The Joint Distribution of Bivariate Exponential Under Linearly Related Model

Norou Diawara
Department of Mathematics and Statistics
Old Dominion University
USA
ndiawara@odu.edu

Kumer Pial Das
Department of Mathematics, Lamar University,
USA
kumer.das@lamar.edu

Abstract

In this paper, fundamental results of the joint distribution of the bivariate exponential distributions are established. The positive support multivariate distribution theory is important in reliability and survival analysis, and we applied it to the case where more than one failure or survival is observed in a given study. Usually, the multivariate distribution is restricted to those with marginal distributions of a specified and familiar lifetime family. The family of exponential distribution contains the absolutely continuous and discrete case models with a nonzero probability on a set of measure zero. Examples are given, and estimators are developed and applied to simulated data. Our findings generalize substantially known results in the literature, provide flexible and novel approach for modeling related events that can occur simultaneously from one based event.

Keywords and phrases: bivariate exponential, Dirac delta, reliability models, survival analysis.

1. Introduction

The multivariate exponential distribution plays an important role in survival and reliability analysis as suggested in Walker and Stephens (1999), Kotz et al. (2000), Lawless (2003), Marshall and Olkin (1967), Joe (1997), Ghosh and Gelfand (1998), and Hougaard (2000), to mention a few. Such researchers have presented many problems related to the bivariate exponential distribution. This study provides generalization of substantially known results, with a flexible and novel approach for modeling related events that can occur proportionally from one based event. For example, in medical science, our model provides a valid response to an important question in the estimation and modeling of data from patients with left and right kidney failures after a given event such as gastric surgery or transplant. In reliability analysis, the model can be used for failure times of machinery components. So the paper is concerned with a class of multivariate data of systems with multiple components. The multiple components share the same onset factor. Simple variable transformations are not sufficient to achieve tractable and efficient modeling and estimation result. Normal and independence assumptions will also be inefficient as we consider the sample sizes not large enough to consider such assumptions. In this article, we propose a bivariate exponential that will greatly improve precision in the estimation of parameters. After describing the model in Section 2, we study some of its associated properties. In Section 3, we present the estimation technique. In many cases, the bivariate distributions are considered as the results are more meaningful. In that sense, we explore the bivariate exponential distributions and we provide a simulation example with results in higher dimensions in Section 4 with the distribution having exponential marginals. A final discussion concludes the article.

2. The Bivariate Exponential Model

In this section, we focus on the bivariate exponential distribution, where at each level a linearly associated exponential survival distribution with specified exponential marginal distribution model is considered. More specifically, we consider the bivariate X_1 and X_2 be fixed marginally as exponential random variables with hazard rates λ_1 and λ_2, respectively. Then by introducing two types of latent non-negative variables, X_0, and Z_1 and Z_2, statistically independent between themselves and of X_0, a linear relationship is formed between X_0 and X_1, X_2 as follows:

$$X_i = a_i X_0 + Z_i \tag{1}$$

where a_i's are fixed nonnegative constants and $Z_{i'}s$ are independent of X_0 for $i = 1, 2$. Note that Z_1 *and* Z_2 are considered latent, unobservable random variables, that generate the observable bivariate vector $X = (X_1, X_2)'$.

Our goal is to characterize Z_1 and Z_2, and use their form to write the join density of the distribution of X_0, X_1 and X_2, using the fact that X_1 and X_2 follow an exponential distribution. Once the joint density is obtained, then we can deduce estimator for X_0.

By computing the mean square error (MSE) for the parameter associated with X_0, we show the improvement obtained by comparing it with with estimators in the exponential case as proposed in Lawless (2003).

This parametrization is a convenient method to describe the relation between right and left kidney failures (X_1 and X_2) after the occurrence of gastric surgery (X_0) as we have motivated in the previous section. The dependence structure is given through the joint density function. Our objective is on the description, derivation and characterization of the joint distribution for (X_1, X_2) based on the class of distributions for Z_1 and Z_2, that produce the specified marginal distributions of X_1 and X_2, respectively. We derive and study estimators of the parameters under the resulting dependence structure.

Note that our model possesses the property of conditional independence given a random latent effect or frailty models, commonly used in describing dependent events. Hougaard (2000) considered the random latent effect as parameter. However, we are treating this random effect as latent variable where there is a nonzero probability of simultaneous or proportional occurrence capturing the idea of fatal shock. This idea is not new. Marshall and Olkin (1967) proposed a multivariate exponential distribution that is not absolutely continuous.

Our intent is to describe a general class of linearly related bivariate model, and one could lay out the higher variate version. We study the statistical properties of the bivariate MLE's. The models are formulated with independent exponential distributions.

Ghosh and Gelfand (1998) described a multivariate time to event data model. Focusing on the bivariate case, they use a Bayesian approach for inference using simulation. We focus on the exponential class of distributions because of its importance in the literature. This class is large and includes the continuous and discontinuous cases. We also study the bivariate case approach where X_0 is unobserved and missing. The intent is to increase the applicability of the method, as we know that failure of kidneys could occur without the occurrence of a predetermined event as gastric surgery. Moreover, the multivariate lifetime distribution by Hougaard (2000), has a dependence created by an unobservable quantity. Hougaard (1986) proposed a continuous multivariate lifetime distribution where the marginal distributions are Weibull (continuous) whose form does not allow the property of simultaneous or proportional failures of individuals or components. We want to retain that property in our model.We also lay out the joint density function along with the joint survival function. Estimators for the parameters associated with the model have also been developed. Carpenter et al. (2006) defined a similar approach at the univariate level. They characterize it through Laplace transforms, the distribution of the latent variable in the exponential case as mixture of a point mass at zero and an exponential with hazard rate λ_i. Note that when $Z_i = 0,$ there is a positive probability that X_i is proportional to X_0 with proportionality constant a_i, i.e. $P(X_i = a_i X_0) > 0$, for $i = 1, 2$.

The proposed bivariate model has many resemblance with multivariate models that are suggested. However, as mentioned by Karlis (2003), extensions of univariate distributions have not been applied in many practical situations mainly due to a non straightforward method of generalizing univariate to multivariate models and a lack of inferential procedures. The different multivariate exponential procedures that have been proposed are summarized in Kotz et al. (2000). They include the so called Gumbel distribution and the Marshall-Olkin distribution. Statistical inference is complicated by the fact that there is no burden free density forms.

The bivariate (and multivariate) survival data of the experiment gives multiple events and involves several members or components in a system. There is no simple expression in the density as given in Mathai and Moschopoulos (1992).

Definition 2.1 Let X_0, X_1, X_2 be exponential random variables as in (1) with scale parameters λ_i for $i = 0,1,2.$ Let $Z_i,\ \ i = 1,2,$ be independent random variables satisfying (1) with fixed positive constants a_1 and a_2. We define the joint distribution of $X = (X_1, X_2)$ as the bivariate exponential distribution. Our goal is to understand the distribution of Z_i, and the joint distribution of X_1 and X_2.

Thus, for $i = 1,2$:

$$X_i = \begin{cases} a_i X_0, & \textit{with probability } p_i, \quad i.e\ Z_i = 0, \\ a_i X_0 + Z_i, & \textit{with probability } 1 - p_i, \quad i.e\ Z_i > 0. \end{cases}$$

The mean of X is given as

$$E(X) = (1\lambda_1, 1\lambda_2)', \qquad (2)$$

and its variance/covariance matrix is given as

$$\Sigma = 1\lambda_0^2 \begin{pmatrix} \lambda_0^2\lambda_1^2 & a_1 a_2 \\ a_2 a_1 & \lambda_0^2\lambda_2^2 \end{pmatrix}. \qquad (3)$$

From Carpenter et al. (2006) and Equation (1), the LST of Z_i is

$$L_{Z_i}(s) = p_i + (1 - p_i) L_{X_i}(s), \quad \textit{with} \quad p_i = a_i \lambda_i \lambda_0, \quad i = 1,2.$$

That is Z_i is a mixture of a Bernoulli random variable with probability p_i and an exponential random variable with parameter λ_i for i=1,2.

To describe the exact form of the distribution of $Z_i, i = 1,2,$ we introduce the Dirac delta function at the point $c \in \mathbb{R}$, as a point mass distribution denoted δ_c, and we say that a random variable X has point mass δ_c distribution at c if its pmf is given by

$$f(x \mid c) = \delta_c(x) = \delta(x - c) = 0 \quad \textit{if} \quad x \neq c, \quad \textit{and} \quad \int_{-\infty}^{\infty} f(x \mid c)\,dx = 1. \qquad (4)$$

More details on the Dirac function are given in Khuri (2004), Au and Tam (1999) and Pazman and Pronzato (1996). It is well known that the Heaviside step function is an antiderivative of the Dirac distribution. The Heaviside step function, also called unit step function, see for example Abramowitz and Stegun (1972), is a discontinuous function defined as

$$H(x) = \int_{-\infty}^{x} \delta(t)\,dt = \begin{cases} 0, & \textit{if } x \le 0; \\ 1, & \textit{if } x > 0. \end{cases} \qquad (5)$$

Iyer et al. (2002) studied the case of positive fixed a_i's in the one dimensional case assuming that X_0 and X_i are exponential with parameters λ_0 and $\lambda_i, \quad , i = 1, 2,$ respectively.

Carpenter et al. (2006) showed that the density of Z_i is then expressed as:

$$f_{Z_i}(z) = p_i \delta(z) + (1 - p_i) f_{X_i}(z) I_{(z>0)}, \tag{6}$$

where $p_i = P(Z_i = 0) = a_i \lambda_i \lambda_0, \quad i = 1, 2,$ and δ is the Dirac delta function defined as in (4).

From Equation (??), the conditional survival function of X_i given X_0, is obtained as

$$S(t \mid x_0) = P(X_i > t \mid x_0) = P(a_i x_0 + Z_i > t \mid x_0) = P(Z_i > t - a_i x_0 \mid x_0)$$

$$= \int_{t-a_i x_0}^{\infty} f_{Z_i}(z) dz = \int_{t-a_i x_0}^{\infty} [p_i \delta(z) + (1 - p_i) \lambda_i e^{-\lambda_i z}] dz$$

$$= p_i \int_{t-a_i x_0}^{\infty} \delta(z) dz + (1 - p_i) \int_{t-a_i x_0}^{\infty} e^{-\lambda_i z} dz$$

$$= p_i (1 - H(t - a_i x_0)) + (1 - p_i) e^{-\lambda_i (t - a_i x_0)}, \quad i = 1, 2,$$

where H is the Heaviside function defined in Equation (5).

Also using the fact that

$$f(x_0, z_1, z_2) = f(x_0) f(z_1) f(z_2), \quad \textit{we have that}$$

$$f(x_0, x_1, x_2) = f(x_0) f_{Z_1}(x_1 - a_1 x_0) f_{Z_2}(x_2 - a_2 x_0), \quad \textit{or} \tag{7}$$

$$f(x_1, x_2 \mid x_0) = f_{Z_1}(x_1 \mid x_0) f_{Z_2}(x_2 \mid x_0), \tag{8}$$

and then (X_1, X_2) is conditionally independent given X_0.

From property (8), the joint conditional survival function is

$$S(x_1, x_2 \mid x_0) = \prod_{i=1}^{2} [p_i (1 - H(x_i - a_i x_0)) + (1 - p_i) e^{-\lambda_i (x_i - a_i x_0)}]. \tag{9}$$

The distribution of the minimum lifetime distribution $X_{(1)} = min\{X_1 a_1, X_2 a_2\}$ can be derived directly from (??) and from properties of the Heaviside function in Equation (5). It is given as

$$P(X_{(1)} > t \mid x_0) = \prod_{i=1}^{2} P(X_i a_i > t \mid x_0) = \prod_{i=1}^{2} P(X_i > a_i t \mid x_0)$$

$$= \prod_{i=1}^{2} [p_i (1 - H(x_i - x_0)) + (1 - p_i) e^{-\lambda_i a_i (x_i - x_0)}].$$

Note that,

$$S(0,0)=1 \quad and \quad S(\infty,\infty)=0.$$

Taking the derivative of the i^{th} survival function in Equation (??), and using Equation (5), the conditional density is given by

$$f_{X_i|X_0}(t) = p_i\delta(t-a_i x_0)+(1-p_i)\lambda_i e^{-\lambda_i(t-a_i x_0)}, \quad i=1,2. \tag{10}$$

and from (4), the conditional expectation is

$$E_{X_i|X_0}(X_i)=\int_0^\infty t f_{X_i|X_0}(t)dt$$

$$= p_i\int_0^\infty t\delta(t-a_i x_0)dt+(1-p_i)\int_{a_i x_0}^\infty \lambda_i t e^{-\lambda_i(t-a_i x_0)}dt$$

$$= p_i a_i x_0+(1-p_i)(1\lambda_i+a_i x_0)=a_i x_0+(1-p_i)1\lambda_i, \quad i=1,2.$$

Note that taking the expectation of the above with respect to X_0 gives

$$E(X_i)=E_{X_0}E_{X_i|X_0}(X_i)=E_{X_0}[p_i a_i x_0+(1-p_i)(1\lambda_i+a_i x_0)]$$

$$= p_i a_i\lambda_0+(1-p_i)(1\lambda_i+a_i\lambda_0)$$

$$= a_i\lambda_0+(1-p_i)1\lambda_i$$

$$=1\lambda_i, \quad i=1,2,$$

since $p_i=a_i\lambda_i\lambda_0$, confirming earlier results in Equation (1) and in Carpenter et al. (2006).

Although the joint density of $(X_0,X_1,X_2)'$ was easily found in (??), the density of $(X_1,X_2)'$ is not obvious. However, we can study the density of $(X_0,X_1,X_2)'$ through the latent variables $Z_1,Z_2,$ with relative ease. Using the independence of the Z_i's, $i=1,2$, between each other and of X_0, and result (??), we have that:

$$f(x_0,x_1,x_2)=\lambda_0 e^{-\lambda_0 x_0}\prod_{i=1}^{2}[p_i\delta(x_i-a_i x_0)+(1-p_i)\lambda_i e^{-\lambda_i(x_i-a_i x_0)}I_{(x_i>a_i x_0)}]. \tag{11}$$

Hence (??) can be written as:

$$\begin{aligned} f(x_0,x_1,x_2) &= \lambda_0 e^{-\lambda_0 x_0}[p_1p_2\delta(x_1-a_1x_0)\delta(x_2-a_2x_0) \\ &+(1-p_1)p_2\lambda_1 e^{-\lambda_1(x_1-a_1x_0)}I_{(x_1a_1>x_0=x_2a_2)} \\ &+p_1(1-p_2)\lambda_2 e^{-\lambda_2(x_2-a_2x_0)}I_{(x_2a_2>x_0=x_1a_1)} \\ &+(1-p_1)(1-p_2)\lambda_1\lambda_2 e^{-\lambda_1x_1}e^{-\lambda_2x_2}e^{-x_0(a_1\lambda_1+a_2\lambda_2)}I_{(\varphi>x_0)}] \end{aligned}$$

where $\varphi=min\{x_1a_1,x_2a_2\}$.

Therefore the joint density of the bivariate exponential $(X_1, X_2)'$ is obtained by integrating the above expression with respect to x_0, giving

$$f(x_1,x_2) = p_1 p_2 \lambda_0 a_1 a_2 e^{-\lambda_0 x_2 a_2} \delta_{(x_1 a_1 - x_2 a_2)}$$

$$+(1-p_1) p_2 1 a_2 \lambda_0 \lambda_1 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-\lambda^* \varphi} I_{(x_1 a_1 > x_2 a_2)}$$

$$+p_1 (1-p_2) 1 a_1 \lambda_0 \lambda_2 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-\lambda^* \varphi} I_{(x_2 a_2 > x_1 a_1)}$$

$$+(1-p_1)(1-p_2) \lambda_0 \lambda_1 \lambda_2 \lambda^* e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} (1 - e^{-\lambda^* \varphi}),$$

$$where \quad \varphi = min(x_1 a_1, x_2 a_2) \quad and \quad \lambda^* = \lambda_0 - a_1 \lambda_1 - a_2 \lambda_2 .$$

The proof of the above result is shown in the Appendix section.

Similarly, to get the unconditional survival function, one would derive it from (??), and therefore compute

$$S(x_1,x_2) = \int_0^\infty \prod_{i=1}^{2} S(x_i \mid x_0) f_{X_0}(x_0) dx_0$$

$$= \lambda_0 \int_0^\infty \prod_{i=1}^{2} [p_i (1 - H(x_i - a_i x_0)) + (1-p_i) e^{-\lambda_i (x_i - a_i x_0)}] e^{-\lambda_0 x_0} dx_0 .$$

As we can see, we cannot interchange integration and product in the above expression. This problem will be discussed at in the next section.

3. Estimation Technique for the Bivariate Exponential Model

In this section, we study the estimation techniques for the bivariate exponential distribution of the previous section. The bivariate exponential distribution from Definition 2.1 can be expressed as:

$$\begin{cases} X_1 = a_1 X_0 + Z_1 & ; \\ X_2 = a_2 X_0 + Z_2 & . \end{cases} \tag{12}$$

Then, from Carpenter et al. (2006), the joint density of (X_0, X_1) is given by

$$f(x_0,x_1) = \begin{cases} p_1 f_{X_0}(x_0) \delta(x_1 = a_1 x_0) \\ (1-p_1) f_{X_0}(x_0) f_{X_1}(x_1 - a_1 x_0) I_{(x_1 > a_1 x_0)} \end{cases},$$

$$= \begin{cases} p_1 \lambda_0 e^{-\lambda_0 x_0}, & if \ x_0 = x_1 \ a_1 \\ (1-p_1) \lambda_0 \lambda_1 e^{-\lambda_0 x_0} e^{-\lambda_1 (x_1 - a_1 x_0)}, & if \ x_0 < x_1 \ a_1 \end{cases},$$

where $p_1 = a_1 \lambda_1 \lambda_0$. Similarly based on the expression $X_2 = a_2 X_0 + Z_2$, we have for $p_2 = a_2 \lambda_2 \lambda_0$,

$$f(x_0, x_2) = \begin{cases} p_2 f_{X_0}(x_0)\delta(x_2 = a_2 x_0) \\ (1-p_2) f_{X_0}(x_0) f_{X_2}(x_2 - a_2 x_0) I_{(x_2 > a_2 x_0)} \end{cases},$$

$$= \begin{cases} p_2 \lambda_0 e^{-\lambda_0 x_0}, & if\ x_0 = x_2\ a_2 \\ (1-p_2)\lambda_0\lambda_2 e^{-\lambda_0 x_0} e^{-\lambda_2(x_2 - a_2 x_0)}, & if\ x_0 < x_2\ a_2 . \end{cases}$$

Hence, using the independence between X_0, Z_1, Z_2, the joint density of (X_0, X_1, X_2) based on (??), is shown in the Appendix and is given by:

$$f(x_0, x_1, x_2) = p_1 p_2 \lambda_0 e^{-\lambda_0 x_0} \delta_{(x_1 - a_1 x_0)} \delta_{(x_2 - a_2 x_0)}$$

$$+ p_1(1-p_2)\lambda_0 e^{-\lambda_0 x_0} f_{X_2}(x_2 - a_2 x_0)\delta_{(x_1 - a_1 x_0)}$$

$$+(1-p_1) p_2 \lambda_0 e^{-\lambda_0 x_0} f_{X_1}(x_1 - a_1 x_0)\delta_{(x_2 - a_2 x_0)}$$

$$+(1-p_1)(1-p_2)\lambda_0 e^{-\lambda_0 x_0} f_{X_1}(x_1 - a_1 x_0) f_{X_2}(x_2 - a_2 x_0) I_{(x_1 > a_1 x_0, x_2 > a_2 x_0)}.$$

The expression $f(x_0, x_1, x_2)$ is one way to obtain an estimate for x_0 or the parameter associated with it, λ_0. The above likelihood equations can be used to estimate λ_0, λ_1 and λ_2 if the x_{0i}'s, $1 \le i \le n$, were known. We develop estimators of these latent terms. It is worth noting that no approximations has been used here.

To develop the unconditional estimators, we would like to avoid the problem of the large sample sizes needed raised in Bowman and Shenton (2002). We integrate out x_0 from the joint density $f(x_0, x_1, x_2)$.

Based on a random sample of size n denoted $(x_{11}, x_{21}), (x_{12}, x_{22}), \cdots, (x_{1n}, x_{2n})$, let's define for $1 \le i \le n$,

$$\varphi_i = min(x_{1i} a_1, x_{2i} a_2) \qquad and$$

$$r_i^{(1)} = \begin{cases} 1, & if\ x_{0i} = x_{1i}\ a_1 \le x_{2i}\ a_2\ ; \\ 0, & if\ x_{1i}\ a_1 > x_{2i}\ a_2\ , \end{cases} \quad \text{and } r_i^{(2)} = \begin{cases} 1, & if\ x_{0i} = x_{2i}\ a_2 \le x_{1i}\ a_1\ ; \\ 0, & if\ x_{2i}\ a_2 > x_{1i}\ a_1\ . \end{cases}$$

Then

$$1\lambda_0 = 1a_2 \sum_i x_{2i}(1 - r_i^{(1)} r_i^{(2)}) n + \sum_i \varphi_i r_i^{(1)} r_i^{(2)} n - 1a_2\lambda_2 \sum_i (1 - r_i^{(2)}) n.$$

The details are given in the Appendix. This estimator is for the parameter associated with the unknown latent variable x_0 based on the minimum of x_1 / a_1 and x_2 / a_2.

We have proposed a very general bivariate exponential class of distributions. We have described the form of the joint distribution functions. Estimations of the parameters are given based on the likelihood equation. We have done all that retaining the form of the marginal exponential distribution, and the fatal shock idea as in Marshall and Olkin (1967).

In the next section, we examine a simulated example to illustrate our proposed model.

4. Simulation Example

In this section, we perform, as in Minhajuddin et al. (2003), a simulation study of the bivariate exponential to examine the properties of estimator of the parameter from the latent distribution, λ_0. We focus on λ_0 because it is an important portion of the correlation structure, and all of the parameters associated with X_1 and X_2 can be easily estimated marginally, since they are the observed diseases or events that occurred after the primary event X_0. We assess the performance of the proposed model by computing the errors in the estimation differences.

Simulation Design

Based on 10,000 replications of sample size 50 each, of $(X_1, X_2)'$, we choose all a_i's to be 1, λ_0 to be 1, and solving for λ_i w.r.t. ρ, we have that $\lambda_i = \rho a_i \lambda_0 = \rho$, for $i = 1, 2$. Separate simulations are done for $\rho = 0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90 \quad and \quad 0.95$.

Results for the bias and MSE are presented in Figure 1 and Figure 2, respectively. $Bias0_2$ and $Mse0_2$ represent the bias and MSE for $\hat{\lambda}_0 = 1/\bar{x}_0$, where $\bar{x}_0 = \sum_{i=1}^{n} x_{0i}$, if the latent unobservable values, $x_{01}, \ldots, x_{0n}$, were actually known. It is important to point out that $\hat{\lambda}_0$ is not observable. However, if these values were observable, then $\hat{\lambda}_0$ would be MLE and the best unbiased estimator for λ_0. Therefore, the performance of $\hat{\lambda}_0$ serves as a good benchmark. More precisely, if we denote x_{min} to be the minimum of $\{x_1/a_1, x_2/a_2\}$ as we suggested on (??), then:

$$Bias0 = 1\bar{x}_0 - \lambda_0 \quad and \quad Bias1 = 1\bar{x}_{min} - \lambda_0, \quad .$$

The MSE are

$$Mse0 = Bias0^2 \quad and \quad Mse1 = Bias1^2$$

Simulation Results

Figure 1 summarizes the results of our simulation. Our proposed estimator compares very well with the true value, in both univariate and bivariate case. As the correlation ρ increases, the bivariate case bias improves significantly. Also, as ρ increases, the estimate of x_0 becomes more efficient. It is also observable that the bias becomes satisfactory with higher correlation for the bivariate.

Figure 1: Bias of λ_0 for different correlations

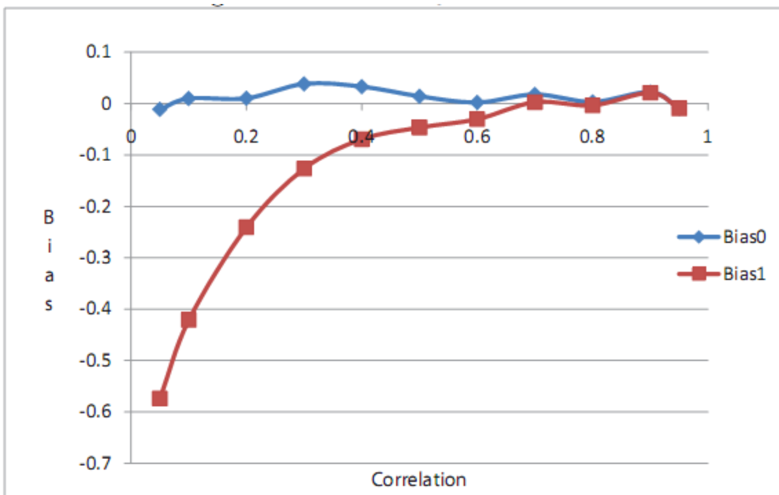
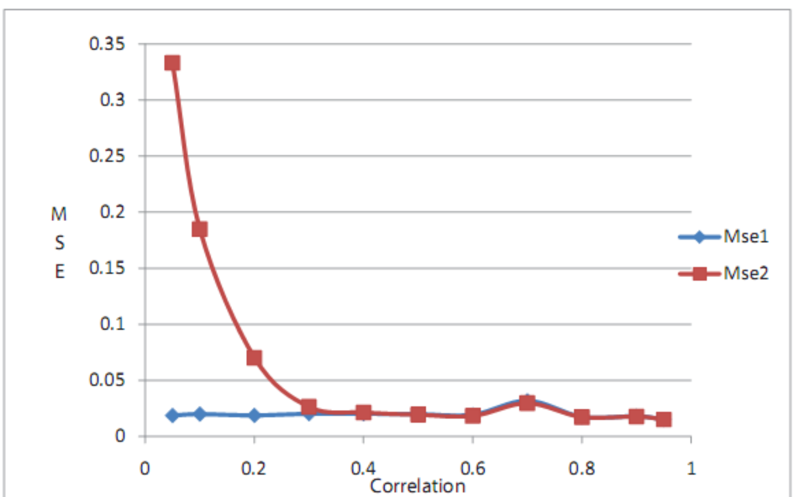


Figure 2: MSE of λ_0 for different correlations



We also present the MSE of the estimates in Figure 2. These estimated values show the effectiveness of the proposed estimation techniques developed. As we see from the Figure 2 of MSE, the difference does appear to be consistently small, although the high values of correlations do appear to give lower MSE's. The algorithm of the proposed estimation was implemented using the $SAS^{®}$ program.

5. Conclusion

In this article, we have defined and characterized a new bivariate generalized exponential distribution with potential applications in survival and reliability modeling. This family possesses exponential marginals and it contains absolutely continuous classes, as well as, the Marshall Olkin type of distributions with a positive probability mass on a set of measure zero. The variables making up the bivariate vector were made linearly related indirectly through a collection of latent random variables. Also, the bivariate distribution is not necessarily restricted to those with exponential marginal distributions. Estimators based on the EM algorithm idea were proposed for the unknown parameters, and, in addition, methods were given to estimate the latent terms in the model. We have shown that our approach generalizes cases of the models proposed by Iyer et al. (2002). The possible implication of this work is enormous. It takes into account the non identically and independently distributed properties of small sample size data. Assuming that a_i's are unknown in their structures will add a lot more applications to the model. This and other related issues are topics for further research.

Acknowledgements

The first author's research was sponsored by the Office of Research at Old Dominion University, in Norfolk, Virginia under the Summer Research Fellowship contract project number ODU-RF 993036.

The second author's research was partially sponsored by the Research Enhancement Grant-2009, Lamar University, Beaumont, TX-77710.

References

1. Abramowitz and Stegun (1972), Abramowitz, M and Stegun, I.A. (1972), Handbook of Mathematical Functions, Selected Government Publications, Chapter 6, New York, Dover.
2. Au and Tam (1999), Au, C. and Tam, J. (1999), Transforming Variables using the Dirac Generalized Function, The American Statistician, Vol. 53, 3, pp. 270-273.
3. Carpenter et al. (2006), Carpenter, M., Diawara, N. and Han, Yi, (2006), A New Class of Bivariate Weibull Survival Distributions, American Journal of Mathematical and Management Sciences, Vol. 26 (1 & 2), pp. 164-184.
4. Dudewicz and Mishra (1988), Dudewicz, E.J. and Mishra, S. (1988), Modern Mathematical Statistics, Wiley Series in Probability and Mathematical Statistics.
5. Ghosh and Gelfand (1998), Ghosh, S.K. and Gelfand, A.E. (1998), Latent Waiting Time Models for Bivariate Event Times with Censoring, Sankya, Vol. 60, Series A, pp. 31-47.
6. Hougaard (1986), Hougaard, P. (1986), A class of multivariate failure time distributions, Biometrika, 73, 3, pp. 671-678.

7. Hougaard (2000), Hougaard, P. (2000), Analysis of Multivariate Survival Data, Springer.
8. Iyer et al. (2002), Iyer, Srilanth K., Manjunath, D. and Manivasakan, R. (2002), Bivariate Exponential Distributions Using Linear Structures, Sankhya, V. 64, Series A, pp. 156-166.
9. Joe (1997), Joe, H. (1997), Multivariate Models and Dependence Concepts, Chapman & Hall.
10. Karlis (2003), Karlis, D. (2003), ML Estimation for Multivariate Shock Models via an EM Algorithm, Ann. Inst. Statist. Math., Vol. 55, 4, pp. 817-830.
11. Khuri (2004), Khuri, A. (2004), Applications of Dirac's Delta Function in Statistics, Int. J. Math. Educ. Sci. Technol., 35, 2, pp. 185-195.
12. Kotz et al. (2000), Kotz, S., Balakrishnan, N. and Johnson, N. (2000), Continuous Multivariate Distributions, Volume 1 Wiley Series in Probability and Statistics.
13. Lawless (2003), Lawless, J. F. (2003), Statistical Models and Methods for Lifetime Data, 2^{nd} edition Wiley Series in Probability and Statistics.
14. Marshall and Olkin (1967), Marshall, A.W. and Olkin, I. (1967), A Multivariate Exponential Distribution, J. Amer. Stat. Assoc., 63, pp. 30-44.
15. Mathai and Moschopoulos (1992), Mathai, A.M. and Moschopoulos, P.G. (1992), A Form of Multivariate Gamma Distribution, Annals of the Institute of Statistical Mathematics, 44, pp. 97-106.
16. Minhajuddin et al. (2003), Minhajuddin, A.T.M., Harris, I.R., and Schucany, W.R. (2003), Simulating Multivariate Distributions with Specific Correlations, J. Statist. Comp. Simul., Vol. 74, pp. 599-607.
17. Pazman and Pronzato (1996), Pazman, A. and Pronzato, L. (1996), A Dirac-function method for densities of nonlinear statistics and for marginals densities in nonlinear regression, Statistics and Probability Letters, Vol 26, pp. 159-167.
18. Walker and Stephens (1999), Walker, S.G. and Stephens, D.A. (1999), A Multivariate Family of Distributions on $(0,\infty)^p$, Biometrika, Vol. 86, No. 3, pp. 703-709.
19. Bowman and Shenton (2002), Bowman, K.O. and Shenton, L.R. (2002), Problems with Maximum Likelihood Estimation and the 3-Parameter Gamma Distribution, J. Statist. Comp. Simul. Vol. 72 (5), pp. 391-401.

Appendix

Based on the independence between X_0, Z_1, Z_2 and using (??), the joint density of (X_0, X_1, X_2) is given by:

$$f(x_0,x_1,x_2) = \lambda_0 e^{-\lambda_0 x_0}[p_1\delta(x_1 - a_1x_0) + (1-p_1)f_{X_1}(x_1 - a_1x_0)I_{(x_1>a_1x_0)}]$$

$$\times[p_2\delta(x_2 - a_2x_0) + (1-p_2)f_{X_2}(x_2 - a_2x_0)I_{(x_2>a_2x_0)}]$$

$$= p_1p_2\lambda_0 e^{-\lambda_0 x_0}\delta_{(x_1-a_1x_0)}\delta_{(x_2-a_2x_0)}$$

$$+p_1(1-p_2)\lambda_0 e^{-\lambda_0 x_0} f_{X_2}(x_2-a_2x_0)\delta_{(x_1-a_1x_0)}$$

$$+(1-p_1)p_2\lambda_0 e^{-\lambda_0 x_0} f_{X_1}(x_1-a_1x_0)\delta_{(x_2-a_2x_0)}$$

$$+(1-p_1)(1-p_2)\lambda_0 e^{-\lambda_0 x_0} f_{X_1}(x_1-a_1x_0) f_{X_2}(x_2-a_2x_0) I_{(x_1>a_1x_0,x_2>a_2x_0)}.$$

The expression $f(x_0,x_1,x_2)$ is one way to obtain an estimate for x_0 or the parameter associated with it, λ_0. Let's assume that $\varphi = min(x_1a_1, x_2a_2) = x_1a_1 \le x_2a_2$.

Also, set $r_i^{(1)} = I_{(z_{1i}=0)} = I_{(x_1=a_1x_0)} = \begin{cases} 1 & , if\ z_{1i}=0 \\ 0 & , if\ z_{1i} \ne 0, \end{cases}$

and $r_i^{(2)} = I_{(z_{2i}=0)} = I_{(x_2=a_2x_0)} = \begin{cases} 1 & , if\ z_{2i}=0 \\ 0 & , if\ z_{2i} \ne 0. \end{cases}$

Then, the full likelihood function based on a random sample of size n is the product of n contributed likelihoods and is given as:

$$L(\lambda_0,\lambda_1,\lambda_2) = \prod_{i=1}^{n}[p_1p_2\lambda_0 e^{-\lambda_0 x_{0i}}]^{r_i^{(1)}r_i^{(2)}}$$

$$\times[p_1(1-p_2)\lambda_0\lambda_2 e^{-(\lambda_0-a_2\lambda_2)x_{0i}} e^{-\lambda_2 x_{2i}}]^{(1-r_i^{(1)})r_i^{(2)}}$$

$$\times[(1-p_1)p_2\lambda_0\lambda_1 e^{-(\lambda_0-a_1\lambda_1)x_{0i}} e^{-\lambda_1 x_{1i}}]^{r_i^{(1)}(1-r_i^{(2)})}$$

$$\times[(1-p_1)(1-p_2)\lambda_0\lambda_1\lambda_2 e^{-(\lambda_0-a_1\lambda_1-a_2\lambda_2)x_{0i}} e^{-\lambda_1 x_{1i}} e^{-\lambda_2 x_{2i}}]^{(1-r_i^{(1)})(1-r_i^{(2)})}$$

$$=\prod_{i=1}^{n}[a_1a_2\lambda_1\lambda_2\lambda_0 e^{-\lambda_0 x_{0i}}]^{r_i^{(1)}r_i^{(2)}}$$

$$\times[a_1\lambda_1\lambda_2\lambda_0(\lambda_0-a_2\lambda_2)e^{-(\lambda_0-a_2\lambda_2)x_{0i}} e^{-\lambda_2 x_{2i}}]^{(1-r_i^{(1)})r_i^{(2)}}$$

$$\times[a_2\lambda_1\lambda_2\lambda_0(\lambda_0-a_1\lambda_1)e^{-(\lambda_0-a_1\lambda_1)x_{0i}} e^{-\lambda_1 x_{1i}}]^{r_i^{(1)}(1-r_i^{(2)})}$$

$$\times[\lambda_1\lambda_2\lambda_0(\lambda_0-a_1\lambda_1)(\lambda_2-a_2\lambda_2)$$

$$\times e^{-(\lambda_0-a_1\lambda_1-a_2\lambda_2)x_{0i}} e^{-\lambda_1 x_{1i}} e^{-\lambda_2 x_{2i}}]^{(1-r_i^{(1)})(1-r_i^{(2)})}$$

$$=(a_1a_2\lambda_1\lambda_2\lambda_0)^{\sum_i r_i^{(1)}r_i^{(2)}} e^{-\lambda_0\sum_i x_{0i}r_i^{(1)}r_i^{(2)}}$$

$$\times(a_1a_2\lambda_1\lambda_2\lambda_0)^{\sum_i (1-r_i^{(1)})r_i^{(2)}} (\lambda_0-a_2\lambda_2)^{\sum_i (1-r_i^{(1)})} e^{-(\lambda_0-a_2\lambda_2)\sum_i x_{0i}(1-r_i^{(1)})r_i^{(2)}}$$

$$\times(a_1a_2\lambda_1\lambda_2\lambda_0)^{\sum_i r_i^{(1)}(1-r_i^{(2)})}(\lambda_0-a_1\lambda_1)^{\sum_i(1-r_i^{(2)})}e^{-(\lambda_0-a_1\lambda_1)\sum_i x_{0i}r_i^{(1)}(1-r_i^{(2)})}$$

$$\times(\lambda_1\lambda_2\lambda_0)^{\sum_i(1-r_i^{(1)})(1-r_i^{(2)})}e^{-\lambda_1\sum_i x_{1i}(1-r_i^{(2)})}e^{-\lambda_2\sum_i x_{2i}(1-r_i^{(1)})}$$

$$\times e^{-(\lambda_0-a_1\lambda_1-a_2\lambda_2)\sum_i x_{0i}(1-r_i^{(1)})(1-r_i^{(2)})}.$$

Hence,

$$L(\lambda_0,\lambda_1,\lambda_2)=a_1^{\sum_i r_i^{(2)}}a_2^{\sum_i r_i^{(1)}}(\lambda_1\lambda_2\lambda_0)^n e^{-\lambda_0\sum_i x_{0i}r_i^{(1)}r_i^{(2)}}$$

$$(\lambda_0-a_2\lambda_2)^{\sum_i(1-r_i^{(1)})}e^{-(\lambda_0-a_2\lambda_2)\sum_i x_{0i}(1-r_i^{(1)})r_i^{(2)}}$$

$$(\lambda_0-a_1\lambda_1)^{\sum_i(1-r_i^{(2)})}e^{-(\lambda_0-a_1\lambda_1)\sum_i x_{0i}r_i^{(1)}(1-r_i^{(2)})}$$

$$e^{-\lambda_1\sum_i x_{1i}(1-r_i^{(2)})}e^{-\lambda_2\sum_i x_{2i}(1-r_i^{(1)})}$$

$$e^{-(\lambda_0-a_1\lambda_1-a_2\lambda_2)\sum_i x_{0i}(1-r_i^{(1)})(1-r_i^{(2)})}.$$

Hence the log likelihood is

$$LL(\lambda_0,\lambda_1,\lambda_2)=\log(a_1)\sum_i r_i^{(2)}+\log(a_2)\sum_i r_i^{(1)}+n\log(\lambda_1\lambda_2\lambda_0)$$

$$+\log(\lambda_0-a_1\lambda_1)\sum_i(1-r_i^{(2)})+\log(\lambda_0-a_2\lambda_2)\sum_i(1-r_i^{(1)})$$

$$-(\lambda_0-a_1\lambda_1)\sum x_{0i}r_i^{(1)}(1-r_i^{(2)})-(\lambda_0-a_1\lambda_1)\sum x_{0i}(1-r_i^{(1)})r_i^{(2)}$$

$$-\lambda_1\sum_i x_{1i}(1-r_i^{(2)})-\lambda_2\sum_i x_{2i}(1-r_i^{(1)})$$

$$-(\lambda_0-a_1\lambda_1-a_2\lambda_2)\sum_i x_{0i}(1-r_i^{(1)})(1-r_i^{(2)}),$$

and $\partial LL\partial\lambda_0=-n\lambda_0+\sum_i(1-r_i^{(2)})\lambda_0-a_1\lambda_1+\sum_i(1-r_i^{(1)})\lambda_0-a_2\lambda_2$

$$-\sum_i x_{0i}r_i^{(1)}(1-r_i^{(2)})-\sum_i x_{0i}(1-r_i^{(1)})r_i^{(2)}$$

$$-\sum_i x_{0i}(1-r_i^{(1)})(1-r_i^{(2)})$$

$$=-n\lambda_0+\sum_i(1-r_i^{(2)})\lambda_0-a_1\lambda_1+\sum_i(1-r_i^{(1)})\lambda_0-a_2\lambda_2-\sum_i x_{0i}(1-r_i^{(1)}r_i^{(2)}).$$

Similarly

$$\partial LL\partial\lambda_1 = n\lambda_1 - a_1\sum_i(1-r_i^{(2)})\lambda_0 - a_1\lambda_1 + a_1\sum_i x_{0i}(1-r_i^{(2)}) - \sum_i x_{1i}(1-r_i^{(2)})$$

So setting $\partial LL\partial\lambda_1 = 0$ gives

$$\sum_i(1-r_i^{(2)})\lambda_0 - a_1\lambda_1 = na_1\lambda_1 + \sum_i x_{0i}(1-r_i^{(2)}) - \sum_i x_{1i}(1-r_i^{(2)})a_1,$$

$$and \quad \partial LL\partial\lambda_2 = n\lambda_2 - a_2\sum_i(1-r_i^{(1)})\lambda_0 - a_2\lambda_2 + a_2\sum_i x_{0i}(1-r_i^{(1)}) - \sum_i x_{2i}(1-r_i^{(1)}).$$

So setting $\partial LL\partial\lambda_2 = 0$ gives

$$\sum_i(1-r_i^{(1)})\lambda_0 - a_2\lambda_2 = na_2\lambda_2 + \sum_i x_{0i}(1-r_i^{(1)}) - \sum_i x_{2i}(1-r_i^{(1)})a_2.$$

Now setting $\partial LL\partial\lambda_0 = 0$, and substituting values for $\sum_i(1-r_i^{(2)})\lambda_0 - a_1\lambda_1$ and $\sum_i(1-r_i^{(1)})\lambda_0 - a_2\lambda_2$ gives

$$1\lambda_0 = 1a_1\lambda_1 + 1a_2\lambda_2 - \sum_i\{x_{2i}(1-r_i^{(1)}) - a_2x_{0i}(1-r_i^{(1)})\}a_2n$$

$$-\sum_i\{x_{1i}(1-r_i^{(2)}) - a_1x_{0i}(1-r_i^{(2)})\}a_1n - \sum_i x_{0i}(1-r_i^{(1)}r_i^{(2)})n$$

$$= 1a_1\lambda_1 + 1a_2\lambda_2 - \sum_i z_{1i}(1-r_i^{(2)})a_1n - \sum_i z_{2i}(1-r_i^{(1)})a_2n - \sum_i x_{0i}(1-r_i^{(1)}r_i^{(2)})n$$

$$= 1a_1(1\lambda_1 - \sum_i z_{1i}(1-r_i^{(2)})n) + 1a_2(1\lambda_2 - \sum_i z_{2i}(1-r_i^{(1)})n) - \sum_i x_{0i}(1-r_i^{(1)}r_i^{(2)})n.$$

The above likelihood equations can be used to estimate λ_0, λ_1 and λ_2 if the x_{0i}'s, $1 \le i \le n$, were known. We develop estimators of these latent terms. It is worth noting that no approximations has been used here.

To develop the unconditional estimators, we would like to avoid the problem of the large sample sizes needed raised in Bowman and Shenton (2002). We integrate out x_0 from the joint density $f(x_0, x_1, x_2)$, and we have that:

$$f(x_1,x_2) = \int_{x_0} f(x_0,x_1,x_2)dx_0$$

$$= p_1p_2\int\lambda_0 e^{-\lambda_0x_0}\delta_{(x_1-a_1x_0)}\delta_{(x_2-a_2x_0)}dx_0$$

$$+p_1(1-p_2)\int\lambda_0 e^{-\lambda_0x_0} f_{X_2}(x_2-a_2x_0)I_{(x_2>a_2x_0)}\delta_{(x_1-a_1x_0)}dx_0$$

$$+(1-p_1)p_2\int\lambda_0 e^{-\lambda_0x_0} f_{X_1}(x_1-a_1x_0)I_{(x_1>a_1x_0)}\delta_{(x_2-a_2x_0)}dx_0$$

$$+(1-p_1)(1-p_2)\int\lambda_0 e^{-\lambda_0x_0} f_{X_1}(x_1-a_1x_0)f_{X_2}(x_2-a_2x_0)I_{(x_1a_1>x_0,x_2a_2>x_0)}dx_0$$

$$= p_1 p_2 PartA_1 + p_1(1-p_2)PartA_2$$

$$+(1-p_1)p_2 PartA_3 + (1-p_1)(1-p_2)PartA_4,$$

where $PartA_1 = \int \lambda_0 e^{-\lambda_0 x_0} \delta_{(x_1 - a_1 x_0)} \delta_{(x_2 - a_2 x_0)} dx_0$

$$= 1a_1 a_2 \lambda_0 \int e^{-\lambda_0 x_0} \delta_{(x_1 a_1 - x_0)} \delta_{(x_2 a_2 - x_0)} dx_0$$

$$= \begin{cases} 1a_1 a_2 \lambda_0 e^{-\lambda_0 x_2 a_2} \delta_{(x_1 a_1 - x_2 a_2)} & ; \\ or & \\ 1a_1 a_2 \lambda_0 e^{-\lambda_0 x_1 a_1} \delta_{(x_1 a_1 - x_2 a_2)} & . \end{cases}$$

$$PartA_2 = \int \lambda_0 e^{-\lambda_0 x_0} f_{X_2}(x_2 - a_2 x_0) I_{(x_2 > a_2 x_0)} \delta_{(x_1 - a_1 x_0)} dx_0$$

$$= 1a_1 \int \lambda_0 e^{-\lambda_0 x_0} f_{X_2}(x_2 - a_2 x_0) I_{(x_2 > a_2 x_0)} \delta_{(x_1 a_1 - x_0)} dx_0$$

$$= 1a_1 \lambda_0 e^{-\lambda_0 x_1 a_1} f_{X_2}(x_2 - a_2 x_1 a_1) I_{(x_2 > a_2 x_1 a_1)}$$

$$= 1a_1 \lambda_0 \lambda_2 e^{-\lambda_0 x_1 a_1} e^{-\lambda_2 (x_2 - a_2 x_1 a_1)} I_{(x_2 > a_2 x_1 a_1)}$$

$$= 1a_1 \lambda_0 \lambda_2 e^{-x_1 a_1 (\lambda_0 - a_2 \lambda_2)} e^{-\lambda_2 x_2} I_{(x_2 > a_2 x_1 a_1)}$$

$$= 1a_1 \lambda_0 \lambda_2 e^{-x_1 a_1 (\lambda_0 - a_2 \lambda_2)} e^{-\lambda_2 x_2} I_{(x_2 a_2 > x_1 a_1)}$$

$$= 1a_1 \lambda_0 \lambda_2 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-x_1 a_1 (\lambda_0 - a_1 \lambda_1 - a_2 \lambda_2)} I_{(x_2 a_2 > x_1 a_1)}$$

$$= 1a_1 \lambda_0 \lambda_2 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-\lambda^* \varphi} I_{(x_2 a_2 > x_1 a_1)},$$

where $\varphi = min(x_1 a_1, x_2 a_2)$ *and* $\lambda^* = \lambda_0 - a_1 \lambda_1 - a_2 \lambda_2$.

Similarly, $PartA_3 = \int \lambda_0 e^{-\lambda_0 x_0} f_{X_1}(x_1 - a_1 x_0) I_{(x_1 > a_1 x_0)} \delta_{(x_2 - a_2 x_0)} dx_0$

$$= 1a_2 \int \lambda_0 e^{-\lambda_0 x_0} f_{X_1}(x_1 - a_1 x_0) I_{(x_1 > a_1 x_0)} \delta_{(x_2 a_2 - x_0)} dx_0$$

$$= 1a_2 \lambda_0 e^{-\lambda_0 x_2 a_2} f_{X_1}(x_1 - a_1 x_2 a_2) I_{(x_1 > a_1 x_2 a_2)}$$

$$= 1a_2 \lambda_0 \lambda_1 e^{-\lambda_0 x_2 a_2} e^{-\lambda_1 (x_1 - a_1 x_2 a_2)} I_{(x_1 > a_1 x_2 a_2)}$$

$$= 1a_2 \lambda_0 \lambda_1 e^{-x_2 a_2 (\lambda_0 - a_1 \lambda_1)} e^{-\lambda_1 x_1} I_{(x_1 > a_1 x_2 a_2)}$$

$$= 1a_2 \lambda_0 \lambda_1 e^{-x_2 a_2 (\lambda_0 - a_1 \lambda_1)} e^{-\lambda_1 x_1} I_{(x_1 a_1 > x_2 a_2)}$$

$$= 1a_2\lambda_0\lambda_1 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-x_2 a_2(\lambda_0 - a_1\lambda_1 - a_2\lambda_2)} I_{(x_1 a_1 > x_2 a_2)}$$

$$= 1a_2\lambda_0\lambda_1 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-\lambda^* \varphi} I_{(x_1 a_1 > x_2 a_2)},$$

$$\text{and} \quad PartA_4 = \int \lambda_0 e^{-\lambda_0 x_0} f_{X_1}(x_1 - a_1 x_0) f_{X_2}(x_2 - a_2 x_0) I_{(x_1 > a_1 x_0, x_2 > a_2 x_0)} dx_0$$

$$= \int_0^{\varphi} \lambda_0\lambda_1\lambda_2 e^{-\lambda_0 x_0} e^{-\lambda_1(x_1 - a_1 x_0)} e^{-\lambda_2(x_2 - a_2 x_0)} dx_0$$

$$= \int_0^{\varphi} \lambda_0\lambda_1\lambda_2 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-(\lambda_0 - a_1\lambda_1 - a_2\lambda_2)x_0} dx_0$$

$$= \lambda_0\lambda_1\lambda_2 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} \int_0^{\varphi} e^{-\lambda^* x_0} dx_0$$

$$= \lambda_0\lambda_1\lambda_2\lambda^* e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} (1 - e^{-\lambda^* \varphi}).$$

Hence the expression for the joint density becomes:

$$f(x_1, x_2) = p_1 p_2 \lambda_0 a_1 a_2 e^{-\lambda_0 x_2 a_2} \delta_{(x_1 a_1 - x_2 a_2)}$$

$$+ p_1(1 - p_2) 1 a_1 \lambda_0 \lambda_2 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-\lambda^* \varphi} I_{(x_2 a_2 > x_1 a_1)}$$

$$+ (1 - p_1) p_2 1 a_2 \lambda_0 \lambda_1 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-\lambda^* \varphi} I_{(x_1 a_1 > x_2 a_2)}$$

$$+ (1 - p_1)(1 - p_2) \lambda_0\lambda_1\lambda_2\lambda^* e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} (1 - e^{-\lambda^* \varphi}),$$

$$\text{where} \quad \varphi = min(x_1 a_1, x_2 a_2) \quad \text{and} \quad \lambda^* = \lambda_0 - a_1\lambda_1 - a_2\lambda_2. \tag{13}$$

Based on a random sample of size n denoted $(x_{11}, x_{21}), (x_{12}, x_{22}), \cdots, (x_{1n}, x_{2n})$, let's define

$$r_j^{(1)} = \begin{cases} 1, & if \ x_{0j} = x_{1j}\ a_1 \le x_{2j}\ a_2\ ; \\ 0, & if \ x_{1j}\ a_1 > x_{2j}\ a_2 \end{cases} \quad \text{and} \quad r_j^{(2)} = \begin{cases} 1, & if \ x_{0j} = x_{2j}\ a_2 \le x_{1j}\ a_1\ ; \\ 0, & if \ x_{2j}\ a_2 > x_{1j}\ a_1\ . \end{cases}$$

Then

$$L(\lambda_0, \lambda_1, \lambda_2) = \prod_{j=1}^{n} f(x_{1j}, x_{2j})$$

$$= \prod_{j=1}^{n} [p_1 p_2 \lambda_0 a_1 a_2 e^{-\lambda_0 x_2 a_2}]^{r_j^{(1)} r_j^{(2)}}$$

$$[p_1(1 - p_2) 1 a_1 \lambda_0 \lambda_2 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-\lambda^* \varphi}]^{r_j^{(1)}(1 - r_j^{(2)})}$$

$$[(1 - p_1) p_2 1 a_2 \lambda_0 \lambda_1 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-\lambda^* \varphi}]^{(1 - r_j^{(1)}) r_j^{(2)}}$$

$$[(1 - p_1)(1 - p_2) \lambda_0\lambda_1\lambda_2\lambda^* e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} (1 - e^{-\lambda^* \varphi})]^{(1 - r_j^{(1)})(1 - r_j^{(2)})}.$$

In order to obtain estimators, the log likelihood is:

$$l(\lambda_0,\lambda_1,\lambda_2) = \log L(\lambda_0,\lambda_1,\lambda_2)$$

$$= \sum_i r_i^{(1)} r_i^{(2)} [\log(p_1 p_2) + \log(\lambda_0) - \log(a_1 a_2) - \lambda_0 \varphi_i]$$

$$+ r_i^{(1)}(1-r_i^{(2)})[\log p_1(1-p_2) + \log\lambda_0 + \log\lambda_2 - \log a_1 - \lambda_1 x_{1i} - \lambda_2 x_{2i} - \lambda^* \varphi_i]$$

$$+ (1-r_i^{(1)}) r_i^{(2)} [\log(1-p_1) p_2 + \log\lambda_0 + \log\lambda_1 - \log a_2 - \lambda_1 x_{1i} - \lambda_2 x_{2i} - \lambda^* \varphi_i]$$

$$+ (1-r_i^{(1)})(1-r_i^{(2)})[\log(1-p_1)(1-p_2) + \log\lambda_0 + \log\lambda_1\lambda_2 - \log\lambda^*$$

$$-\lambda_1 x_{1i} - \lambda_2 x_{2i} + \log(1-e^{-\lambda^* \varphi_i})].$$

Then

$$\partial l \partial \lambda_0 = \sum_i r_i^{(1)} r_i^{(2)} [1\lambda_0 - \varphi_i] + r_i^{(1)}(1-r_i^{(2)})[1\lambda_0 - \varphi_i]$$

$$+ (1-r_i^{(1)}) r_i^{(2)} [1\lambda_0 - \varphi_i] + (1-r_i^{(1)})(1-r_i^{(2)})[1\lambda_0 - 1\lambda^* - \varphi e^{-\lambda^* \varphi_i} 1 - e^{-\lambda^* \varphi_i}]$$

$$= \sum_i [1\lambda_0 - \varphi_i (r_i^{(1)} + r_i^{(2)} - r_i^{(1)} r_i^{(2)})$$

$$-(1-r_i^{(1)})(1-r_i^{(2)})(1\lambda^* + \varphi_i e^{-\lambda^* \varphi_i} 1 - e^{-\lambda^* \varphi_i})]$$

$$= n\lambda_0 - \sum_i \varphi_i (r_i^{(1)} + r_i^{(2)} - r_i^{(1)} r_i^{(2)})$$

$$-\sum_i (1-r_i^{(1)})(1-r_i^{(2)})(1\lambda^* + \varphi_i e^{-\lambda^* \varphi_i} 1 - e^{-\lambda^* \varphi_i}).$$

$$\textit{Similarly,} \quad \partial l \partial \lambda_1 = \sum_i r_i^{(1)}(1-r_i^2)[-x_{1i} + a_1 \varphi_i] + (1-r_i^{(1)}) r_i^{(2)} [1\lambda_1 - x_{1i} + a_1 \varphi_i]$$

$$+ (1-r_i^{(1)})(1-r_i^{(2)})[1\lambda_1 + a_1 \lambda^* - x_{1i} + a_i \varphi_i e^{-\lambda^* \varphi_i} 1 - e^{-\lambda^* \varphi_i}]$$

$$= \sum_i -x_{1i}(1-r_i^{(1)} r_i^{(2)}) + a_1 \varphi_i [r_i^{(1)}(1-r_i^{(2)}) + (1-r_i^{(1)}) r_i^{(2)}]$$

$$+ 1\lambda_1 (1-r_i^{(1)}) + a_1 (1-r_i^{(1)})(1-r_i^{(2)})(1\lambda^* + \varphi_i e^{-\lambda^* \varphi_i} 1 - e^{-\lambda^* \varphi_i}).$$

So setting $\partial l \partial \lambda_1 = 0$ and $\partial l \partial \lambda_0 = 0$ gives

$$a_1 \sum_i (1-r_i^{(1)})(1-r_i^{(2)})(1\lambda^* + \varphi_i e^{-\lambda^* \varphi_i} 1 - e^{-\lambda^* \varphi_i}) =$$

$$\sum_i x_{1i}(1-r_i^{(1)} r_i^{(2)}) - a_1 \varphi_i [r_i^{(1)}(1-r_i^{(2)}) + (1-r_i^{(1)}) r_i^{(2)}] - 1\lambda_1 (1-r_i^{(1)}),$$

$$\text{and hence} \quad n\lambda_0 = 1a_1\sum_i x_{1i}(1-r_i^{(1)}r_i^{(2)}) + \sum_i \varphi_i r_i^{(1)} r_i^{(2)} - 1a_1\lambda_1\sum_i(1-r_i^{(1)}),$$

$$\text{or} \quad 1\lambda_0 = 1a_1\sum_i x_{1i}(1-r_i^{(1)}r_i^{(2)})n + \sum_i \varphi_i r_i^{(1)} r_i^{(2)} n - 1a_1\lambda_1\sum_i(1-r_i^{(1)})n.$$

A similar formula can be obtained by taking $\partial l \partial \lambda_2$ and setting it equal to zero.

$$\partial l \partial \lambda_2 = \sum_i r_i^{(1)}(1-r_i^2)[1\lambda_2 - x_{2i} + a_2\varphi_i] + (1-r_i^{(1)})r_i^{(2)}[-x_{2i} + a_2\varphi_i]$$

$$+(1-r_i^{(1)})(1-r_i^{(2)})[1\lambda_2 + a_2\lambda^* - x_{2i} + a_2\varphi_i e^{-\lambda^*\varphi_i} 1 - e^{-\lambda^*\varphi_i}]$$

$$= \sum_i - x_{1i}(1-r_i^{(1)}r_i^{(2)}) + a_2\varphi_i[r_i^{(1)}(1-r_i^{(2)}) + (1-r_i^{(1)})r_i^{(2)}]$$

$$+1\lambda_2(1-r_i^{(2)}) + a_2(1-r_i^{(1)})(1-r_i^{(2)})(1\lambda^* + \varphi_i e^{-\lambda^*\varphi_i} 1 - e^{-\lambda^*\varphi_i}).$$

$$\text{So,} \quad a_2\sum_i(1-r_i^{(1)})(1-r_i^{(2)})(1\lambda^* + \varphi_i e^{-\lambda^*\varphi_i} 1 - e^{-\lambda^*\varphi_i}) =$$

$$\sum_i x_{2i}(1-r_i^{(1)}r_i^{(2)}) - a_2\varphi_i[r_i^{(1)}(1-r_i^{(2)}) + (1-r_i^{(1)})r_i^{(2)}] - 1\lambda_2(1-r_i^{(2)})$$

$$\text{and hence,} \quad n\lambda_0 = 1a_2\sum_i x_{2i}(1-r_i^{(1)}r_i^{(2)}) + \sum_i \varphi_i r_i^{(1)} r_i^{(2)} - 1a_2\lambda_2\sum_i(1-r_i^{(2)}),$$

$$\text{or} \quad 1\lambda_0 = 1a_2\sum_i x_{2i}(1-r_i^{(1)}r_i^{(2)})n + \sum_i \varphi_i r_i^{(1)} r_i^{(2)} n - 1a_2\lambda_2\sum_i(1-r_i^{(2)})n.$$