Applied Stochastic Models in Business and Industry

Research Article

Received 21 June 2010, Revised 26 September 2011, Accepted 28 September 2011 Published online 22 February 2012 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/asmb.935

Some properties of the bivariate lognormal distribution for reliability applications

Pushpa L. Gupta and Ramesh C. Gupta*†

In this paper, we study the bivariate lognormal distribution from a reliability point of view. The conditional distribution of X given $Y > y$ is found to be log-skew normal. The monotonicity of the hazard rates of the univariate as well as the conditional distributions is discussed. Clayton's association measure is obtained in terms of the hazard gradient, and its value in the case of our model is derived. The probability distributions, in the case of series and parallel systems, are derived, and the monotonicity of their failure rates is discussed. Three real applications of the bivariate lognormal distribution are provided, two from financial economics and one from reliability. Copyright © 2012 John Wiley & Sons, Ltd.

Keywords: association measure; failure rate; hazard components; log-skew normal distribution; monotonicity; series and parallel systems

1. Introduction

It has been observed that, in many practical problems, the lognormal distribution is widely used to model a positive random variable that exhibits skewness in disciplines ranging from economics to medicine. In order to test the equality of means of healthcare costs in a paired design, Zhou *et al.* [1] showed that a bivariate lognormal distribution is appropriate to model the outpatient costs, 6 months before and after the Medicaid policy change in the state of Indiana; they presented five tests for the equality of two lognormal means. Hawkins [2] discussed an example dealing with 56 assay pairs of cyclosporin from blood samples of organ transplant recipients obtained by two different methods.

In this paper, we are interested in studying the class of bivariate lognormal distributions from a reliability point of view. More specifically, we study the association between the variables and obtain conditions for which this class of distributions is totally positive of order 2 (TP_2) or reverse rule of order 2 (RR_2). This enables us to study the dependence properties of the reliability model. We also study components of the hazard gradient in the sense of Johnson and Kotz [3]. For instance, the conditional distribution of X given $Y > y$ is found to be log-skew normal, a class not widely studied in the literature. An association measure $\theta(x, y)$ defined by Oakes [4] is investigated for this class of bivariate distributions. Some of the results presented here are general and would be useful in studying the association in other classes of bivariate distributions as well.

The organization of this paper is as follows. In Section 2, we present some general results for bivariate distributions. Also, we give some definitions and background of reliability functions. Univariate lognormal and log-skew normal distributions and the monotonicity of their failure rates are also studied in this section. The monotonicity of the failure rates of the conditional distributions of the bivariate lognormal distribution is discussed in Section 3. We investigate the relationship between some dependence notions in reliability and an association measure Clayton [5] and evaluate the measure for our model. Section 4 discusses the distributions relating to the series and parallel systems and investigates the monotonicity of their failure rates. Three real applications of the bivariate lognormal distribution are provided in Section 5. Two of the applications deal with financial economics and one deals with the reliability of a series system.

2. Review and examples of concepts in reliability

A random variable (X, Y) is said to have a bivariate lognormal distribution if $(\ln X, \ln Y)$ has a bivariate normal distribution. It is clear that the marginal and conditional distributions are univariate lognormal.

Department of Mathematics and Statistics, University of Maine, Orono, ME 04469-5752, USA

**Correspondence to: Ramesh C. Gupta, Department of Mathematics and Statistics, University of Maine, Orono, ME 04469-5752, USA.*

†E-mail: rcgupta@maine.edu

In order to study the dependence between the two variables X and Y, we define a local dependence function

$$\gamma_f(x,y) = \frac{\partial^2}{\partial x \partial y} \ln f(x,y),$$

where $f(x,y)$ is the joint PDF of (X,Y) (see Holland and Wang [6] for the definition and the properties of the local dependence function). One important property of the local dependence function is the TP_2 (RR_2) property defined as follows.

Definition 2.1
A function $f(x,y)$ is said to be TP_2 or RR_2 if

$$f(x_1,y_1)f(x_2,y_2) \geqslant (\leqslant) f(x_1,y_2)f(x_2,y_1), x_1 < x_2, y_1 < y_2.$$

The following theorem ties the local dependence function and the TP_2 (RR_2) property.

Theorem 2.2
The density of (X,Y) is TP_2 or RR_2 according to the local dependence function $\gamma_f(x,y) > (<) 0$.

Proof
See Theorem 7.1 of Holland and Wang [6]. □

For the bivariate normal distribution, it is well known that the joint density is TP_2 according to $\rho > (<) 0$ (see for example Joe [7]).

We now state the following result.

Theorem 2.3
Suppose the random variable (X,Y) has the TP_2 (RR_2) property. Let $h(.)$ be an increasing function. Then, $(h(X), h(Y))$ has the TP_2 (RR_2) property.

Proof
See Shaked [8]. □

Using the above result, we can conclude that the bivariate lognormal distribution has the TP_2 (RR_2) property according to $\rho > (<) 0$, where ρ is the correlation coefficient between $\ln X$ and $\ln Y$.

2.1. Some definitions and background of reliability functions

Let T be a nonnegative random variable denoting the life length of a component having distribution function $F(t)$ with $F(0) = 0$ and the PDF $f(t)$. Then, the failure rate of T is given by $r(t) = f(t)/R(t)$, where $R(t) = 1 - F(t)$ is the survival (reliability) function of T. We also assume that $f(t)$ is continuous and twice differentiable on $(0,\infty)$.

Let $h : R_+ \to R_+$ be a real-valued differentiable function. Then, $h(t)$ is said to be

(1) Increasing if $h'(t) > 0$ for all t and is denoted by I.
(2) Decreasing if $h'(t) < 0$ for all t and is denoted by D.
(3) Bathtub shaped if $h'(t) < 0$ for $t \in (0, t_0)$, $h'(t_0) = 0$, $h'(t) > 0$ for $t > t_0$ and is denoted by B.
(4) Upside down bathtub shaped if $h'(t) > 0$ for $t \in (0, t_0)$, $h'(t_0) = 0$, $h'(t) < 0$ for $t > t_0$ and is denoted by U.

For definitions, see Gupta and Warren [9] and Barlow and Proschan [10]. Also, see Barlow *et al.* [11] for some properties of probability distributions with a monotone hazard rate.

In order to determine the monotonicity of the failure rates, we proceed as follows.

Define

$$\eta(t) = -f'(t)/f(t). \tag{2.1}$$

The shape of $\eta(t)$ (I, D, B, etc.) often determines the shape of the failure rate. The relation between $r(t)$ and $\eta(t)$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} \ln r(t) = r(t) - \eta(t). \tag{2.2}$$

For more relations between $\eta(t)$ and $r(t)$, see Marshall and Olkin [12].

In order to determine the monotonicity of $r(t)$, we present a modification, due to Marshall and Olkin [12], of Glaser [14], which helps us determine the shape of the failure rates of the four types described above.

Theorem 2.4
Let f be a density strictly positive and differentiable on $(0,\infty)$ such that $\lim_{x\to\infty} f(x)=0$. Then,

(a) If $\eta(t)\in I$, then $r(t)\in I$ (increasing failure rate).
(b) If $\eta(t)\in D$, then $r(t)\in D$ (decreasing failure rate).
(c) If $\eta(t)\in B$, then $r(t)\in B$ (bathtub-shaped failure rate).
(d) If $\eta(t)\in U$, then $r(t)\in U$ (upside-bathtub-shaped failure rate).

Proof
See Marshall and Olkin [12]. □

2.2. Lognormal and log-skew normal distributions

2.2.1. Lognormal distribution. Suppose V has a lognormal distribution so that $\ln V \sim N(\mu,\sigma^2)$. The PDF and the reliability functions of V are given by

$$f_V(t)=\frac{1}{t\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(\ln t-\mu)^2}, t>0, \tag{2.3}$$

$$\overline{F}_V(t)=1-\Phi\left[\frac{\ln at}{\sigma}\right], \tag{2.4}$$

where $a=e^{-\mu}$ and $\Phi(.)$ is the cumulative distribution function of a standard normal distribution.

The failure rate is given by

$$r_V(t)=\frac{\left(1/\sqrt{2\pi}t\sigma\right)e^{-(\ln t-\mu)^2/2\sigma^2}}{\left[1-\Phi\left[\frac{\ln at}{\sigma}\right]\right]}. \tag{2.5}$$

The mean residual life is given by

$$\mu_V(t)=e^{\mu+\sigma^2/2}\frac{[1-\Phi((\ln t-\mu-\sigma^2)/\sigma)]}{1-\Phi((\ln t-\mu)/\sigma))}. \tag{2.6}$$

For details, see Gupta *et al.* [13].

By using Glaser's [14] method and its modification, described earlier, we have observed that $r(t)$ is of the type U. This means that it increases initially to a maximum and then decreases to 0 as time approaches infinity. On the other hand, the mean residual life function is of the type B. This means that it decreases to a minimum and then steadily increases.

2.2.2. Log-skew normal distribution and the monotonicity of its failure rate. Suppose Y has a skew normal density given by

$$f_Y(y)=\frac{(1/\sigma)\phi\left((y-\mu)/\sigma)\right)\Phi\left[\lambda_0+\lambda_1((y-\mu)/\sigma)\right]}{\Phi(\lambda_0/\sqrt{1+\lambda_1^2})}, y\in R, \lambda_0,\lambda_1\in R, \tag{2.7}$$

where $\phi(.)$ and $\Phi(.)$ are the PDF and the cumulative distribution function, respectively, of a standard normal variable Arnold [15].

Let $Y=\ln X$. Then, the PDF of X, called the log-skew normal distribution, is given by

$$f_X(x)=\frac{(1/\kappa v\sigma)\phi((\ln x-\mu)/\sigma))\,\Phi(\lambda_0+\lambda_1(\ln x-\mu)/\sigma)}{\Phi(\lambda_0/\sqrt{1+\lambda_1^2})}. \tag{2.8}$$

Note that a variant of the log-skew normal distribution given by Equation (2.8) has been studied by Azzalini *et al.* [16]. Recently, Lin and Stoyanov [17] have studied the problem of moments for this class of distributions, showing that they are all moment indeterminate. Chai and Bailey [18] have demonstrated the potential use of this model through an actual data set involving coronary artery classification scores from an atherogenesis study.

In order to study the monotonicity of its failure rate, we form Glaser's [14] eta function

$$\begin{aligned}\eta_X(x) &= -\frac{f_X'(x)}{f_X(x)} = -\frac{f_{\mathrm{N_L}}'(x)}{f_{\mathrm{N_L}}(x)} - \frac{\lambda_1}{x}\frac{\phi(-\lambda_0-\lambda_1\ln x)}{[1-\Phi(-\lambda_0-\lambda_1\ln x)]} \\ &= \eta_{\mathrm{N_L}}(x) - \frac{\lambda_1}{x} r_{\mathrm{N}}(-\lambda_0-\lambda_1\ln x),\end{aligned} \tag{2.9}$$

where $f_{\mathrm{N_L}}(x)$ is the density function of a lognormal distribution and $r_{\mathrm{N}}(.)$ is the hazard rate of a standard normal distribution. This gives

$$\eta_X'(x) = \eta_{\mathrm{N_L}}'(x) + \frac{\lambda_1}{x^2}\left[r_{\mathrm{N}}(-\lambda_0-\lambda_1\ln x) + \lambda_1 r_{\mathrm{N}}'(-\lambda_0-\lambda_1\ln x)\right]. \tag{2.10}$$

Because $\eta_{\mathrm{N_L}}(x)$ is of the type U and $\lim_{x\to\infty} f(x)=0$, we conclude that $\eta_X(x)$ is also of the type U. With the use of Glaser's modification, the failure rate of X is of the type U.

3. Bivariate lognormal distribution-reliability properties and dependence

A bivariate random vector (X,Y) is said to have a bivariate lognormal distribution if $(\ln X, \ln Y)$ has a bivariate normal distribution. It is evident that the conditional distribution of X given $Y=y$ or of Y given $X=x$ is univariate lognormal. Therefore, their reliability functions can be easily obtained by using formulas (2.4)–(2.6) and appropriate values of the parameters. The monotonicity of the failure rate and of the mean residual life function will follow the same pattern as of the univariate lognormal distribution.

To study the monotonicity of the failure rate of the conditional distribution of X given $Y=y$ as a function of y, we employ the following result due to Shaked [8].

Lemma 3.1
If $f(x,y)$ is TP_2 (RR_2), the conditional failure rate of X given $Y=y$ is decreasing (increasing) in y.

Recalling Theorems 2.2 and 2.3 and using the above result, we conclude that the failure rate of the conditional distribution of X given $Y=y$ is decreasing (increasing) in y according to $\rho > (<)\ 0$.

We now discuss the conditional distribution of X given $Y>y$.

Conditional distribution of $X|Y>y$

Suppose that $(X,Y)\sim BVLN(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$. Then, the conditional density of X given $Y>y$ is given by

$$\begin{aligned}f_{X|Y>y}(x|Y>y) &= -\frac{\partial}{\partial x}P(X>x|Y>y) \\ &= \int_y^\infty f(x,v)\mathrm{d}v/P(Y>y) \\ &= \frac{\int_y^\infty \frac{1}{2\pi\sigma_1\sigma_2 x v\sqrt{1-\rho^2}}\exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{\ln x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{\ln x-\mu_1}{\sigma_1}\right)\left(\frac{\ln v-\mu_2}{\sigma_2}\right) + \left(\frac{\ln v-\mu_2}{\sigma_2}\right)^2\right]\right\}\mathrm{d}v}{P(Y>y)} \\ &= \frac{\frac{1}{x\sigma_1}\phi\left(\frac{\ln x-\mu_1}{\sigma_1}\right)\left[1-\Phi\left(\frac{\ln y-[\mu_2+(\rho\sigma_2/\sigma_1)(\ln x-\mu_1)]}{\sigma_2\sqrt{1-\rho^2}}\right)\right]}{1-\Phi\left(\frac{\ln y-\mu_2}{\sigma_2}\right)}.\end{aligned}$$

This is of the same form as Equation (2.8) with $\lambda_0 = -\left(\ln y-\mu_2/\sigma_2\right)/\sqrt{1-\rho^2}$, $\lambda_1=\rho/\sqrt{1-\rho^2}$.

Special case

When $\mu_1=\mu_2=0$ and $\sigma_1^2=\sigma_2^2=1$, the above reduces to

$$f_{X|Y>y}(x|Y>y) = f_X(x)\left[1-\Phi\left(\frac{\ln y-\rho\ln x}{\sqrt{1-\rho^2}}\right)\right]/P(Y>y), \tag{3.1}$$

where $f_X(x)$ is the marginal density function of a standard lognormal distribution. Notice that the above density is of the same form as Equation (2.8). Thus, X given $Y > y$ has a log-skew normal distribution. Note that this result is true for standard as well as for nonstandard bivariate lognormal vectors, see Arnold [15].

Hazard components and their monotonicity

For the study of hazard components and their monotonicity, it is enough to consider the standard bivariate lognormal model.

The hazard rate of the conditional distribution of X given $Y > y$ is given by

$$\begin{aligned} h_1(x,y) &= -\frac{\mathrm{d}}{\mathrm{d}x}\ln \overline{F}_\rho(x,y) \\ &= \int_y^\infty f(x,v)\mathrm{d}v / \overline{F}(x,y) \\ &= f_X(x)\left[1-\Phi\left(\frac{\ln y-\rho\ln x}{\sqrt{1-\rho^2}}\right)\right] / \overline{F}_\rho(x,y), \end{aligned} \tag{3.2}$$

using equation Equation (3.1). Because the distribution of X given $Y > y$ is log-skew normal, its hazard rate is of the type U.

In a similar manner, we can obtain the conditional density of Y given $X > x$ and the hazard component $h_2(x,y)$.

The monotonicity of the conditional hazard of X given $Y > y$ as a function of y can be determined by employing the following result due to Shaked [8].

Lemma 3.2

If $f(x,y)$ is TP_2 (RR_2), the conditional failure rate of X given $Y > y$ is decreasing (increasing) in y.

As noticed before, the bivariate lognormal distribution is TP_2 (RR_2) according to $\rho > (<)\ 0$. Thus, in the case of a bivariate lognormal distribution, $h_1(x,y)$ is decreasing (increasing) in y according to $\rho > (<)\ 0$. A similar statement can be made regarding the other hazard component.

3.1. Association measures

In the context of bivariate survival models induced by frailties, Oakes [4] studied the following association measure:

$$\theta(x,y) = \frac{SS_{12}}{S_1S_2},$$

where $S = S(x,y)$ is the survival function, $S_{12} = \partial^2 S(x,y)/\partial x\partial y$, $S_1 = (\partial/\partial x)S(x,y)$, and $S_2 = (\partial/\partial y)S(x,y)$ (see also Clayton [5]).

Clayton [5] presented the above association measure, deriving from the Cox model, in a study of the association between the life spans of fathers and their sons.

It can be easily seen that

$$\theta(x,y) = \frac{r(x|Y=y)}{h_1(x,y)}.$$

The numerator is the hazard rate for sons at time x given that their fathers died at y. The denominator is the hazard rate for sons at time x given that their fathers live past y. Also,

$$r(x|Y=y) = -S_{12}/S_2 \quad \text{and} \quad h_1(x,y) = -S_1/S.$$

It can now be verified that

$$\begin{aligned} \frac{\partial}{\partial y}h_1(x,y) &= \frac{S_2}{S}\left[-h_1(x,y) + r(x|Y=y)\right] \\ &= \frac{S_2}{S}h_1(x,y)(\theta-1), \end{aligned}$$

suppressing the argument of θ.

Because $S_2 < 0$, $\theta > (<)\ 1$ is equivalent to $(\partial/\partial y)h_1(x,y) < (>)\ 0$.

Thus, in the case of a bivariate lognormal distribution, $\theta > (<)\ 1$ according to $\rho > (<)\ 0$.

3.1.1. Effect of the association measure. The deviation of the ratio of these hazards from 1 characterizes the measure of mutual dependence of the respective life spans. The stronger the dependence between X and Y is, the higher is the value of $\theta(x, y)$. If $\theta(x, y)$ decreases to 1 when x and y tend to $+\infty$, then the dependence of the pair (X, Y) is at least $DTP(0, 1)$ or $DTP(1, 0)$. This means that the conditional hazard of Y given $X = x$ and the conditional hazard of X given $Y = y$ decrease in x and y, respectively (see Shaked [8] for the definitions of $DTP(0, 1)$ and $DTP(1, 0)$).

Clayton [5] proposed this measure by assuming that an association arises because the two members of a pair share some common influence and not because one event influences the other. Thus, $\theta(x, y)$ explains an association between two non-negative survival times with continuous joint distribution by their common dependence on an unobserved random variable. This unobserved random variable is commonly known as frailty or environmental effect (see Oakes [4] and Manatunga and Oakes [19] for more details on frailty models).

Clayton [5] described the estimation of the parameter $\theta(x, y)$ from longitudinal studies. He also provided a numerical example in case–control studies. The estimation of $\theta(x, y)$ in a discrete form of the model is considered by Oakes [4] even in the presence of censoring in either or both components, and it is shown that these estimates can be used to test the independence of the two variables. He uses the data of Hanley and Parnes [20] to illustrate this methodology.

3.1.2. Derivation of $\theta(x, y)$. Using the definitions given in the previous section, we can verify that

$$\begin{aligned}\frac{\partial^2}{\partial x \partial y} \ln S(x, y) &= \frac{S_1 S_2}{S^2}(\theta - 1) \\ &= h_1(x, y) h_2(x, y)(\theta - 1),\end{aligned}$$

where

$$h_2(x, y) = -\frac{\partial}{\partial y} \ln S(x, y).$$

This gives

$$\theta = 1 - \frac{\frac{\partial}{\partial y} h_1(x, y)}{h_1(x, y) h_2(x, y)}.$$

By symmetry, we also have

$$\theta(x, y) = 1 - \frac{\frac{\partial}{\partial x} h_2(x, y)}{h_1(x, y) h_2(x, y)}.$$

Remark 3.1
It can be proven that X and Y are independent if and only if $\theta = 1$.

Using the values of $h_1(x, y)$ and $h_2(x, y)$ derived earlier, we can obtain $\theta(x, y)$ (Equation (3.2)).

4. Series and parallel systems of two components

In this section, we shall obtain the density functions of $T_1 = \min(X, Y)$ and $T_2 = \max(X, Y)$. Also, we study the monotonicity of the failure rates of T_1 and T_2.

We know that for any bivariate vector (X, Y), the density functions of T_1 and T_2 are given by the following Gupta and Gupta [21]:

$$f_{T_1}(t) = f_X(t) P(Y > t | X = t) + f_Y(t) P(X > t | Y = t) \tag{4.1}$$

and

$$f_{T_2}(t) = f_X(t) P(Y < t | X = t) + f_Y(t) P(X < t | Y = t). \tag{4.2}$$

For the bivariate lognormal distribution, with the use of Equation (4.1), it can be verified that

$$f_{T_1}(t) = \frac{1}{\sigma_1 t}\phi\left(\frac{\ln t-\mu_1}{\sigma_1}\right)\left[1-\Phi\left(\frac{\ln t\left(\frac{1}{\sigma_2}-\frac{\rho}{\sigma_1}\right)}{\sqrt{1-\rho^2}}+\frac{\frac{\mu_1\rho}{\sigma_1}-\frac{\mu_2}{\sigma_2}}{\sqrt{1-\rho^2}}\right)\right]$$
$$+\frac{1}{\sigma_2 t}\phi\left(\frac{\ln t-\mu_2}{\sigma_2}\right)\left[1-\Phi\left(\frac{\ln t\left(\frac{1}{\sigma_1}-\frac{\rho}{\sigma_2}\right)}{\sqrt{1-\rho^2}}+\frac{\frac{\mu_2\rho}{\sigma_2}-\frac{\mu_1}{\sigma 1}}{\sqrt{1-\rho^2}}\right)\right]. \tag{4.3}$$

For the standard bivariate lognormal distribution, it reduces to

$$f_{T_1}(t) = \frac{2}{t}\phi(\ln t)\left[1-\Phi\left(\ln t\sqrt{\frac{1-\rho}{1+\rho}}\right)\right]. \tag{4.4}$$

Note that Equation (4.4) is of the same form as Equation (2.8). Hence, T_1 has a log-skew normal distribution. With the use of the result established earlier, the failure rate of T_1 is of the type U.

It can be verified that the turning point of $\eta_{T_1}(t)$ is given by the solution of the equation

$$\ln t + \lambda r_{\mathrm{N}}(\lambda \ln t) = \lambda^2 r_{\mathrm{N}}'(\lambda \ln t), \tag{4.5}$$

where $\lambda = \sqrt{(1-\rho)/\,(1+\rho)}$ and $r_{\mathrm{N}}(.)$ is the failure rate of the standard normal distribution.

For example, for $\lambda = \sqrt{\pi/2}$, the maximum value of $\eta_{T_1}(t)$ is achieved at $t=1$.

Similarly, with Equation (4.2), it can be verified that

$$f_{T_2}(t) = \frac{1}{\sigma_1 t}\phi\left(\frac{\ln t-\mu_1}{\sigma_1}\right)\left[\Phi\left(\frac{\ln t(\frac{1}{\sigma_2}-\frac{\rho}{\sigma_1})}{\sqrt{1-\rho^2}}+\frac{\frac{\mu_1\rho}{\sigma_1}-\frac{\mu_2}{\sigma_2}}{\sqrt{1-\rho^2}}\right)\right.$$
$$\left.+\frac{1}{\sigma_2 t}\phi\left(\frac{\ln t-\mu_2}{\sigma_2}\right)\left[\Phi\left(\frac{\ln t\left(\frac{1}{\sigma_1}-\frac{\rho}{\sigma_2}\right)}{\sqrt{1-\rho^2}}+\frac{\frac{\mu_2\rho}{\sigma_2}-\frac{\mu_1}{\sigma 1}}{\sqrt{1-\rho^2}}\right)\right]\right]. \tag{4.6}$$

For the standard bivariate lognormal distribution, the above expression reduces to

$$f_{T_2}(t) = \frac{2}{t}\phi(\ln t)\left[\Phi\left(\ln t\sqrt{\frac{1-\rho}{1+\rho}}\right)\right]. \tag{4.7}$$

Again, T_2 has a log-skew normal distribution, and its failure rate is of the type U.

5. Some real applications

As indicated earlier, the lognormal distribution is used to approximate right-skewed data arising in a wide range of disciplines. In this section, we shall present three real examples where a bivariate lognormal distribution has been employed. The first two examples are from health economics, and the third example deals with the series and parallel system in reliability studies. The bivariate lognormal distribution has been employed in other disciplines as well, including economics, environmental science, and reliability. The references include Rappaport and Selvin [22], Crow and Shimzu [23], Krishnamoorthy *et al.* [24], Fletcher [25], and Zou *et al.* [26].

Example 5.1

In this example, we describe the use of a bivariate lognormal distribution to test the equality of means of healthcare costs in a paired design [1].

Subjects for the study were Medicaid patients in Indianapolis who visited an urban public hospital, its emerging department, or any of its network of primary care centers during two periods: January 1–June 30, 1993, and January 1–June 30, 1994. A sample of patients who had outpatient cost in both periods was chosen. The mean costs for the two periods were \$2080.43 and \$4886.37, and standard deviations were \$3927.7 and \$17,009.9, respectively. The distributions of the two cost data sets were skewed towards high costs. The Shapiro–Wilk's test for normality of the log-transformed data yielded a p-value of 0.46 and 0.6 for the two periods under study and a correlation coefficient of 0.45. In addition, the probability plots suggested that the data followed a bivariate lognormal distribution approximately.

The log-transformed data have $\mu_X^{\wedge} = 6.41$, $\mu_Y^{\wedge} = 6.50$, $S_X^2 = 2.73$, and $S_Y^2 = 3.48$, and the correlation coefficient of the log-transformed data is 0.45.

Zhou *et al.* [1] applied different tests to the above data and concluded that the hypothesis that there have been no changes in the mean hospital costs prior to and after the reimbursement policy be accepted.

More recently, Zou *et al.* [26] considered the estimation of the costs of the two periods and obtained 95% confidence intervals of (1511.15, 4266.23) and (2195.38, 7770.97), respectively. The 95% confidence interval for the cost difference is given by (−5362.49, 881.57).

Example 5.2
Between August 1998 and February 1999, a Methodist hospital in Indianapolis initiated a hospitalist program. The program was operated under a referral model where physicians, as hospitalists in the program, care for their hospitalized patients. One question in this study was whether the hospitalist program reduces the healthcare cost, compared with a historical control group that consists of patients with similar admitting diagnosis during the year before the start of the hospitalist program. For details, see Zhou *et al.* [1].

The hospitalist group has 110 patients, and the control group has 1537 patients. The distribution of the two cost data sets were skewed towards high costs. The normal quantile–quantile plots suggested approximate marginal lognormal distribution for both groups. However, Mardia's test for multivariate normality, based on measures of skewness, gives a p-value of 0.0299. Thus, the bivariate normality assumption may be questionable. Analyzing that data and taking into consideration the covariates, Zhou *et al.* [1] concluded that there was a nonsignificant difference on the means of healthcare costs between the hospital group and the matched control groups.

Example 5.3
A series structure with k components works if and only if all components work. Examples of such systems include chains, high-voltage multicell batteries, inexpensive computer systems, and inexpensive decorative tree lights using low-voltage bulbs. In a survival analysis, if a person has multiple causes of death, the death takes place because of a cause that occurs first. So we observe the minimum of the potential death times.

Meeker and Escobar [27, p. 373] considered a two-component series system with dependent components having a bivariate lognormal distribution. Assuming that the means and variances for both components are the same, they plot the reliability of the system as a function of the reliability of the individual component. Their plot shows that, when there is a positive correlation between the failure times of the individual components, the actual reliability of the system exceeds that predicted by the independent component system.

In our investigation, we are interested in the monotonicity of the hazard components relative to the monotonicity of the hazard rates of the individual components. We show that the hazard components, in the bivariate lognormal case, preserve the monotonicity of the failure rates of the individual components.

Acknowledgements

The authors are thankful to the Editor, Associate Editor, and two referees for some useful comments that enhanced the paper. P. Gupta's research was partially funded by the University of Maine's NSF ADVANCE project titled 'Increasing the Participation and Advancement of Women in Academic Science and Engineering Careers,' award # HDR-1008498.

References

1. Zhou XH, Li C, Gao S, Tierney WM. Methods for testing equality of means of health care costs in a paired design study. *Statistics in Medicine* 2001; **20**:1703–1720.
2. Hawkins DH. Dignostics of conformity of paired quantitative mesurements. *Statistics in Medicine* 2002; **21**:1913–1935.
3. Johnson NL, Kotz S. A vector valued multivariate hazard rate. *Journal of Multivariate Analysis* 1975; **5**:53–66.
4. Oakes D. Bivariate survival models induced by fralities. *Journal of the American Statistical Association* 1989; **84**:487–493.
5. Clayton DG. A model for association in bivariate life tables and its applications in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 1978; **65**:141–151.
6. Holland PW, Wang YJ. Dependence function for continuous bivariate densities. *Communications in Statistics: Theory and Methods* 1987; **16**:863–876.
7. Joe H. *Multivariate Models and Dependence Concepts*. Chapman and Hall: New York, 1997.
8. Shaked M. A family of concepts of dependence for bivariate distributions. *Journal of the American Statistical Association* 1977; **72**:642–650.
9. Gupta RC, Warren W. Determination of change point of non-monotonic failure rates. *Communications in Statistics: Theory and Methods* 2001; **30**(8–9):1903–1920.
10. Barlow RE, Proschan F. *Statistical Theory and Life Testing*. Holt, Rinehart and Winston, Inc.: New York, 1975.
11. Barlow RE, Marshall AW, Proschan F. Properties of probability distributions with monotone hazard rate. *Annals of Mathematical Statistics* 1963; **34**:375–389.
12. Marshall AW, Olkin I. *Life Distributions*. Springer Verlag: Heidelberg, 2007.

13. Gupta RC, Kannan N, Raychaudhuri A. Analysis of lognormal survival data. *Mathematical Biosciences* 1997; **139**:103–115.
14. Glaser RE. Bath tub and related failure rate characterizations. *Journal of the American Statistical Association* 1980; **75**:667–672.
15. Arnold BC. Flexible univariate and multivariate models based on hidden truncation. *Journal of Statistical Planning and Inference* 2009; **139**:3741–3749.
16. Azzalini A, Dal Cappelo T, Kotz S. Log skew normal and log skew t distributions. *Journal of Income Distribution* 2003; **11**:12–20.
17. Lin GD, Stoyanov J. The logarithmic skew-normal distributions are moment indeterminate. *Journal of Applied Probability* 2009; **46**:909–916.
18. Chai HS, Bailey KR. Use of log-skew normal distribution in analysis of continuous data with a discrete component at zero. *Statistics in Medicine* 2008; **27**:3643–3655.
19. Manatunga AK, Oakes D. A measure of association for bivariate frailty distributions. *Journal of Multivariate Analysis* 1996; **56**:60–74.
20. Hanley JA, Parnes MN. Nonparametric estimation of mutivariate distribution in the presence of censoring. *Biometrika* 1983; **39**:129–139.
21. Gupta PL, Gupta RC. Failure rate of the minimum and maximum of a multivariate normal distribution. *Metrika* 2001; **53**:39–49.
22. Rappaport SM, Selvin S. A method of evaluating the mean exposure from a log-normal distribution. *American Industrial Hygiene Journal* 1987; **48**:374–379.
23. Crow EL, Shimzu K. *The Log-normal Distributions. Theory and Applications*. Dekker: New York, 1988.
24. Krishnamoorthy K, Mathew T, Ramachandran G. Generalized p values and confidence intervals: a novel approach for analyzing lognormally distribted exposure data. *Journal of Occupational and Environmental Hygiene* 2006; **3**:642–650.
25. Fletcher D. Confidence intervals for the mean of the delta-lognormal distribution. *Environmental and Ecological Statistics* 2008; **15**:175–189.
26. Zou GY, Taleban J, Huo CY. Confidence interval estimation for log-normal data with application to health economics. *Computational Statistics and Data Analysis* 2009; **53**:3755–3764.
27. Meeker WQ, Escobar LA. *Statistical Methods of Reliability Data*. John Wiley & Sons: New York, 1998.