2006

RELIABILITY FOR SOME BIVARIATE EXPONENTIAL DISTRIBUTIONS

Saralees Nadarajah
University of Nebraska - Lincoln, saralees.nadarajah@manchester.ac.uk

Samuel Kotz
The George Washington University, kotz@gwu.edu

RELIABILITY FOR SOME BIVARIATE
EXPONENTIAL DISTRIBUTIONS

SARALEES NADARAJAH AND SAMUEL KOTZ

In the area of stress-strength models, there has been a large amount of work as regards estimation of the reliability $R = \Pr(X < Y)$. The algebraic form for $R = \Pr(X < Y)$ has been worked out for the vast majority of the well-known distributions when X and Y are independent random variables belonging to the same univariate family. In this paper, forms of R are considered when (X, Y) follow bivariate distributions with dependence between X and Y. In particular, explicit expressions for R are derived when the joint distribution is bivariate exponential. The calculations involve the use of special functions. An application of the results is also provided.

1. Introduction

Without a doubt, exponential distributions are the most popular and the most applied "life time" models in many areas, including life testing and telecommunications. In the context of reliability, the stress-strength model describes the life of a component which has a random strength Y and is subjected to random stress X. The component fails at the instant that the stress applied to it exceeds the strength, and the component will function satisfactorily whenever $Y > X$. Thus, $R = \Pr(X < Y)$ is a measure of component reliability. It has many applications especially in engineering concepts such as structures, deterioration of rocket motors, static fatigue of ceramic components, fatigue failure of aircraft structures, and the aging of concrete pressure vessels. Some examples are as follows:

(i) if X represents the maximum chamber pressure generated by ignition of a solid propellant and Y represents the strength of the rocket chamber, then R is the probability of successful firing of the engine;

(ii) if X represents the diameter of a shaft and Y represents the diameter of a bearing that is to be mounted on the shaft, then R is the probability that the bearing fits without interference;

(iii) let Y and X be the remission times of two chemicals when they are administered to two kinds of mechanical systems. Inferences about R present a comparison of the effectiveness of the two chemicals;

(iv) if X and Y are future observations on the stability of an engineering design, then R would be predictive probability that X is less than Y. Similarly, if X and Y represent lifetimes of two electronic devices, then R is the probability that one fails before the other;

(v) if Y represents the distance of a pyrotechnic igniter from its adjacent pellet and X represents its *ignition distance*, then R is the probability that the igniter succeeds to bridge the gap in the pyrotechnic chain;

(vi) a certain unit—be it a receptor in a human eye, ear, or any other organ (including sexual)—operates only if it is stimulated by source of random magnitude Y and the stimulus exceeds a lower threshold X specific for that unit. In this case, R is the probability that the unit functions;

(vii) in military warfare, R could be interpreted as the probability that a given round of ammunition will penetrate its target;

(viii) in quality control, the use of process capability indices is motivated by a desire to have an index related to the probability that an attribute (Z) of a component (size, density, elastic strength, etc.) falls within fixed specification limits. However, in some circumstances it may be desirable to have an "index" allowing for possibly varying limits—T_L or T_U, say, for lower and upper limits, respectively. One is then interested in $\Pr(T_L < Z < T_U)$. If only one of the limits is finite, then this probability reduces to calculations of $R = \Pr(X < Y)$ type.

Because of these applications, the calculation and the estimation of $R = \Pr(X < Y)$ is important for the class of bivariate exponential distributions. The calculation of R has been extensively investigated in the literature when X and Y are independent random variables belonging to the same univariate family of distributions. The algebraic form for R has been worked out for the majority of the well-known distributions, including normal, uniform, exponential, gamma, beta, extreme value, Weibull, Laplace, logistic, and the Pareto distributions (Nadarajah, [12–16]; Nadarajah and Kotz [17]). In this paper, we calculate R when X and Y are dependent random variables from six flexible families of bivariate exponential distributions (Sections 2 to 7). We also provide an application of the results to receiver operating characteristic curves (Section 8).

We will assume throughout this paper that (X, Y) has a bivariate exponential distribution with joint probability density function (pdf) f and joint survivor function $\bar{F}$. One can write

$$R = \int_0^\infty \int_x^\infty f(x, y)\, dy\, dx. \tag{1.1}$$

Our calculations of R make use of a number of special functions. They are the complementary incomplete gamma function defined by

$$\Gamma(a, x) = \int_x^\infty t^{a-1} \exp(-t)\, dt, \tag{1.2}$$

the exponential integral function defined by

$$\mathrm{Ei}(x) = -\int_{-x}^{\infty} \frac{\exp(-t)}{t}\,dt, \tag{1.3}$$

and the modified Bessel function of the first kind of order zero defined by

$$I_0(z) = \sum_{k=0}^{\infty} \frac{z^{2k}}{4^k (k!)^2}. \tag{1.4}$$

The properties of these special functions being used can be found in Prudnikov et al. [19–21] and Gradshteyn and Ryzhik [5].

2. Gumbel's bivariate exponential distribution

Gumbel's [6] bivariate exponential distribution has the joint survivor function and joint pdf specified by

$$\bar{F}(x, y) = \exp\{-(\alpha x + \beta y + \theta \alpha \beta x y)\},$$
$$f(x, y) = \{(1 - \theta)\alpha\beta + \theta\alpha^2\beta x + \theta\alpha\beta^2 y + \theta^2\alpha^2\beta^2 x y\}\bar{F}(x, y), \tag{2.1}$$

respectively, for $x > 0$, $y > 0$, $\alpha \geq 0$, $\beta \geq 0$, and $0 < \theta < 1$. This is the earliest and the simplest known bivariate exponential distribution. It has received applications in many areas, including competing risks, extreme values, failure times, regional analyses of precipitation, and reliability. The marginal distributions of X and Y are exponential with parameters α and β, respectively; so, in particular, $E(X) = 1/\alpha$ and $E(Y) = 1/\beta$. The correlation coefficient $\rho = \mathrm{Corr}(X, Y)$ is given by

$$\rho = 1 - \frac{1}{\theta}\exp\left(\frac{1}{\theta}\right)\mathrm{Ei}\left(\frac{1}{\theta}\right). \tag{2.2}$$

The correlation is, of course, zero when $\theta = 0$ (the case of independence between X and Y) and it decreases to -0.40365 as θ increases to 1. The reliability in (1.1) can be expressed as

$$R = \int_0^{\infty} \{(1 - \theta)\alpha\beta + \theta\alpha^2\beta x\} \int_x^{\infty} \bar{F}(x, y)\,dy\,dx$$
$$+ \int_0^{\infty} \{\theta\alpha\beta^2 + \theta^2\alpha^2\beta^2 x\} \int_x^{\infty} y\bar{F}(x, y)\,dy\,dx$$
$$= \int_0^{\infty} \frac{(1 - \theta)\alpha + \theta\alpha^2 x}{1 + \theta\alpha x}\bar{F}(x, x)\,dx + \int_0^{\infty} \frac{\theta\alpha + \theta^2\alpha^2 x}{(1 + \theta\alpha x)^2}\bar{F}(x, x)\,dx$$
$$= \alpha I_1 + \theta\alpha\beta I_2, \tag{2.3}$$

where

$$I_1 = \int_0^\infty \bar{F}(x,x)dx,$$
$$I_2 = \int_0^\infty x\bar{F}(x,x)dx. \tag{2.4}$$

Application of Lemma A.1 shows that one can reduce

$$I_1 = \sqrt{\frac{\pi}{\theta\alpha\beta}}\exp\left\{\frac{(\alpha+\beta)^2}{4\theta\alpha\beta}\right\}\left[1 - \Phi\left(\frac{\alpha+\beta}{\sqrt{2\theta\alpha\beta}}\right)\right],$$

$$I_2 = \frac{1}{2\theta\alpha\beta} - \frac{\sqrt{\pi}(\alpha+\beta)}{2(\theta\alpha\beta)^{3/2}}\exp\left\{\frac{(\alpha+\beta)^2}{4\theta\alpha\beta}\right\}\left[1 - \Phi\left(\frac{\alpha+\beta}{\sqrt{2\theta\alpha\beta}}\right)\right], \tag{2.5}$$

where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution. Thus, it follows using (2.3) that the form of R for Gumbel's bivariate exponential distribution is given by

$$R = \frac{1}{2} + \frac{\sqrt{\pi}(\alpha-\beta)}{2\sqrt{\theta\alpha\beta}}\exp\left\{\frac{(\alpha+\beta)^2}{4\theta\alpha\beta}\right\}\left[1 - \Phi\left(\frac{\alpha+\beta}{\sqrt{2\theta\alpha\beta}}\right)\right]. \tag{2.6}$$

Note that if $\alpha < \beta$, then $R < 1/2$; and if $\alpha > \beta$, then $R > 1/2$. Moreover, if $\alpha = \beta$, then $R = 1/2$.

3. Hougaard's bivariate exponential distribution

Hougaard's [7] bivariate exponential distribution has the joint survivor function and joint pdf specified by

$$\bar{F}(x,y) = \exp\left[-\left\{\left(\frac{x}{\theta}\right)^r + \left(\frac{y}{\phi}\right)^r\right\}^{1/r}\right],$$

$$f(x,y) = \frac{(xy)^{r-1}}{(\theta\phi)^r}\left\{\left(\frac{x}{\theta}\right)^r + \left(\frac{y}{\phi}\right)^r\right\}^{1/r-2}\left[r - 1 + \left\{\left(\frac{x}{\theta}\right)^r + \left(\frac{y}{\phi}\right)^r\right\}^{1/r}\right]\bar{F}(x,y), \tag{3.1}$$

respectively, for $x > 0$, $y > 0$, $\theta \geq 0$, $\phi \geq 0$, and $r > 0$. This distribution has been quite popular as a *frailty model*. The marginal distributions of X and Y are exponential with parameters $1/\theta$ and $1/\phi$, respectively; so, in particular $E(X) = \theta$ and $E(Y) = \phi$. The correlation coefficient $\rho = \text{Corr}(X, Y)$ is given by

$$\rho = \frac{\Gamma^2(1/r)}{r\Gamma(2/r)} - 1. \tag{3.2}$$

Note that if one transforms $(U, V) = (X/\theta, Y/\phi)$, then the joint survivor function and joint pdf of (U, V) take the simpler forms

$$\bar{F}(u,v) = \exp\left\{-(u^r + v^r)^{1/r}\right\},$$

$$f(u,v) = (uv)^{r-1}(u^r + v^r)^{1/r-2}\left\{r - 1 + (u^r + v^r)^{1/r}\right\}\bar{F}(u,v). \tag{3.3}$$

Thus, the form of R corresponding to (3.1) can be computed as

$$
\begin{aligned}
R &= \Pr(\theta U/\phi < V) \\
&= \int_0^\infty u^{r-1} \int_{\theta u/\phi}^\infty v^{r-1}(u^r + v^r)^{1/r-2}\{r - 1 + (u^r + v^r)^{1/r}\}\bar{F}(u,v)dv\,du \\
&= \int_0^\infty u^{r-1} \int_{(\theta^r + \phi^r)^{1/r}u/\phi}^\infty z^{-r}(r - 1 + z)\exp(-z)dz\,du \\
&= (r - 1)I_1 + I_2,
\end{aligned}
\tag{3.4}
$$

where the transformation $z = (u^r + v^r)^{1/r}$ has been applied, and

$$
\begin{aligned}
I_1 &= \int_0^\infty u^{r-1}\Gamma\left(1 - r, \left(\frac{\theta^r}{\phi^r} + 1\right)^{1/r} u\right)du, \\
I_2 &= \int_0^\infty u^{r-1}\Gamma\left(2 - r, \left(\frac{\theta^r}{\phi^r} + 1\right)^{1/r} u\right)du.
\end{aligned}
\tag{3.5}
$$

Application of Lemma A.2 shows that one can reduce

$$
\begin{aligned}
I_1 &= \frac{\phi^r}{r(\theta^r + \phi^r)}, \\
I_2 &= \frac{\phi^r}{r(\theta^r + \phi^r)}.
\end{aligned}
\tag{3.6}
$$

Thus, it follows from (3.4) that the form of (1.1) for Houggard's bivariate exponential distribution takes the simple form

$$
R = \frac{\phi^r}{\theta^r + \phi^r}.
\tag{3.7}
$$

Note that if $\theta = \phi$, then $R = 1/2$.

4. Downton's bivariate exponential distribution

Downton's [3] bivariate exponential distribution has the joint pdf specified by

$$
f(x,y) = \frac{\mu_1 \mu_2}{1 - \rho}\exp\left(-\frac{\mu_1 x + \mu_2 y}{1 - \rho}\right)I_0\left\{\frac{2\sqrt{\rho\mu_1\mu_2 xy}}{1 - \rho}\right\}
\tag{4.1}
$$

for $x > 0$, $y > 0$, $\mu_1 > 0$, $\mu_2 > 0$, and $0 \leq \rho < 1$. This distribution arises from "shocks" causing various types of failure to components which have geometric distributions. It has received applications in queueing systems and hydrology and has been used a model for Wold's Markov dependent processes, intensity and duration of rainfall, and height of water waves. The marginal distributions of X and Y are exponential with parameters μ_1 and μ_2, respectively; so, in particular, $E(X) = 1/\mu_1$ and $E(Y) = 1/\mu_2$. The parameter ρ is the correlation coefficient between X and Y with independence corresponding to $\rho = 0$.

Using the definition of $I_0(\cdot)$ given in **Section 1**, the corresponding form of R can be calculated as

$$
\begin{aligned}
R &= \frac{\mu_1\mu_2}{1-\rho} \sum_{k=0}^{\infty} \frac{(\rho\mu_1\mu_2)^k}{(1-\rho)^{2k}(k!)^2} \int_0^{\infty} \int_x^{\infty} x^k y^k \exp\left(-\frac{\mu_1 x + \mu_2 y}{1-\rho}\right) dy\, dx \\
&= \frac{\mu_1\mu_2}{1-\rho} \sum_{k=0}^{\infty} \frac{(\rho\mu_1\mu_2)^k}{(1-\rho)^{2k}(k!)^2} \int_0^{\infty} x^k \exp\left(-\frac{\mu_1 x}{1-\rho}\right) \int_x^{\infty} y^k \exp\left(-\frac{\mu_2 y}{1-\rho}\right) dy\, dx \\
&= \mu_1 \sum_{k=0}^{\infty} \frac{(\rho\mu_1)^k}{(1-\rho)^k k!} \int_0^{\infty} x^k \exp\left\{-\frac{(\mu_1+\mu_2)x}{1-\rho}\right\} \sum_{l=0}^{k} \frac{1}{l!}\left(\frac{\mu_2 x}{1-\rho}\right)^l dx \\
&= \mu_1 \sum_{k=0}^{\infty} \frac{(\rho\mu_1)^k}{(1-\rho)^k k!} \sum_{l=0}^{k} \frac{1}{l!}\left(\frac{\mu_2}{1-\rho}\right)^l \int_0^{\infty} x^{k+l} \exp\left\{-\frac{(\mu_1+\mu_2)x}{1-\rho}\right\} dx \\
&= \mu_1 \sum_{k=0}^{\infty} \frac{(\rho\mu_1)^k}{(1-\rho)^k k!} \sum_{l=0}^{k} \frac{1}{l!}\left(\frac{\mu_2}{1-\rho}\right)^l \left(\frac{1-\rho}{\mu_1+\mu_2}\right)^{k+l+1} (k+l)! \\
&= (1-\rho) \sum_{k=0}^{\infty} \frac{\rho^k}{k!}\left(\frac{\mu_1}{\mu_1+\mu_2}\right)^{k+1} \sum_{l=0}^{k} \frac{(k+l)!}{l!}\left(\frac{\mu_2}{\mu_1+\mu_2}\right)^l.
\end{aligned} \tag{4.2}
$$

In the particular case $\mu_1 = \mu_2$, one can show that the above reduces to $R = 1/2$.

5. Arnold and Strauss' bivariate exponential distribution

Arnold and Strauss [1] bivariate exponential distribution has the joint pdf specified by

$$
f(x,y) = K \exp\{-(ax+by+cxy)\} \tag{5.1}
$$

for $x > 0$, $y > 0$, $a > 0$, $b > 0$, and $c > 0$, where K denotes the normalizing constant given by

$$
\frac{1}{K} = -\frac{1}{c}\exp\left(\frac{ab}{c}\right)\mathrm{Ei}\left(\frac{ab}{c}\right). \tag{5.2}
$$

Note that the marginal pdfs of X and Y are not exponential and are given by

$$
\begin{aligned}
f_X(x) &= \frac{\exp(-ax)}{b+cx}, \\
f_Y(y) &= \frac{\exp(-by)}{a+cy},
\end{aligned} \tag{5.3}
$$

respectively, with the expectations

$$
\begin{aligned}
E(X) &= \frac{b}{c}\left(\frac{K}{ab}-1\right), \\
E(Y) &= \frac{a}{c}\left(\frac{K}{ab}-1\right).
\end{aligned} \tag{5.4}
$$

However, the conditional distributions of $Y \mid X = x$ and $X \mid Y = y$ are exponential with parameters $b + cx$ and $a + cy$, respectively. Arnold and Strauss [1] motivated the use of (5.1) by observing that it often happens that a researcher has better insight into the forms of conditional distributions rather than the joint distribution. The distribution has been used to model such variables as blood counts and survival times of patients. The correlation coefficient $\rho = \mathrm{Corr}(X, Y)$ is given by

$$\rho = \frac{abc + abK - K^2}{K(ab + c - K)}. \tag{5.5}$$

One can show that the correlation is always negative and is bounded from below by the value -0.32. The form of R can be derived as

$$
\begin{aligned}
R &= K \int_0^\infty \exp(-ax) \int_x^\infty \exp\{-(b + cx)y\} \, dy \, dx \\
&= K \int_0^\infty \frac{\exp\{-(a + b + cx)x\}}{b + cx} \, dx \\
&= \frac{K}{c} \exp\left\{\frac{(a + b)^2}{4c}\right\} \int_{(a+b)/2}^\infty \frac{\exp(-z^2/c)}{z + (b - a)/2} \, dz \\
&= \frac{K}{c} \exp\left\{\frac{(a + b)^2}{4c}\right\} \int_{(a+b)/2}^\infty \exp(-z^2/c) \sum_{k=0}^\infty (-1)^k \left(\frac{b - a}{2}\right)^k z^{-(k+1)} \, dz \\
&= \frac{K}{c} \exp\left\{\frac{(a + b)^2}{4c}\right\} \sum_{k=0}^\infty (-1)^k \left(\frac{b - a}{2}\right)^k \int_{(a+b)/2}^\infty z^{-(k+1)} \exp(-z^2/c) \, dz \\
&= \frac{K}{2c} \exp\left\{\frac{(a + b)^2}{4c}\right\} \sum_{k=0}^\infty (-1)^k \left(\frac{b - a}{2\sqrt{c}}\right)^k \int_{(a+b)^2/(4c)}^\infty w^{-(k/2+1)} \exp(-w) \, dw \\
&= \frac{K}{2c} \exp\left\{\frac{(a + b)^2}{4c}\right\} \sum_{k=0}^\infty (-1)^k \left(\frac{b - a}{2\sqrt{c}}\right)^k \Gamma\left(-\frac{k}{2}, \frac{(a + b)^2}{4c}\right),
\end{aligned}
\tag{5.6}
$$

where the transformations $z = cx + (a + b)/2$ and $w = z^2/c$ have been applied, and the series expansion

$$\frac{1}{z + d} = \sum_{k=0}^\infty (-1)^k d^k z^{-(k+1)} \tag{5.7}$$

used. Thus, one has an expression for R which is an infinite sum of incomplete gamma functions. In the particular case $a = b$, one can show that $R = 1/2$.

6. Freund's bivariate exponential distribution

Freund's [4] bivariate exponential distribution has the joint pdf specified by

$$f(x, y) = \begin{cases} \alpha_1 \beta_2 \exp\{-\beta_2 y - (\alpha_1 + \alpha_2 - \beta_2)x\}, & \text{if } x < y, \\ \alpha_2 \beta_1 \exp\{-\beta_1 x - (\alpha_1 + \alpha_2 - \beta_1)y\}, & \text{if } y < x \end{cases} \tag{6.1}$$

for $x > 0$, $y > 0$, $\alpha_1 > 0$, $\alpha_2 > 0$, $\beta_1 > 0$, and $\beta_2 > 0$. This distribution arises in the following setting: X and Y are the lifetimes of two components assumed to be independent exponentials with parameters α_1 and α_2, respectively; but, a dependence between X and Y is introduced by taking a failure of either component to change the parameter of the life distribution of the other component; if component 1 fails, the parameter for Y changed to β_2; and if component 2 fails, the parameter for X changed to β_1. As in the previous section, the marginal pdfs of X and Y are not exponential. They take the form of exponential mixtures given by

$$f_X(x) = \frac{1}{\alpha_1 + \alpha_2 - \beta_1} \left[(\alpha_1 - \beta_1)(\alpha_1 + \alpha_2) \exp\left\{ -(\alpha_1 + \alpha_2)x \right\} + \beta_1 \alpha_2 \exp\left(-\beta_1 x \right) \right],$$

$$f_Y(y) = \frac{1}{\alpha_1 + \alpha_2 - \beta_2} \left[(\alpha_2 - \beta_2)(\alpha_1 + \alpha_2) \exp\left\{ -(\alpha_1 + \alpha_2)y \right\} + \beta_2 \alpha_1 \exp\left(-\beta_2 y \right) \right],$$

$$\text{(6.2)}$$

respectively, with the expectations

$$E(X) = \frac{\beta_1 + \alpha_2}{\beta_1(\alpha_1 + \alpha_2)},$$

$$E(Y) = \frac{\beta_2 + \alpha_1}{\beta_2(\alpha_1 + \alpha_2)}.$$

$$\text{(6.3)}$$

The correlation coefficient $\rho = \operatorname{Corr}(X, Y)$ is given by

$$\rho = \frac{\beta_1 \beta_2 - \alpha_1 \alpha_2}{\sqrt{\beta_1^2 + 2\alpha_1\alpha_2 + \alpha_2^2}\sqrt{\beta_2^2 + 2\alpha_1\alpha_2 + \alpha_1^2}}. \tag{6.4}$$

One can show that $-1/3 < \rho < 1$. The form of R is given by

$$\begin{aligned}
R &= \alpha_1 \beta_2 \int_0^\infty \int_x^\infty \exp\left\{ -\beta_2 y - (\alpha_1 + \alpha_2 - \beta_2)x \right\} dy\, dx \\
&= \alpha_1 \int_0^\infty \exp\left\{ -(\alpha_1 + \alpha_2)x \right\} dx \\
&= \frac{\alpha_1}{\alpha_1 + \alpha_2}.
\end{aligned} \tag{6.5}$$

This expression is independent of β_1 and β_2 because of the context described above: since X and Y are independent exponentials with parameters α_1 and α_2, one must have $X < Y$ with probability $\alpha_1/(\alpha_1 + \alpha_2)$.

7. Marshall and Olkin's bivariate exponential distribution

Marshall and Olkin's [9, 10] bivariate exponential distribution has the joint pdf specified by

$$f(x, y) = \begin{cases} \theta_1(\theta_2 + \theta_3) \exp\left\{ -\theta_1 x - (\theta_2 + \theta_3)y \right\}, & \text{if } x < y, \\ \theta_2(\theta_1 + \theta_3) \exp\left\{ -\theta_2 y - (\theta_1 + \theta_3)x \right\}, & \text{if } y < x, \\ \theta_3 \exp\left\{ -(\theta_1 + \theta_2 + \theta_3)y \right\}, & \text{if } x = y \end{cases} \tag{7.1}$$

for $x > 0$, $y > 0$, $\theta_1 > 0$, $\theta_2 > 0$, and $\theta_3 > 0$. This distribution arises in the following context: X and Y are the lifetimes of two components subjected to three kinds of shocks; these shocks are assumed to be governed by independent Poisson processes with parameters θ_1, θ_2, and θ_3, according as the shock applies to component 1 only, component 2 only, or both components. The distribution has received wide applicability in nuclear reactor safety, competing risks, reliability and in quantal response contexts. The marginal pdfs of X and Y are exponential with parameters $\theta_1 + \theta_3$ and $\theta_2 + \theta_3$, respectively; so, in particular,

$$
\begin{aligned}
E(X) &= \frac{1}{\theta_1 + \theta_3}, \\
E(Y) &= \frac{1}{\theta_2 + \theta_3}.
\end{aligned}
\tag{7.2}
$$

The correlation coefficient $\rho = \mathrm{Corr}(X, Y)$ is given by

$$
\rho = \frac{\theta_3}{\theta_1 + \theta_2 + \theta_3}.
\tag{7.3}
$$

Solving the three preceding equations, one can express θ_1, θ_2, and θ_3 as

$$
\begin{aligned}
\theta_1 &= \frac{1}{E(X)} - \theta_3, \\
\theta_2 &= \frac{1}{E(Y)} - \theta_3, \\
\theta_3 &= \frac{\{E(X) + E(Y)\}\rho}{E(X)E(Y)(1 + \rho)}.
\end{aligned}
\tag{7.4}
$$

It is easily seen that R can be expressed as

$$
\begin{aligned}
R &= \theta_1(\theta_2 + \theta_3) \int_0^\infty \int_x^\infty \exp\left\{-\theta_1 x - (\theta_2 + \theta_3)y\right\} dy\, dx \\
&= \theta_1 \int_0^\infty \exp\left\{-(\theta_1 + \theta_2 + \theta_3)x\right\} d \\
&= \frac{\theta_1}{\theta_1 + \theta_2 + \theta_3}.
\end{aligned}
\tag{7.5}
$$

8. Application

For a good part of the 20th century, the assumption of independent random samples from continuous distributions dominated applications of statistical methodology. From the middle of the eighties of the 20th century, we are beginning to observe deviations from this setup, mainly because real-world sources of data were not conforming to the i.i.d. continuous model. In fact, a substantial amount of categorized data plays an important role, especially in medical-orientated applications. One of the developments of this type is the analysis of receiver operating characteristic (ROC) curves. This topic was a real hit in the last decade with a large number of publications appearing.

ROC curve is a particular type of an ordinal dominance (OD) graph. Consider random variables X and Y, and for every real number c plot a point $T(c)$ in a Cartesian coordinate system with the coordinates $(\Pr(X \le c), \Pr(Y \le c))$. The collection of the points $T(c)$ form a ROC graph. Note that the coordinates of $T(c)$ lie between 0 and 1, so that the ROC graph is always located within the unit square $\{(x,y) : 0 \le x \le 1,\ 0 \le y \le 1\}$. Moreover, by letting $c = \pm\infty$, we conclude that ROC graph always starts at $(0,0)$ and ends at $(1,1)$.

The relation between OD graphs and $R = \Pr(X < Y)$ was originally pointed out by Bamber [2] and brought a variety of methods developed for an inference about $\Pr(X < Y)$ into analysis of ROC curves. Bamber [2] observed that the area above the OD graph for continuous X and Y is equal to

$$
\begin{aligned}
A(X,Y) &= \int_0^1 \Pr(X \le x)\, d\Pr(Y \le c) \\
&= \int_0^1 \Pr(X \le x) f_Y(c)\, dc \\
&= \Pr(X \le Y) \\
&= R.
\end{aligned}
\tag{8.1}
$$

In view of this relation, the area $A(X,Y)$ can be utilized as a measure of the size difference between two populations with $A(X,Y) = 1$ if and only if the distribution of X lies entirely below the distribution of Y. On the other hand, if X and Y are identically distributed, $A(X,Y) = 1/2$.

ROC curve analysis has been used in various fields of medical imaging, radiology, psychiatry, nondestructive, and manufacturing inspection systems (see, e.g., Hsiao et al. [8], Metz [11], Nockermann et al. [18], Reiser [22], Swets [23], and Swets and Pickett [24]). As an example, consider the "yes-no" signal detection experiment. In this experiment, the observer is told to respond "yes" if he/she thinks that the signal was presented on the trial, and to respond "no" otherwise. It is assumed that the observer performs this task as follows. First, he/she adopts (often subjectively) an impression strength criterion, say c. Then, on each trial if the impression strength reaches or exceeds the criterion, he/she responds "yes," and responds "no" otherwise. Let I_s and I_n be continuous random variables denoting the strengths of sensory impressions aroused by signal and noise events, respectively. Then, $\Pr(\text{yes} \mid \text{signal}) = \Pr(I_s \ge c)$ and $\Pr(\text{yes} \mid \text{noise}) = \Pr(I_n \ge c)$. ROC curve is then a collection of points $(\Pr(I_n \ge c), \Pr(I_s \ge c))$ in a unit square. If numerical data on I_n, I_s, and yes-no responses is available, one can construct a sample ROC curve by estimating probabilities $\Pr(I_n \ge c)$ and $\Pr(I_s \ge c)$ under certain parametric assumptions.

Figure 8.1 illustrates the variation of $A(X,Y) = \Pr(X < Y) = R$ with respect to ρ, $E(X)$, and $E(Y)$ for the six bivariate exponential distributions discussed in the previous sections. The general pattern of variation is $R > 1/2$ and is an increasing function of ρ when $E(X) > E(Y)$; $R < 1/2$ and is a decreasing function of ρ when $E(X) < E(Y)$; and $R = 1/2$ when $E(X) = E(Y)$. The variations can be grouped into four classes as follows:

(1) Arnold and Strauss and Freund into the first class showing the least sensitivity to ρ. For these two models, one can compute R without much error by assuming that X and Y are independent. Of these two, Freund is the simpler since the
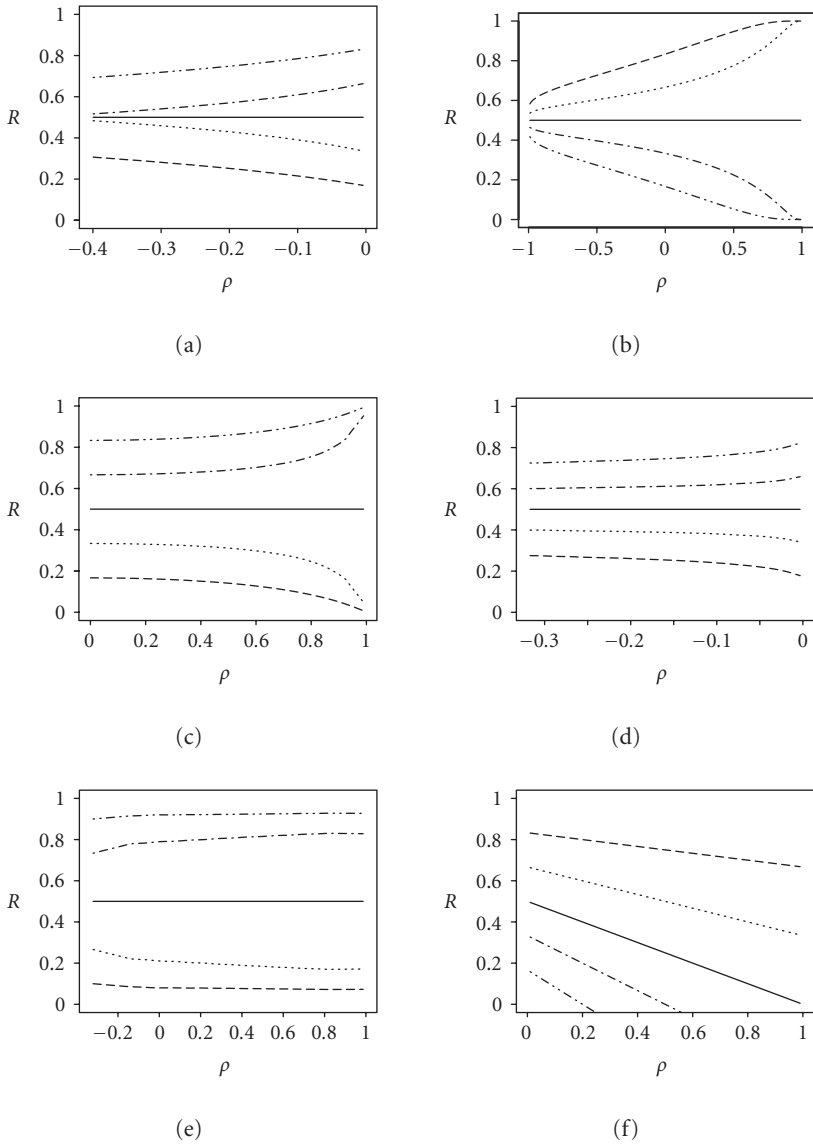
Figure 8.1. Variation of R versus ρ for the six bivariate exponential distributions and for selected values of $E(X)$ and $E(Y)$: Gumbel (a); Hougaard (b); Downton (c); Arnold and Strauss (d); Freund (e); Marshall and Olkin (f); the four curves in each plot from (a) to (f) correspond to $(E(X), E(Y)) = (5,1), (2,1), (1,2),$ and $(1,5)$.

expressions for it are elementary and do not involve special functions. Further-more, Freund exhibits a wider range of values for ρ ($-0.32 < \rho < 0$ for Arnold and Strauss, and $-1/3 < \rho < 1$ for Freund) and attains larger values for R.

(2) Gumbel and Downton into the second class showing moderate sensitivity to ρ. Of these two, although Gumbel is the earliest and simplest known model, one might choose Downton because it has a wider range of values for ρ ($-0.40365 < \rho < 0$ for Gumbel, and $0 \le \rho < 1$ for Downton) and attains larger values for R. Also, the limits of R are attained as $\rho \to 1$ in the case of Downton.

(3) Hougaard into the third class which shows the highest sensitivity to ρ. Hougaard exhibits the widest range of values for ρ ($-1 < \rho < 1$) and attains the limits of R as $\rho \to 1$.

(4) The case of Marshall and Olkin stands out from the rest because of the presence of singularity along the axis $x = y$ (none of the other models have this). Here, R is a decreasing function of ρ for all values of $E(X)$ and $E(Y)$, so the system is most reliable when X and Y are independent. The amount of sensitivity to ρ is comparable to that of Hougaard.

Based on the above discussion, the best model to choose would be that due to Hougaard since it gives the widest range of values for both ρ and R. However, if one is not interested in the dependence, then the model due to Freund might give as good a result. When selecting a model, one should also take into account the physical contexts described in Sections 4, 6, and 7.

9. Conclusions

We have calculated the forms of $R = \Pr(X < Y)$ for six flexible families of bivariate exponential distributions and discussed their utility to ROC curve analysis. It would be of interest to emulate this work for other continuous bivariate distributions, including bivariate beta distributions, bivariate gamma distributions, and bivariate Pareto distributions. It would also be of interest to extend this work for continuous multivariate distributions. We hope to address some of these issues in a future paper.

Appendix

Some technical lemmas required for the calculations above are noted below.

LEMMA A.1 (Prudnikov et al. [19, (2.3.15.7)]). *For $p > 0$,*

$$\int_0^\infty x^n \exp\left\{ -(px^2 + qx) \right\} dx = (-1)^n \sqrt{\frac{\pi}{p}} \frac{\partial^n}{\partial q^n} \left\{ \exp\left(\frac{q^2}{4p} \right) \Phi\left(-\frac{q}{\sqrt{2p}} \right) \right\}, \qquad \text{(A.1)}$$

where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution.

LEMMA A.2 (Prudnikov et al. [19, (2.10.2.1)]). *For $\alpha > 0$,*

$$\int_0^\infty x^{\alpha-1} \Gamma(\nu, cx) dx = \frac{\Gamma(\alpha + \nu)}{\alpha c^\alpha}. \qquad \text{(A.2)}$$
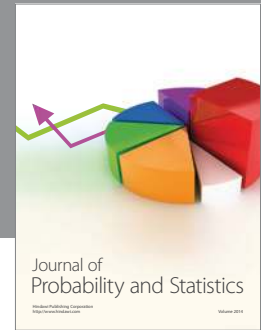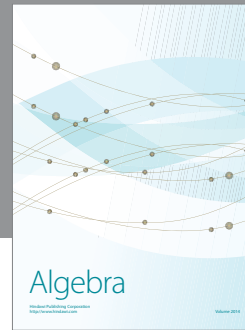
Acknowledgments

References

[1]   B. C. Arnold and D. Strauss, *Bivariate distributions with exponential conditionals*, Journal of the American Statistical Association **83** (1988), no. 402, 522–527.

[2]   D. Bamber, *The area above the ordinal dominance graph and the area below the receiver operating characteristic graph*, Journal of Mathematical Psychology **12** (1975), no. 4, 387–415.

[3]   F. Downton, *Bivariate exponential distributions in reliability theory*, Journal of the Royal Statistical Society. Series B. **32** (1970), 408–417.

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

[5]   I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 6th ed., Academic Press, California, 2000.

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

[7]   P. Hougaard, *A class of multivariate failure time distributions*, Biometrika **73** (1986), no. 3, 671–678.

[8]   J. K. Hsiao, J. J. Bartko, and W. Z. Potter, *Diagnosing diagnoses*, Archives of General Psychiatry **46** (1989), no. 7, 664–667.

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

[10]   ⸻, *A multivariate exponential distribution*, Journal of the American Statistical Association **62** (1967), 30–44.

[11]   C. E. Metz, *Some practical issues of experimental design and data analysis in radiological ROC studies*, Investigative Radiology **24** (1989), no. 3, 234–245.

[12]   S. Nadarajah, *Reliability for beta models*, Serdica. Mathematical Journal **28** (2002), no. 3, 267–282.

[13]   ⸻, *Reliability for extreme value distributions*, Mathematical and Computer Modelling **37** (2003), no. 9-10, 915–922.

[14]   ⸻, *Reliability for lifetime distributions*, Mathematical and Computer Modelling **37** (2003), no. 7-8, 683–688.

[15]   ⸻, *Reliability for Laplace distributions*, Mathematical Problems in Engineering **2004** (2004), no. 2, 169–183.

[16]   ⸻, *Reliability for Logistic distributions*, Engineering Simulation **26** (2004), 81–98.

[17]   S. Nadarajah and S. Kotz, *Reliability for Pareto models*, Metron **61** (2003), no. 2, 191–204.

[18]   C. Nockermann, H. Heidt, and N. Thomsen, *Reliability in NTD: ROC study of radiographic weld inspections*, NDT & E Internatioal **24** (1991), no. 5, 235–245.

[19]   A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series. Vol. 1*, Gordon & Breach, New York, 1986.

[20]   ⸻, *Integrals and Series. Vol. 2*, Gordon & Breach, New York, 1986.

[21]   ⸻, *Integrals and Series. Vol. 3*, Gordon & Breach, New York, 1986.

[22]   B. Reiser, *Measuring the effectiveness of diagnostic markers in the presence of measurement error through the use of ROC curves*, Statistics in Medicine **19** (2000), no. 16, 2115–2129.

[23] J. A. Swets, *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics: Collected Papers*, Lawrence Erlbaum Associates, New Jersey, 1996.

[24] J. A. Swets and R. M. Pickett, *Evaluation of Diagnostic Systems: Methods from Signal Detection Theory*, Academic Press, New York, 1982.

Saralees Nadarajah: Department of Statistics, University of Nebraska, Lincoln, NE 68583, USA
E-mail address: snadaraj@unlserve.unl.edu

Samuel Kotz: Department of Engineering Management and Systems Engineering,
The George Washington University, Washington, DC 20052, USA
E-mail address: kotz@gwu.edu