مجلّه‌ی پژوهش‌های آماری ایران

J. Statist. Res. Iran **8** (2011): 29–47

دوره‌ی ۸، شماره‌ی ۱، بهار و تابستان ۱۳۹۰، صص ۲۹–۴۷

پژوهشکده‌ی آمار

A Note on the Bivariate Maximum Entropy Modeling

S. Ashrafi and M. Asadi*

University of Isfahan

Abstract. Let $\mathbf{X} = (X_1, X_2)$ be a continuous random vector. Under the assumption that the marginal distributions of X_1 and X_2 are given, we develop models for vector $\mathbf{X}$ when there is partial information about the dependence structure between X_1 and X_2. The models which are obtained based on well-known Principle of Maximum Entropy are called the maximum entropy (ME) models. Our results lead to characterization of some well-known bivariate distributions such as Generalized Gumbel, Farlie-Gumbel-Morgenstern and Clayton bivariate distributions. The relationship between ME models and some well known dependence notions are studied. Conditions under which the mixture of bivariate distributions are ME models are also investigated.

Keywords. Fréchet class of distributions; hazard gradient; dependence; total positive of order 2.

MSC 2010: 60E05, 62E15.

1 Introduction

Let $\mathbf{X} = (X_1, X_2)$ be a continuous random vector with distribution function $F(x_1, x_2) = P(X_1 \leqslant x_1, X_2 \leqslant x_2)$, $x_1, x_2 \in R$. Specification of a bivariate distribution requires full information about marginal distributions as well as dependence structure between X_1 and X_2. There are many situations in which the marginal distributions are known but the complete information about the dependence structure between X_1 and X_2 is unknown. The

* Corresponding author

problem of interest, in such situations, is to make inference about joint distribution based on constraints on some specifications of the population. A well known approach to characterize a model for the data generating distribution is the maximum entropy method. In this approach, insufficient knowledge about the data generating distribution is formulated in terms of a set of information constraints. Usually the constraints are made on the moments of the model and then the aim is to find the model that maximizes Shannon entropy under these constraints.

In reliability engineering and survival analysis, there are several criterion which play the central role in analyzing the lifetime data in both univariate and multivariate cases. Among the well known measures, the hazard rate and mean residual life function are of particular interest. The present paper provides a solution for problem of specification of bivariate models using the well-known Principle of Maximum Entropy; especially when partial information about the dependence structures between X_1 and X_2 are available and the constraints are made based on hazard gradient or reversed hazard gradient. In the univariate case, when the constraints are based on hazard rate and mean residual life function, Asadi et al. (2004) studied a concept of maximum dynamic entropy and Asadi et al. (2005) introduced a notion of minimum dynamic discrimination information and obtained various univariate lifetime distributions as maximum dynamic entropy and minimum dynamic discrimination models. Recently, Asadi et al. (2010) have studied a concept of bivariate dynamic ME model and derived several bivariate distributions when partial information is available on hazard gradient.

This communication is a continuation of the work by Asadi et al. (2010). In the first part of the paper, we develop several bivariate models based on partial information on the hazard gradient and reversed hazard gradient. This enables us to characterize several well-known bivariate distributions, with given marginal distributions, as maximum entropy (ME) models when partial information about hazard gradient or reversed hazard gradient are formulated in some inequality constraints. In the second part of the paper, we characterize mixture of bivariate distributions as ME models when partial information is available on the hazard rate of mixing distribution. The rest of the paper is organized as follows. Section 2 gives some preliminary results which are useful in subsequent sections. Section 3 gives results about ME models based on some dependence concepts such as total positivity of order 2 (TP_2) and reversed regularity of order 2 (RR_2). Some well known bivariate distributions are characterized as ME models under partial information about the hazard and reversed hazard gradient. In Section 4, we discuss methods for characterizing bivariate mixtures as ME models when partial information about the distribution function of dependence parameter is available.

2 Preliminaries

In this section we give some definitions and preliminary results which are used in the subsequent sections. Let $\mathcal{M}(F_1, F_2)$ be the Fréchet class of absolutely continuous bivariate distribution functions (BDF) with given marginal distribution functions F_1 and F_2. The Kullback-Leibler discrimination information function between the BDF $F(x_1, x_2)$ and reference BDF $F_0(x_1, x_2)$ is defined by

$$K(F : F_0) = \int\int f(x_1, x_2) \log \frac{f(x_1, x_2)}{f_0(x_1, x_2)} dx_1 dx_2 \geqslant 0,$$

where F is absolutely continuous with respect to F_0 and f and f_0 denote the probability density functions (PDFs) of F and F_0, respectively. Note that $K(F : F_0) = 0$ if and only if $f_0(x_1, x_2) = f(x_1, x_2)$ with probability 1. The joint entropy of F, denoted by $H(\mathbf{X})$, is defined to be

$$H(\mathbf{X}) = -\int\int f(x_1, x_2) \log f(x_1, x_2) dx_1 dx_2.$$

Definition 1. Let Ω_F be the set of all bivariate distributions in $\mathcal{M}(F_1, F_2)$ that satisfy some partial information (some constraints). The ME model in Ω_F is a BDF $F^* \in \Omega_F$ such that

$$F^* = \arg \max_{F \in \Omega_F} H(F).$$

That is, the ME model is the one that its PDF maximizes joint entropy among all distributions in Ω_F (Jaynes, 1957).

In the univariate case, the hazard rate of a continuous distribution F with density f is defined as $\lambda(t) = \frac{f(t)}{\bar{F}(t)}$, $t > 0$, where $\bar{F}(t) = 1 - F(t)$ is the survival function. The hazard rate plays a central role in the study of lifetime random variables. In the literature there are various extensions of hazard rate $\lambda(t)$ to the multivariate. An extension which is defined in the bivariate case as follows is called the hazard gradient. Assume that the bivariate random vector $\mathbf{X}$ has the survival function $\bar{F}(x_1, x_2) = P(X_1 > x, X_2 > x_2)$. The vector of hazard gradient of $\mathbf{X}$ is defined as

$$\begin{aligned}\Lambda_F(x_1, x_2) &= -\left\{\frac{\partial \log \bar{F}(x_1, x_2)}{\partial x_1}, \frac{\partial \log \bar{F}(x_1, x_2)}{\partial x_2}\right\} \\ &\equiv \left\{\lambda_{F,1}(x_1, x_2), \lambda_{F,2}(x_1, x_2)\right\}.\end{aligned}$$

Note that $\lambda_{F,i}(x_1, x_2)$ can be interpreted as the conditional hazard rate of X_i evaluated at x_i, given that $X_j > x_j$ for all $i, j = 1, 2, i \neq j$. That is,

$$\lambda_{F,i}(x_1, x_2) = \frac{f_i(x_i | X_j > x_j, i \neq j)}{\bar{F}_i(x_i | X_j > x_j, i \neq j)}$$

where $f_i(\cdot | X_j > x_j, j \neq i)$ and $\bar{F}_i(\cdot | X_j > x_j, j \neq i)$ are, respectively, the conditional density and survival functions of X_i, given that $X_j > x_j$ for $i, j = 1, 2, i \neq j$. The hazard gradient $\Lambda_F(x_1, x_2)$ has the property that its relation to BDF $F(x_1, x_2)$ is one-to-one; see, Johnson and Kotz (1975) and Marshall (1975).

Another measure which is important in reliability and survival analysis is reversed hazard rate which is defined as $r(t) = \frac{f(t)}{F(t)}$. In the bivariate case the reversed hazard gradient of $\mathbf{X}$ is defined as a vector similar to the hazard gradient. In fact, the reversed hazard gradient of $\mathbf{X}$ is

$$\begin{aligned}R_F(x_1, x_2) &= \left\{\frac{\partial \log F(x_1, x_2)}{\partial x_1}, \frac{\partial \log F(x_1, x_2)}{\partial x_2}\right\} \\ &\equiv \left\{r_{F,1}(x_1, x_2), r_{F,2}(x_1, x_2)\right\},\end{aligned}$$

where $r_{F,i}(x_1, x_2)$ is the reversed hazard gradient of $i$ts component.

Definition 2. Let the bivariate random vectors $\mathbf{X}$ and $\mathbf{Y}$ have survival functions $\bar{F}$ and $\bar{G}$ and distribution functions F and G, respectively.

- Suppose that F and G belong to $\mathcal{M}(F_1, F_2)$. $\mathbf{X}$ is said to be smaller than $\mathbf{Y}$ in positive quadrant dependent order (denoted by $\mathbf{X} \leqslant_{PQD} \mathbf{Y}$) if

$$\bar{F}(x_1, x_2) \leqslant \bar{G}(x_1, x_2) \quad \text{for all } (x_1, x_2).$$

- $\mathbf{X}$ is said to be smaller than $\mathbf{Y}$ in the weak bivariate hazard rate order (denoted by $\mathbf{X} \leqslant_{whr} \mathbf{Y}$) if

$$\frac{\bar{G}(\mathbf{x})}{\bar{F}(\mathbf{x})} \uparrow \text{for all } \mathbf{x} \in \{\mathbf{x} : \bar{G}(\mathbf{x}) > 0\}.$$

Let $\Lambda_F(\mathbf{x})$ and $\Lambda_G(\mathbf{x})$ denote the hazard gradients of $\mathbf{X}$ and $\mathbf{Y}$, respectively. Then it can be easily seen that

$$\mathbf{X} \leqslant_{whr} \mathbf{Y} \Longleftrightarrow \lambda_{F,i}(\mathbf{x}) \geqslant \lambda_{G,i}(\mathbf{x}), \qquad i = 1, 2,\ \mathbf{x} \in R^2;$$

see Shaked and Shanthikumar (2007), p. 291.
In order to give the next definition, we need the concept of supermodular. A function $\phi : R^2 \to R$ is said to be supermodular if for any $\mathbf{x}, \mathbf{y} \in R^2$ it satisfies

$$\phi(\mathbf{x}) + \phi(\mathbf{y}) \leqslant \phi(\mathbf{x} \wedge \mathbf{y}) + \phi(\mathbf{x} \vee \mathbf{y}),$$

where the operators $\wedge$ and $\vee$ denote coordinatewise minimum and maximum, respectively; see Shaked and Shanthikumar (2007), p. 335.

Definition 3. Let $\mathbf{X}$ and $\mathbf{Y}$ be two bivariate random vectors such that

$$E[\phi(\mathbf{X})] \leqslant E[\phi(\mathbf{Y})] \quad \text{for all supermodular functions } \phi : R^2 \to R,$$

provided the expectations exist. Then $\mathbf{X}$ is said to be smaller than $\mathbf{Y}$ in supermodular order (denoted by $\mathbf{X} \leqslant_{sm} \mathbf{Y}$); see Shaked and Shanthikumar (2007), p. 395.

In statistical literature there have been defined various notions of dependency between two random variables. Two concepts of dependency which are used in this paper are the concepts of total positivity of order 2 (TP_2) and right tail increasing (RTI). These are defined as follows.

Definition 4. (a) Let X_1 and X_2 have the joint density function $f(x_1, x_2)$. $f(x_1, x_2)$ is said to be total positive of order 2 if

$$f(x_1, x_2)f(y_1, y_2) \geqslant f(x_1, y_2)f(y_1, x_2) \quad \text{for all } (x_1, x_2) < (y_1, y_2). \qquad (1)$$

(b) X_2 is said to be right-tail increasing in X_1 (written as $RTI(X_2|X_1)$) if

$$P(X_2 > x_2|X_1 > x_1) \text{ is increasing in } x_1 \text{ for all } x_2. \qquad (2)$$

The dual of (1) and (2) are respectively called reverse regular of order $2(RR_2)$ and right tail decreasing (RTD).

3 ME Models with Given Marginal Distributions

In this section we study conditions under which a distribution function is an ME model in Ω_F. The conditions are based on partial information about hazard gradient and reversed hazard gradient. Before giving the main results we need the following Lemma.

***Lemma* 1.** *Consider bivariate random vectors* $\mathbf{X}$ *and* $\mathbf{Y}$ *with distribution functions* F *and* G *and hazard gradients* $\Lambda_F(\mathbf{x}) = (\lambda_{F,1}(\mathbf{x}), \lambda_{F,2}(\mathbf{x}))$ *and* $\Lambda_G(\mathbf{x}) = (\lambda_{G,1}(\mathbf{x}), \lambda_{G,2}(\mathbf{x}))$*, respectively. Assume that* $F, G \in \mathcal{M}(F_1, F_2)$*. If for all* $\mathbf{x}$*,* $\lambda_{F,i}(\mathbf{x}) \geqslant \lambda_{G,i}(\mathbf{x}),\ i = 1, 2$ *then* $\mathbf{X} \leqslant_{PQD} \mathbf{Y}$*.*

Proof. Let $\mathbf{X}$ and $\mathbf{Y}$ have survival functions $\bar{F}(\mathbf{x})$ and $\bar{G}(\mathbf{x})$, respectively. The condition that $\lambda_{F,i}(\mathbf{x}) \geqslant \lambda_{G,i}(\mathbf{x}),\ i = 1, 2$ is equivalent to $\frac{\bar{G}(\mathbf{x})}{\bar{F}(\mathbf{x})}$ is increasing in $\mathbf{x}$. Hence $\bar{G}(\mathbf{x}) \geqslant \bar{F}(\mathbf{x})$ for all $\mathbf{x}$. Using this and the assumption that $F, G \in \mathcal{M}(F_1, F_2)$, we have $\mathbf{X} \leqslant_{PQD} \mathbf{Y}$. □

Remark 1. One can show that the result of Lemma 1 is valid if the reversed hazard gradients are ordered. That is, under the condition that $F, G \in \mathcal{M}(F_1, F_2)$, if $r_{F,i}(\mathbf{x}) \geqslant r_{G,i}(\mathbf{x}),\ i = 1, 2$, then $\mathbf{X} \leqslant_{PQD} \mathbf{Y}$.

Now, we are ready to prove the following theorem.

***Theorem* 1.** *Let* $\Omega_F = \{F(x_1, x_2) \in \mathcal{M}(F_1, F_2) : \Lambda_F(x_1, x_2) \leqslant (\geqslant) Q(x_1, x_2)\}$ *be a set of distributions in* $\mathcal{M}(F_1, F_2)$ *having hazard gradient* Λ_F*. Suppose that there exists a distribution* $F^* \in \Omega_F$ *with PDF* f^* *such that* $\Lambda_{F^*}(x_1, x_2) = Q(x_1, x_2)$*. Then*

(a) F^* *is ME in* Ω_F *with* $\Lambda_F(x_1, x_2) \leqslant Q(x_1, x_2)$ *if* $f^*(x_1, x_2)$ *is* TP_2*.*

(b) F^* *is ME in* Ω_F *with* $\Lambda_F(x_1, x_2) \geqslant Q(x_1, x_2)$ *if* $f^*(x_1, x_2)$ *is* RR_2*.*

Proof. We prove part (a) of the theorem. Part (b) can be proved similarly. Let $\mathbf{X}$ and $\mathbf{Y}$ denote two bivariate random vectors with distribution functions F^* and F in $\mathcal{M}(F_1, F_2)$. Since $\Lambda_F(x_1, x_2) \leqslant \Lambda_{F^*}(x_1, x_2)$ for all $\mathbf{x}$ and F^* and F are in $\mathcal{M}(F_1, F_2)$, from Lemma 1, we have $\mathbf{X} \leqslant_{PQD} \mathbf{Y}$ which is equivalent to say that $\mathbf{X} \leqslant_{sm} \mathbf{Y}$ (see Shaked and Shanthikumar, 2007, p. 395). Also it is easily seen that if $f^*(x_1, x_2)$ is TP_2 then $\log f^*(x_1, x_2)$ is supermodular. Thus

$$E[\log f^*(\mathbf{X})] \leqslant E[\log f^*(\mathbf{Y})]. \tag{3}$$

On the other hand

$$K(F : F^*) = -H(\mathbf{Y}) - E[\log f^*(\mathbf{Y})] \geqslant 0.$$

Using this and (3) we obtain $H(\mathbf{X}) \geqslant H(\mathbf{Y})$. This completes the proof of the theorem. □

Remark 2. One way to check that a bivariate density $f^*(x_1, x_2)$ is TP_2 or RR_2 is to use a result of Holland and Wang (1987). They showed that if $f^*(x_1, x_2)$ has second partial derivative, then it is TP_2 (RR_2) if and only if $\frac{\partial^2}{\partial x_1 x_2} \log f^*(x_1, x_2) \geqslant (\leqslant) 0$.

Remark 3. Let f_1 and f_2 be marginal PDFs of distributions in $\mathcal{M}(F_1, F_2)$. Assume that $f_0(x_1, x_2) = f_1(x_1)f_2(x_2)$ is a reference distribution. Using Theorem 1 it can be proved that the ME model in Ω_F is a distribution that minimizes Kullback-Leibler discrimination information function with respect to reference distribution $f_0(x_1, x_2)$. A criterion which is used in information literature to measure the dependency between F_1 and F_2 is mutual information. If $M(X_1, X_2)$ denotes the mutual information between X_1 and X_2, then we have

$$\begin{aligned} M(X_1, X_2) &= H(X_1) - H(X_1|X_2) \\ &= K(F : F_1F_2), \end{aligned}$$

where $H(X_1|X_2)$ denotes the entropy of conditional density of X_1 given X_2. Using this, we conclude that the ME model in Ω_F is a distribution which has the minimum dependency. Also it worth to note that if there is not available any partial information about dependency of X_1 and X_2 i.e. $\Omega_F = \mathcal{M}(F_1, F_2)$, then the independent model $F(x_1, x_2) = F_1(x_1)F_2(x_2)$ is ME. This is so, because for all $F \in \Omega_F$ we have

$$K(F : F_1F_2) = -H(F) + H(F_1F_2) \geqslant 0$$

and equality occurs if and only if $F = F_1F_2$.

In the following we characterize some bivariate distributions as ME models where the constraints are based on hazard gradient. The key function on constructing the constraints is the ratio of ith element of hazard gradient $\Lambda(x_1, x_2) = (\lambda_1(x_1, x_2), \lambda_2(x_1, x_2))$ over the hazard rate of marginal distribution functions F_i, $i = 1, 2$. That is, to characterize the ME model we assume that partial information is available on $\eta_i(x_1, x_2)$, where

$$\eta_i(x_1, x_2) = \frac{\lambda_i(x_1, x_2)}{\lambda_i(x_i)}, \qquad i = 1, 2. \tag{4}$$

The function $\eta_i(x_1, x_2)$ is closely related to well known concepts of dependency. In other words, assume that (X_1, X_2) has bivariate survival function $\bar{F}$. Then it is known that $\bar{F}$ is both $RTI(X_1|X_2)$ and $RTI(X_2|X_1)$ ($RTD(X_1|X_2)$ and $RTD(X_2|X_1)$) if and only if, for all x_1, x_2

$$\eta_i(x_1, x_2) \leqslant (\geqslant) 1, \qquad i = 1, 2. \tag{5}$$

Also it can be shown that if $\bar{F}$ is TP_2 (RR_2) then (5) holds (see, Khaledi and Kochar, 2005). This discussion leads to the following theorem.

***Theorem* 2.** *Let* $\Omega_F = \{F(x_1, x_2) \in \mathcal{M}(F_1, F_2) : \Lambda_F(x_1, x_2) \leqslant (\geqslant) Q(x_1, x_2)\}$. *Assume that there exists a* $F^* \in \Omega_F$ *such that* $\Lambda_{F^*}(x_1, x_2) = Q(x_1, x_2)$ *and* f^* *is* $TP_2(RR_2)$. *Then any* $F \in \Omega_F$ *is both* $RTI(X_1|X_2)$ *and* $RTI(X_2|X_1)$ *(*$RTD(X_1|X_2)$ *and* $RTD(X_2|X_1)$*). That is* $\eta_i(x_1, x_2) \leqslant (\geqslant) 1, \quad i = 1, 2$.

Proof. Let $\Lambda_F(\mathbf{x}) = (\lambda_{F,1}(\mathbf{x}), \lambda_{F,2}(\mathbf{x}))$ be hazard gradient of a distribution F in Ω_F. First assume that f^* is TP_2. Then it can be shown that $\bar{F}^*$ is also TP_2. This implies that $\lambda_{F^*,i}(x_1, x_2) \leqslant \lambda_i(x_i)$ $i = 1, 2$. Thus the assumption that $\lambda_{F,i}(x_1, x_2) \leqslant \lambda_{F^*,i}(x_1, x_2)$ gives $\lambda_{F,i}(x_1, x_2) \leqslant \lambda_i(x_i)$. Hence we have that F is both $RTI(X_1|X_2)$ and $RTI(X_2|X_1)$. When f^* is RR_2, the same arguments show that any $F \in \Omega$ is $RTD(X_1|X_2)$ and $RTD(X_2|X_1)$. This completes the proof. □

The following theorem explores the relationship between the covariance of the elements of ME model and the covariance of the elements of any other distribution in Ω_F.

***Theorem* 3.** *Let* (X_1, X_2) *and* (Y_1, Y_2) *be two bivariate vectors with PDF* $f^*(\mathbf{x})$ *and* $f(\mathbf{x})$ *and hazard gradients* $\Lambda_{F^*}(\mathbf{x}) = (\lambda_{F^*,1}(\mathbf{x}), \lambda_{F^*,2}(\mathbf{x}))$ *and* $\Lambda_F(\mathbf{x}) = (\lambda_{F,1}(\mathbf{x}), \lambda_{F,2}(\mathbf{x}))$*, respectively.*

(a) If $f^*(\mathbf{x})$ *is* TP_2 *and* $\Lambda_{F^*}(\mathbf{x}) \geqslant \Lambda_F(\mathbf{x})$ *then* $0 \leqslant \mathrm{cov}(X_1, X_2) \leqslant \mathrm{cov}(Y_1, Y_2)$.

(b) If $f^*(\mathbf{x})$ *is* RR_2 *and* $\Lambda_{F^*}(\mathbf{x}) \leqslant \Lambda_F(\mathbf{x})$ *then* $\mathrm{cov}(Y_1, Y_2) \leqslant \mathrm{cov}(X_1, X_2) \leqslant 0$.

Proof. We prove part (a). The proof of part (b) is similar. It is known that when $f^*(\mathbf{x})$ is TP_2 then $\text{cov}(X_1, X_2) \geqslant 0$ (see Lai and Xie, 2006, p. 267). Also from Lemma 1, if $\Lambda_{F^*}(\mathbf{x}) \geqslant \Lambda_F(\mathbf{x})$ then $(X_1, X_2) \leqslant_{PQD} (Y_1, Y_2)$ which implies $\text{cov}(X_1, X_2) \leqslant \text{cov}(Y_1, Y_2)$. This completes the proof. □

The result of this theorem shows that when $\text{cov}(X_1, X_2)$ and $\text{cov}(Y_1, Y_2)$ are covariances between the elements of ME model and elements of any other distribution in Ω_F, respectively, if $f^*(\mathbf{x})$ is ME in $\Omega_F = \{F(\mathbf{x}) \in \mathcal{M}(F_1, F_2) : \Lambda_F(\mathbf{x}) \leqslant (\geqslant)\Lambda_{F^*}(\mathbf{x})\}$ then $|\text{cov}(X_1, X_2)| \leqslant |\text{cov}(Y_1, Y_2)|$. This means that X_1 and X_2, the elements of ME model, have the minimum absolute value of the linear dependency in the class.

In the sequel, we give some examples.

Example 1. Let Ω_F be a subset of $\mathcal{M}(F_1, F_2)$ consisting of all bivariate distributions with hazard gradient $\Lambda(x_1, x_2) = (\lambda_1(x_1, x_2), \lambda_2(x_1, x_2))$ that satisfies the following inequalities

$$\eta_i(x_1, x_2) \geqslant (1 - \delta \log \bar{F}_j(x_j)), \quad 0 \leqslant \delta \leqslant 1, \quad i, j = 1, 2, \quad i \neq j,$$

where $\eta_i(x_1, x_2)$ is defined as (4). Then, the ME model in Ω_F is Generalized Gumbel distribution with PDF

$$\begin{aligned} f^*(x_1, x_2) = {} & f_1(x_1) f_2(x_2)[\{1 - \delta \log \bar{F}_1(x_1)\}\{1 - \delta \log \bar{F}_2(x_2)\} - \delta] \\ & \times \exp\{-\delta \log \bar{F}_1(x_1) \log \bar{F}_2(x_2)\}, \end{aligned}$$

in which $0 \leqslant \delta \leqslant 1$ and f_1, f_2 are marginal PDFs of distributions in $\mathcal{M}(F_1, F_2)$ and $\bar{F}_1, \bar{F}_2$ are survival functions associated to f_1, f_2, respectively. The validity of this result follows from part (b) of Theorem 1. To see this, one can easily show that

$$\frac{\partial^2}{\partial x_1 \partial x_2} \log f^*(x_1, x_2) \leqslant 0$$

which is equivalent to say that f^* is RR_2. Also it is easy to see that $\lambda_i(x_i)(1 - \delta \log \bar{F}_j(x_j))$ is the ith element of hazard gradient of $f^*(x_1, x_2)$.

Example 2. Sarmanov (1966) introduced a family of bivariate densities of the form

$$f^*(x_1, x_2) = f_1(x_1) f_2(x_2)\{1 + \omega \phi_1(x_1) \phi_2(x_2)\}, \quad x_1, x_2 \in R \tag{6}$$

where $\omega (\in R)$, ϕ_1 and ϕ_2 satisfy the following conditions

$$\int_{-\infty}^{\infty} \phi_i(u) f_i(u) du = 0, \; i = 1, 2 \;\text{ and }\; 1 + \omega \phi_1(x_1) \phi_2(x_2) \geqslant 0 \text{ for all } x_1, x_2. \tag{7}$$

It can be easily shown that $f^*(x_1, x_2)$ is $TP_2(RR_2)$ if

$$\omega\phi_1'(x_1)\phi_2'(x_2) \geqslant (\leqslant) 0 \quad \text{for all } x_1, x_2.$$

Let Ω_F be a set of bivariate distributions with marginal PDFs f_1 and f_2 whose hazard gradients $\Lambda(x_1, x_2) = (\lambda_1(x_1, x_2), \lambda_2(x_1, x_2))$ satisfy the following inequalities

$$\eta_i(x_1, x_2) \leqslant (\geqslant) \frac{1 + \omega\phi_i(x_i)\psi_j(x_j)}{1 + \omega\psi_1(x_1)\psi_2(x_2)}, \qquad i, j = 1, 2, i \neq j, \tag{8}$$

where

$$\psi_i(x) = E\{\phi_i(X_i)|X_i > x\}, \qquad i = 1, 2,$$

and ω, ϕ_1, ϕ_2 satisfy (7) and $\omega\phi_1'(x_1)\phi_2'(x_2) \geqslant (\leqslant) 0$. Using Theorem 1, f^* is ME in Ω_F since the ith element of hazard gradient of f^* is equal to

$$\lambda_i(x) \frac{1 + \omega\phi_i(x_i)\psi_j(x_j)}{1 + \omega\psi_1(x_1)\psi_2(x_2)}, \qquad i, j = 1, 2, i \neq j$$

where $\lambda_i(x), i = 1, 2$ is hazard function with respect to $f_i(x)$.

There are members of family (6) for which the condition $\omega\phi_1'(x_1)\phi_2'(x_2) \geqslant (\leqslant) 0$ holds. We give two examples here.

(a) The Farlie-Gumbel-Morgenstern (FGM) bivariate distribution with PDF

$$f^*(x_1, x_2) = f_1(x_1)f_2(x_2)[1 + \alpha\{1 - 2F_1(x_1)\}\{1 - 2F_2(x_2)\}], \quad -1 \leqslant \alpha \leqslant 1,$$

is a well known family of distributions with applications in various branches of statistics. For FGM model it can be shown that

$$\frac{\partial^2}{\partial x_1 \partial x_2} \log f^*(x_1, x_2) \geqslant 0 \ (\leqslant 0) \quad \text{if} \quad 0 \leqslant \alpha \leqslant 1 \ (-1 \leqslant \alpha \leqslant 0).$$

In other words, f^* is $TP_2(RR_2)$ if $0 \leqslant \alpha \leqslant 1(-1 \leqslant \alpha \leqslant 0)$. Let Ω_F be a set of bivariate distributions with marginal PDFs $f_1(x)$ and $f_2(x)$ and hazard gradient $\Lambda(x_1, x_2) = (\lambda_1(x_1, x_2), \lambda_2(x_1, x_2))$ that satisfies the following inequalities

$$\eta_i(x_1, x_2) \leqslant (\geqslant)[1 + \alpha F_j(x_j)(2F_i(x_i) - 1)][1 + \alpha F_1(x_1)F_2(x_2)]^{-1}, \quad 0 \leqslant \alpha \leqslant 1(-1 \leqslant \alpha \leqslant 0), \ i \neq j, \ i, j = 1, 2.$$

Under this constraint and the fact that $f^*(x_1, x_2)$ is $TP_2(RR_2)$ we get, from Theorem 1, that $f^*(x_1, x_2)$ is ME in Ω_F.

(b) Lee (1996) derived a bivariate exponential distribution with PDF

$$f(x_1,x_2)=\lambda_1\lambda_2 e^{-(\lambda_1 x_1+\lambda_2 x_2)}\left[1+\omega(e^{-x_1}-\frac{\lambda_1}{1+\lambda_1})(e^{-x_2}-\frac{\lambda_2}{1+\lambda_2})\right]$$

where

$$\frac{-(1+\lambda_1)(1+\lambda_2)}{\max(\lambda_1\lambda_2,1)}\leqslant\omega\leqslant\frac{(1+\lambda_1)(1+\lambda_2)}{\max(\lambda_1,\lambda_2)}$$

and $\phi_i(x_i)=e^{-x_i}-\frac{\lambda_i}{1+\lambda_i},\ i=1,2$. Therefore

$$\omega\phi_1'(x_1)\phi_2'(x_2)=\begin{cases}\geqslant 0 & 0\leqslant\omega\leqslant\frac{(1+\lambda_1)(1+\lambda_2)}{\max(\lambda_1,\lambda_2)}\\ \leqslant 0 & \frac{-(1+\lambda_1)(1+\lambda_2)}{\max(\lambda_1\lambda_2,1)}\leqslant\omega\leqslant 0,\end{cases}$$

From this we get

$$\begin{cases}f(x_1,x_2)\text{ is }TP_2 & \text{if}\quad 0\leqslant\omega\leqslant\frac{(1+\lambda_1)(1+\lambda_2)}{\max(\lambda_1,\lambda_2)}\\ f(x_1,x_2)\text{ is }RR_2 & \text{if}\quad \frac{-(1+\lambda_1)(1+\lambda_2)}{\max(\lambda_1\lambda_2,1)}\leqslant\omega\leqslant 0.\end{cases}$$

For some distributions, constraints based on reversed hazard gradient are more simple than hazard gradient. The following theorem gives ME models in $\mathcal{M}(F_1,F_2)$, in which the constraints are made on reversed hazard gradient. The proof of the theorem, which is similar to the proof of Theorem 1, is based on Remark 1 and hence is omitted.

Theorem 4. *Let* $\Omega_F=\{F(x_1,x_2)\in\mathcal{M}(F_1,F_2): R_F(x_1,x_2)\geqslant(\leqslant)R(x_1,x_2)\}$ *be a set of distributions in* $\mathcal{M}(F_1,F_2)$ *having reversed hazard gradient* R_F. *Suppose that there exists a distribution* $F^*\in\Omega_F$ *with PDF* f^**such that* $R_{F^*}(x_1,x_2)=R(x_1,x_2)$. *Then*

(a) F^* *is ME in* Ω_F *with* $R_F(x_1,x_2)\leqslant R(x_1,x_2)$ *if* $f^*(x_1,x_2)$ *is* TP_2.

(b) F^* *is ME in* Ω_F *with* $R_F(x_1,x_2)\geqslant R(x_1,x_2)$ *if* $f^*(x_1,x_2)$ *is* RR_2.

Example 3. Consider the Clayton's bivariate distribution with PDF

$$f^*(x_1,x_2)=\frac{(\theta+1)f_1(x_1)f_2(x_2)\{F_1(x_1)F_2(x_2)\}^{-\theta-1}}{[\{F_1(x_1)\}^{-\theta}+\{F_2(x_2)\}^{-\theta}-1]^{\frac{1}{\theta}+2}},\qquad\theta>0,$$

in which f_1 and f_2 are marginal densities with distribution functions F_1 and F_2, respectively. For this distribution it can be shown that for all $x_1, x_2, \theta > 0$

$$\frac{\partial^2}{\partial x_1 \partial x_2} \log f^*(x_1, x_2) \geqslant 0.$$

Consider Ω_F as a set of bivariate distributions with marginal distribution functions F_1 and F_2 and reversed hazard gradient $R(x_1, x_2) = (r_1(x_1, x_2), r_2(x_1, x_2))$ satisfying the following inequalities

$$\beta_i(x_1, x_2) \leqslant \frac{\{F_i(x_i)\}^{-\theta}}{[\{F_1(x_1)\}^{-\theta} + \{F_2(x_2)\}^{-\theta} - 1]}, \quad \theta > 0, \ i = 1, 2,$$

where $\beta_i(x_1, x_2) = \frac{r_i(x_1, x_2)}{r_i(x_i)}, i = 1, 2$ and $r_i(x), i = 1, 2$ is reversed hazard function of $F_i(x)$. Clayton's bivariate distribution is ME in Ω_F. The result follows from part (a) of Theorem 4 because $r_i(x_i)\{F_i(x_i)\}^{-\theta}[\{F_1(x_1)\}^{-\theta} + \{F_2(x_2)\}^{-\theta} - 1]^{-1}$ is ith element of reversed hazard gradient of Clayton's bivariate distribution.

Example 4. Consider Gumbel-Hougaard distribution with PDF

$$f^*(x_1, x_2) = \frac{f_1(x_1) f_2(x_2)}{F_1(x_1) F_2(x_2)} \{-\ln F_1(x_1)\}^{\theta-1} \{-\ln F_2(x_2)\}^{\theta-1} \{\kappa(x_1, x_2)\}^{\frac{1}{\theta}-2}$$
$$\times [\{\kappa(x_1, x_2)\}^{\frac{1}{\theta}} + \theta - 1] \exp\left[-\{\kappa(x_1, x_2)\}^{\frac{1}{\theta}}\right], \quad \theta \geqslant 1$$

in which $\kappa(x_1, x_2) = [\{-\ln F_1(x_1)\}^{\theta} + \{-\ln F_2(x_2)\}^{\theta}]$ and f_1 and f_2 are PDFs of distribution functions F_1 and F_2, respectively. It can be shown that

$$\frac{\partial^2}{\partial x_1 \partial x_2} \log f^*(x_1, x_2) \geqslant 0.$$

Let Ω_F be a set of bivariate distributions with marginal distribution functions F_1 and F_2 and reversed hazard gradient $R(x_1, x_2) = (r_1(x_1, x_2), r_2(x_1, x_2))$ satisfying the following inequalities

$$\beta_i(x_1, x_2) \leqslant \{-\ln F_i(x_i)\}^{\theta-1} \{\kappa(x_1, x_2)\}^{\frac{1}{\theta}-1}, \qquad \theta \geqslant 1, \ i = 1, 2,$$

where $\beta_i(x_1, x_2) = \frac{r_i(x_1, x_2)}{r_i(x_i)}, i = 1, 2$ and $r_i(x_i)$ is reversed hazard function of $F_i(x_i), i = 1, 2$. The Gumbel-Hougaard distribution is ME in Ω_F. The validity of this result follows from part (a) of Theorem 4 because $r_i(x_i)\{-\ln F_i(x_i)\}^{\theta-1}\{\kappa(x_1, x_2)\}^{\frac{1}{\theta}-1}$ is ith element of reversed hazard gradient of Gumbel-Hougaard distribution.

4 ME Models for Mixtures

Let $\mathcal{G} = \{G_\theta(x_1, x_2) \in \mathcal{M}(G_1, G_2), \theta \in \chi\}$ be a family of bivariate distribution functions, where χ is a subset of the real line and θ is dependence parameter between G_1 and G_2. In the Bayesian context the parameter θ is assumed to be the realization of a random variable Θ with support in χ. In the following we assume that Θ has distribution function H. The distribution H is known as the prior distribution. The mixture of $\mathcal{G}$ with respect to prior distribution H, which is also known as the predictive distribution function, is defined as

$$F(x_1, x_2) = \int_\chi G_\theta(x_1, x_2) dH(\theta), \quad (x_1, x_2) \in R^2. \tag{9}$$

In this section we study the ME models in class of predictive models in bivariate setup for which the constraints are made on hazard rate (reversed hazard rate) of prior distribution $H(\theta)$. The key result is given in the following theorem.

***Theorem* 5.** *Let $\Omega_H = \{H : \lambda_H(\theta) \leqslant (\geqslant) q(\theta)\}$ be the set of proper prior distributions with support in χ and hazard function $\lambda_H(\theta)$. Consider Ω_F as a set of predictive distributions of the form (9) in which $H(\theta) \in \Omega_H$ and $G_\theta(x_1, x_2) \in \mathcal{G}$ is given. Suppose that there exists a prior distribution $H^* \in \Omega_H$ such that $\lambda_{H^*}(\theta) = q(\theta)$ and let F^* with PDF f^* is predictive distribution with respect to H^*. If elements of hazard gradient G_θ are decreasing in θ then*

(a) F^* is ME in Ω_F relative to Ω_H with $\lambda_H(\theta) \leqslant q(\theta)$ if $f^*(x_1, x_2)$ is TP_2.

(b) F^* is ME in Ω_F relative to Ω_H with $\lambda_H(\theta) \geqslant q(\theta)$ if $f^*(x_1, x_2)$ is RR_2.

Proof. We prove part (a). The proof of (b) is similar. Let $F(x_1, x_2)$ and $F^*(x_1, x_2)$ be of the form (9) in which prior distributions have hazard functions $\lambda_H(\theta)$ and $\lambda_{H^*}(\theta)$, respectively. If $\lambda_H(\theta) \leqslant \lambda_{H^*}(\theta)$ and elements of hazard gradient G_θ are decreasing in θ then from Theorem 6.D.5 of Shaked and Shanthikumar (2007),

$$\Lambda_F(x_1, x_2) \leqslant \Lambda_{F^*}(x_1, x_2), \text{ for all } (x_1, x_2),$$

where $\Lambda_F(x_1, x_2)$ and $\Lambda_{F^*}(x_1, x_2)$ are hazard gradient of F and F^*, respectively. Now we show that $F(x_1, x_2) \in \mathcal{M}(G_1, G_2)$. Let $f_1(x)$ and $f_2(x)$ be marginal PDFs of $F(x_1, x_2)$ and $g_i(x)$ be PDF of $G_i(x)$, $i = 1, 2$. Then using Fubini's Theorem, we have

$$\begin{aligned} f_i(x_i) &\equiv \int \int g_\theta(x_1, x_2) dH(\theta) dx_j \\ &= \int \int g_\theta(x_1, x_2) dx_j dH(\theta) \\ &= g_i(x_i), \qquad i \neq j, i, j = 1, 2. \end{aligned}$$

If $f^*(x_1, x_2)$ is TP_2 then using part (a) of Theorem 1 the proof is complete. □

An application of Theorem 5 is given in the following example.

Example 5.

(a) Let Ω_H be the set of prior distributions with support in $(0, 1)$ having hazard function $\lambda_H(\theta)$ such that

$$\lambda_H(\theta) \leqslant \frac{1}{1-\theta}, \qquad \theta \in (0, 1).$$

Consider FGM bivariate distribution with $\alpha = \theta$ and $0 < \alpha < 1$. Let Ω_F be the class of mixtures of the family of FGM bivariate distributions with respect to prior distributions $H(\theta) \in \Omega_H$. Uniform mixture of FGM bivariate distribution, which is FGM bivariate distribution with $\alpha = \frac{1}{2}$ is ME in Ω_F. This result is obtained by noting that $\frac{1}{1-\theta}$ is the hazard function of uniform distribution on $(0, 1)$. Also it can be easily seen that elements of hazard gradient of FGM distribution are decreasing in α and PDF of FGM distribution is TP_2 for $\alpha = \frac{1}{2}$. Thus, part (a) of Theorem 5 gives the result.

(b) Using the same arguments used to prove part (a) and based on part (b) of Theorem 5, if $\Omega_H = \{H(\theta) : \lambda_H(\theta) \geqslant \frac{1}{1-\theta}, \quad \theta \in (-1, 0)\}$ then FGM distribution with $\alpha = \frac{-1}{2}$ is ME in Ω_F.

There are bivariate distributions that their hazard gradients are increasing in dependence parameter. To characterize the mixture of these distributions, in the following we present some results. Before giving the main result we need the following Lemma.

***Lemma* 2.** *Let $\{G_\theta, \theta \in \chi\}$ be a family of bivariate distribution functions. Suppose Θ_1 and Θ_2 as two random variables with supports in χ and distribution functions H_1 and H_2, respectively. Let $\mathbf{Y}_1$ and $\mathbf{Y}_2$ be two bivariate random vectors such that $\mathbf{Y}_i =_{st} \mathbf{X}(\Theta_i), i = 1, 2$. That is, suppose that the survival function of $\mathbf{Y}_i$ is given by*

$$\bar{F}_i(\mathbf{x}) = \int_\chi \bar{G}_\theta(\mathbf{x}) dH_i(\theta), \quad \mathbf{x} \in R^2, \ i = 1, 2.$$

If

$$\mathbf{X}(\theta) \geqslant_{whr} \mathbf{X}(\theta') \quad \textit{whenever} \quad \theta \leqslant \theta' \tag{10}$$

and if

$$\Theta_1 \leqslant_{rh} \Theta_2, \tag{11}$$

then

$$\mathbf{Y}_1 \geqslant_{whr} \mathbf{Y}_2.$$

Proof. Assumption (10) means that for $i = 1, 2$, $\bar{G}_\theta(x_1, x_2)$, as a function of $\theta \in \chi$ and $x_i \in R$ is RR_2. Assumption (11) means that $H_i(\theta)$, as a function of $i \in \{1, 2\}$ and θ is TP_2. Also from (10), $\bar{G}_\theta(y)$ is decreasing in θ. Therefore using Remark of Theorem 2.1 of Joag-Dev et al. (1995), $\bar{F}_i(x_1, x_2)$ is RR_2 in $i \in \{1, 2\}$ and in $x_i, \ i = 1, 2$. That is

$$\frac{\bar{F}_1(x_1, x_2)}{\bar{F}_2(x_1, x_2)} \text{ is increasing in } (x_1, x_2)$$

which implies $\mathbf{Y}_1 \geqslant_{whr} \mathbf{Y}_2$. This completes the proof of the Lemma. □

The same as Theorem 5, using Lemma 2, we have the following result in which $r_H(\theta)$ is reversed hazard function of any distribution in Ω_H and $r(\theta)$ is reversed hazard function of H^*.

***Theorem* 6.** *Let elements of hazard gradient G_θ be increasing in θ.*

(a) If $f^*(x_1, x_2)$ is TP_2 then F^* is ME in Ω_F relative to $\Omega_H = \{H : r_H(\theta) \leqslant r(\theta)\}$.

(b) If $f^*(x_1, x_2)$ is RR_2 then F^* is ME in Ω_F relative to $\Omega_H = \{H : r_H(\theta) \geqslant r(\theta)\}$.

Example 6. Consider the survival function of uniform mixture of Generalized Gumbel distribution as follows

$$\begin{aligned}\bar{F}^*(x_1, x_2) &= \int_0^1 \bar{F}_1(x_1)\bar{F}_2(x_2)e^{-\delta \ln \bar{F}_1(x_1) \ln \bar{F}_2(x_2)} d\delta \\ &= \frac{\bar{F}_1(x_1)\bar{F}_2(x_2)\{1 - e^{-\ln \bar{F}_1(x_1) \ln \bar{F}_2(x_2)}\}}{\ln \bar{F}_1(x_1) \ln \bar{F}_2(x_2)}.\end{aligned}$$

It is easy to see that elements of hazard gradient of Generalized Gumbel distribution are increasing in δ. Also it can be shown that the PDF of $F^*(x_1, x_2)$ is RR_2. Let Ω_H be the set of prior distributions with reversed hazard functions r_H that satisfy

$$r_H(\theta) \geqslant \theta^{-1}.$$

Suppose that Ω_F is the set of mixtures of Generalized Gumbel distribution with respect to priors in Ω_H. Then uniform mixture of Generalized Gumbel distribution is ME in Ω_F. The validity of this result follows from part (b) of Theorem 6 because θ^{-1} is reversed hazard function of uniform distribution on (0,1).

Example 7. Lee (1996) considered Sarmanov distribution as follows

$$f^*(x_1, x_2) = f_1(x_1)f_2(x_2)\{1 + \theta(x_1 - \mu_1)(x_2 - \mu_2)\} \tag{12}$$

where $\mu_1 = E(X_1),\ \mu_2 = E(X_2)$ such that X_1, X_2 are random variables with respect to PDFs f_1, f_2. In this case the range of θ is

$$\max\left\{\frac{-1}{\mu_1\mu_2}, \frac{-1}{(1-\mu_1)(1-\mu_2)}\right\} \leqslant \theta \leqslant \min\left\{\frac{1}{\mu_1(1-\mu_2)}, \frac{1}{\mu_2(1-\mu_1)}\right\}.$$

Also the right-hand side of inequality (8) is

$$\frac{1 + \theta(x_i - \mu_i)\{\mu_j(x_j) - \mu_j\}}{1 + \theta\{\mu_1(x_1) - \mu_1\}\{\mu_2(x_2) - \mu_2\}}, \quad j = 1, 2$$

in which $\mu_j(x) = E(X_j|X_j > x), j = 1, 2$. It can be easily seen that $f^*(x_1, x_2)$ is TP_2 if

$$\mu_1 \geqslant 1,\ \mu_2 < 0 \quad \text{or} \quad \mu_2 \geqslant 1,\ \mu_1 < 0.$$

Also $f^*(x_1, x_2)$ is RR_2 if

$$\mu_1,\ \mu_2 > 1 \quad \text{or} \quad \mu_1,\ \mu_2 < 0.$$

The elements of hazard gradient of Sarmanov distribution with PDF (12) are decreasing if

$$\{\mu_i(x_i) - \mu_i\}\{\mu_j(x_j) - x_j\} > 0, \quad i \neq j, i, j = 1, 2 \text{ for all } x_i, x_j$$

and are increasing if

$$\{\mu_i(x_i) - \mu_i\}\{\mu_j(x_j) - x_j\} < 0, \quad i \neq j, i, j = 1, 2 \text{ for all } x_i, x_j,$$

in which $\mu_i(x) - x = E(X_i - x|X_i > x), i = 1, 2$ is mean residual lifetime function of X_i. It is easy to see that $\mu_i(x_i) - x_i \geqslant 0$. Then elements of hazard gradient of Sarmanov distribution are decreasing (increasing) if $\{\mu_i(x_i) - \mu_i\} > (<) 0, i = 1, 2$. Let Ω_H be the set of prior distributions $H(\theta)$ whose hazard functions $\lambda_H(\theta)$ or reversed hazard functions $r_H(\theta)$ satisfy some constraints. Also let $H(\theta)$ have support in $\{a(\mu_1, \mu_2), b(\mu_1, \mu_2)\}$ where

$$a(\mu_1, \mu_2) = \max\left\{\frac{-1}{\mu_1\mu_2}, \frac{-1}{(1-\mu_1)(1-\mu_2)}\right\}$$

and

$$b(\mu_1, \mu_2) = \min\left\{\frac{1}{\mu_1(1-\mu_2)}, \frac{1}{\mu_2(1-\mu_1)}\right\}.$$

Let Ω_F be class of mixtures of (12) with respect to prior distribution $H \in \Omega_H$.

(a) Suppose that

$$\lambda_H(\theta) \geqslant \frac{1}{\{b(\mu_1, \mu_2) - \theta\}}, \quad a(\mu_1, \mu_2) < \theta < b(\mu_1, \mu_2) \tag{13}$$

in which $\mu_1, \mu_2 > 1$. Uniform mixture of Sarmanov distribution, in which $\{\mu_i(x_i) - \mu_i\} > 0, i = 1, 2$ and $\mu_1, \mu_2 > 1$, with PDF

$$f^*(x_1, x_2) = f_1(x_1) f_2(x_2)\Big\{1 + \frac{a(\mu_1, \mu_2) + b(\mu_1, \mu_2)}{2} \times (x_1 - \mu_1)(x_2 - \mu_2)\Big\} \tag{14}$$

is ME in Ω_F. This result is obtained from part (b) of Theorem 5. To see this it can be shown that right-hand side of (13) is hazard function of uniform distribution on $\{a(\mu_1, \mu_2), b(\mu_1, \mu_2)\}$.

(b) Suppose that

$$r_H(\theta) \geqslant \frac{1}{\{\theta - a(\mu_1, \mu_2)\}}, \quad a(\mu_1, \mu_2) < \theta < b(\mu_1, \mu_2) \tag{15}$$

in which $\mu_1, \mu_2 < 0$. Uniform mixture of Sarmanov distribution with PDF (14), in which $\{\mu_i(x_i) - \mu_i\} < 0, i = 1, 2$, is ME in Ω_F. This result is obtained from part (b) of Theorem 6 because right-hand side of (15) is reversed hazard function of uniform distribution on $\{a(\mu_1, \mu_2), b(\mu_1, \mu_2)\}$.

Acknowledgement

The authors thank the editor and two referees for helpful and constructive comments which have improved the presentation of the paper. This work is supported in part by the Ordered and Spacial Data Center of Excellence of Ferdowsi University of Mashhad.

References

Asadi, M., Ashrafi, S., Ebrahimi, N. and Soofi, E.S. (2010). Models based on partial information about survival and hazard gradient. *Probability in the Engineering and Informational Sciences.* **24**, 561-584.

Asadi, M., Ebrahimi, N., Hamedani, G.G. and Soofi, E.S. (2004). Maximum dynamic entropy models. *J. Appl. Prob.* **41**, 379-390.

Asadi, M., Ebrahimi, N., Hamedani, G.G. and Soofi, E.S. (2005). Minimum dynamic discrimination information models. *J. Appl. Prob.* **42**, 643-660.

Holland, P.W. and Wang, Y.J. (1987). Dependence function for continuous bivariate densities. *Communications in Statistics Theory and Methods.* **16**, 863-876.

Jaynes, E.T. (1957). Information theory and statistical mechanics. *Physics Rev.* **106**, 620-630.

Joag-Dev, K., Kochar, S. and Proschan, F. (1995). A general composition theorem and its applications to certain partial orderings of distributions. *Statistics and Probability Letters.* **22**, 111-119.

Johnson, N.L. and Kotz, S. (1975). A vector multivariate hazard rate. *Journal of Multivariate Analysis.* **5**, 53-66.

Khaledi, B. and Kochar, S. (2005). Dependence, dispersiveness, and multivariate hazard rate ordering. *Probability in the Engineering and Information Sciences*. **19**, 427-446.

Lai, C.D. and Xie, M. (2006). *Stochastic Ageing and Dependence for Reliability*. Springer.

Lee, M.L.T. (1996). Properties and applications of the Sarmanov family of bivariate distributions. *Communications in Statistics Theory and Methods*. **25**, 1207-1222.

Marshall, A.W. (1975). Some comments on the hazard gradient. *Stochastic Processes and Their Applications*. **3**, 293-300.

Sarmanov, O.V. (1966). Generalized normal correlation and two-dimensional Frechét classes. *Doklady (Soviet Mathematics)*. **168**, 596-599.

Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*. Springer.

S. Ashrafi
Department of Statistics,
University of Isfahan,
Isfahan, 81744,
Iran,
email: *s.ashrafi@sci.ui.ac.ir*

M. Asadi
Department of Statistics,
University of Isfahan,
Isfahan, 81744,
Iran,
email: *m.asadi@sci.ui.ac.ir*