$ . The corresponding p.d.f. is expressed as

$$(15) \quad f(x, y) = \begin{cases} \frac{\beta_1 \beta (\beta_2 + \beta_{12})}{\beta_1 + \beta_2} e^{-\beta_1 x - (\beta_2 + \beta_{12}) y}, & \text{if } x < y; \\ \frac{\beta_2 \beta (\beta_1 + \beta_{12})}{\beta_1 + \beta_2} e^{-(\beta_1 + \beta_{12}) x - \beta_2 y}, & \text{if } x > y. \end{cases}$$

Such a distribution possess the loss of memory property (see [2]), and, in addition, (i)  $\min(X, Y)$  is exponentially distributed; (ii) the distributions of  $X - Y$  and  $|X - Y|$  are well determined; and, (iii)  $\min(X, Y)$  and  $X - Y$  (as well as  $|X - Y|$ ) are independent.

For the BBBED, the coefficient of correlation can be computed as

$$(16) \quad \rho_{BBBED} = \frac{\beta_{12} [(\beta_1^2 + \beta_2^2) \beta + \beta_1 \beta_2 \beta_{12}]}{[(\beta_1 + \beta_2)^2 (\beta_1 + \beta_{12})^2 + \beta_2 (\beta_2 + 2\beta_1) \beta^2] [(\beta_1 + \beta_2)^2 (\beta_2 + \beta_{12})^2 + (\beta_1 + 2\beta_2) \beta^2]},$$

and hence, for this distribution we also have only positive association between  $X$  and  $Y$ ; moreover, [2] show that  $\rho_{BBBED} \in [0, 1]$ . Finally, for  $\beta_{12} = 0$ , not only  $\rho_{BBBED} = 0$  but also  $X$  and  $Y$  are independent.

**4. Parametric intensity processes.** In this section we obtain analytic expression for the three bivariate models defined in the previous section.

#### 4.1. Intensity for the GBED.

Considering the general expression for the GBED presented in (10) we can compute the conditional hazard functions present in expression (6). For component 1 this implies in computing  $p_1^{(1)}(t)$  and  $p_2^{(1)}(t)$ . Hence, for  $\bar{F}(t) = 1 - F(t)$  representing the survival function at time  $t$ ,

$$\begin{aligned} p_1^{(1)}(t) &= \lim_{\Delta \rightarrow 0} \Delta^{-1} \frac{\Pr\{X \in [t, t + \Delta]; Y < t\}}{\Pr\{X \geq t; Y < t\}} = \lim_{\Delta \rightarrow 0} \Delta^{-1} \frac{\int_t^{t+\Delta} \int_0^t f(x, y) dy dx}{\int_t^\infty \int_0^t f(x, y) dy dx} \\ &= \frac{f_X(t) \{1 + \alpha(\bar{F}_X(t) - F_X(t)) \bar{F}_Y(t)\}}{\bar{F}_X(t) \{1 - \alpha F_X(t) \bar{F}_Y(t)\}} = h_X(t) \frac{1 + \alpha(\bar{F}_X(t) - F_X(t)) \bar{F}_Y(t)}{1 - \alpha F_X(t) \bar{F}_Y(t)} \\ &= h_X(t) \left[ 1 + \alpha \frac{\bar{F}_X(t) \bar{F}_Y(t)}{1 - \alpha F_X(t) \bar{F}_Y(t)} \right], \end{aligned}$$

where  $h_X(t)$  is the marginal hazard function for  $X$ . As for  $p_2^{(1)}(t)$ , we have

$$\begin{aligned} p_2^{(1)}(t) &= \lim_{\Delta \rightarrow 0} \Delta^{-1} \frac{\Pr\{X \in [t, t + \Delta]; Y \geq t\}}{\Pr\{X \geq t; Y \geq t\}} = \lim_{\Delta \rightarrow 0} \Delta^{-1} \frac{\int_t^{t+\Delta} \int_t^\infty f(x, y) dy dx}{\int_t^\infty \int_t^\infty f(x, y) dy dx} \\ &= \frac{f_X(t) \{1 - \alpha(\bar{F}_X(t) - F_X(t)) F_Y(t)\}}{\bar{F}_X(t) \{1 + \alpha F_X(t) F_Y(t)\}} = h_X(t) \frac{1 - \alpha(\bar{F}_X(t) - F_X(t)) F_Y(t)}{1 + \alpha F_X(t) F_Y(t)} \\ &= h_X(t) \left[ 1 - \alpha \frac{\bar{F}_X(t) F_Y(t)}{1 + \alpha F_X(t) F_Y(t)} \right], \end{aligned}$$

Hence, the intensity process for the first component for this general setting is given by [cf. (6)].

$$\lambda_1(t) = h_X(t)Y_1(t) \left\{ 1 + \alpha \frac{\bar{F}_X(t)}{1 - \alpha \bar{F}_X(t)\bar{F}_Y(t)} \left( \bar{F}_Y(t) - \frac{Y_2(t)}{1 + \alpha \bar{F}_X(t)\bar{F}_Y(t)} \right) \right\},$$

and note that whenever  $X$  and  $Y$  are independent,  $\alpha = 0$  and  $\lambda_1(t) = h_X(t)Y_1(t)$ , so that the marginal intensity process equals the marginal hazard function for  $X$ , times the predictable process  $Y_1(t)$ , as expected.

When we consider that  $X$  and  $Y$  have exponential distribution with parameters  $\theta_1$  and  $\theta_2$ , we get

$$(17) \quad \lambda_1(t) = \theta_1 Y_1(t) \left\{ 1 + \alpha \frac{e^{-\theta_1 t}}{1 - \alpha A(\theta_1 t)e^{-\theta_2 t}} \left( e^{-\theta_2 t} - \frac{Y_2(t)}{1 + \alpha A(\theta_1 t)A(\theta_2 t)} \right) \right\},$$

that reduces to  $\theta_1$  when  $X$  and  $Y$  are independent. Similar expressions can be obtained for the other component.

In order to visualize the behavior of the intensity in (17), we set  $\theta_2 = k\theta_1$  and  $\theta_1 = 1$ , in which case the intensity process reduces to

$$\lambda_1(t) = Y_1(t) \left\{ 1 + \alpha \frac{e^{-t}}{1 - \alpha A(t)e^{-kt}} \left( e^{-kt} - \frac{Y_2(t)}{1 + \alpha A(t)A(kt)} \right) \right\}.$$

Figures 1 and 2 contain several curves for this intensity, for  $k = 0.2$  and  $k = 2$  respectively, along with some values of  $\alpha$  and  $t_F$ , the instant of failure for the second component.

*4.2. Intensity for the SBED.* Using the expression shown in (13), we can also obtain an expression for the intensity process  $\lambda(t)$ , similarly. However, in this cases the computations are somewhat more involving and hence, we present the derivations with some details. Here also we consider the intensity for the first component only, as the expression for the other element can be similarly obtained by convenient changes in the indices.

The numerator of the conditional hazard function  $p_1^{(1)}(t)$  is given by

$$\lim_{\Delta t \rightarrow 0} (\Delta t)^{-1} \Pr\{T_1 \in [t, t + \Delta t); T_2 < t\} = \int_0^t f(t, v) dv = e^{-(\beta_1 + \beta_{12})t} \{\beta_2 \gamma + \beta_1(1 + \gamma)A(\beta_2 t)\}.$$

As for the denominator a more elaborated manipulation is needed. In this case,

$$\begin{aligned} \Pr\{T_1 \geq t; T_2 < t\} &= \int_t^\infty \int_0^t f(u, v) dv \\ &= (1 + \gamma) \int_t^\infty e^{-(\beta_1 + \beta_{12})u} \{\beta_2 \gamma + \beta_1(1 + \gamma)A(\beta_2 u)\} [A(\beta_2 u)]^{-(1+\gamma)} du \\ &\quad \times \int_0^t \beta_2 e^{-\beta_2 v} [A(\beta_2 v)]^\gamma dv. \end{aligned}$$

The second integral in the expression above is equal to  $[A(\beta_2 t)]^{\gamma+1}/(\gamma + 1)$ . The first integral can be expressed as

$$(18) \quad \beta_2 \gamma \int_t^\infty e^{-(\beta_1 + \beta_{12})u} [A(\beta_2 u)]^{-(1+\gamma)} du + \beta_1(1 + \gamma) \int_t^\infty e^{-(\beta_1 + \beta_{12})u} [A(\beta_2 u)]^{-\gamma} du.$$

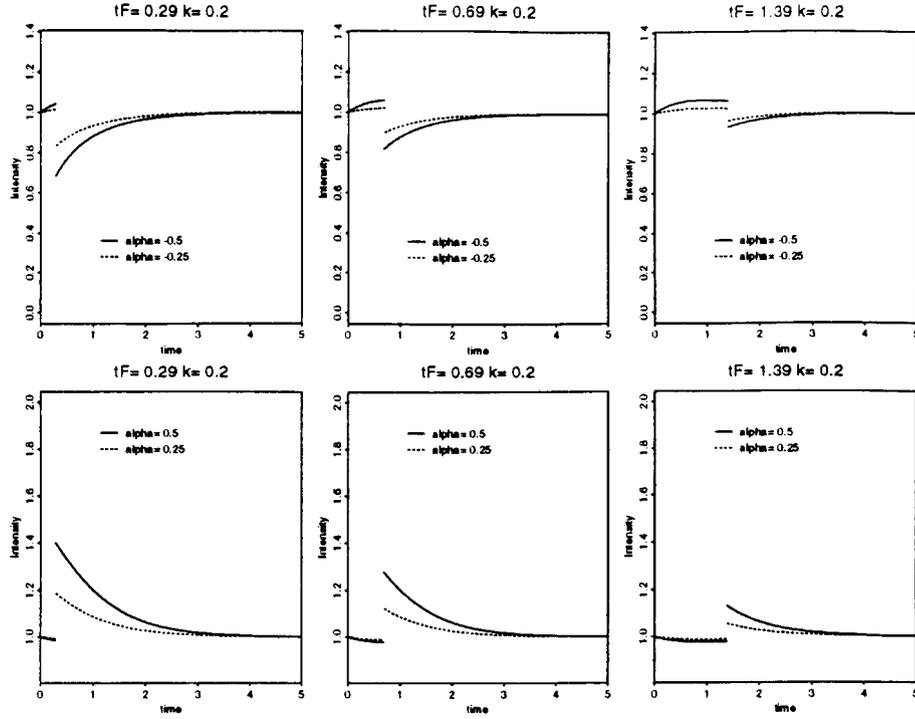


Figure 1: Gumbel bivariate exponential intensity function ( $k = 0.2$ )

Using the relation  $(1 - t)^a = \sum_{k \geq 0} \binom{a}{k} (-t)^k$ , for any number  $a$  and  $t \in (-1, 1)$  [see, e.g., [3], page 51], then by the dominated convergence theorem we can write for the first factor in (18)

$$\begin{aligned}
 & \int_t^\infty e^{-(\beta_1 + \beta_{12})u} \sum_{k \geq 0} (-1)^k \binom{-(1+\gamma)}{k} (e^{-\beta_2 u})^k du \\
 &= \sum_{k \geq 0} (-1)^k \binom{-(1+\gamma)}{k} \int_t^\infty e^{-(\beta_1 - k\beta_2 + \beta_{12})u} du \\
 &= e^{-(\beta_1 + \beta_{12})t} \sum_{k \geq 0} (-1)^k \frac{\binom{-(1+\gamma)}{k} e^{k\beta_2 t}}{\beta_1 + k\beta_2 + \beta_{12}}.
 \end{aligned}$$

Similarly, the second integral in (18) is

$$\int_t^\infty e^{-(\beta_1 + \beta_{12})u} [A(\beta_2 u)]^{-\gamma} = e^{-(\beta_1 + \beta_{12})t} \sum_{k \geq 0} (-1)^k \frac{\binom{-\gamma}{k} e^{k\beta_2 t}}{\beta_1 + k\beta_2 + \beta_{12}}.$$

Combining both terms and using that for  $a > 0$ ,  $\binom{-a}{k} = (-1)^k \binom{a+k-1}{k}$  [[3], page 63], expression (18) is equal to

$$\begin{aligned}
 & e^{-(\beta_1 + \beta_{12})t} \sum_{k \geq 0} (-1)^k \frac{e^{-k\beta_2 t}}{\beta_1 + k\beta_2 + \beta_{12}} \left\{ \beta_2 \gamma \binom{-(1+\gamma)}{k} + \beta_1 (1+\gamma) \binom{-\gamma}{k} \right\} \\
 &= e^{-(\beta_1 + \beta_{12})t} \sum_{k \geq 0} (-1)^k \binom{-\gamma}{k} (e^{-\beta_2 t})^k = e^{-(\beta_1 + \beta_{12})t} (1 - e^{-\beta_2 t})^{-\gamma}.
 \end{aligned}$$

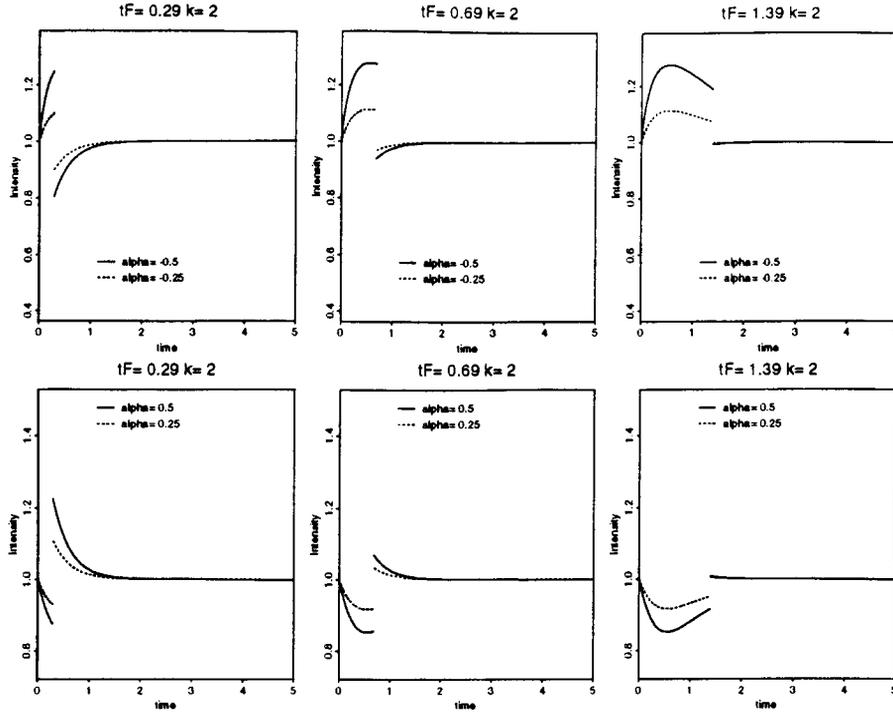


Figure 2: Gumbel bivariate exponential intensity function ( $k = 2$ )

Hence,

$$\Pr\{T_1 \geq t; T_2 < t\} = e^{-(\beta_1 + \beta_{12})t} A(\beta_2 t),$$

so that

$$p_1^{(1)}(t) = \beta_1(1 + \gamma) + \gamma \frac{\beta_2}{A(\beta_2 t)}.$$

The same type of manipulations applies to determining  $p_2^{(1)}(t)$ . In this case the numerator will be given by

$$\lim_{\Delta t \rightarrow 0} (\Delta t)^{-1} \Pr\{T_1 \in [t, t + \Delta t); T_2 \geq t\} = \int_t^\infty f(t, v) dv = \beta_1(1 + \gamma)e^{-\beta t},$$

and the denominator is given by the survival function (12) in the case  $u \leq v$ , so that

$$p_2^{(1)}(t) = \frac{\beta_1(1 + \gamma)e^{-\beta t}}{e^{-\beta t}} = \beta_1(1 + \gamma).$$

Thus,

$$\begin{aligned} \lambda_1(t) &= p_1^{(1)}(t)Y_1(t) + [p_2^{(1)}(t) - p_1^{(1)}(t)]Y_1(t)Y_2(t) \\ &= \left\{ \beta_1(1 + \gamma) + \gamma \frac{\beta_2}{A(\beta_2 t)} \right\} Y_1(t) - \gamma \frac{\beta_2}{A(\beta_2 t)} Y_1(t)Y_2(t). \end{aligned}$$

Considering now  $\beta_2 = k\beta_1$ ,  $\beta_1 = 1$ , we may study the behavior of  $\lambda_1(t)$  given above. In this case the expression turns out to be

$$\lambda_1(t) = \left(1 + \frac{\beta_{12}}{1+k} + k \frac{\beta_{12}}{(1+k)(1-e^{-kt})}\right) Y_1(t)[1 - Y_2(t)] + \left(1 + \frac{\beta_{12}}{1+k}\right) Y_1(t)Y_2(t).$$

Figure 3 contains graphics of  $\lambda_1(t)$  for several values of  $\beta_{12}$ ,  $k$ , and  $t_F$ , the failure time for the second component.

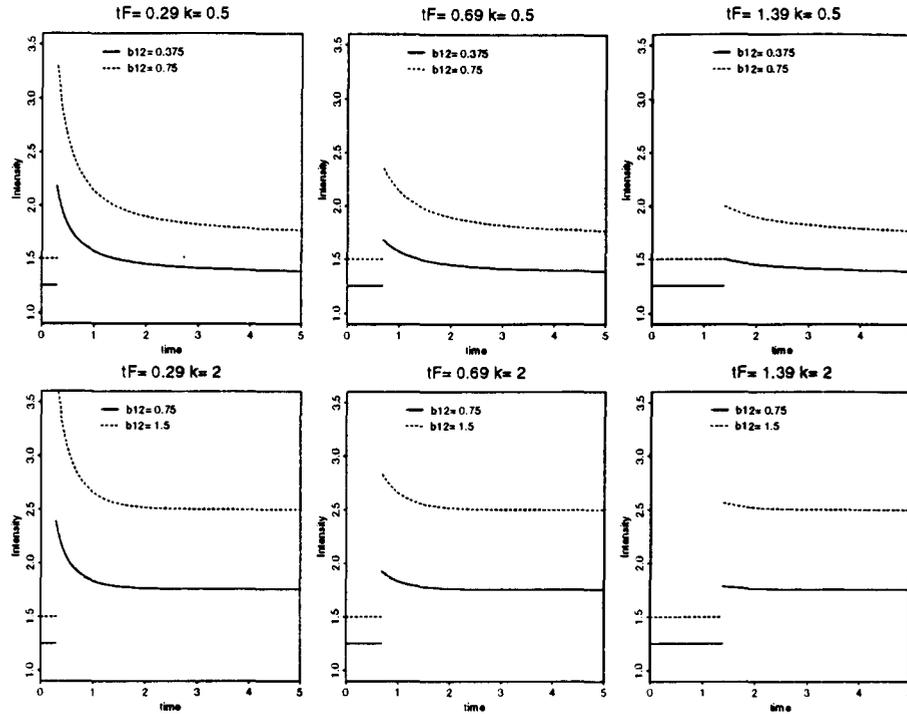


Figure 3: Sarkar bivariate exponential intensity function

4.3. *Intensity for the BBED.* The characterization for the matrix-valued model when the underlying model is the BBED can be performed similarly to the earlier models, and here the computations are simpler. Hence, in this case it is easy to show that, for component 1,

$$p_1^{(1)}(t) = \alpha_1 + \alpha_{12}$$

and

$$p_2^{(1)}(t) = \alpha_1 \frac{\alpha}{\alpha_1 + \alpha_2}.$$

Hence, in this case the curves have always a similar pattern, as shown in Figure 4, where  $\alpha_1 = 1$ ,  $\alpha_2 = k = 0.2$ ,  $t_F = 1.39$ , and  $\alpha_{12}$  assuming values 0.5 and 0.25.

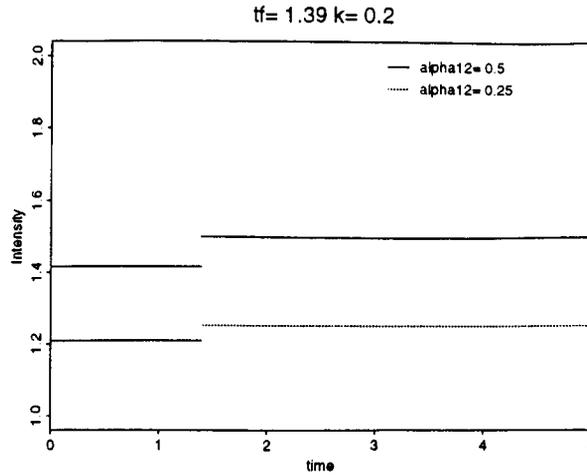


Figure 4: Block-Basu bivariate exponential intensity function

**5. Discussion.** The interrelations of joint density functions and intensity processes for their dual matrix-valued counting processes provide valuable insights from theoretical as well as applicational perspectives.

For all the three absolutely continuous bivariate exponential models considered in the previous sections, as well as the Gumbel model for general marginal distributions  $F_X$  and  $F_Y$ , the intensity processes for the dual counting processes characterize a continuous time parameter Markovian structure, with four states which can be represented by the four possible combinations of  $(Y_1(t), Y_2(t))$ , each coordinate having two possible values  $\{0, 1\}$ . Note that the state represented by  $Y_1(t) = Y_2(t) = 0$  is an absorbing state. Nevertheless, excepting the Block-Basu model, for all other cases the Markovian structures are time-dependent (nonstationary) and hence, there may not be any advantage in incorporating such Markovian structures for statistical analysis. In the general case of a matrix-valued counting process, treated in [7], a similar Markovian property holds, but with arbitrary nonstationary intensities and an absorbing state. For this reason, an alternative martingale based approach formulated in [7] works out more conveniently.

Our study centers around two endpoints, represented by the occurrence of two events. In biomedical studies or clinical trials, such events are essentially related to the development of some physical phenomenon such as the incidence of cancer in two sites (e.g., uterus and breast for females or prostate and lungs for males), inadequate functioning of some particular organs such as kidney (right and left), loss of visual acuity on either eye, loss of hearing, etc. Such events are not necessarily associated with death or failure of the system; if it were so, we would have a competing risks model, where failure of a

single component terminates the functioning of the system and different methodologies would be required to handle such situations. In this context, we may identify the two endpoints associated with the same organ or relating to two components of a system. If they relate to two different components such as the uterus and the breast, the incidence of a disease in one of them requiring medical therapy may help the other component too. Therefore, it is possible to have a negative association in the two failure times. Alternatively, if the inactivity of a component leads to additional responsibility for the other one, it may generally have nonnegative association. Thus we may encounter both negative or nonnegative association schemes in such studies.

In view of the discussion made above, it may be wiser to look at the association pattern for all the three bivariate exponential distributions. Whereas for both the Basu-Block and Sarkar models we can only have nonnegative association, the Gumbel model allows some negative pattern though the underlying correlation coefficient can not be smaller than  $-0.25$ . Therefore, if biological or physiological considerations lead us to anticipate nonpositive association, and we want to make use of a parametric bivariate exponential model, the Gumbel model is the only choice, though this model also accommodates nonnegative associations, again, up to a limited range. Note that a similar picture holds for the Gumbel model when the underlying marginals are not necessarily exponentials. As an illustration, consider the problem of lumps in the breast. The two endpoints relate to the occurrence of these in the two breasts. If one of the two has it first, then medical therapies can not only reduce the risk in that particular site but also concur a similar reduction on the other one, leading to an anticipated nonpositive association. On the contrary, as is the case with the kidneys, inactivity of one of the components may lead to an increase of activity for the other one and thereby increases its risk of failure. Therefore, we anticipate nonnegative association of the two endpoints. In view of that, in a parametric setup a model selection should carefully be made in the light of underlying biomechanistics. In a majority of the cases, a nonnegative association may be justified wherein all the three models remain potentially adoptable. Therefore, we proceed to examine their intensity structures in further details.

The simplest situation arises with the Block-Basu model. As Figure 4 illustrates, the intensity for one of the components jumps to a constant higher level right after the occurrence of the event for the other component. The picture for the Sarkar model (Figure 3) is somewhat different, where we have a constant risk of failure for one component up to the occurrence of the event for the other one, when the intensity jumps to a higher level, and then monotonically decreases to an asymptote generally above the original level. We may note that the amplitude of this jump is time dependent, becoming smaller

as the time for the other failure becomes larger. Also, the intensity curve moves upward and the jump becomes larger with an increase of the value of the association parameter  $\beta_{12}$ . For the Gumbel exponential model, looking at Figures 1 and 2 we may note that whenever the association is nonpositive the intensity for one component initially increases, but not necessarily monotonically, until the other component fails, where followed by a drop, it monotonically increases to an asymptotic value that equals the initial level. For nonnegative association, the picture is exactly the opposite: initially decreasing but not necessarily monotonically, jumping to a height at the time of occurrence of the other event, and then monotonically decreasing to an asymptote equaling the initial level. Further, we note that when the two events are stochastically independent, the intensities are stationary, unaffected by the occurrence of a failure of the other events. Finally, we note that for all the three bivariate exponential models, as  $t$  increases, the intensity functions approach a finite asymptotic level. There are situations in multi-endpoint survival analysis where the hazards may not have a finite asymptote (as  $t \rightarrow \infty$ .) While this is difficult to cover with the class of exponential models considered here, the Gumbel formulation, explained before (10), allows such types for both positive and nonpositive association patterns. For example, for the marginal  $F_X(t)$  and/or  $F_Y(t)$ , instead of taking an exponential distribution, we may take an increasing failure rate (IFR) distribution which permits such unbounded hazards. An important example of such an IFR distribution with unbounded hazard (as  $t \rightarrow \infty$ ) is the Weibull distribution, for which

$$\bar{F}_X(t) = \exp\{-\theta t^\gamma\}, \quad t > 0, \theta > 0, \gamma > 0.$$

Here, for  $\gamma = 1$  we have the exponential distribution, while for  $\gamma > 1$  the corresponding hazard rate  $h_X(t) = \theta\gamma t^{\gamma-1}$  is monotonically increasing in  $t$  and tends to  $\infty$  as  $t \rightarrow \infty$ . In this case, our conclusion on the related intensity processes in the counting process model are similar to the exponential (Gumbel) model, excepting that here the asymptotes (as  $t \rightarrow \infty$ ) are unbounded. If instead of the Weibull we use the Gamma family with density

$$f(t) = \frac{1}{c^p \Gamma(p)} t^{p-1} e^{-t/c},$$

we may also have an IFR (when  $p > 1$ ), but the marginal hazard is bounded. In this way, the Gumbel prescription offers a wider range of choices, as long as the association is not too strong, and has the maximum advantage.

## References

- [1] Aalen, O. O. (1978). Nonparametric inference for a family of counting processes, *The Annals of*

*Statistics* **6**: 701–726.

- [2] Block, W. B. and Basu, A. P. (1974). A continuous bivariate exponential extension, *Journal of the American Statistical Association* **69**: 1031–1037.
- [3] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*. Vol. I, 3rd edn, Springer-Verlag.
- [4] Gumbel, E. J. (1959). Multivariate distributions with given margins. *Revista da Faculdade de Ciências. Série A, Lisboa*. **VII**: 179–218.
- [5] Gumbel, E. J. (1960). Bivariate exponential distributions, *Journal of the American Statistical Association* **55**: 698–707.
- [6] Marshal, A. W. and Olkin, I. (1967). A multivariate exponential distribution, *Journal of the American Statistical Association* **62**: 30–44.
- [7] Pedroso de Lima, A. C. and Sen, P. K. (1996). *A Matrix-Valued Counting Process Model with First Order Interactive Intensity*. RTMAE 9604, Instituto de Matemática e Estatística, Universidade de São Paulo, Brasil.
- [8] Sarkar, S. K. (1987). A continuous bivariate exponential distribution, *Journal of the American Statistical Association* **82**: 667–695.

DEPARTAMENTO DE ESTATÍSTICA  
UNIVERSIDADE DE SÃO PAULO  
CAIXA POSTAL 66281  
05389-970 SÃO PAULO, S.P., BRAZIL

DEPARTMENT OF BIostatISTICS  
UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL  
CHAPEL HILL, NORTH CAROLINA 27599-7400