MULTIVARIATE DISTRIBUTIONS GENERATED FROM MIXTURES OF CONVOLUTION AND PRODUCT FAMILIES

BY ALBERT W. MARSHALL[1,2] AND INGRAM OLKIN[1]

University of British Columbia and Western Washington University, and Stanford University

A number of standard univariate distributions can be represented as mixtures of other standard distributions. In this paper such mixture representations are exploited to generate families of multivariate distributions with given marginals. Attention is confined to mixtures of parametric families where the parameter appears as the order of a convolution or as a power of the distribution or survival function. The mixture structure yields properties of the generated multivariate distributions such as total positivity, association and infinite divisibility. Examples obtained include the bivariate Poisson, binomial, negative binomial, normal, chi-square, logistic and Pareto distributions.

1. Introduction. For any given parametric family of distributions $F(\cdot \mid \theta)$, it is possible to regard the parameter θ as the value of a random variable Θ with distribution G, say. Then $F(\cdot \mid \theta)$ is a conditional distribution given $\Theta = \theta$ and the corresponding unconditional distribution

$$H(x) = \int F(x \mid \theta) dG(\theta) \tag{1}$$

is a mixture.

Here, both x and θ can be vectors, often of different dimensions. Many examples arise in which θ is a scalar and $F(\cdot \mid \theta)$ is the product of its marginals. Then (1) takes the form

$$H(\mathbf{x}) = \int \Pi F_i(x_i \mid \theta) dG(\theta), \tag{2}$$

[1]Supported in part by the National Science Foundation, and

[2] The Natural Sciences and Engineering Research Council of Canada.

AMS Classification. 60E05, 62E10.

Key words and phrases. Associated random variables, total positivity, multivariate distributions with given marginals, infinite divisibility, bivariate Poisson distribution, bivariate binomial distribution, bivariate negative binomial distribution, bivariate Pareto distribution.

where G and each F_i is a univariate distribution function. Clearly, multivariate distributions H of the form (1) or (2) have univariate marginals of the form (2). Of course many univariate examples are well known.

This paper is concerned with properties and examples of mixtures of the form (1) or (2) for two kinds of parametric families $\{F(\cdot \mid \theta) : \theta \in A\}$ which we call "convolution families" and "product families." These families arise in a natural fashion which is here described in a univariate setting.

Suppose that $X_1, \cdots, X_n$ are independent random variables with common distribution F, and suppose that $\Theta \geq 0$ is a random nonnegative integer having distribution G. Denote the θ-th convolution of F with itself by $F^{\theta *}$. With the conventions that $F^{\theta *}$ is degenerate at 0 and that an empty sum is 0, the random variable $U = X_1 + \cdots + X_\Theta$ has the distribution function H given by

$$H(x) = \int_0^\infty F^{\theta *}(x) dG(\theta). \tag{3}$$

In case F is infinitely divisible, (3) has meaning and H is a distribution function whenever G satisfies $G(0-) = 0$. Distributions of the form (3) are often called *compound distributions* (see, e.g., Feller, 1968, p. 286).

Again, suppose that $X_1, X_2, \cdots$ are independent random variables with common distribution F and suppose this time that $\Theta \geq 1$ is a random positive integer having distribution G. Then the random variables

$$V = \min(X_1, \cdots, X_\Theta), \quad W = \max(X_1, \cdots, X_\Theta)$$

have respective distributions H and K given by

$$\bar{H}(x) = \int \bar{F}^\theta(x) dG(\theta), \tag{4}$$

$$K(x) = \int F^\theta(x) dG(\theta), \tag{5}$$

where for any distribution function L, $\bar{L}$ is the corresponding survival function. The mixtures (3)–(5) give rise to the following definition.

DEFINITION 1.1. Let $\mathcal{F} = \{F(\cdot \mid \theta) : \theta \in A\}$ be an indexed family of n-dimensional distributions with index set $A \subset \mathcal{R}^k$ satisfying

$$\alpha \in A,\ \beta \in A \Rightarrow \alpha + \beta \in A. \tag{6}$$

$\mathcal{F}$ is said to be a *convolution family* if

$$F(\cdot \mid \alpha) * F(\cdot \mid \beta) = F(\cdot \mid \alpha + \beta), \quad \alpha, \beta \in A; \tag{7}$$

$\mathcal{F}$ is said to be a *survival product family* if

$$\bar{F}(\cdot \mid \alpha)\, \bar{F}(\cdot \mid \beta) = \bar{F}(\cdot \mid \alpha + \beta), \quad \alpha, \beta \in A; \tag{8}$$

finally $\mathcal{F}$ is said to be a *distribution product family* if

$$F(\cdot \mid \alpha)\, F(\cdot \mid \beta) = F(\cdot \mid \alpha + \beta), \quad \alpha, \beta \in A. \tag{9}$$

Convolution, survival product and distribution product families can be defined as semi-groups under the appropriate operation without the aid of an index set, but in this paper the index plays an important role.

EXAMPLE 1.2. For any distribution F, $\{F^{k*} : k = 0, 1, \ldots\}$ is a convolution family. In case F is a Bernoulli distribution with parameter p, F^{k*} is a binomial distribution with parameters k and p, $k = 0, 1, \ldots$.

More generally, if $F_1, \ldots, F_\ell$ is a finite collection of distributions each having a support in $\mathcal{R}^n$, then the set of all distributions of the form

$$F_1^{k_1*} * \cdots * F_\ell^{k_\ell*}$$

is a convolution family.

EXAMPLE 1.3. The prototype survival product family is the family of univariate exponential distributions. For any univariate distribution function F, $\{\bar{F}^\theta \mid \theta > 0\}$ is a survival product family of distributions with proportional hazards. Some, but not all, bivariate distributions can be used in the same way to generate a survival product family of bivariate distributions (see Theorem 3.4).

In the study of both convolution families (Section 2) and product families (Section 3), the notions of total positivity and association play an important role. Some results concerning these notions are reviewed in an Appendix (Section 4).

2. Convolution Families.

Convolution families involve infinite divisibility as well as the dependency property of total positivity (see Section 4). These properties sometimes carry over to mixtures and sometimes can be easily obtained from mixture representations.

Convolution families combine in obvious ways to give new convolution families.

OBSERVATION 2.1. If $\{F_{(i)}^{\theta*} : \theta \in A_i\}$ is a convolution family of k_i-variate distributions, $i = 1, 2$, then the distributions of the form

$$F(\mathbf{x}_1, \mathbf{x}_2 \mid \theta_1, \theta_2) = F_{(1)}^{\theta_1*}(\mathbf{x}_1)\, F_{(2)}^{\theta_2*}(\mathbf{x}_2), \quad \theta_1 \in A_1, \ \theta_2 \in A_2,$$

constitute a convolution family of $(k_1 + k_2)$-variate distributions.

OBSERVATION 2.2. If $\{F_{(i)}^{\theta*} : \theta \in A_i\}$ is a convolution family of n-variate distributions, $i = 1, 2$, then the distributions of the form

$$F(\mathbf{x} \mid \theta_1, \theta_2) = (F_{(1)}^{\theta_1*} * F_{(2)}^{\theta_2*})(\mathbf{x}), \quad \theta_1 \in A_1, \ \theta_2 \in A_2$$

form a convolution family of n-variate distributions.

For the study of mixtures of convolution families, a basic fact is that a convolution of mixtures is a mixture of convolutions.

LEMMA 2.3. *If* $H_{(i)}(\mathbf{x}) = \int F_{(i)}(\mathbf{x} \mid \theta) dG_i(\theta), \quad i = 1, 2$ *and*

$$F(\mathbf{x} \mid \theta, \eta) = \int F_{(1)}(\mathbf{x} - \mathbf{t} \mid \theta) dF_{(2)}(\mathbf{t} \mid \eta),$$

then

$$(H_{(1)} * H_{(2)})(\mathbf{x}) = \int \int F(\mathbf{x} \mid \theta, \eta)\, dG_{(1)}(\theta) dG_{(2)}(\eta).$$

PROOF.

$$\begin{aligned}
& (H_{(1)} * H_{(2)})(\mathbf{x}) \\
= & \int H_{(1)}(\mathbf{x} - \mathbf{z})\, dH_{(2)}(\mathbf{z}) = \int_{\mathbf{z}} \int_{\theta} F_{(1)}(\mathbf{x} - \mathbf{z} \mid \theta)\, dG_{(1)}(\theta)\, dH_{(2)}(\mathbf{z}) \\
= & \int_{\theta} \int_{\mathbf{z}} F_{(1)}(\mathbf{x} - \mathbf{z} \mid \theta)\, dH_{(2)}(\mathbf{z})\, dG_{(1)}(\theta) = \int_{\theta} \int_{\mathbf{z}} H_{(2)}(\mathbf{x} - \mathbf{z})\, dF_{(1)}(\mathbf{z} \mid \theta)\, dG_{(1)}(\theta) \\
= & \int_{\theta} \int_{\mathbf{z}} \int_{\eta} F_{(2)}(\mathbf{x} - \mathbf{z} \mid \eta)\, dG_{(2)}(\eta)\, dF_{(1)}(\mathbf{z} \mid \theta)\, dG_{(1)}(\theta) \\
= & \int_{\theta} \int_{\eta} F(\mathbf{x} \mid \theta, \eta)\, dG_{(2)}(\eta)\, dG_{(1)}(\theta). \quad \|
\end{aligned}$$

Infinite Divisibility in Convolution Families.

LEMMA 2.4. *If* $\{F(\cdot \mid \theta) : \theta > 0\}$ *is a convolution family then* $F(\cdot \mid 1)$ *is infinitely divisible and*

$$F(\cdot \mid \theta) = F^{\theta *}(\cdot \mid 1).$$

PROOF. Let ϕ_θ be the characteristic function of $F(\cdot \mid \theta), \theta > 0$. From (7) it follows that

$$\phi_\theta\, \phi_\eta = \phi_{\theta+\eta}, \quad \theta, \eta > 0.$$

This functional equation has the solution $\phi_\theta = \phi_1^\theta$ (Aczél, 1966, p. 36). ||

In the following theorem, the assumption is made that the index set A is a convex cone. A subset $\mathcal{T}$ of a convex cone is said to be a *frame* for the cone if $\mathcal{T}$, but no proper subset of $\mathcal{T}$, spans the cone positively.

THEOREM 2.5. *If* $\{F(\cdot \mid \theta) : \theta \in A\}$ *is a convolution family indexed by a convex cone* A*, then for all* $\theta \in A$*,* $F(\cdot \mid \theta)$ *is infinitely divisible. If the convex cone* A *has a finite frame* $\mathcal{T} = \{t_1, \ldots, t_\ell\}$*, then* F *is of the form*

$$F(\cdot \mid \theta) = F^{\theta_1 *}(\cdot \mid t_1) * \cdots * F^{\theta_\ell *}(\cdot \mid t_\ell),$$

where $\theta = \sum_1^\ell \theta_i t_i$.

PROOF. This result follows from (7) and Lemma 2.4. ||

Various versions of the following results can be found in the literature. Lemma 2.6 is given for the univariate case by Keilson and Steutel (1974, p. 116). Theorem 2.8 in one dimension is due to Feller (1971, p. 538). With the assumption that $F(\cdot \mid \theta)$ is infinitely divisible, Theorem 2.8 is given in a very general setting by Kent (1981), who also lists some additional relevant references.

LEMMA 2.6. *If* $\{F(\cdot \mid \theta) : \theta \in A\}$ *is a convolution family of multivariate distributions and*

$$H_{(i)}(\mathbf{x}) = \int F(\mathbf{x} \mid \theta)\, dG_{(i)}(\theta), \quad i = 1, 2,$$

then

$$(H_1 * H_2)(\mathbf{x}) = \int F(\mathbf{x} \mid \theta)\, d(G_1 * G_2)(\theta).$$

PROOF. In Lemma 2.3, take $F_{(1)} = F_{(2)} = F$. Since $\{F(\cdot \mid \theta) : \theta \in A\}$ is a convolution family, the distribution $F(\cdot \mid \theta, \eta)$ of Lemma 2.3 is just $F(\cdot \mid \theta + \eta)$ and consequently the result follows from Lemma 2.3. ||

THEOREM 2.7. *If* $\{F(\cdot \mid \theta) : \theta \in A\}$ *and* $\{G(\cdot \mid \alpha) : \alpha \in B\}$ *are convolution families, and if*

$$H(\mathbf{x} \mid \alpha) = \int F(\mathbf{x} \mid \theta)\, dG(\theta \mid \alpha), \quad \alpha \in B, \; \mathbf{x} \in \mathcal{R}^n, \tag{10}$$

then $\{H(\cdot \mid \alpha), \alpha \in B\}$ *is a convolution family.*

PROOF. This is immediate from Lemma 2.6. ||

THEOREM 2.8. *If* $\{F(\cdot \mid \theta) : \theta \in A\}$ *is a convolution family and* G *is an infinitely divisible distribution, then* H *given by (10) is infinitely divisible. Moreover,*

$$H^{\alpha *}(\mathbf{x}) = \int F(\mathbf{x} \mid \theta)\, dG^{\alpha *}(\theta), \quad \alpha > 0, \; \mathbf{x} \in \mathcal{R}^n. \tag{11}$$

PROOF. Suppose that for some positive integer m, $\alpha = 1/m$ so that $(G^{\alpha *})^{m *} = G$. If $H^{\alpha *}$ is defined by (11) then by Lemma 2.6 $(H^{\alpha *})^{m *} = H$. This proves that H is infinitely divisible and (11) is satisfied for $\alpha = 1/m$, $m = 1, 2, \ldots$. Again from Lemma 2.6, it follows immediately that (11) holds for rational α and the proof is completed by a limiting argument. ||

Total Positivity in Convolution Families.

The following theorem shows how the positive dependency notion of total positivity (see Section 4) arises in the context of convolution families. There, $\tilde{F}$ can be taken to be either F or $\bar{F}$. For typographical simplicity, $\overline{F^{\theta *}}$ is written $\bar{F}^{\theta *}$.

THEOREM 2.9.

(i) If $\tilde{F}(x \mid \theta) = \tilde{F}^{\theta *}(x)$, $\theta = 0,1,2,\ldots$ where F is a univariate distribution function such that $F(0-) = 0$ and $\tilde{F}(x-y)$ is TP_2 in x and y, then $\tilde{F}(x \mid \theta)$ is TP_2 in x and θ.

(ii) If $\tilde{F}(x \mid \theta) = \tilde{F}^{\theta *}(x)$, $\theta \geq 0$, where F is a univariate infinitely divisible distribution function such that $F(0-) = 0$ and $\tilde{F}^{\theta *}(x-y)$ is TP_2 in x and y for all θ, then $\tilde{F}(x \mid \theta)$ is TP_2 in x and θ.

(iii) If $F(x \mid \theta) = F^{\theta *}(x)$, $\theta = 0,1,2,\ldots$ where F is a univariate distribution function with a density f such that $f(x-y)$ is TP_2 in x and y, then $F(\cdot \mid \theta)$ has a density $f(\cdot \mid \theta)$ which is TP_2 in x and $\theta = 0,1,\ldots$.

PROOF. Let $x_1 < x_2$ and $\theta_1 < \theta_2$, and suppose that $\tilde{F}$ is $\bar{F}$. Then

$$\begin{vmatrix} \bar{F}(x_1 \mid \theta_1) & \bar{F}(x_1 \mid \theta_2) \\ \bar{F}(x_2 \mid \theta_1) & \bar{F}(x_2 \mid \theta_2) \end{vmatrix} = \begin{vmatrix} \bar{F}^{\theta_1 *}(x_1) & \bar{F}^{\theta_2 *}(x_1) \\ \bar{F}^{\theta_1 *}(x_2) & \bar{F}^{\theta_2 *}(x_2) \end{vmatrix}$$

$$= \int_0^\infty \begin{vmatrix} \bar{F}^{\theta_1 *}(x_1) & \bar{F}^{\theta_1 *}(x_1 - u) \\ \bar{F}^{\theta_1 *}(x_2) & \bar{F}^{\theta_1 *}(x_2 - u) \end{vmatrix} df^{(\theta_2 - \theta_1)*}(u) \geq 0$$

because $\bar{F}(x-y)$ is TP_2 in x and y implies that the same is true for $\bar{F}^{\theta *}$, $\theta = 0,1,\ldots$ (Barlow and Proschan, 1975, p. 100), and this means the integrand is nonnegative. The proofs for other cases are similar. For a proof of (iii), see Karlin (1968, p. 150). ||

THEOREM 2.10. *Let*

$$\tilde{H}(\mathbf{x}) = \int \Pi \tilde{F}^{\theta *}(x_i)\, dG(\theta) \tag{12}$$

where each F_i is a univariate distribution function such that $F_i(0-) = 0$.

(i) If for each i, $\tilde{F}_i(x-y)$ is TP_2 in x and y, then $\tilde{H}$ is MTP_2.

(ii) If for each i, $f_i(x-y)$ is TP_2 in x and y, then H has a density h that is MTP_2.

PROOF. This is immediate from Theorem 2.9 and Theorem 4.15. ||

Examples

EXAMPLE 2.11. BIVARIATE POISSON DISTRIBUTION. In the mixture (2), let F_i be a binomial distribution with parameters (θ, p_i), where p_i is fixed, and suppose that G is a Poisson distribution with parameter η. It is well known that this mixture has Poisson marginals. With $q_i = 1 - p_i, i = 1, 2$, the probability mass function of this mixture is given by

$$h(k,\ell) = \sum_{\theta=0}^{\infty} \binom{\theta}{k} p_1^k q_1^{\theta-k} \binom{\theta}{\ell} p_2^\ell q_2^{\theta-\ell} e^{-\eta} \frac{\eta^\theta}{\theta!}, \quad k, \ell = 0, 1, \ldots . \tag{13}$$

Since $\binom{x}{y}$ is TP_∞ in x and $y = 0, 1, \ldots$ (Karlin, 1968, p. 137), it follows from the basic composition theorem for totally positive functions that $h(k, \ell)$ is TP_∞ in k and $\ell = 0, 1, \ldots$. Consequently, by Corollary 4.8 it follows that variables having the probability mass function (13) are associated. From Theorem 2.8, it follows that the bivariate Poisson distribution of (13) is also infinitely divisible.

A different construction of a bivariate Poisson distribution starts with independent random variables U_1, U_2 and Θ having Poisson distributions with respective parameters λ_1, λ_2 and λ_{12}. If $X_i = U_i + \Theta$, $i = 1, 2$, then (X_1, X_2) has the bivariate Poisson distribution of M'Kendrick (see Marshall and Olkin, 1985). The joint probability mass function of (X_1, X_2) is

$$\begin{aligned} h(k,\ell) &= \sum_{\theta=0}^{\infty} P(U_1 + \theta = k)\, P(U_2 + \theta = \ell)\, e^{-\lambda_{12}} \frac{\lambda_{12}^\theta}{\theta!} \\ &= \sum_{\theta} \frac{\lambda_{12}^\theta \lambda_1^{k-\theta} \lambda_2^{\ell-\theta}}{\theta!(k-\theta)!(\ell-\theta)!} e^{-(\lambda_1+\lambda_2+\lambda_{12})}, \quad k, \ell = 0, 1, \ldots . \end{aligned} \tag{14}$$

Dwass and Teicher (1957) show that this is the only infinitely divisible bivariate Poisson distribution, so it is reassuring to note by comparing Laplace transforms that when $\lambda_1 = \eta p_1 q_2$, $\lambda_2 = \eta p_2 q_1$ and $\lambda_{12} = \eta p_1 p_2$, (13) and (14) define the same distribution.

EXAMPLE 2.12. BIVARIATE NEGATIVE BINOMIAL DISTRIBUTION. It is well known that if F_1 and F_2 are Poisson distributions with respective parameters $\alpha\theta$ and $\beta\theta$, and if G is a gamma distribution with shape parameter r and scale parameter $\lambda = 1$, then the mixture (2) has negative binomial marginals and probability mass function

$$
\begin{aligned}
h(k,\ell \mid r) &= \int_0^\infty \frac{(\alpha\theta)^k e^{-\alpha\theta}}{k!} \frac{(\beta\theta)^\ell e^{-\beta\theta}}{\ell!} \frac{\theta^{r-1}e^{-\theta}}{\Gamma(r)} d\theta \\
&= \frac{\Gamma(k+\ell+r)}{k!\,\ell!\,\Gamma(r)} \left(\frac{\alpha}{\alpha+\beta+1}\right)^k \left(\frac{\beta}{\alpha+\beta+1}\right)^\ell \left(\frac{1}{\alpha+\beta+1}\right)^r, \\
& \qquad k,\ell = 0,1,\ldots .
\end{aligned}
$$

The above derivation of this bivariate negative binomial distribution is due to Arbous and Kerrich (1951). It follows from Theorem 2.10 that $h(k,\ell \mid r)$ is MTP_2; by Corollary 4.8, this means the distribution is associated, so that the correlation is non-negative. By Theorem 2.8, $h(k,\ell \mid r)$ is infinitely divisible and moreover

$$h(\cdot \mid r_1) * h(\cdot \mid r_2) = h(\cdot \mid r_1 + r_2), \quad r_1, r_2 > 0.$$

EXAMPLE 2.13. BIVARIATE NORMAL DISTRIBUTION. If F_1 and F_2 are normal distributions with means μ and $\alpha\mu$, ($\alpha = \pm 1$) and variances σ_1^2 and σ_2^2, respectively, and if μ has a $\mathcal{N}(0,\sigma_0^2)$ distribution, then the mixture (2) has density function

$$
\begin{aligned}
h(x,y) &= \frac{1}{(2\pi)^{3/2}\sigma_0\sigma_1\sigma_2} \int_{-\infty}^\infty [\exp -\frac{1}{2}\{\frac{(x-\mu)^2}{\sigma_1^2} + \frac{(y-\alpha\mu)^2}{\sigma_2^2} + \frac{\mu^2}{\sigma_0^2}\}] d\mu \\
&= \frac{|\Sigma|^{-\frac{1}{2}}}{(2\pi)} \exp[-\frac{1}{2}(x,y)\Sigma^{-1}(x,y)'],
\end{aligned}
$$

where

$$\Sigma = \frac{1}{d}\begin{pmatrix} \sigma_0^2 + \sigma_1^2 & \alpha\sigma_0^2 \\ \alpha\sigma_0^2 & \sigma_0^2 + \sigma_2^2 \end{pmatrix}$$

and $d = (\sigma_1^2\sigma_2^2 + \sigma_0^2\sigma_1^2 + \sigma_0^2\sigma_2^2)/[(\sigma_0^2+\sigma_1^2)(\sigma_0^2+\sigma_2^2) - \alpha^2\sigma_0^4]$.

The choice $\alpha = +1$ yields a positive correlation and $\alpha = -1$ yields a negative correlation.

EXAMPLE 2.14. A BIVARIATE CHI-SQUARE DISTRIBUTION. If F_1 and F_2 are gamma distributions with common shape parameter $m+2\theta$ and with respective scale parameters $\frac{1}{2}\lambda_1$ and $\frac{1}{2}\lambda_2$ and G is a negative binomial distribution, then the mixture (2) has density function

$$
\begin{aligned}
& h(x,y \mid m) \\
&= \sum_{\ell=0}^\infty \left(\frac{\lambda_1^{m+\ell} x^{m+\ell-1} e^{-\frac{1}{2}\lambda_1 x}}{2^{m+\ell}\Gamma(m+\ell)}\right) \left(\frac{\lambda_2^{m+\ell} y^{m+\ell-1} e^{-\frac{1}{2}\lambda_2 y}}{2^{m+\ell}\Gamma(m+\ell)}\right) \frac{\Gamma(m+\ell)(\rho^2)^\ell(1-\rho^2)^m}{\Gamma(m)\ell!}.
\end{aligned}
$$

This is the joint density of the sample variances s_{11} and s_{12} from a sample of size $n = m/2$ from a bivariate normal distribution with inverse covariance matrix $\Sigma^{-1} = (\sigma^{ij})$, $i, j = 1, 2$. Here $\lambda_1 = \sigma^{11}, \lambda_2 = \sigma^{22}$ and $\rho = \sigma^{12}/\sqrt{\sigma^{11}\sigma^{22}}$ is the correlation.

Since the gamma density is TP_2, it follows from Theorem 2.10 that $h(x, y \mid m)$ is MTP_2, and hence, by Corollary 4.8, is associated. Because the negative binomial distribution is infinitely divisible, by Theorem 2.8 $h(x, y \mid m)$ is infinitely divisible.

EXAMPLE 2.15. BIVARIATE NON-CENTRAL CHI-SQUARE DISTRIBUTIONS. Because a non-central chi-square distribution is a Poisson mixture of central chi-square distributions, a bivariate non-central chi-square distribution can formally be obtained using (2). The corresponding density is

$$h(x,y) = \sum_{\theta=0}^{\infty} \frac{x^{\frac{n}{2}+\theta-1}e^{-\frac{x}{2}}}{\Gamma(\frac{n}{2}+\theta)2^{\frac{n}{2}+\theta}} \frac{y^{\frac{m}{2}+\theta-1}e^{-\frac{y}{2}}}{\Gamma(\frac{m}{2}+\theta)2^{\frac{m}{2}+\theta}} \frac{\alpha^\theta e^{-\alpha}}{\theta!}, \quad x, y \geq 0.$$

It follows from Theorem 2.10 that $h(x, y)$ is MTP_2 and hence (by Corollary 4.8) is associated.

A possibly more meaningful bivariate non-central chi-square distribution is obtained from the representation

$$X = U_1^2 + Z_1, \quad Y = U_2^2 + Z_2,$$

where U_1, U_2 and (Z_1, Z_2) are independently distributed with $U_i \sim \mathcal{N}(\mu_i, 1/\lambda_{ii}), i = 1, 2$ and (Z_1, Z_2) has the bivariate chi-square distribution of Example 2.14. The corresponding bivariate chi-square density function is

$$h(x,y \mid n) = \sum_{\ell=0}^{\infty} \sum_{\theta_1=0}^{\infty} \sum_{\theta_2=0}^{\infty} f_1(x \mid \ell, \theta_1)\, f_2(y \mid \ell, \theta_2)\, g(\ell \mid \rho^2) \frac{e^{-\alpha_1}\alpha_1^{\theta_1}}{\theta_1!} \frac{e^{-\alpha_2}\alpha_2^{\theta_2}}{\theta_2!},$$

where $\alpha_i = \lambda_{ii}\mu_i^2/2$, $i = 1, 2, g(\ell \mid \rho^2)$ is negative binomial distribution of 2.14, and

$$f_i(t \mid \ell, \theta_i) = \frac{\lambda_{ii}^{(n+2\ell+2\theta_i+1)/2} t^{(n+2\ell+2\theta_i+1)/2} e^{-\lambda_{ii}t/2}}{2^{(n+2\ell+2\theta_i)/2}\Gamma((n+2\ell+2\theta_i)/2)},$$

$i = 1, 2$. When $\rho = 0$ we obtain the independent case with X and Y each having a non-central chi-square distribution.

The association of $h(x, y)$ follows from the representation of X and Y given above, or from Theorem 2.10. By Theorem 2.8, $h(x, y \mid n)$ is infinitely divisible.

Multivariate Extension of Convolution Families.

Let $\mathcal{F}$ be an indexed family of distributions and let $\mathcal{S}$ be the class of nonempty subsets of $\{1, \ldots, n\}$. For each $S \in \mathcal{S}$, let U_S have a distribution $F(\cdot \mid \theta_S)$ in $\mathcal{F}$. Suppose that the random variables $U_S, S \in \mathcal{S}$ are independent, and let

$$X_i = \sum_{S:i \in S} U_S, \quad i = 1, \ldots, n. \tag{15}$$

Denote by $\mathcal{F}^{n*}$ the family of n-dimensional distributions for random vectors of the form $(X_1, \ldots, X_n)$. Clearly distributions in $\mathcal{F}^{n*}$ have $n-1$ dimensional marginals in $\mathcal{F}^{(n-1)*}$. If $\mathcal{F}$ is a convolution family, then distributions in $\mathcal{F}^{n*}$ have one dimensional marginals in $\mathcal{F}$; in particular, X_i has distribution $F(\cdot \mid \sum_{S:i \in S} \theta_S)$, $i = 1, \ldots, n$. From (15) it is clear that distributions in $\mathcal{F}^{n*}$ are associated. The family $\mathcal{F}^{n*}$ has various other desirable properties (see Marshall and Shaked, 1986).

THEOREM 2.16. *If $\mathcal{F} = \{F(\cdot \mid \theta) : \theta \in A\}$ is a convolution family of univariate distributions, then $\mathcal{F}^{n*}$ is a convolution family (indexed by A^{2^n-1}).*

PROOF. Clearly A^{2^n-1} satisfies (6). Suppose X and Y are independent random vectors with respective distributions $F(\cdot \mid \alpha)$ and $F(\cdot \mid \beta)$ in $\mathcal{F}^{n*}$. Then there exist independent random variables $U_S, V_S, S \in \mathcal{S}$, such that

$$X_i = \sum_{S:i \in S} U_S, \;\; Y_i = \sum_{S:i \in S} V_S, \;\; X_i + Y_i = \sum_{S:i \in S} (U_S + V_S), \;\; i = 1, \ldots, n.$$

If U_S and V_S have respective distributions $F(\cdot \mid \alpha_S)$ and $F(\cdot \mid \beta_S)$ in $\mathcal{F}$, then by (7) $U_S + V_S$ has the distribution $F(\cdot \mid \alpha_S + \beta_S)$. Thus

$$F(\cdot \mid \alpha) * F(\cdot \mid \beta) = F(\cdot \mid \alpha + \beta). \quad \|$$

THEOREM 2.17. *If F is infinitely divisible and $\mathcal{F} = \{F^{\theta *} : \theta \geq 0\}$, then distributions in $\mathcal{F}^{n*}$ are infinitely divisible.*

The proof of this result is similar to the proof of Theorem 2.16.

EXAMPLE 2.18. BIVARIATE BINOMIAL AND POISSON DISTRIBUTIONS. Let $X_i = U_i + U_{12}, i = 1, 2$, where U_1, U_2 and U_{12} are independent random variables with distributions in $\mathcal{F}$. Then the joint distribution of X_1 and X_2 is in $\mathcal{F}^{2*}$. When $\mathcal{F}$ consists of binomial distributions, (X_1, X_2) has the bivariate binomial distribution of Wicksell (see Marshall and Olkin, 1985); it follows from Theorem 4.13 that such distributions form a convolution family. When $\mathcal{F}$ consists of the Poisson distributions then (X_1, X_2) has the bivariate Poisson distribution of Example 2.11; it follows from Theorem 2.17 that such distributions are infinitely divisible. See also Example 2.20.

Mixtures of distributions in $\mathcal{F}^{n*}$ can take various forms, but discussion here is confined to the case that the parameters $\theta_S, S \in \mathcal{S}$ are independent random variables with respective distributions $G(\cdot \mid \alpha_S)$, $S \in \mathcal{S}$. Denote the joint distribution of $X_1, \ldots X_n$ given $\theta_S, S \in \mathcal{S}$, by $F^*(\cdot \mid \theta_S,\ S \in \mathcal{S})$.

The next observation says that the operations of mixing and of extending $\mathcal{F}$ to $\mathcal{F}^{n*}$ commute. Alternatively, it can be viewed as saying that the structure exhibited in (15) is preserved under mixing.

PROPOSITION 2.19. *Let $\mathcal{F} = \{F(\cdot \mid \theta) : \theta \in A\}$, and let $\mathcal{H} = \{H(\cdot \mid \alpha) : \alpha \in B\}$ be the family of distributions of the form*

$$H(x \mid \alpha) = \int F(x \mid \theta) dG(\theta \mid \alpha),$$

where $F \in \mathcal{F}$ and $\mathcal{G} = \{G(\cdot \mid \alpha) : \alpha \in B\}$ is a family of distributions having support contained in A. Then $\mathcal{H}^{n*}$ consists of distributions having the form

$$H^{n*}(\mathbf{x} \mid \alpha_T, T \in \mathcal{S}) = \int F^*(\mathbf{x} \mid \theta_S,\ S \in \mathcal{S}) \prod_{S \in \mathcal{S}} dG(\theta_S \mid \alpha_T). \tag{16}$$

PROOF. Let $U_S, S \in \mathcal{S}$ be independent random variables with distributions in $\mathcal{F}$ such that

$$X_\ell = \sum_{\ell \in S} U_S, \quad \ell = 1, \ldots, n$$

have joint distribution F^*. Denote the characteristic function of U_S by ϕ_S. Then $X_1, \ldots, X_n$ have joint characteristic function

$$Ee^{i\Sigma t_\ell X_\ell} = Ee^{i\Sigma_{S \in \mathcal{S}} \tau_S U_S} = \Pi_{S \in \mathcal{S}} \phi_S(\tau_S),$$

where $\tau_S = \Sigma_{\ell : \ell \in S} t_\ell$. Consequently, the characteristic function of H^{n*} is

$$\int \prod_{S \in \mathcal{S}} \phi_S(\tau_S) \prod_{S \in \mathcal{S}} dG(\theta_S \mid \alpha_T) = \prod_{S \in \mathcal{S}} \int \phi_S(\tau_S)\, dG(\theta_S \mid \alpha_T). \tag{17}$$

Now let $V_T, T \in \mathcal{S}$ be independent random variables such that V_T has the distribution $H(\cdot \mid \alpha_T)$ and let

$$Y_j = \sum_{T : j \in T} V_T, \ j = 1, \ldots, n.$$

Then $Y_1, \ldots, Y_n$ have a joint distribution in $\mathcal{H}^{n*}$ and joint characteristic function given by (17). ||

EXAMPLE 2.20. A BIVARIATE NEGATIVE BINOMIAL DISTRIBUTION. Let $X_1 = U_1 + U_{12}, X_2 = U_2 + U_{12}$, where U_1, U_2 and U_{12} are independent random variables having Poisson distributions with respective parameters θ_{10}, θ_{01} and θ_{11} (i.e., X_1 and X_2 have the bivariate Poisson distribution of Example 2.11). Suppose that θ_{10}, θ_{01} and θ_{11} are independent random variables having gamma distributions with respective parameters $(\alpha_{10}, \lambda), (\alpha_{01}, \lambda)$ and (α_{11}, λ). Then for $k, \ell = 0, 1, \ldots,$

$$h(k,\ell) = \int\int\int\sum_j \frac{\theta_{11}^j \theta_{10}^{k-j} \theta_{01}^{\ell-j}}{j!(k-j)!(\ell-j)!} e^{-(\theta_{10}+\theta_{01}+\theta_{11})}$$
$$\frac{\lambda^{\alpha_{10}}\theta_{10}^{\alpha_{10}-1}e^{-\lambda\theta_{10}}}{\Gamma(\alpha_{10})} \frac{\lambda^{\alpha_{01}}\theta_{01}^{\alpha_{01}-1}e^{-\lambda\theta_{01}}}{\Gamma(\alpha_{01})} \frac{\lambda^{\alpha_{11}}\theta_{11}^{\alpha_{11}-1}e^{-\lambda\theta_{11}}}{\Gamma(\alpha_{11})} d\theta_{01}d\theta_{10}d\theta_{11}$$

$$= \sum_j \frac{\lambda^\alpha}{\Gamma(\alpha_{10})\Gamma(\alpha_{01})\Gamma(\alpha_{11})} \frac{\Gamma(\alpha_{10}+k-j-1)\Gamma(\alpha_{01}+\ell-j-1)\Gamma(\alpha_{11}+j-1)}{j!(k-j)!(\ell-j)!(\lambda+1)^{\alpha+k+\ell-j}}$$
$$= \sum_j \frac{\Gamma(\alpha_{10}+k-j-1)}{\Gamma(\alpha_{10})(k-j)!} \frac{\Gamma(\alpha_{01}+\ell-j-1)}{\Gamma(\alpha_{01})(\ell-j)!} \frac{\Gamma(\alpha_{11}+j-1)}{\Gamma(\alpha_{11})j!} p^\alpha (1-p)^{k+\ell-j},$$

where $\alpha = \alpha_{11}+\alpha_{10}+\alpha_{01}$, $p = \lambda/(\lambda+1)$.

This distribution has negative binomial marginals; with $\alpha_{10} = \alpha_{01} = 1-\alpha_{11}$, this is a bivariate geometric distribution.

From Proposition 2.19, it follows that this negative binomial distribution is in $\mathcal{F}^{2*}$ when $\mathcal{F}$ consists of negative binomial distributions with fixed parameter p.

3. Product Families. Results which follow are stated for survival product families, but it should be understood that parallel results hold for distribution product families.

OBSERVATION 3.1. If $\{F_{(i)}(\cdot \mid \theta) : \theta \in A_i\}$ is a survival product family, $i = 1, 2$, then distributions of the form

$$\bar{F}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \mid \theta_1, \theta_2) = \bar{F}_1(\mathbf{x}_1, \mathbf{x}_2 \mid \theta_1)\, \bar{F}_2(\mathbf{x}_2, \mathbf{x}_3 \mid \theta_2), \quad \theta_1 \in A_1,\ \theta_2 \in A_2$$

constitute a survival product family.

LEMMA 3.2. *A survival product family of distributions $F(\cdot \mid \theta)$ indexed by $A = (0, \infty)$ must be a proportional hazard family, that is,*

$$\bar{F}(\cdot \mid \theta) = \bar{F}^\theta(\cdot \mid 1), \quad \theta > 0. \tag{18}$$

PROOF. Fix $\mathbf{x}$ and let $\phi(\theta) = \bar{F}(\mathbf{x} \mid \theta)$. From (8) it follows that

$$\phi(\alpha+\beta) = \phi(\alpha)\, \phi(\beta), \quad \alpha, \beta > 0.$$

Since ϕ is bounded, it follows that for some real number $\gamma, \phi(\alpha) = e^{-\gamma\alpha}$, that is $\phi(\alpha) = [\phi(1)]^\alpha$. But this is (18). ||

THEOREM 3.3. *If $\{F(\cdot \mid \theta) : \theta \in A\}$ is a survival product family indexed by a convex cone $A \subset \mathcal{R}^k$ with finite frame $\mathcal{T} = \{t_1, \ldots, t_\ell\}$ then*

$$\bar{F}(\mathbf{x} \mid \theta) = \prod_{i=1}^{\ell} \bar{F}^{\theta_i}(\mathbf{x} \mid t_i),$$

where $\theta = \sum_1^\ell \theta_i t_i$.

PROOF. This is immediate from (8) and Lemma 3.2. $\|$

Proportional Hazard Families in Higher Dimensions.

If $\bar{F}$ is a univariate survival function, then for all $\theta > 0$, $\bar{F}^\theta$ is also a univariate survival function. In the multivariate case, $\bar{F}^k$, $k = 1, 2, \ldots$ is always a survival function, but $\bar{F}^\theta$ is a survival function for all $\theta > 0$ only in special circumstances.

THEOREM 3.4. *Let $\bar{F}$ be a bivariate survival function. Then $\bar{F}^\theta$ is a bivariate survival function for all $\theta > 0$ if and only if $\bar{F}(x,y)$ is TP_2 in (x,y).*

PROOF. Suppose first that $\bar{F}^\theta$ is a survival function for all $\theta > 0$. Then for all $\epsilon, \delta \geq 0$,

$$\xi(\theta) = \bar{F}^\theta(x,y) - \bar{F}^\theta(x, y+\epsilon) - \bar{F}^\theta(x+\delta, y) + \bar{F}^\theta(x+\delta, y+\epsilon) \geq 0.$$

Since $\xi(\theta) \geq 0$ for all $\theta > 0$ and $\xi(0) = 0$, it follows that the derivative $\xi'(0) \geq 0$; but this is just the condition

$$\frac{\bar{F}(x,y)\bar{F}(x+\delta, y+\epsilon)}{\bar{F}(x, y+\epsilon)\bar{F}(x+\delta, y)} \geq 1$$

that $\bar{F}$ is TP_2.

Next, suppose that $\bar{F}$ is TP_2. With $R(x,y) = -\log \bar{F}(x,y)$, this condition can be written in the form

$$R(x+\delta, y+\epsilon) \leq R(x, y+\epsilon) + R(x+\delta, y) - R(x,y) \quad \text{for all } \epsilon, \delta \geq 0 \text{ and all } x, y. \tag{19}$$

Note that $R(x+\delta, y) - R(x,y) \geq 0$ and write

$$R(x, y+\epsilon) - R(x,y) = [R(x, y+\epsilon) + R(x+\delta, y) - R(x,y)] - R(x+\delta, y).$$

Since $\psi(x) = e^{-x}$ is decreasing and convex, it follows that for all $\theta > 0$,

$$\begin{aligned} e^{-\theta R(x,y)} - e^{-\theta R(x,y+\epsilon)} &\geq e^{-\theta R(x+\delta,y)} - e^{-\theta[R(x,y+\epsilon)+R(x+\delta,y)-R(x,y)]} \\ &\geq e^{-\theta R(x+\delta,y)} - e^{-\theta R(x+\delta,y+\epsilon)}; \end{aligned}$$

the last inequality because ψ is decreasing and (19) holds. But this says that $\bar{F}^\theta$ is a survival function. ||

As noted in Section 4, the condition that $\bar{F}$ is TP_2 is a positive dependency property which implies that the correlation is non-negative (when it exists).

REMARK 3.5. If F is a bivariate distribution with density f such that $f(x,y)$ is TP_2 in x and y, then it follows from Theorem 4.9 that $\bar{F}(x,y)$ is also TP_2 in x and y. Several examples from Section 2 have this property, including 2.11, 2.12. See also Examples 3.8 and 3.9.

Mixtures of Product Families.

LEMMA 3.6. *If $\bar{F}_i^\theta$ is a survival function for all θ in the support of G_i and $\bar{H}_i(\mathbf{x}) = \int \bar{F}_i^\theta(\mathbf{x})dG_i(\theta)$, $i = 1, 2$, then*

$$\bar{H}_1(\mathbf{x})\,\bar{H}_2(\mathbf{y}) = \int \bar{F}^\theta(\mathbf{x})\,d(G_1 * G_2)(\theta).$$

PROOF. Write $F^\theta(\mathbf{x}) = e^{-\theta R(\mathbf{x})}$, where $R(\mathbf{x}) = -\log \bar{F}(\mathbf{x})$. Then the result is easily seen to be a reflection of the fact that the Laplace transform of a convolution is the product of Laplace transforms. ||

THEOREM 3.7. *Let $\{G(\cdot \mid \alpha) : \alpha \in B\}$ be a convolution family of distributions such that for each α, $G(\cdot \mid \alpha)$ has support contained in A. Let $\{F(\cdot \mid \theta) : \theta \in A\}$ be a survival product family of distributions such that $F(x \mid \theta)$ is measurable in θ for each fixed x.*

If

$$H(x \mid \alpha) = \int F(x \mid \theta)\,dG(\theta \mid \alpha), \quad \alpha \in B,$$

then

$$\bar{H}(x \mid \alpha + \beta) = \bar{H}(x \mid \alpha)\,\bar{H}(x \mid \beta) \text{ for all } \alpha, \beta \in B, \; -\infty < x < \infty. \tag{20}$$

PROOF. This is immediate from Lemma 3.6. ||

EXAMPLE 3.8. MULTIVARIATE LOGISTIC DISTRIBUTION. If F_i are iterated exponential extreme value distributions for minima, that is, $\bar{F}_i(x_i \mid \theta) = \exp\{-\theta e^{x_i}\}, -\infty < x_i < \infty$, $\theta > 0$, $i = 1, \ldots, n$, and if G is a gamma distribution with shape parameter r and scale parameter λ, then the distribution H of (2) takes the form

$$\bar{H}(\mathbf{x} \mid r) = \lambda^r / (\lambda + \sum_1^n e^{x_i})^r, \quad \lambda, r > 0.$$

EXAMPLE 3.9. MULTIVARIATE PARETO DISTRIBUTIONS. If

$$\bar{F}_i(x_i) = \exp\left\{-\theta\left(\frac{x_i - \mu_i}{\sigma_i}\right)^{1/\gamma_i}\right\}, \quad x_i \geq \mu_i, \quad i = 1, \ldots, n$$

are Weibull survival functions and if G is a $\text{Gam}(\alpha, 1)$ distribution then the mixture (2) is a Type IV multivariate Pareto distribution as defined by Arnold (1983). This mixture has survival function

$$\text{(21)} \qquad \bar{H}(\mathbf{x} \mid \alpha) = \left[1 + \sum_{i=1}^{n}\left(\frac{x_i - \mu_i}{\sigma_i}\right)^{1/\gamma_i}\right]^{-\alpha}, \quad x_i \geq \mu_i, \quad i = 1, \ldots, n.$$

The representation of (21) as a mixture was used with a minor variation by Takahasi (1965) to define a multivariate Burr distribution.

It follows from Theorem 4.15 that the density corresponding to H is TP_∞ in each pair of arguments, the other arguments being fixed. Consequently the distribution is associated.

Lemma 3.6 and Theorem 3.7 can be generalized to allow the F_i to involve different sets of variables.

LEMMA 3.10. *Let $\mathbf{x}^{(i)}$ be a subvector of $\mathbf{x}$ of dimension n_i, and let F_i be a distribution fuction of dimension n_i such that $\bar{F}_i^\theta$ is a survival function for all $\theta \in A_i$, $i = 1, \ldots, k$. Let G_j be a distribution of dimension k such that $G_j(A_1 \times \cdots \times A_k) = 1$, $j = 1, 2$. If*

$$\bar{H}_j(\mathbf{x}) = \int \prod_{i=1}^{k} \bar{F}_i^{\theta_i}(\mathbf{x}^{(i)})\, dG_j(\theta), \quad j = 1, 2,$$

then

$$\bar{H}_1(\mathbf{x})\, \bar{H}_2(\mathbf{x}) = \int \prod_{i=1}^{k} \bar{F}_i^{\theta_i}(\mathbf{x}^{(i)})\, d(G_1 * G_2)(\theta).$$

PROOF. Let $R_i(\mathbf{x}^{(i)}) = -\log \bar{F}_i(\mathbf{x}^{(i)})$ so that

$$\bar{H}_j(\mathbf{x}) = \int \exp\{-\Sigma_{i=1}^{k}\theta_i R_i(\mathbf{x}^{(i)})\}\, dG_j(\theta).$$

Then

$$\begin{aligned} \bar{H}_1(\mathbf{x})\, \bar{H}_2(\mathbf{x}) &= \int\int \exp\{-\sum_{i=1}^{k}(\theta_i + \eta_i)R_i(\mathbf{x}^{(i)})\}dG_1(\theta)dG_2(\eta) \\ &= \int \exp\{-\sum_{i=1}^{k}\theta_i R_i(\mathbf{x}^{(i)})\}d(G_1 * G_2)(\theta). \quad \| \end{aligned}$$

Special cases of particular interest include

(i) $k = 1$ and $\mathbf{x}^{(1)} = \mathbf{x}$,

(ii) $k = n$ and $\mathbf{x}^{(i)} = x_i$ where $\mathbf{x} = (x_1, \ldots, x_n)$.

(iii) $k = 2^n - 1$ and $\mathbf{x}^{(i)}$ ranges through the nonempty subvectors of $\mathbf{x}$.

THEOREM 3.11. *Let $\mathcal{G} = \{G(\cdot \mid \alpha),\ \alpha \in A\}$ be a convolution family of k-dimensional distribution functions such that*

$$\bar{G}(0 \mid \alpha) = 1 \quad \textit{for all} \quad \alpha \in A. \tag{22}$$

With the notation of Lemma 3.10, let

$$\bar{H}(\mathbf{x} \mid \alpha) = \int \prod_{i=1}^{k} \bar{F}_i^{\theta_i}(\mathbf{x}^{(i)})\, dG(\theta \mid \alpha), \quad \mathbf{x} \in \mathcal{R}^n,\ \alpha \in A.$$

Then $\mathcal{H} = \{H(\cdot \mid \alpha),\ \alpha \in A\}$ is a survival product family.

PROOF. This follows from Lemma 3.10. ||

Multivariate Extensions of Survival Product Families.

Let $\mathcal{F}$ be an indexed family of distributions and for each $S \in \mathcal{S}$, the non-empty subsets of $\{1, 2, \ldots, n\}$, let U_S have distribution $F(\cdot \mid U_S) \in \mathcal{F}$. Suppose that the random variables U_S, $S \in \mathcal{S}$ are independent, and let

$$X_i = \min_{S: i \in S} U_S, \quad i = 1, \ldots, n. \tag{23}$$

Then the joint survival function of the X_i's is given by

$$\bar{F}_{(n)}(\mathbf{x} \mid \theta_S,\ S \in \mathcal{S}) = \prod_{S \in \mathcal{S}} \bar{F}(\max_{j \in S} x_j \mid \theta_S). \tag{24}$$

The family $\bar{\mathcal{F}}_{(n)}$ of such distributions has various desirable properties (see Marshall and Shaked, 1986); in particular if $\mathcal{F}$ is a survival product family, then distributions in $\bar{\mathcal{F}}_{(n)}$ have univariate marginals in $\mathcal{F}$. If $\mathcal{F}$ is a survival product family indexed by $(0, \infty)$, then Lemma 3.2 applies and

$$\bar{F}_{(n)}(\mathbf{x} \mid \theta_S,\ S \in \mathcal{S}) = \prod_{S \in \mathcal{S}} \bar{F}^{\theta_S}(\max_{j \in S} x_j \mid 1). \tag{25}$$

EXAMPLE 3.12. If $X_1 = \min(U_1, Z)$ and $X_2 = \min(U_2, Z)$, where U_1, U_2 and Z are independent random variables with distributions in the family $\mathcal{F}$ then (X_1, X_2) has a distribution in $\bar{\mathcal{F}}_{(2)}$. When $\mathcal{F}$ consists of the exponential distributions, $\bar{\mathcal{F}}_{(2)}$ consists of the bivariate exponential distributions of Marshall and Olkin (1967).

For distribution product families, "min" in (23) is replaced by "max" and (24) is replaced by

$$F(\mathbf{x} \mid \theta) \equiv F_{(n)}(\mathbf{x} \mid \theta) = \prod_{S \in \mathcal{S}} F(\min_{j \in S} x_j \mid \theta_S), \tag{26}$$

and the family of such distributions is denoted by $\mathcal{F}_{(n)}$.

THEOREM 3.13. *If $\mathcal{F} = \{F(\cdot \mid \theta) : \theta \in A\}$ is a survival product family then $\bar{\mathcal{F}}_{(n)}$ is a survival product family; if $\mathcal{F}$ is a distribution product family, then $\mathcal{F}_{(n)}$ is a distribution product family.*

The next observation says that the operations of mixing and of extending $\mathcal{F}$ to $\bar{\mathcal{F}}_{(n)}$ commute (cf. Proposition 2.19).

PROPOSITION 3.14. *Let $\mathcal{F} = \{F(\cdot \mid \theta) : \theta \in A\}$, and let $\mathcal{H} = \{H(\cdot \mid \alpha) : \alpha \in B\}$ be the family of distributions of the form*

$$H(x \mid \alpha) = \int F(x \mid \theta)\, dG(\theta \mid \alpha),$$

where $F \in \mathcal{F}$ and $\mathcal{G} = \{G(\cdot \mid \alpha) : \alpha \in B\}$ is a family of distributions having support contained in A. Then $\bar{\mathcal{H}}_{(n)}$ consists of distributions having the form

$$\bar{H}_{(n)}(\mathbf{x} \mid \alpha_S, S \in \mathcal{S}) = \int \bar{F}_{(n)}(\mathbf{x} \mid \theta_S, S \in \mathcal{S}) \prod_{S \in \mathcal{S}} dG(\theta_S \mid \alpha_S).$$

PROOF.

$$\begin{aligned} \bar{H}_{(n)}(\mathbf{x} \mid \alpha_S, S \in \mathcal{S}) &= \int \prod_{S \in \mathcal{S}} \bar{F}(\max_{j \in S} x_j \mid \theta_S) \prod_{S \in \mathcal{S}} dG(\theta_S \mid \alpha_S) \\ &= \prod_{S \in \mathcal{S}} \int \bar{F}(\max_{j \in S} x_j \mid \theta_S)\, dG(\theta_S \mid \alpha_S) = \prod_{S \in \mathcal{S}} \bar{H}(\max_{j \in S} x_j \mid \alpha_S). \quad \| \end{aligned}$$

If in the above proposition, $\mathcal{F}$ is a survival product family and $\mathcal{G}$ is a convolution family, then it follows from Theorem 3.7 and Theorem 3.13 that $\bar{\mathcal{H}}_{(n)}$ is a survival product family, and of course this is the most interesting case.

EXAMPLE 3.15. MULTIVARIATE LOGISTIC DISTRIBUTIONS. Let $\mathcal{F}$ consist of the iterated exponential extreme value distributions for minima as in Example 3.8. Then $\mathcal{F}$ is a survival product family and $\bar{\mathcal{F}}_{(n)}$ consists of distributions of the form

$$\bar{F}(\mathbf{x} \mid \theta_S, S \in \mathcal{S}) = \exp\left[-\sum_{S \in \mathcal{S}} \theta_S \exp(\max_{j \in S} x_j)\right].$$

If the θ_S are independent and have gamma distributions with respective shape and scale parameters r_S and λ_S, then the mixture H of (1) is given by

$$\bar{H}(\mathbf{x} \mid r_S, \lambda_S, S \in \mathcal{S}) = \prod_{S \in \mathcal{S}} \left[\frac{\lambda_S}{\lambda_S + \exp(\max_{i \in S} x_i)} \right]^{r_S}.$$

If the parameters λ_S, $S \in \mathcal{S}$ are either 0 or are equal to $\lambda > 0$, say, then the mixing gamma distributions form a convolution family, and in this case the distribution $\bar{H}(\cdot \mid r_S, \lambda, S \in \mathcal{S})$ form a survival product family as expected.

4. Appendix: Association and Total Positivity. The following result of Ahmed, León and Proschan (1978) shows that the positive dependency property of association is preserved under mixing.

THEOREM 4.1. *Let H be a mixture given by* (4). *If*

(27) *for each fixed $\theta, F(x \mid \theta)$ is associated,*

(28) *G is associated,*

(29) *$\int \xi(\mathbf{x}) dF(\mathbf{x} \mid \theta)$ is increasing in θ for all increasing $\xi : \mathcal{R}^n \to \mathcal{R}$ such that the integral exists,*

then H is associated.

LEMMA 4.2. *Let H be a mixture given by* (2). *If*

$$\int \xi(x) dF_i(x \mid \theta) \text{ is increasing in } \theta \text{ for all increasing } \xi : \mathcal{R} \to \mathcal{R} \tag{30}$$

such that each integral exists, $i = 1, \ldots, n$, then (29) *holds where $F(\mathbf{x} \mid \theta) = \prod F_i(x_i \mid \theta)$.*

THEOREM 4.3. *Let H be a mixture given by* (2). *If G is associated and if condition (30) holds, then H is associated.*

COROLLARY 4.4. *Let H be a mixture given by* (2). *If*

$$F_i(x_i \mid \theta) \text{ is decreasing in } \theta \text{ for all } x_i, i = 1, \ldots, n, \tag{31}$$

then H is associated.

Total Positivity in Mixtures.

Total positivity is often encountered in mixtures (e.g., see Marshall and Olkin, 1979, Example 18.A.12). In the multivariate setting, multivariate total positivity in the following sense arises.

DEFINITION 4.5. (Karlin and Rinott, 1980). Let $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ where each $\mathcal{X}_i$ is totally ordered. For $\mathbf{x}, \mathbf{y} \in \mathcal{X}$, let

$$\mathbf{x} \vee \mathbf{y} = (\max(x_1, y_1), \ldots, \max(x_n, y_n)), \quad \mathbf{x} \wedge \mathbf{y} = (\min(x_1, y_1), \ldots, \min(x_n, y_n)).$$

A function $\psi : \mathcal{X} \to [0, \infty)$ is said to be *multivariate totally positive of order* 2 (MTP_2) if

$$\psi(\mathbf{x} \vee \mathbf{y})\, \psi(\mathbf{x} \wedge \mathbf{y}) \geq \psi(\mathbf{x})\, \psi(\mathbf{y}) \text{ for all } \mathbf{x}, \mathbf{y} \in \mathcal{X}.$$

PROPOSITION 4.6. (Kemperman, 1977; Karlin and Rinott, 1980, Proposition 2.1). *Suppose that $\psi : \mathcal{X} \to [0, \infty)$ is totally positive of order 2 (TP_2) in each pair of arguments, the remaining arguments being fixed. Suppose also that $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ and $\psi(\mathbf{x})\psi(\mathbf{y}) > 0$ implies $\psi(\mathbf{u}) > 0$ for all $\mathbf{u}$ such that $\mathbf{x} \wedge \mathbf{y} \leq \mathbf{u} \leq \mathbf{x} \vee \mathbf{y}$. Then ψ is MTP_2 on $\mathcal{X}$.*

THEOREM 4.7. (Karlin and Rinott, 1980, p. 472). *If $\mathbf{X} = (X_1 \ldots, X_n)$ has a joint density that is MTP_2 and if A and B are upper Borel sets in $\mathcal{R}^n$ (i.e., $a \in A$ and $a' \geq a \Rightarrow a' \in A$, and similarly for B), then*

$$P\{\mathbf{X} \in A \vee B\}\, P\{\mathbf{X} \in A \wedge B\} \geq P\{\mathbf{X} \in A\}\, P\{\mathbf{X} \in B\}, \tag{32}$$

where $A \vee B = \{\mathbf{u} = \mathbf{a} \vee \mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\}$ and $A \wedge B = \{\mathbf{u} = \mathbf{a} \wedge \mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\}$.

According to Theorem 3.1 of Esary, Proschan and Walkup (1967), random variables $X_1, \ldots, X_n$ are associated if and only if for all upper Borel sets in $\mathcal{R}^n$,

$$P\{\mathbf{X} \in A \cap B\} \geq P\{\mathbf{X} \in A\} P\{\mathbf{X} \in B\}. \tag{33}$$

Note that $A \cap B = A \wedge B$. A comparison of (32) and (33) provides a proof of the following.

COROLLARY 4.8. (Fortuin, Kastelyn and Ginibre, 1971; Karlin and Rinott, 1980, Theorem 4.2). *If $\mathbf{X} = (X_1, \ldots, X_n)$ has an MTP_2 joint density, then $X_1, \ldots, X_n$ are associated.*

To a large extent, joint densities which are MTP_2 arise in mixtures as a consequence of the following theorem.

THEOREM 4.9. (Karlin and Rinott, 1980, Proposition 3.4). *Let $\mathcal{X} = \Pi_{i=1}^n \mathcal{X}_i$, $\mathcal{Y} = \Pi_{i=1}^n \mathcal{Y}_i$ and $\mathcal{Z} = \Pi_{i=1}^n \mathcal{Z}_i$, where each $\mathcal{X}_i, \mathcal{Y}_i$ and $\mathcal{Z}_i$ is totally ordered. If ψ_1 is MTP_2 on $\mathcal{X} \times \mathcal{Y}$, ψ_2 is MTP_2 on $\mathcal{Y} \times \mathcal{Z}$ and if*

$$\psi(\mathbf{x}, \mathbf{z}) = \int_{\mathcal{Y}} \psi_1(\mathbf{x}, \mathbf{y})\, \psi_2(\mathbf{y}, \mathbf{z})\, \Pi d\sigma_i(y_i),$$

where each σ_i is σ-finite, then ψ is MTP_2 on $\mathcal{X} \times \mathcal{Z}$.

PROPOSITION 4.10.

(i) *If* $f(\mathbf{x} \mid \theta)$ *is* MTP_2 *in* $(\mathbf{x}, \theta)$ *and* g *is* MTP_2, *then*

$$h(\mathbf{x}) = \int f(\mathbf{x} \mid \theta) g(\theta) d\theta$$

is MPT_2.

(ii) *If* $f(\mathbf{x} \mid \theta)$ *is* MTP_2 *in* $(\mathbf{x}, \theta)$ *and if* $g(\theta \mid \alpha)$ *is* MTP_2 *in* (θ, α), *then*

$$h(\mathbf{x} \mid \alpha) = \int f(\mathbf{x} \mid \theta) g(\theta \mid \alpha) d\theta$$

is MTP_2 *in* $(\mathbf{x}, \alpha)$.

PROOF. This is an immediate consequence of Theorem 4.9.

PROPOSITION 4.11.

(i) *If* $h(\mathbf{x}) = \int \Pi_{i=1}^{n} f_i(x_i \mid \theta_i) g(\boldsymbol{\theta}) d\boldsymbol{\theta}$, *where each* $f_i(x_i \mid \theta_i)$ *is* TP_2 *in* (x_i, θ_i) *and if* g *is* MTP_2, *then* h *is* MTP_2.

(ii) *If* $h(\mathbf{x} \mid \boldsymbol{\alpha}) = \int \Pi_{i=1}^{n} f_i(x_i \mid \theta_i) g(\boldsymbol{\theta} \mid \boldsymbol{\alpha}) d\boldsymbol{\theta}$, *where each* $f_i(x_i \mid \theta_i)$ *is* TP_2 *in* x_i, θ_i, *and if* $g(\boldsymbol{\theta} \mid \boldsymbol{\alpha})$ *is* MTP_2 *in* $(\boldsymbol{\theta}, \boldsymbol{\alpha})$, *then* h *is* MTP_2 *in* $(\mathbf{x}, \boldsymbol{\alpha})$.

PROOF. By Proposition 4.6, $\Pi_i f_i(x_i \mid \theta_i)$ is MTP_2 in $(\mathbf{x}, \boldsymbol{\theta})$ so this result follows from Proposition 4.10. ||

PROPOSITION 4.12. *Let* $h(\mathbf{x}) = \int f(\mathbf{x} \mid \boldsymbol{\theta}) \Pi dG_i(\theta_i)$. *If* $f(\mathbf{x} \mid \boldsymbol{\theta})$ *is* MTP_2 *in* $(\mathbf{x}, \boldsymbol{\theta})$, *then* h *is* MTP_2.

PROOF. This follows directly from Theorem 4.9 or from Proposition 4.11. ||

Just as MTP_2 densities arise so do distribution functions and survival functions.

PROPOSITION 4.13.

(i) *Let* $H(\mathbf{x}) = \int F(\mathbf{x} \mid \boldsymbol{\theta}) g(\boldsymbol{\theta}) d\boldsymbol{\theta}$ *and suppose that* g *is* MTP_2. *If* $F(\mathbf{x} \mid \boldsymbol{\theta})$ *is* MTP_2 *in* $(\mathbf{x}, \boldsymbol{\theta})$, *then* H *is* MTP_2; *if* $\bar{F}(\mathbf{x} \mid \boldsymbol{\theta})$ *is* MTP_2 *in* $(\mathbf{x}, \boldsymbol{\theta})$, *then* $\bar{H}$ *is* MTP_2.

(ii) *Let* $H(\mathbf{x} \mid \boldsymbol{\alpha}) = \int F(\mathbf{x} \mid \boldsymbol{\theta})\, g(\boldsymbol{\theta} \mid \boldsymbol{\alpha}) d\boldsymbol{\theta}$ *and suppose that* $g(\boldsymbol{\theta} \mid \boldsymbol{\alpha})$ *is* MTP_2 *in* $(\boldsymbol{\theta}, \boldsymbol{\alpha})$. *If* $F(\mathbf{x} \mid \boldsymbol{\theta})$ *is* MTP_2 *in* $(\mathbf{x}, \boldsymbol{\theta})$, *then* $H(\mathbf{x} \mid \boldsymbol{\alpha})$ *is* MTP_2 *in* $(\mathbf{x}, \boldsymbol{\alpha})$. *A similar statement holds if* F *and* H *are replaced by* $\bar{F}$ *and* $\bar{H}$.

PROOF. These results follow from Theorem 4.9. ||

PROPOSITION 4.14. *Let* $H(\mathbf{x}) = \int F(\mathbf{x} \mid \boldsymbol{\theta})\Pi dG_i(\theta_i)$. *If* $F(\mathbf{x} \mid \boldsymbol{\theta})$ *is* MTP_2 *in* $(\mathbf{x}, \boldsymbol{\theta})$, *then* H *is* MTP_2. *A similar statement holds with* $\bar{F}$ *and* $\bar{H}$ *in place of* F *and* H.

PROOF. These results follow from Theorem 4.9. ||

Although higher order multivariate total positivity has to our knowledge not been defined, one counterpart to MTP_2 is the condition of higher order total positivity in pairs of arguments.

THEOREM 4.15. *If* $\bar{F}_i(x_i \mid \theta)$ *is totally positive of order* k *(*TP_k*) in* (x_i, θ), $i = 1, \ldots, n$, *then*

$$\bar{H}(\mathbf{x}) = \int \Pi_{i=1}^{n} \bar{F}_i(x_i \mid \theta)\, dG(\theta)$$

is TP_k *in each pair* x_j, x_ℓ, $1 \leq j,\ \ell \leq k$ $(j \neq \ell)$, *the other arguments being fixed. If* $F_i(x_i \mid \theta)$ *is* TP_k *in* (x_i, θ), $i = 1, \ldots, n$, *then*

$$H(\mathbf{x}) = \int \Pi_{i=1}^{n} F_i(x_i \mid \theta)\, dG(\theta)$$

is TP_k *in each pair* $x_j, x_\ell, j \neq \ell$, *the other arguments being fixed.*

If F_i *has a density* f_i *with respect to some measure that is* TP_k *in* (x_i, θ), $i = 1, \ldots, n$, *and if*

$$h(\mathbf{x}) = \int \prod_{i=1}^{n} f_i(x_i \mid \theta)\, dG(\theta),$$

then h *is* TP_k *in each pair* $x_j, x_\ell, j \neq \ell$, *the other arguments being fixed.*

PROOF. This is an immediate consequence of the basic composition formula (Karlin, 1968, p. 17). ||

It is not difficult to show that if h is TP_k in pairs of its arguments, then H and $\bar{H}$ both have this property.

When $n = 2$, even TP_2 is known to have useful implications.

A random vector (X_1, X_2) is said to be *right corner set increasing* (RCSI) if

$$P\{X_1 > x_1,\ X_2 > x_2 \mid X_1 > x_1',\ X_2 > x_2'\}$$

is increasing in x_1' and x_2' for all x_1, x_2. Shaked (1977) shows that the survival function of (X_1, X_2) is TP_2 if and only if (X_1, X_2) is RCSI. Barlow and Proschan (1975) show that if (X_1, X_2) is RCSI, then X_1 and X_2 are associated. By analogous arguments or by applying these results to $(-X_1, -X_2)$ it can be shown the distribution function of (X_1, X_2) is TP_2 if and only if (X_1, X_2) is *left corner set decreasing* (LCSD), i.e.,

$$P\{X_1 \le x_1,\ X_2 \le x_2 \mid X_1 \le x_1',\ X_2 \le x_2'\}$$

is decreasing in x_1' and x_2' for all x_1, x_2. Moreover, if (X_1, X_2) is LCSD then X_1 and X_2 are associated.

Thus we see that when $n = 2$, TP_2 of either the distribution function or the survival function implies association. For $n > 2$, corresponding results are false (C. Newman, 1986, private communication).

REFERENCES

ACZÉL, J. (1966). *Lectures on Functional Equations and Their Applications.* Academic Press, New York.

AHMED, L., LEÓN, R.V., and PROSCHAN, F. (1978). Generalized association with applications in multivariate statistics. Technical Report, Florida State University.

ARBOUS, A.G. and KERRICH, J.E. (1951). Accident statistics and the concept of accident proneness. *Biometrics* **7** 340–432.

ARNOLD, B.C. (1983). *Pareto Distributions.* International Co-operative Publishing House, Fairland, Maryland.

BARLOW, R.E. and PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing.* Holt, Rinehart and Winston, New York.

DWASS, M. and TEICHER, H. (1957). On infinitely divisible random vectors. *Ann. Math. Statist.* **28** 461–470.

ESARY, J.D., PROSCHAN, F. and WALKUP, D. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **38** 1466–1474.

FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications.* Vol. I (third edition), Wiley, New York.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications.* Vol. II (second edition), Wiley, New York.

FORTUIN, C.M., KASTELYN, P.W. and GINIBRE, J. (1971). Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.* **22** 89–103.

KARLIN, S. (1968). *Total Positivity,* Vol. 1. Stanford University Press, Stanford, CA.

KARLIN, S. and RINOTT, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Mult. Anal.* **10** 467–498.

KEILSON, J. and STEUTEL, F.W. (1974). Mixtures of distributions, moment inequalities and measures of exponentiality and normality. *Ann. Prob.* **2** 112–130.

KEMPERMAN, J.H.B. (1977). On the FKG-inequality for measures on a partially ordered space. *Indag. Math.* **39** 313–331.

KENT, J.T. (1981). Convolution mixtures of infinitely divisible distributions. *Math. Proc. Camb. Phil. Soc.* **90** 141–153.

MARSHALL, A.W. and OLKIN, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62** 30–44.

MARSHALL, A.W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.

MARSHALL, A.W. and OLKIN, I. (1985). A family of bivariate distributions generated by the bivariate Bernoulli distribution. *J. Amer. Statist. Assoc.* **80** 332–338.

MARSHALL, A.W. and SHAKED, M. (1986). Multivariate new better than used distributions. *Math. Oper. Res.* **11** 110–116.

SHAKED, M. (1977). A family of concepts of dependence for bivariate distributions. *J. Amer. Statist. Assoc.* **72** 642–650.

SHAKED, M. (1977). A concept of positive for exchangeable random variables. *Ann. Statist.* **5** 505–515.

TAKAHASI, K. (1965). Note on the multivariate Burr's distribution. *Ann. Inst. Statist. Math.* **17** 257–260.

DEPARTMENT OF STATISTICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, B.C.
CANADA V6T 1W5

STATISTICS AND EDUCATION DEPARTMENT
STANFORD UNIVERSITY
STANFORD, CA 94305