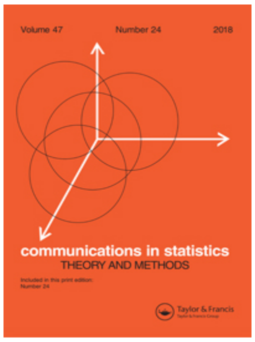

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Properties and applications of the sarmanov family of bivariate distributions

Mei-Ling Ting Lee

COMMUN. STATIST.—THEORY METH., 25(6), 1207-1222 (1996)

PROPERTIES AND APPLICATIONS OF THE SARMANOV FAMILY OF BIVARIATE DISTRIBUTIONS

Mei-Ling Ting Lee

Channing Laboratory, Harvard Medical School,
180 Longwood Avenue, Boston, Massachusetts 02115, U.S.A.

Keywords: Bayesian method, bivariate beta, bivariate Cauchy, bivariate exponential, bivariate Poisson, bivariate gamma, conjugate priors, correlated binary data, Farlie-Gumbel-Morgenstern distributions, bivariate proportional hazards distributions

ABSTRACT

We discuss properties of the bivariate family of distributions introduced by Sarmanov (1966). It is shown that correlation coefficients of this family of distributions have wider range than those of the Farlie-Gumbel-Morgenstern distributions. Possible applications of this family of bivariate distributions as prior distributions in Bayesian inference are discussed. The density of the bivariate Sarmanov distributions with beta marginals can be expressed as a linear combination of products of independent beta densities. This pseudo-conjugate property greatly reduces the complexity of posterior computations when this bivariate beta distribution is used as a prior. Multivariate extensions are derived.

1. INTRODUCTION

In this article we discuss properties of the Sarmanov (1966) bivariate density functions. This family of bivariate densities include some of the Farlie-Gumbel-Morgenstern (FGM; Farlie, 1960) distributions as special cases. Unlike uniterated FGM distributions, whose correlation coefficients are limited to the interval (-1/3, 1/3), we show that correlation coefficients of the Sarmanov's family of bivariate distributions have wider ranges. For example, the bivariate beta distribution can have correlation coefficients close to 1. Some general methods to construct bivariate densities are provided in section 4. Bivariate distributions with marginals belonging to the natural exponential family of distributions are discussed in section 5. Some interesting bivariate distributions with marginals distributed as Cauchy, exponential, normal, and Poisson are given in section 6. Applications for correlated binary data are discussed in section 7. It is shown that the Sarmanov's family of bivariate beta distribution can be expressed as a linear combination of products of univariate beta densities. Therefore the posterior distributions have the pseudo-conjugate property if this family of distributions are used as priors. There posterior computations in Bayesian inferences can be greatly simplified. In section 8 we consider multivariate extensions of the Sarmanov family of distributions.

2. THE SARMANOV FAMILY OF BIVARIATE DISTRIBUTIONS

Sarmanov (1966) described a family of bivariate densities. The family has not been studied further, perhaps because the work appears in a somewhat inaccessible journal. Assume that $f_1(x_1)$ and $f_2(x_2)$ are univariate probability density functions (p.d.f.), or probability mass functions (p.m.f.), with supports defined on $A_1 \subseteq R$, $A_2 \subseteq R$, respectively. Let $\phi_i(t)$, $i = 1, 2$, be bounded nonconstant functions such that $\int_{-\infty}^{\infty} \phi_i(t) f_i(t) dt = 0$. Then, the function defined by

$$h(x_1, x_2) = f_1(x_1) f_2(x_2) \{1 + \omega \phi_1(x_1) \phi_2(x_2)\} \tag{1}$$

is a bivariate joint density (or a probability mass function) with specified marginals $f_1(x_1)$ and $f_2(x_2)$, provided ω is a real number which satisfies the condition that $1 + \omega\phi_1(x_1)\phi_2(x_2) \geq 0$ for all x_1 and x_2. Sarmanov's family is a special case of Cohen's (1984) construction.

It is of interest to investigate properties of the Sarmanov family of distributions. In the following sections, we denote, for any given univariate density $f_i(x_i)$, $\mu_i = \int_{-\infty}^{\infty} t f_i(t)dt$, $\sigma_i^2 = \int_{-\infty}^{\infty}(t-\mu_i)^2 f_i(t)dt$, $\nu_i = \int_{-\infty}^{\infty} t\phi_i(t)f_i(t)dt$, and $\eta_i = \int_{-\infty}^{\infty} t^2\phi_i(t)f_i(t)dt$, if they exist, for $i = 1, \ldots, n$. As a result, moments and regressions of this family of distributions can be easily derived as below.

Theorem 1:

Assume that the vector (X_1, X_2) has a joint p.d.f. $h(x_1, x_2)$ as defined in (1).

(a) *The product moment is given by*

$$E[X_1X_2] = \mu_1\mu_2 + \omega\nu_1\nu_2 \tag{2}$$

(b) *The conditional distribution function of X_2 given $X_1 = x_1$ is*

$$P(X_2 \leq x_2 | X_1 = x_1) = F_2(x_2) - \omega\phi_1(x_1)\int_{x_2}^{\infty} f_2(t)\phi_2(t)dt. \tag{3}$$

where $F_2(x_2) = P(X_2 \leq x_2)$ is the marginal distribution function of X_2.

(c) The regression of X_2 on X_1 is given by

$$E[X_2 | X_1 = x_1] = \mu_2 + \omega\nu_2\phi_1(x_1) \tag{4}$$

Results in (b) show that the monotonicity of ϕ_i determines whether variables X_j and X_i are regression dependent, a concept introduced by Lehmann (1966).

3. CORRELATIONS AND DEPENDENCE

It is well known that correlation coefficients of FGM distributions lies between -1/3 and 1/3 (Schucany, Parr and Boyer, 1978). Huang and Kotz (1984) show that the range of correlation coefficients can be widened by considering the iterated generalization of FGM distribution proposed by Johnson and Kotz (1977). Theorem 2 below shows that the range of correlation coefficients of the Sarmanov family of distributions is determined by both marginal distributions and their mixing functions ϕ and therefore the range may be wider than that of uniterated FGM distributions. It is shown in section 3.2 that the local correlation function is proportional to the derivative of the mixing function ϕ.

3.1 Pearson Correlation Coefficients

Theorem 2:

Let (X_1, X_2) be a bivariate random vector with p.d.f. defined in (1).

(a) The correlation coefficient of X_1 and X_2, if it exists, is given by

$$\rho = Corr[X_1, X_2] = \frac{\omega\nu_1\nu_2}{\sigma_1\sigma_2} \tag{5}$$

Thus X_1 and X_2 are independent if $\omega = 0$.

(b) The correlation coefficient ρ of X_1 and X_2 is bounded by

$$|\rho| \leq |\omega|\sqrt{E[\phi_1^2(X_1)]E[\phi_2^2(X_2)]} \tag{6}$$

Proof: (a) Follows from Theorem 1(a). (b) Result (a) implies that product moments are determined by ν_1 and ν_2. It then follows from the assumption that $\int \phi_i(t)f_i(t)dt = 0$ and the Cauchy-Schwarz inequality

$$\begin{aligned}
\nu_i^2 &= \{\int t\ \phi_i(t)\ f(t)\ dt\}^2 \\
&= \{\int (t-\mu_i)\ \phi_i(t)\ dF_i(t)\}^2 \\
&\leq \{\int (t-\mu_i)^2 dF_i(t)\}\{\int \phi_i^2(t) dF_i(t)\} \\
&= \sigma_i^2\{\int \phi_i^2(t)\ dF_i(t)\}
\end{aligned} \tag{7}$$

Notice that in the case where the support of $f_i, i = 1, 2$, is contained in the unit interval [0,1], and $\phi_i(x_i) = x_i - \mu_i$ is used in contructing the bivariate density, then we have $\nu_i = {\sigma_i}^2$ and the bound can be attained.

3.2 Likelihood Ratio Dependence and Sign Regularity Properties

Lehmann (1966) defined a bivariate random vector (X_1, X_2) as likelihood ratio dependent if the joint p.d.f (or p.m.f.) $f(x_1, x_2)$ satisfies the condition that

$$f(x_1, x_2)f({x_1}', {x_2}') \geq f({x_1}', x_2)f(x_1, {x_2}') \tag{8}$$

for any $x_1 \leq x_1'$ and $x_2 \leq x_2'$. Likelihood ratio dependence is also called dependence by total positivity of order 2 (TP2). The random vector (X_1, X_2) is called dependent by reverse regular rule of order 2 (RR2) if the inequality in (8) is reversed for all $x_1 \leq x_1'$ and $x_2 \leq x_2'$. See Karlin (1968) for sign regularity properties of TP2 or RR2 densities.

Theorem 3:

Let (X_1, X_2) be a bivariate random vector with p.d.f. (or p.m.f.) $f(x_1, x_2)$ defined in (1). Then (X_1, X_2) is TP2 if $\omega\phi_1'(x_1)\phi_2'(x_2) \geq 0$ for all x_1 and x_2, and RR2 if $\omega\phi_1'(x_1)\phi_2'(x_2) \leq 0$ for all x_1 and x_2, where ϕ_i' is the first derivative of the function ϕ_i.

Therefore, if monotone mixing functions ϕ_i are used to construct the bivariate family of distributions as described in Theorem 1, the resulting density is TP2 or RR2 depending upon the sign of ω.

4. SOME GENERAL METHODS FOR FINDING ϕ

We propose some general methods to construct mixing functions $\phi_i(x_i)$ for different types of marginals. Notice that different types of mixing functions can be used to yield different multivariate distributions with the same set of marginals. Moreover, using a combination of different marginals and functions $\phi_i(x_i)$, one can construct multivariate densities with marginals belonging to different families of distributions as well.

Corollary 1:

(a) *When A_i, the support of $f_i(x_i)$, is contained in $[0, 1]$, we can consider the function $\phi_i(x_i) = x_i - \mu_i$, where $\mu_i = \int_0^1 t f_i(t)dt$, for $i = 1, 2$. Then the function $h(x_1, x_2)$ defined in (1) is a bivariate density or probability mass function with designated marginals $f_1(x_1)$ and $f_2(x_2)$. The correlation coefficient of X_1 and X_2, if it exists, is given by $\rho = \omega\sigma_1\sigma_2$ where ω satisfies the condition that*

$$max\left(\frac{-1}{\mu_1\mu_2}, \frac{-1}{(1-\mu_1)(1-\mu_2)}\right) \leq \omega \leq min\left(\frac{1}{\mu_1(1-\mu_2)}, \frac{1}{\mu_2(1-\mu_1)}\right) \quad (9)$$

(b) *Results in(a) can be easily generalized to cases where A_i is contained in any bounded set.*

(c) *Results in (a) can also be generalized by considering higher moments, if they exist. That is, we can consider the mixing function* $\phi_i(x_i) = x_i{}^k - \mu_i{}^{(k)}$, *where* $\mu_i{}^{(k)} = \int t^k f_i(t)dt$, *for* $i = 1, 2$ *and integer* $k \geq 1$.

Therefore X_1 and X_2 are positively dependent if $\omega > 0$; independent if $\omega = 0$; and negatively dependent if $\omega < 0$. If marginal densities f_1 and f_2 in Corollary 1 contain no unknown parameters, then the maximum likelihood estimator for the unknown parameter ω is given by $\hat{\omega} = \hat{\rho}/(\hat{\sigma}_1\hat{\sigma}_2)$ where $\hat{\rho}$, $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are sample estimates of the correlation coefficient ρ, and standard deviations σ_1, and σ_2, respectively.

Corollary 2:

Assume that $f_i(x_i)$ *are defined on* $[0, \infty)$. *Let* $L_i(t) = \int_0^\infty \exp(-tx_i)f_i(x_i)dx_i$ *denote the Laplace transform of* f_i, *for* $i = 1, 2$. *Define* $\phi_i(x_i) = \exp(-x_i) - L_i(1)$, *for* $x_i \geq 0$. *Then the function* $h(x_1, x_2)$ *defined in (1) is a bivariate density with designated marginals* $f_i(x_i)$, $i = 1, 2$. *The correlataion coefficient of* X_1 *and* X_2, *if it exists, is given by*

$$\rho = Corr[X_1, X_2] = \frac{\omega[-L_1'(1) - L_1(1)\mu_1][-L_2'(1) - L_2(1)\mu_2]}{\sigma_1\sigma_2} \qquad (10)$$

where ω *is a real number which is bounded below by the value*
$-1/max\{L_1(1)L_2(1), (1 - L_1(1))(1 - L_2(1))\}$ *and bounded above by*
$1/max\{L_1(1)(1 - L_2(1)), L_2(1)(1 - L_1(1))\}$.

Corollary 3:

Let $f_i(x_i)$ *be any univariate density function, for* $i = 1, 2$, *and consider*

$$\phi_i(x_i) = f_i(x_i) - \int_{-\infty}^{\infty} f_i^2(t)dt. \qquad (11)$$

Then (1) is a bivariate density function with given marginal p.d.f. $f_i(x_i)$.

Corollary 4:

Let $F_i(x_i)$ *denote the distribution function of* $f_i(x_i)$ *and consider* $\phi_i(x_i) = 1 - 2F(x_i)$, *for* $i = 1, 2$, *then* $\int \phi_i(t)f_i(t)dt = 0$ *and* $|\phi_i(x_i)| \leq 1$ *for any* $x_i \in R$. *In this case the function* $h(x_1, x_2)$ *defined in Theorem 1 gives the p.d.f. of a FGM distribution.*

5. THE NATURAL EXPONENTIAL MARGINALS

Assume that density function $f(x)$ belongs to the one-parameter exponential family of distributions such that $f(x) = \exp\{c(\theta)T(x) + d(\theta) + Q(x)\}I_A(x)$, where the set A does not depend on θ. Using reparametrization with $\eta = c(\theta)$ and $d_0(\eta) = d(\theta)$, then $d_0(\eta) = d(\theta) = -\log \int_A \exp[\eta T(x) + Q(x)]dx < \infty$, and $f(x)$ can be represented as $f(x) = \exp\{\eta T(x) + d_0(\eta) + Q(x)\}I_A(x)$.

Case 1: $T(x) \geq 0$, for all $x \in A$

The Laplace transform of $T(X)$ is $L[s] = E[\exp\{-sT(X)\}] = \int_A \exp\{(\eta - s)T(x) + d_0(\eta) + Q(x)\}I_A(x)dx$ which equals $\exp[d_0(\eta) - d_0(\eta - s)]$. For any f_i, $i = 1, 2$, belonging to the natural exponential family, one can construct bivariate densities as defined in Theorem 1 with mixing function

$$\phi_i(x_i) = \exp[-T_i(x_i)] - \exp[d_{0i}(\eta_i) - d_{0i}(\eta_i - 1)].$$

Examples of bivariate distributions with natural exponential family marginals are given in section 6.

Case 2: $T(x)$ may be negative, but $f(x)$ is bounded

For those distributions, such as the normal, where the sufficient statistic $T(x)$ may have negative values but the density function is bounded, one may apply corollary 3 to construct bivariate densities with specified marginals.

6. EXAMPLES

6.1 Bivariate distributions with beta marginals

Let

$$f_i(x_i|a_i, b_i) = \frac{x_i^{a_i-1}(1 - x_i)^{b_i-1}}{B(a_i, b_i)} \tag{12}$$

where $B(a_i, b_i) = \Gamma(a_i)\Gamma(b_i)/\Gamma(a_i + b_i)$, for $i = 1, 2$. Using corollary 1, define mixing functions $\phi_i(x_i) = x_i - \mu_i$, where $\mu_i = a_i/(a_i + b_i)$. Then, the function defined by

$$h(x_1, x_2) = f_1(x_1|a_1, b_1)f_2(x_2|a_2, b_2)\{1 + \omega(x_1 - \frac{a_1}{a_1 + b_1})(x_2 - \frac{a_2}{a_2 + b_2})\} \tag{13}$$

is a bivariate density function with beta marginals, where

$$\frac{-(a_1+b_1)(a_2+b_2)}{max(a_1a_2,b_1b_2)} \le \omega \le \frac{(a_1+b_1)(a_2+b_2)}{max(a_1b_2,a_2b_1)} \tag{14}$$

The correlation coefficient of (13) is given by $\rho = \omega\sigma_1\sigma_2$. When $a_1 = a_2 = a$ and $b_1 = b_2 = b$, it can be shown that the correlation coefficient has a range

$$max\left\{\frac{-b}{a(1+a+b)}, \frac{-a}{b(1+a+b)}\right\} \le \rho \le \frac{1}{(1+a+b)} \tag{15}$$

Hence ρ tends to 1 (or -1) as a and b both tend to zero and ω tends to 4 (or -4). For example, if $a = b = 1/n$ and $-4 \le \omega \le 4$ then

$$-\frac{n}{n+2} \le \rho \le \frac{n}{n+2}$$

A good property of the proposed family of bivariate beta distributions is that it can be expressed as a linear combination of products of univariate beta densities as follows.

$$\begin{aligned} h(x_1,x_2) &= (1+\frac{\omega a_1a_2}{(a_1+b_1)(a_2+b_2)})f_1(x_1|a_1,b_1)f_2(x_2|a_2,b_2) \\ &-\frac{\omega a_1 B(a_2+1,b_2)}{(a_1+b_1)B(a_2,b_2)}f_1(x_1|a_1,b_1)f_2(x_2|a_2+1,b_2) \\ &-\frac{\omega a_2 B(a_1+1,b_1)}{(a_2+b_2)B(a_1,b_1)}f_2(x_2|a_2,b_2)f_1(x_1|a_1+1,b_1) \\ &+\frac{\omega B(a_1+1,b_1)B(a_2+1,b_2)}{B(a_1,b_1)B(a_2,b_2)}f_2(x_2|a_2+1,b_2)f_1(x_1|a_1+1,b_1) \end{aligned} \tag{16}$$

As a result, if the proposed bivariate beta density is adopted as a prior in Bayesian analysis, the posterior is pseudo-conjugate to the prior. Possible applications are discussed in section 7.

6.2. Bivariate distributions with binomial marginals

Let $P_i(X_i = x_i|N_i,\theta_i)$ be the p.m.f. of a binomial distribution with parameters N_i and θ_i . Using corollary 2, for any $x_1 = 0,1,\dots,N_1$, $x_2 = 0,1,\dots,N_2$,

$$h(x_1,x_2) = \binom{N_1}{x_1}\binom{N_2}{x_2}\theta_1{}^{x_1}(1-\theta_1)^{N_1-x_1}\theta_2{}^{x_2}(1-\theta_2)^{N_2-x_2}\{1+\omega\phi_1(x_1)\phi_2(x_2)\}$$

is a p.m.f. with binomial marginals, where $\phi_i(x_i) = e^{-x_i} - \{\theta_i e^{-1} + (1 - \theta_i)\}^{N_i}$ for $i = 1, 2$. Correlation coefficients can be derived from equation (10). For example, if $N_1 = N_2 = 10$, $\theta_1 = \theta_2 = 0.1$, and parameter $-3.6913 \leq \omega \leq 4.0067$, the correlation coefficients of the proposed bivariate binomial distribution has the range $-0.41 \leq \rho \leq 0.45$.

6.3. Bivariate distributions with Poisson marginals.

Let $P_i(X_i = x_i | \lambda_i)$ be the p.m.f. of a univariate Poisson distribution with parameter λ_i. The Laplace transform of P_i is given by $L_i(t) = \exp\{\lambda_i \exp(-t) - \lambda_i\}$. Using mixing functions $\phi_i(x_i) = e^{-x_i} - \exp\{\lambda_i e^{-1} - \lambda_i\}$ as discussed in Corollary 2, we can construct bivariate distributions with Poisson marginals.

6.4. Bivariate distributions with gamma marginals

Let $f_i(x_i) = \lambda^\alpha x_i{}^{\alpha-1} e^{-\lambda_i x_i}/\Gamma(\alpha)$, for $i = 1, 2$. The Laplace transform of f_i has the form $L(t) = (1 + t/\lambda_i)^{-\alpha_i}$. Using the mixing function $\phi_i(x_i) = e^{-x} - (1 + \lambda_i{}^{-1})^{-\alpha_i}$, we can construct bivariate gamma densities. Similar to the bivariate beta case, this bivariate gamma density can be readily expressed as a linear combination of products of univariate gamma densities.

6.5 Bivariate distributions with exponential marginals

Since the exponential distribution is a special case of the gamma distribution, bivariate densities with exponential marginals can be easily derived from the results in section 6.4. As a result we have

$$\lambda_1 e^{-\lambda_1 x_1} \lambda_2 e^{-\lambda_2 x_2} \left\{ 1 + \omega(e^{-x_1} - \frac{\lambda_1}{1+\lambda_1})(e^{-x_2} - \frac{\lambda_2}{1+\lambda_2}) \right\} \tag{17}$$

as a bivariate exponential density which is different from those discussed by Marshall and Olkin (1967) or Block and Basu (1974).

6.6 Bivariate distribution with normal marginals

Consider standard normal densities $f_i(x_i) = \exp(-x_i^2/2)/\sqrt{2\pi}$, and let

$$\phi_i(x_i) = \exp(-x_i^2 - 2x_i) - e^{2/3}/\sqrt{3} \tag{18}$$

for $i = 1, 2$, substituting (18) into (1) and we get a nonnormal bivariate density which has standard normal marginals.

6.7 Bivariate distributions with Cauchy marginals

Let $f_i(x_i) = [\pi(1+x_i^2)]^{-1}$. Then $\int f_i(t)^2 dt = (2\pi)^{-1}$. Consider $\phi_i(x_i) = [\pi(1+x_i^2)]^{-1} - (2\pi)^{-1}$ as discussed in Corollary 3, then bivariate densities with Cauchy marginals can be derived.

6.8 Bivariate Distributions with Proportional Hazards Marginals

Dependent lifetime distributions are useful in modeling paired survival data. Let X_{ki} denote the lifetime random variable of individual i in group k, and let $\mathbf{z}_{ki}$ be the corresponding covariate vector, where $i = 1, 2$, and $k = 1, \ldots, n$. We assume that X_{ki} follows the proportional hazards (PH) model introduced by Cox (1972), for $i = 1, 2$. The hazard function of X_{ki}, given $\mathbf{z}_{ki}$, is of the form $\lambda_i(x_{ki}|\mathbf{z}_{ki}) = \lambda_{i,0}(x_{ki}) \exp[\beta \mathbf{z}_{ki}]$, where β is a vector of regression coefficients, and the density function is

$$f_i(x_{ki}) = \lambda_{i,0}(x_{ki}) e^{\beta \mathbf{z}_{ki}} \exp[-\Lambda_{i,0}(x_{ki}) e^{\beta \mathbf{z}_{ki}}]. \tag{19}$$

where $\Lambda_{i,0}(x_{ki})$ denotes the cumulative hazard function. Using corollary 4, a bivariate distribution with PH marginals can be derived from (1). Thus the sample likelihood is given by

$$\prod_{k=1}^{n} f_1(x_{k1}) f_2(x_{k2}) \{1 + \omega \phi_1(x_{k1}) \phi_2(x_{k2})\} \tag{20}$$

where $\phi_i(x_{ki}) = \exp[-\Lambda_0(x_{ki}) e^{\beta \mathbf{z}}_{ki}] - 0.5$. This family of bivariate PH distributions can model both positive and negative association.

Similar to the above examples, bivariate distributions with Weibull or inverse Gaussian marginals can be easily constructed.

7. APPLICATIONS

Bivariate densities with given marginals are often needed in statistical models. We discuss possible applications in Bayesian inference and survival analysis.

7.1 Bayesian methods for Correlated Binary Data

When a bivariate prior is needed in Bayesian inference, a family of distributions which can model both positive and negative associations and also allow one to easily compute the posterior density will be preferred. In this section we demonstrate the advantages of using the Sarmanov family of distributions as priors in analyzing correlated binary data. Bayesian computations can be easily carried out because the joint densities of this family of bivariate beta distributions can be expressed as a linear combination of products of independent betas.

Correlated binary data occur in many applications. Assume, for the moment, that each cluster contains a pair of binary observations (X_1, X_2) which has the joint likelihood

$$f(x_1, x_2|\theta_1, \theta_2) \propto \theta_1^{x_1}(1-\theta_1)^{1-x_1}\theta_2^{x_2}(1-\theta_2)^{1-x_2}. \tag{21}$$

Much research has been done using beta-binomial analysis where θ_1 and θ_2 are assumed to be independently distributed as univariate beta distributions. In general, however, θ_1 and θ_2 may be dependent, such as dichotomous data collected from fathers and sons. Thus a dependent prior distribution with beta marginals introduced in section 6.1 is considered. Let

$$h_{\alpha,\beta}(\theta_1, \theta_2|\alpha_1, \beta_1; \alpha_2, \beta_2; \omega) = f_1(\theta_1|\alpha_1, \beta_1)f_2(\theta_2|\alpha_2, \beta_2)\{1+\omega(\theta_1-\mu_1)(\theta_2-\mu_2)\}$$

with $\mu_i = \alpha_i/(\alpha_i + \beta_i)$. Then

$$f(x_1, x_2|\alpha_1, \beta_1; \alpha_2, \beta_2; \omega) \propto \frac{C(x_1, x_2)}{B(\alpha_1, \beta_1)B(\alpha_2, \beta_2)} \tag{22}$$

and the posterior density of (θ_1, θ_2) can be expressed in a simple form

$$\begin{aligned} &h_{\mathbf{x},\alpha,\beta}(\theta_1, \theta_2|x_1, x_2; \alpha_1, \beta_1; \alpha_2, \beta_2; \omega) \\ &= \{\ (1+\omega\mu_1\mu_2)\theta_1{}^{(x_1+\alpha_1-1)}(1-\theta_1)^{(-x_1+\beta_1)}\theta_2{}^{(x_2+\alpha_2-1)}(1-\theta_2)^{(-x_2+\beta_2)} \\ &+\omega\theta_1{}^{(x_1+\alpha_1)}(1-\theta_1)^{(-x_1+\beta_1)}\theta_2{}^{(x_2+\alpha_2)}(1-\theta_2)^{(-x_2+\beta_2)} \\ &-\omega\mu_1\theta_1{}^{(x_1+\alpha_1-1)}(1-\theta_1)^{(-x_1+\beta_1)}\theta_2{}^{(x_2+\alpha_2)}(1-\theta_2)^{(-x_2+\beta_2)} \\ &-\omega\mu_2\theta_1{}^{(x_1+\alpha_1)}(1-\theta_1)^{(-x_1+\beta_1)}\theta_2{}^{(x_2+\alpha_2-1)}(1-\theta_2)^{(-x_2+\beta_2)}\ \}/C(x_1, x_2) \end{aligned} \tag{23}$$

where $C(x_1, x_2)$ is a constant in the posterior distribution. Therefore the posterior density is a linear combination of products of univariate beta densities. Thus the posterior means of θ_1 and θ_2 can be readily derived.

Cole, Lee, Whitmore and Zaslavsky (1995) demonstrate the usefulness of the proposed bivariate beta distribution as a prior in an example analyzing incompletely observed longitudinal binary store display data. In their article, a new empirical Bayes estimation methodology for the analysis of binary Markov chains having randomly missing observations is developed. The pseudo-conjugate properties of the proposed bivariate beta priors greatly reduced the complexity of posterior computations.

7.2 Marker dependent survival distributions

In order to determine how some of the markers indicating disease status (for example, T4 cell counts for HIV positive patients, or number of tumors of cancer patients) are useful for predicting the subsequent course of a disease, a joint density of the marker and survival distribution can be derived.

In general, let X_1 denote the marker random variable with p.d.f. $f(x_1|\,\theta)$, let X_2 denote the lifetime random variable of an individual, and let $\mathbf{z}$ be a known vector of regressor variables. We assume that X_2 has a PH density given by

$$f_2(x_2) = \lambda_0(x_2)e^{\mathbf{z}\beta}\exp[-\Lambda_0(x_2)e^{\mathbf{z}\beta}]. \tag{24}$$

where β is a vector of regression coefficients, Let $\phi_2(x_2) = \exp[-\Lambda_0(x_2)e^{\mathbf{z}\beta}] - 0.5$. It then follows from (1) and Theorem 2 (b) that, with the marker variable $X_1 = x_1$ given, the conditional survival function of X_2 is

$$P(X_2 > x_2|X_1 = x_1) = \exp[-\Lambda_0(x_2)e^{\mathbf{z}\beta}] + \omega\phi_1(x_1)\int_{x_2}^{\infty} f_2(t)\phi_2(t)dt \tag{25}$$

For a discrete marker one might assume that X_1 has a univariate Poisson distribution with parameter θ and consider $\phi_1(x_1) = e^{-x_1} - \exp[\theta e^{-1} - \theta]$.

8. MULTIVARIATE EXTENSIONS

We extend the Sarmonov's family of distributions to the multivariate case.

Theorem 4:

(a) Assume that $f_1(x_1), \ldots, f_n(x_n)$ are univariate probability density functions (p.d.f.), or probability mass functions (p.m.f.), with supports defined on $A_1 \subseteq$

R, $A_2 \subseteq R, \ldots, A_n \subseteq R$, respectively. Let $\phi_i(t)$, $i = 1, \ldots, n$, be a set of bounded nonconstant functions such that $\int_{-\infty}^{\infty} \phi_i(t) f_i(t) dt = 0$ for all $1 \leq i \leq n$. Then, the function

$$h(x_1, \ldots, x_n) = \{\prod_{i=1}^{n} f_i(x_i)\} \left(1 + R_{\phi_1,\ldots,\phi_n,\Omega_n}(x_1, \ldots, x_n)\right) \tag{26}$$

is a multivariate joint density (or a probability mass function) with specified marginals $f_1(x_1), \ldots, f_n(x_n)$, where

$$R_{\phi_1,\ldots,\phi_n,\Omega_n}(x_1, x_2, .., x_n) = \{\sum_{j_1<}^{n-1} \sum_{j_2}^{n} \omega_{j_1,j_2} \phi_{j_1}(x_{j_1}) \phi_{j_2}(x_{j_2}) +$$
$$\sum_{j_1<}^{n-2} \sum_{j_2<}^{n-1} \sum_{j_3}^{n} \omega_{j_1,j_2,j_3} \phi_{j_1}(x_{j_1}) \phi_{j_2}(x_{j_2}) \phi_{j_3}(x_{j_3}) + \ldots + \omega_{1,2,\ldots,n} \prod_{i=1}^{n} \phi_i(x_i)\} \tag{27}$$

and $\Omega_n = \{\omega_{j_1,j_2}, \omega_{j_1,j_2,j_3}, \cdots, \omega_{1,2,\ldots,n}\}$. The set of real numbers Ω_n is chosen such that $1 + R_{\phi_1,\ldots,\phi_n,\Omega_n}(x_1, x_2, .., x_n) \geq 0$ holds for all $x_i \in R$, and $i = 1, \ldots, n$.
(b) *When all of the ω's equal zero, (26) reduces to the independence case.*

Theorem 5:

Assume that $(X_1, X_2, .., X_n)$ has a p.d.f. h as defined in Theorem 4.
(a) Any subset of $X_1, X_2, \ldots, X_n$ is distributed as a subfamily of distributions defined in Theorem 1. That is, any subvector $(X_{k_1}, \ldots, X_{k_m})$, $1 \leq k_1 < k_2 \ldots < k_m \leq n$, has a joint density of the form

$$h(x_{k_1}, \ldots, x_{k_m}) = \prod_{i=1}^{m} f_{k_i}(x_{k_i}) \left(1 + R_{\phi_{k_1},\ldots,\phi_{k_m},\Omega_m}(x_{k_1}, \ldots, x_{k_m})\right) \tag{28}$$

where $R_{\phi_1,\Omega_1} = 0$, and Ω_m is a subset of Ω_n such that subscripts of ω's involve combinations of integers $k_1, \ldots, k_m$ only.
(b) For any fixed a, the conditional distribution function is

$$P(X_j \leq a | X_i = x_i) = F_j(a) - \omega_{i,j} \phi_i(x_i) \int_a^{\infty} f_j(t) \phi_j(t) dt. \tag{29}$$

where $F_j(a) = P(X_j \leq a)$ is the marginal marginal function of X_j.
(c) The product moment is given by

$$E[\prod_{i=1}^{n} X_i] = \prod_{i=1}^{n} E[X_i] + \{\sum_{j_1<}^{n-1} \sum_{j_2}^{n} \omega_{j_1,j_2} \nu_{j_1} \nu_{j_2} \prod_{i \neq j_1,j_2} E[X_i] +$$
$$\sum_{j_1<}^{n-2} \sum_{j_2<}^{n-1} \sum_{j_3}^{n} \omega_{j_1,j_2,j_3} \nu_{j_1} \nu_{j_2} \nu_{j_3} \prod_{i \neq j_1,j_2,j_3} E[X_i] + \ldots + \omega_{1,2,\ldots,n} \prod_{i=1}^{n} \nu_i\} \tag{30}$$

(d) The regression of X_n on $X_1, \ldots, X_{n-1}$ is given by

$$E[X_n|X_1 = x_1, \ldots, X_{n-1} = x_{n-1}] = \mu_n +$$
$$\mu_n \left(\frac{\sum_{i<n} \omega_{i,n}\phi_i(x_i) + \sum_{i<j}\sum_{<n} \omega_{i,j,n}\phi_i(x_i)\phi_j(x_j) + \ldots + \omega_{1,\ldots,n}\prod_{i=1}^{n-1}\phi_i(x_i)}{1 + R_{\phi_1,\ldots,\phi_{n-1},\Omega_{n-1}}} \right)$$

As a result, we can derive pseudo-conjugate distributions for multivariate distributions with the natural exponential family as marginals. Suppose that observation vector $(X_1, \ldots, X_n)$ has a joint density function of the form

$$f(x_1, \ldots, x_n|\theta_1, \ldots, \theta_n) = \prod_{i=1}^{n} f_i(x_i|\theta_i)\,(1 + R_{\psi_1,\ldots,\psi_n,\Omega_n}(x_1, \ldots, x_n))$$

where marginals $f_i(x_i|\theta_i)$, $i = 1, \ldots, n$, belong to the k-parameter exponential family of distributions such that for $\theta_i=(\theta_{i,1}, \ldots, \theta_{i,k})$,

$$f(x_i|\theta_i) = \exp\{\sum_{j=1}^{k} c_{i,j}(\theta_i)T_{i,j}(x_i) + d_i(\theta_i) + Q_i(x_i)\}I_{A_i}(x_i), \qquad (31)$$

and mixing functions $\psi_i(x_i)$ and $R_{\psi_1,\ldots,\psi_n,\Omega_n}(x_1, \ldots, x_n)$, are defined as in Theorem 4.

Assume that $T_{i,j}(X_i) \geq 0$ for all j and all $x_i \in A_i$, i=1,..., n, and consider the joint prior distribution of $(\theta_1, \ldots, \theta_n)$ with the density

$$\begin{aligned} &\pi_t(\theta_1, \ldots, \theta_n|t_{i,1}, \ldots, t_{i,k+1}, 1 \leq i \leq n) \\ &\quad = \left(\prod_{i=1}^{n} \pi_i(\theta_i|t_{i,1}, \ldots, t_{i,k+1})\right)(1 + R_{\phi_1,\ldots,\phi_n,\Omega_n}(\theta_1, \ldots, \theta_n)) \qquad (32) \end{aligned}$$

where ϕ_i , $i = 1, \ldots, n$, are defined in Theorem 4. Then it follows that the posterior density is pseudo-conjugate to the prior density in the sense that the posterior density is a linear combination of products of densities from the univariate natural exponential family of distributions. This property significantly simplifies computations and derivations of Bayesian inference.

ACKNOWLEDGEMENT

I thank Kjell Doksum, Samuel Kotz, Ingram Olkin, and Alan Zaslavsky for helpful comments on a preliminary draft.

REFERENCES

Bickel and Doksum (1977). *Mathematical Statistics*, Holden-Day, Inc. Oakland, California.

Block, H.W. and Basu, A.P. (1974). A continuous bivariate exponential extension, *JASA*, **V.69**, 1031-1037.

Cohen, L. (1984). Probability distributions with given martivariate marginals. *J. Math. Phys.*, **25**, 2402-2403.

Cole, B.F., Lee, M.-L.T., Whitmore, G.A. and Zaslavsky, A.M. (1995). An empirical Bayes model for Markov-dependent binary sequences with randomly missing observations. *JASA*, **90**, 1364-1372.

Cox, D.R. (1972). Regression models and life tables (with discussion). *JRSS*, **B 34**, 187-202.

Farlie, D.J. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika*, **47**, 307-323.

Huang, J.S. and Kotz, S. (1984). Correlation structure in iterated Farlie-Gumbel- Morgenstern distributions. *Biometrika,* **71**, 633-636.

Johnson, N.L. and Kotz, S. (1975). On some generalized Farlie-Gumbel-Morgenstern distributions. *Comm. Stat.*, **4**, 415-427.

Johnson, N.L. and Kotz, S. (1977). On some generalized Farlie-Gumbel-Morgenstern distributions: regression, correlation and further generalizations. *Comm. Stat.*, **A6**, 485-496.

Karlin, S. (1968). *Total Positivity*, Stanford Univ. Press, Stanford, CA.

Lehmann, E.L. (1966). Some concepts of dependence, *Annals of Mathematical Statistics*, **37**, 1137-1153.

Marshall, A.W. and Olkin, I. (1967). A multivariate exponential distribution, *JASA*, **62**, 30-44.

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

Schucany, W.R., Parr, W.C. and Boyer, J.E. (1978). Correlation structure in Farlie-Gumbel-Morgenstern distribution, *Biometrika*, **65**, 650-653.

Shaked, M. (1975). A note on the exchangeable generalized Farlie-Gumbel-Morgenstern distributions, *Comm. Stat.*, **4(8)**, 711-721.

Received September, 1994; Revised December, 1995.