Characterization of symmetry and exceedance models in multivariate FGM distributions

I. Bairamov and S. Eryilmaz
İzmir University of Economics, Department of Mathematics
Balçova,İzmir, Turkey

Abstract. The set of dependent random variables with multivariate Farlie-Gumbel-Morgenstern (FGM) distribution is considered. For the simple multivariate FGM distributions the admissible range of association parameters are investigated. Distributional properties of some exceedance statistics of finite FGM sequences with respect to a random threshold are studied. Some results characterizing symmetry of the generalized simple FGM distributions through the properties of ranks are presented.

AMS 1991 subject classifications: 60G70, 60E05

Key words and Phrases: Farlie-Gumbel-Morgenstern distribution; random vectors; random threshold; exceedance statistics

1. INTRODUCTION

In this paper we consider Farlie-Gumbel-Morgenstern class of multivariate distributions which was discussed by Morgenstern (1956), Gumbel (1960) and Farlie (1960). The FGM distributions are improved essentially in the works of Johnson and Kotz (1975), (1977), who introduced an additional parameter to achieve stronger correlation structure. Recent discussion of this family are due to Lin (1987), Kotz and Seeger (1991), Cambanis (1993), Huang and Kotz (1999), , Lai and Xie (2000), Bairamov and Kotz (2002).

Because of their simple analytical form, FGM distributions and their modifications have been widely used in many applications, one can see, e.g. Conway (1983), Hutchinson and Lai (1990), Kemp (1995). There are many papers on FGM distributions and their extensions. For more details we refer recent books of Nelsen (1998) and Kotz, Balakrishnan and Johnson (2000). A FGM random sequence $\{X_i\}_{i\geq 1}$ is defined by the univariate marginals $F_i \sim X_i$, $i \geq 1$, and a symmetric function $\alpha(.,.)$ $(\alpha(j,k) = \alpha(k,j))$ such that the joint distribution of $X_{i_1}, X_{i_2}, ..., X_{i_n}$ is

$$H_{i_1,i_2,\ldots,i_n}(\mathbf{x}) = \prod_{k=1}^{n} F_{i_k}(x_k) \left\{ 1 + \sum_{1\leq j<k\leq n} \alpha(i_j, i_k) \bar{F}_{i_j}(x_j) \bar{F}_{i_k}(x_k) \right\} \quad (1)$$

where $\bar{F}(x) = 1 - F(x)$, $\mathbf{x} = (x_1, x_2, ..., x_n)$. The FGM sequence is stationary iff the univariate marginals are all equal, $F_i = F_1, \ i > 1$, and the function $\alpha(j,k)$ depends on j,k only through their difference $\alpha(j,k) = \alpha(j-k)$ for all $j \neq k$.

In this paper we consider some special cases and modifications of FGM distribution H in R^n, for $n \geq 1$, which is defined with respect to given univariate distributions F_i, $i \leq n$, by

$$H(\mathbf{x}) = \prod_{i=1}^{n} F_i(x_i) \left\{ 1 + \sum_{1 \leq j < k \leq n} \alpha_n(j,k) \bar{F}_j(x_j) \bar{F}_k(x_k) \right\} \tag{2}$$

for all vectors $\mathbf{x} = (x_1, x_2, ..., x_n) \in R^n$, where the $n(n-1)/2$ terms $\alpha_n(j,k)$ are suitable constants, such that H is a distribution function. The univariate marginals of H are the F_i. The constants $\alpha_n(j,k)$ are admissible if the inequalities

$$1 + \sum_{1 \leq j < k \leq n} \alpha_n(j,k) \varepsilon_j \varepsilon_k \geq 0 \tag{3}$$

hold, for all $\varepsilon_j = -M_j$ or $1 - m_j$, where $M_j = \sup \{F_j(x), -\infty < x < \infty\} \setminus \{0,1\}$ and $m_j = \inf \{F_j(x), -\infty < x < \infty\} \setminus \{0,1\}$. If F_j is absolutely continuous, then $M_j = 1$ and $m_j = 0$, hence $\varepsilon_j = \pm 1$.

In statistical literature there are many papers appeared in last years dealing with this class of distributions which has form (2). But there not exist results about exact bounds of association parameters $\alpha_n(j,k)$ allowing (2) to be a $n-$dimensional distibution function. The knowledge about the exact bounds of $\alpha_n(j,k)$ is very important from the point of view of applications. Because if we want to use (2) for modeling we must know exact bounds of association parameters $\alpha_n(j,k)$, otherwise the model will not be applicable. In our paper we have considered slightly simple model, where association parameters $\alpha_n(j,k)$ do not depend on j and k, but on n:

$$H(\mathbf{x}) = \prod_{i=1}^{n} F_i(x_i) \left\{ 1 + \alpha_n \sum_{1 \leq j < k \leq n} \bar{F}_j(x_j) \bar{F}_k(x_k) \right\}.$$

In the present paper we show that the constants α_n must satisfy the inequalities

$$-\frac{1}{\binom{n}{2}} \leq \alpha_n \leq \frac{1}{\left[\frac{n}{2}\right]}$$

where $[x]$ denotes the integer part of x.

In section 2, we introduce a simple multivariate FGM distribution and provide the admissible range of the association parameter. In Section 3 we consider ranks of dependent random variables whose finite dimensional distributions are multivariate generalized simple FGM distribution. Let $X_1, X_2, ..., X_n$ be arbitrary dependent random variables and

$$R_{j,n} = \sum_{i=1}^{n} I_{\{X_i \leq X_j\}}$$

be the rank of random variable X_j in a collection $X_1, X_2, ..., X_n$. Define the random variables

$$\xi_j = \begin{cases} 1 & \text{if } R_{j,n} = n \\ 0 & \text{otherwise} \end{cases}, \ j = 1, 2, ..., n$$

and $M_n = \max(X_1, X_2, ..., X_n)$.

In Section 3 of this paper we define multivariate generalized simple FGM distributions and provide some results characterizing symmetry of this distributions using properties of $R_{j,n}, \xi_j$ and M_n.

In Section 4 of this paper we discuss the behaviour of some placement statistics and exceedances for dependent random variables with common multivariate FGM distribution. The importance of these concepts in construction of nonparametric tests of equality (identity) of distributions has been demonstrated by Katzenbeisser (1985), (1986), Matveychuk and Petunin (1990), Johnson and Kotz (1991), (1994) in case of independent random variables. Exceedances are also applicable in reliability studies. The concept of exceedances is closely connected with the negative (inverse) hypergeometric distribution which was developed in the works of Gumbel and Schelling (1950) and Sarkadi (1957). Further detailed discussion can be found in David (1981), Leadbetter et al. (1983), Johnson et al. (1992). A recent interest in the subject lies in the work of Bairamov (1997), Wesolowski and Ahsanullah (1998), Bairamov and Eryılmaz (2000), Bairamov and Kotz (2001), Eryılmaz (2002).

2. MULTIVARIATE SIMPLE FGM DISTRIBUTIONS

Let us define a n variate FGM random vector $(X_1, X_2, ..., X_n)$ by the univariate marginals $F_i \sim X_i$, $i = 1, 2, ..., n$ and a real number α_n such that the joint distribution of $X_1, X_2, ..., X_n$ is given by the FGM distribution

$$H_{1,2,...,n}(\mathbf{x}) = \prod_{i=1}^{n} F_i(x_i) \left\{ 1 + \alpha_n \sum_{1 \le j < l \le n} \bar{F}_j(x_j) \bar{F}_l(x_l) \right\}, \quad (4)$$

where $\mathbf{x} = (x_1, x_2, ..., x_n) \in R^n$, $\bar{F}_j(x_j) = 1 - F_j(x_j)$. For a given n the real number α_n is admissible if

$$1 + \alpha_n \sum_{1 \le j < l \le n} \varepsilon_j \varepsilon_l \ge 0 \quad (5)$$

hold for all $\varepsilon_j = \pm 1$ (as a consequence each parameter must satisfy $|\alpha| \le 1$).

Definition 2.1. We call the random variables whose joint distribution is defined by (4) a simple- FGM (or s-FGM) random variables. (If it will not confuse the readers we will also use the term "finite s-FGM sequence")

It follows from the Lemma 2.1 below that if n tends to infinity, i.e. if we consider infinite sequence of random variables whose finite dimensional distributions are given in (4) then we have only a sequence of independent random variables. We will deal in this paper with finite FGM sequences.

Lemma 2.1. The admissible range of α_n, $n > 1$ allowing (4) to be a n variate distribution function is

$$-\frac{1}{\binom{n}{2}} \leq \alpha_n \leq \frac{1}{\left[\frac{n}{2}\right]}, \tag{6}$$

where $[x]$ denotes the integer part of the real number x.

Proof. The lemma is a consequence of condition (5). Denote

$$D_n^+ = \left\{ (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n) : \varepsilon_i = \pm 1, i = 1, 2, .., n, \sum_{1 \leq i < j \leq n} \varepsilon_i \varepsilon_j \geq 0 \right\}$$

and

$$D_n^- = \left\{ (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n) : \varepsilon_i = \pm 1, i = 1, 2, .., n, \sum_{1 \leq i < j \leq n} \varepsilon_i \varepsilon_j \leq 0 \right\}.$$

Then from (5) it follows that the admissible range for α_n is

$$-\frac{1}{\max\limits_{(\varepsilon_1, \varepsilon_2, ..., \varepsilon_n) \in D_n^+} \sum\limits_{1 \leq i < j \leq n} \varepsilon_i \varepsilon_j} \leq \alpha_n \leq \frac{1}{\max\limits_{(\varepsilon_1, \varepsilon_2, ..., \varepsilon_n) \in D_n^-} (-\sum\limits_{1 \leq i < j \leq n} \varepsilon_i \varepsilon_j)}$$

It is clear that the maximum value of $\sum\limits_{1 \leq i < j \leq n} \varepsilon_i \varepsilon_j$ in D_n^+ is achieved when $\varepsilon_j = 1$, $j = 1, 2, ..., n$ or -1, $j = 1, 2, ..., n$, and equal to $\binom{n}{2}$. Therefore

$$\alpha_n \geq -\frac{1}{\binom{n}{2}}.$$

For the upper bound of α_n we have to investigate the minimum value of $\sum\limits_{1 \leq i < j \leq n} \varepsilon_i \varepsilon_j$ in D_n^-. By symmetry,

$$\begin{aligned} \sum_{1 \leq i < j \leq n} \varepsilon_i \varepsilon_j &= \frac{1}{2} \sum_{i \neq j} \varepsilon_i \varepsilon_j \\ &= \frac{1}{2} [\varepsilon_1(\varepsilon_2 + \varepsilon_3 + ... + \varepsilon_n) \\ &\quad + \varepsilon_2(\varepsilon_1 + \varepsilon_3 + ... + \varepsilon_n) + ... \\ &\quad + \varepsilon_n(\varepsilon_1 + \varepsilon_3 + ... + \varepsilon_{n-1})]. \end{aligned} \tag{7}$$

Let m of ε_i's are equal to 1 and $n-m$ of ε_i's are equal to -1. Then

$$\sum_{i=1}^{n} \varepsilon_i = m - (n-m) = 2m - n. \tag{8}$$

From (7) and (8) we have

$$\sum_{1 \le i < j \le n} \varepsilon_i \varepsilon_j = \frac{1}{2} \left[\varepsilon_1 (2m - n - \varepsilon_1) + \varepsilon_2 (2m - n - \varepsilon_2) + ... \right.$$

$$\left. + \varepsilon_n (2m - n - \varepsilon_n) \right] = \frac{1}{2} \left[(2m-n) \sum_{i=1}^{n} \varepsilon_i - \sum_{i=1}^{n} \varepsilon_i^2 \right] = \frac{1}{2} \left[(2m-n)^2 - n \right]. \tag{9}$$

Then from (9) it follows that

$$\min_{(\varepsilon_1, \varepsilon_2, ..., \varepsilon_n) \in D_n^-} \sum_{1 \le i < j \le n} \varepsilon_i \varepsilon_j = \begin{cases} -\frac{n}{2} & \text{if } \ n \text{ is even} \\ -\frac{n-1}{2} & \text{if } \ n \text{ is odd} \end{cases} = -\left[\frac{n}{2} \right].$$

The lemma is thus proved.

3. CHARACTERIZATION OF SYMMETRY OF THE GENERALIZED s-FGM DISTRIBUTION

Consider s-FGM random variables $X_1, X_2, ..., X_n$ with joint distribution given as in (4). It is clear that if the marginal distributions are all equal then the random variables $X_1, X_2, ..., X_n$ are exchangeable (symmetrically dependent). Now consider the following generalization of uniform s-FGM random variables: let us define a random variables $X_1, X_2, ..., X_n$ by the uniform marginals $F_i \sim U(0,1)$, $i = 1, 2, ..., n$ and a real number α_n such that the joint distribution of $X_1, X_2, ..., X_n$ is given by the n variate distribution

$$H(x_1, x_2, ..., x_n) = \prod_{r=1}^{n} x_r \left\{ 1 + \alpha_n \sum_{1 \le i < j \le n} A(x_i) B(x_j) \right\}, \tag{10}$$

$$0 < x_i < 1, \ \ i = 1, 2, ..., n,$$

where $A(x)$ and $B(x)$, $x \in (0,1)$ are differentiable functions satisfying $\lim_{x \to 1} A(x) = 0$, $\lim_{x \to 1} B(x) = 0$. For the given n the α_n is admissible if

$$-\frac{1}{\max \sum_{1 \le i < j \le n} \hat{A}(x_i) \hat{B}(x_j)} \le \alpha_n \le \frac{1}{\max \left\{ - \sum_{1 \le i < j \le n} \hat{A}(x_i) \hat{B}(x_j) \right\}}, \tag{11}$$

where

$$\hat{A}(x) = \frac{d}{dx}(xA(x)) \quad \text{and} \quad \hat{B}(x) = \frac{d}{dx}(xB(x)). \tag{12}$$

Definition 3.1. The sequence of random variables having finite dimensional distribution given by (10) is called Generalized- s-FGM sequence.

The p.d.f. of $(X_1, X_2, ..., X_n)$ is given by

$$h(x_1, x_2, ..., x_n) = \left\{ 1 + \alpha_n \sum_{1 \le i < j \le n} \hat{A}(x_i)\hat{B}(x_j) \right\}. \tag{13}$$

Example 3.1.

a) Let $A(x) = B(x) = 1 - x^p, 0 < p, x < 1$. In this case $\hat{A}(x) = \hat{B}(x) = 1 - (1+p)x^p$ and

$$\min_{0 \le x \le 1} \hat{A}(x) = \min_{0 \le x \le 1} \hat{B}(x) = -p \quad , \quad \max_{0 \le x \le 1} \hat{A}(x) = \max_{0 \le x \le 1} \hat{B}(x) = 1.$$

Taking $\varepsilon_j = 1$, $j = 1, 2, ..., n$ or $-p$, $j = 1, 2, ..., n$ in (5), we obtain the maximum value of $\sum_{1 \le i < j \le n} \varepsilon_i \varepsilon_j$ when $\varepsilon_j = 1$, $j = 1, 2, ..., n$ since $p < 1$. Therefore, $\alpha_n \ge -\frac{1}{\binom{n}{2}}$. For the upper bound of α_n let m of ε_i's are equal to 1 and $n - m$ of ε_i's are equal to $-p$, by using same arguments in Lemma 2.1 we obtain

$$\sum_{1 \le i < j \le n} \varepsilon_i \varepsilon_j = \frac{1}{2} \left\{ [m - p(n-m)]^2 - \left[m + (n-m)p^2 \right] \right\}.$$

Then it follows that

$$\min \sum_{1 \le i < j \le n} \varepsilon_i \varepsilon_j = -\frac{\left[m + (n-m)p^2 \right]}{2}$$

where $m = \frac{pn}{1+p}$. So

$$-\frac{1}{\binom{n}{2}} \le \alpha_n \le \frac{2}{np}.$$

b) $A(x) = 1 - x^p$, $B(y) = 1 - y^q$. Then repeating the same arguments above one has

$$-\frac{1}{\binom{n}{2}} \le \alpha_n \le \frac{2}{n \max(p, q)}.$$

Definition 3.2. Denote

$$W(x) = \begin{vmatrix} A(x) & B(x) \\ A'(x) & B'(x) \end{vmatrix} = A(x)B'(x) - A'(x)B(x).$$

We say that the Generalized-s-FGM distribution defined by (10) belongs to the class $\Im$ if

1. $A(0) = B(0) = 1$
2. $W(x) \ge 0$ or $W(x) \le 0$ for all $0 < x < 1$.

Let $X_1, X_2, ..., X_n$ be a generalized s-FGM sequence and

$$R_{j,n} = \sum_{i=1}^{n} I_{\{X_i \le X_j\}}$$

be the rank of random variable X_j in a collection $X_1, X_2, ..., X_n$. Define the random variables

$$\xi_j = \begin{cases} 1 & \text{if } R_{j,n} = n \\ 0 & \text{otherwise} \end{cases}, \; j = 1, 2, ..., n$$

and $M_n = \max(X_1, X_2, ..., X_n)$.

Lemma 3.1. It is true that

$$P\{\xi_r = 1\} = \frac{1}{n} + \alpha_n \frac{2r-n-1}{2} E\{X_1^n W(X_1)\}, r = 1, 2, ..., n. \qquad (14)$$

Proof. It easy to see that

$$P\{\xi_r = 1\} = P\{X_1 < X_r, X_2 < X_r, ..., X_{r-1} < X_r, X_{r+1} < X_r, ..., X_n < X_r\}$$

$$= \int_0^1 \int_{x_n=0}^{x_r} ... \int_{x_2=0}^{x_r} \int_{x_1=0}^{x_r} \left[1 + \alpha_n \sum_{1 \le i < j \le n} \hat{A}(x_i)\hat{B}(x_j)\right] dx_1 dx_2 ... dx_n$$

$$= \frac{1}{n} + \alpha_n \int_0^1 \int_0^{x_r} ... \int_0^{x_r} \int_0^{x_r} \sum_{1 \le i < j \le n} \hat{A}(x_i)\hat{B}(x_j) dx_1 dx_2 ... dx_n$$

$$= \frac{1}{n} + \alpha_n \left[\frac{(n-1)(n-2)}{2} \int_0^1 x^{n-1} A(x)B(x)dx \right.$$

$$\left. +(r-1) \int_0^1 x^{n-1} A(x)\hat{B}(x)dx + (n-r) \int_0^1 x^{n-1} \hat{A}(x)B(x)dx\right]$$

$$= \frac{1}{n} + \alpha_n \left[-\frac{n-1}{2} \int_0^1 x^{n-1} A(x)\hat{B}(x)dx - \frac{n-1}{2} \int_0^1 x^{n-1} \hat{A}(x)B(x)dx \right.$$

$$\left. +(r-1) \int_0^1 x^{n-1} A(x)\hat{B}(x)dx + (n-r) \int_0^1 x^{n-1} \hat{A}(x)B(x)dx\right]$$

$$= \frac{1}{n} + \alpha_n \frac{2r-n-1}{2} \left[\int_0^1 x^n \left[A(x)B'(x) - A'(x)B(x)\right] dx\right]$$

$$= \frac{1}{n} + \alpha_n \frac{2r-n-1}{2} E(X_1^n W(X_1)).$$

Hence the result.

Theorem 3.1. Let $X_1, X_2, ..., X_n$ be a Generalized-s-FGM random variables whose joint distribution are given by (10) and belongs to $\Im$. Then the following statements are equivalent:

a) $A(x) = B(x)$, i.e. the Generalized-s- FGM random variables are exchangeable

b) $P\{\xi_r = 1\} = \frac{1}{n}$ for some $r \leq n$.

Proof.

Let $A(x) = B(x)$. It can be easily seen that from (14) $P\{\xi_r = 1\} = \frac{1}{n}$, i.e. *a)* $\Rightarrow$ *b)*.

Now let $P\{\xi_r = 1\} = \frac{1}{n}$, for some $r \leq n$. Then from the Lemma 3.1 one has

$$E(X_1^n W(X_1)) = 0$$

or

$$\int_0^1 x^n \left[A(x)B'(x) - A'(x)B(x)\right] dx = 0. \qquad (15)$$

From (15) one obtains $A(x)B'(x) - A'(x)B(x) = 0$ for all x. Using the condition $A(0) = 1$ and $B(0) = 1$ one has $A(x) = B(x)$ for all x.

Corollary 3.1. Let $X_1, X_2, ..., X_n$ be a Generalized-s-FGM random variables whose joint distribution are given by (10) for $A(x) = (1-x)^{\gamma_1}$ and $B(x) = (1-x)^{\gamma_2}, \gamma_1, \gamma_2 \geq 1$. (For bivariate case see Huang and Kotz (1999)) . The assertion $\gamma_1 = \gamma_2$ is valid, if and only if

$$P\{\xi_r = 1\} = \frac{1}{n}$$

for some $r \leq n$.

Proof. Consider $A(x) = (1-x)^{\gamma_1}$ and $B(x) = (1-x)^{\gamma_2}$. We then have

$$A(x)B'(x) - A'(x)B(x) = (\gamma_1 - \gamma_2)(1-x)^{\gamma_2+\gamma_1-1} \begin{cases} \geq 0 & if \quad \gamma_1 \geq \gamma_2 \\ \leq 0 & if \quad \gamma_1 \leq \gamma_2 \end{cases}$$

for all $0 \leq x \leq 1$. An application of Theorem 3.1 concludes the proof.

In conditions of Theorem 3.1 the restriction asserting that the distribution must belong into class $\Im$ can be changed by the independence of ξ_r and M_n for some $r \leq n$. This fact has been expressed in the following.

Theorem 3.2. Let $X_1, X_2, ..., X_n$ be a Generalized-s-FGM random variables whose joint distribution are given by (10). Then the following statements are equivalent:

a) $A(x) = B(x)$, i.e. $X_1, X_2, ..., X_n$ are symmetrically dependent
b) For some $r \leq n$ the rv's ξ_r and M_n are independent and $P\{\xi_r = 1\} = \frac{1}{n}$.
Proof. For $1 \leq r \leq n$ one can write

$$P\{\xi_r = 1, M_n \leq x\}$$
$$= P\{X_1 < X_r, X_2 < X_r, ..., X_{r-1} < X_r, X_{r+1} < X_r, ..., X_n < X_r, X_r \leq x\}$$
$$= \int\limits_{x_1=0}^{x} \int\limits_{x_2=0}^{x_1} ... \int\limits_{x_n=0}^{x_1} \left[1 + \alpha_n \sum_{1 \leq i < j \leq n} \hat{A}(x_i)\hat{B}(x_j)\right] dx_n...dx_2dx_1 = \frac{x^n}{n} +$$
$$+\alpha_n \left\{\frac{n-1}{2} x^n A(x)B(x) + \frac{2r-n-1}{2}\left[\int\limits_0^x t^n \left[A(t)B'(t) - A'(t)B(t)\right] dt\right]\right\} \tag{16}$$

Let $A(x) = B(x)$. Then consider the r.h.s. of (16)

$$P\{\xi_r = 1, M_n \leq x\}$$
$$= \frac{x^n}{n} + \alpha_n \frac{n-1}{2} x^n A^2(x)$$

Since

$$P\{\xi_r = 1\} P\{M_n \leq x\}$$
$$= \frac{x^n}{n}\left\{1 + \alpha_n \binom{n}{2} A^2(x)\right\} \tag{17}$$

a) $\Rightarrow$ *b)*.

Now let ξ_r and M_n are independent and $P\{\xi_r = 1\} = \frac{1}{n}$. Then from (16) one has

$$\frac{x^n}{n} + \alpha_n \left\{\frac{n-1}{2} x^n A(x)B(x) + \frac{2r-n-1}{2}\left[\int\limits_0^x t^n \left[A(t)B'(t) - A'(t)B(t)\right] dt\right]\right\}$$
$$= \frac{x^n}{n} + \alpha_n \frac{n-1}{2} x^n A(x)B(x), \ 0 \leq x \leq 1. \tag{18}$$

Then it follows from (18) that

$$A(x)B'(x) - A'(x)B(x) = 0,$$

which yields $A(x) = B(x)$.

4. EXCEEDANCE MODELS IN MULTIVARIATE SIMPLE FGM DISTRIBUTIONS

Consider a random variable X with the continuous distribution function F and s-FGM random variables $Y_1, Y_2, ..., Y_m$ independent of X and fitting to the model (4) with equal marginal distributions $G(x_j)$ and pdf $g(x_j)$, $-\infty \leq x_j \leq \infty$, $j = 1, 2, ...m$. The joint pdf of $(Y_1, Y_2, ..., Y_m)$ is given as follows:

$$h(x_1, x_2, ..., x_m) = \prod_{i=1}^{m} g(x_i) \left\{ 1 + \alpha_m \sum_{1 \leq j < l \leq m} (1 - 2G(x_j))(1 - 2G(x_l)) \right\},$$

$$-\infty \leq x_1, x_2, ..., x_m \leq \infty$$

with α_m satisfying

$$-\frac{1}{\binom{m}{2}} \leq \alpha_m \leq \frac{1}{\left[\frac{m}{2}\right]} \tag{19}$$

Definition 4.1. For any integer $m \geq 1$

$$S_m = \# \{k \leq m : Y_k \leq X\}$$

denotes the number of Y's falling below the random threshold X.

The exact distribution of S_m is given in the following theorem.

Theorem 4.1. For any integer $m \geq 1$ and real number α_m satisfying (19)

$$P\{S_m = k\}$$

$$= \binom{m}{k} \Big[E(G^k(X)\bar{G}^{m-k}(X)) + \alpha_m \Big(\frac{k(k-1)}{2} E(G^k(X)\bar{G}^{m-k+2}(X))$$

$$-k(m-k)E(G^{k+1}(X)\bar{G}^{m-k+1}(X)) +$$

$$+ \frac{(m-k)(m-k-1)}{2} E(G^{k+2}(X)\bar{G}^{m-k}(X)) \Big) \Big]$$

$$k = 0, 1, ..., m,$$

where $\bar{G}(x) = 1 - G(x)$.

Proof. From the definition of S_m it immediately follows that

$$P\{S_m = k\} = \sum_{i_1, i_2, ..., i_m} P\left\{A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_k} \cap \bar{A}_{i_{k+1}} \cap \bar{A}_{i_{k+2}} ... \cap \bar{A}_{i_{k+m}}\right\} \tag{20}$$

where $A_{i_j} = \left\{Y_{i_j} \le X\right\}$, $i_j \in \{1,2,...,m\}$; $i_j \ne i_l$ if $j \ne l$ and $\bar{A}_{i_j}$ denotes the complement of event A_{i_j}.

$$P\left\{A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_k} \cap \bar{A}_{i_{k+1}} \cap \bar{A}_{i_{k+2}} ... \cap \bar{A}_{i_{k+m}}\right\}$$

$$= P\left\{Y_1 \le X, ..., Y_k \le X, Y_{k+1} > X, ..., Y_m > X\right\}$$

$$= \int_{-\infty}^{\infty} P\left\{Y_1 \le x, ..., Y_k \le x, Y_{k+1} > x, ..., Y_m > x\right\} dF(x)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{x} ... \int_{-\infty}^{x} \int_{x}^{\infty} ... \int_{x}^{\infty} h_{1,...,k,...,m}(x_1, ..., x_k, ..., x_m) dx_1 ... dx_m dF(x)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{x} ... \int_{-\infty}^{x} \int_{x}^{\infty} ... \int_{x}^{\infty} \left[\prod_{i=1}^{m} g(x_i) \left\{1 + \alpha_m \sum_{1 \le j < k \le m} (1 - 2G(x_j))(1 - 2G(x_k))\right\}\right] \times$$

$$\times dx_1 ... dx_m dF(x)$$

By changing variable $G(x_i) = u_i, i = 1,2,...,m$.

$$= \int_{-\infty}^{\infty} \int_{0}^{G(x)} ... \int_{0}^{G(x)} \int_{G(x)}^{1} ... \int_{G(x)}^{1} \left[1 + \alpha_m \sum_{1 \le j < k \le m} (1 - 2u_j))(1 - 2u_k)\right] du_1 ... du_m dF(x)$$

$$= \int_{-\infty}^{\infty} \Big[G(x)^k (1 - G(x))^{m-k} + \alpha_m \Big(\frac{k(k-1)}{2} G(x)^k (1 - G(x))^{m-k+2}$$

$$- k(m-k) G(x)^{k+1} (1 - G(x))^{m-k+1} + \frac{(m-k)(m-k-1)}{2} \times$$

$$\times G(x)^{k+2} (1 - G(x))^{m-k} \Big) \Big] dF(x). \tag{21}$$

The number of summands in (20) is equal to $\binom{m}{k}$ each having the same probability (21). Hence proof is completed.

Corollary 4.1. Let $F = G$, then

$$P\left\{S_m = k\right\}$$

$$= \binom{m}{k} \left[B(k+1, m-k+1) + \alpha_m \left(\frac{k(k-1)}{2} B(k+1, m-k+3) \right.\right.$$

$$-k(m-k)B(k+2, m-k+2)+$$

$$\left.\left. + \frac{(m-k)(m-k-1)}{2} B(k+3, m-k+1) \right) \right] \tag{22}$$

$$k = 0, 1, ..., m, \qquad -\frac{1}{\binom{m}{2}} \leq \alpha_m \leq \frac{1}{\left[\frac{m}{2}\right]}.$$

where $B(a, b)$ is a beta function.

Remark 4.1. Since $\binom{m}{m-k} = \binom{m}{k}$ and $B(a, b)$ is a symmetric function of its arguments it is not difficult to observe from (22) that $P\{S_m = k\} = P\{S_m = m-k\}$.

Let $X_1, X_2, ..., X_n$ be the finite s-FGM sequence having marginal d.f. F and $Y_1, Y_2, ..., Y_m, ...$ be a sequence of i.i.d. random variables with d.f. G. Let $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ be the order statistics of $X_1, X_2, ..., X_n$.

Definition 4.2. For any integer $m \geq 1$ and $1 \leq r \leq n$

$$\nu_m = \# \{k \leq m : Y_k \leq X_{r:n}\}$$

denotes the number of Y's falling below the random threshold $X_{r:n}$.

The exact distribution of ν_m is given in the following theorem.

Theorem 4.2. If $H_0 : F = G$ is true then for any integer $m \geq 1$ and $1 \leq r \leq n$

$$P\{\nu_m = k\}$$

$$= \binom{m}{k} \sum_{s=r}^{n} (-1)^{s-r} \binom{s-1}{r-1} \binom{n}{s} [sB(s+k, m-k+1)$$

$$\left. + \alpha_n \left(\frac{s^2(s-1)}{2} B(s+k, m-k+3) - s(s-1)B(s+k+1, m-k+2) \right) \right] \tag{23}$$

$$k = 0, 1, ..., m \quad , \quad -\frac{1}{\binom{n}{2}} \leq \alpha_n \leq \frac{1}{\left[\frac{n}{2}\right]} \ .$$

Proof. Analogously to the proof of Theorem 4.1

$$P\left\{\nu_m = k\right\}$$

$$= \binom{m}{k} P\left\{Y_1 \le X_{r:n}, ..., Y_k \le X_{r:n}, Y_{k+1} > X_{r:n}, ..., Y_m > X_{r:n}\right\} \qquad (24)$$

One can write

$$P\left\{Y_1 \le X_{r:n}, ..., Y_k \le X_{r:n}, Y_{k+1} > X_{r:n}, ..., Y_m > X_{r:n}\right\}$$

$$= \int_{-\infty}^{\infty} P\left\{Y_1 \le x, ..., Y_k \le x, Y_{k+1} > x, ..., Y_m > x\right\} dF_{r:n}(x). \qquad (25)$$

It is known that for symmetrically dependent random variables

$$P\left\{X_{r:n} \le x\right\} = \sum_{s=r}^{n} (-1)^{s-r} \binom{s-1}{r-1} \binom{n}{s} P\left\{X_{s:s} \le x\right\}. \qquad (26)$$

(see, e.g. David (1981)).

Since for any s - variate marginals of $X_1, X_2, ..., X_n$ the α_s is equal to α_n one can write $P\left\{X_{s:s} \le x\right\} = P\left\{X_1 \le x, ..., X_s \le x\right\} = F(x)^s \left\{1 + \alpha_n \frac{s(s-1)}{2}(1 - F(x))^2\right\}$. Then from (25) and (26) one obtains

$$P\left\{Y_1 \le X_{r:n}, ..., Y_k \le X_{r:n}, Y_{k+1} > X_{r:n}, ..., Y_m > X_{r:n}\right\}$$

$$= \int_{-\infty}^{\infty} G^k(x)(1 - G(x))^{m-k}$$

$$\times \sum_{s=r}^{n} (-1)^{s-r} \binom{s-1}{r-1} \binom{n}{s} \left\{sF(x)^{s-1} f(x) + \alpha_n \frac{s^2(s-1)}{2} F(x)^{s-1}(1 - F(x))^2 f(x)\right.$$

$$\left. - \alpha_n s(s-1) F(x)^s (1 - F(x)) f(x)\right\} dx$$

Since $F = G$

$$P\left\{Y_1 \le X_{r:n}, ..., Y_k \le X_{r:n}, Y_{k+1} > X_{r:n}, ..., Y_m > X_{r:n}\right\}$$

$$= \sum_{s=r}^{n} (-1)^{s-r} \binom{s-1}{r-1} \binom{n}{s} \left[s \int_{0}^{1} u^{s+k-1}(1-u)^{m-k} du \right.$$

$$+\alpha_n \frac{s^2(s-1)}{2} \int_0^1 u^{s+k-1}(1-u)^{m-k+2}du - \alpha_n s(s-1) \int_0^1 u^{s+k}(1-u)^{m-k+1}du \Bigg] \quad (27)$$

After some algebra using (24) from (27) one obtains (23).
The theorem is thus proved.

Corollary 4.2. Let $r = n$ in Theorem 4.2., in this case the exact distribution of ν_m is

$$P\{\nu_m = k\}$$

$$= \binom{m}{k} \left[nB(n+k, m-k+1) \right.$$

$$\left. +\alpha_n \left(\frac{n^2(n-1)}{2} B(n+k, m-k+3) - n(n-1)B(n+k+1, m-k+2) \right) \right].$$

Below we provide some numerical values for the distributions of S_m and ν_m :

m	α_m	k	$P\{S_m = k\}$	m	α_m	k	$P\{S_m = k\}$
2	-0.25	0	0.325	2	0.50	0	0.350
2	-0.25	1	0.350	2	0.50	1	0.300
2	-0.25	2	0.325	2	0.50	2	0.350
3	0.60	0	0.280	5	0.30	0	0.185
3	0.60	1	0.220	5	0.30	1	0.163
3	0.60	2	0.220	5	0.30	2	0.152
3	0.60	3	0.280	5	0.30	3	0.152
				5	0.30	4	0.163
				5	0.30	5	0.185

Table 1. Some numerical values for $P\{S_m = k\}$ when $F = G$.

n	m	α_n	k	$P\{\nu_m = k\}$
3	2	0.50	0	0.121
3	2	0.50	1	0.307
3	2	0.50	2	0.572
4	2	-0.1	0	0.062
4	2	-0.1	1	0.264
4	2	-0.1	2	0.674
5	3	0.1	0	0.020
5	3	0.1	1	0.094
5	3	0.1	2	0.269
5	3	0.1	3	0.617

Table 2. Some numerical values for $P\{\nu_m = k\}$ when $F = G$ and $r = n$.

Numerical example. Consider a system consisting of m dependent components. Suppose that Y_i $(i = 1, 2, ..., m)$ is the strength of a component subject to a stress X. The component fails if at any moment the applied stress exceeds its strength, i.e. if $Y_i < X$ then the ith component fails. In this case S_m shows the number of failured components.

For a system consisting of $m = 5$ dependent components let $Y_1, Y_2, ..., Y_5$ follow s-FGM distribution with F marginals and $\alpha_m = 0.30$. Assume that X has a distribution function F. In such a system if at least three components fail then system fails. Under these assumptions what is the probability that system does not fail?

By using table 1, the required probability is found as

$$p = 1 - P\{S_m \geq 3\} = 1 - [0.152 + 0.163 + 0.185] = 0.5\,.$$

Remark 4.2. Let in Theorem 4.1 $\alpha_m = 0$ i.e. $Y_1, Y_2, ..., Y_m$ are i.i.d. random variables. Then from Theorem 4.1 one obtains

$$P\{S_m = k\} = \binom{m}{k} E(G^k(X)\bar{G}^{m-k}(X))$$

which coincides with the Theorem 1 of Wesolowski and Ahsanullah (1998). From Theorem 4.2 in this case one has

$$P\{\nu_m = k\} = \frac{\binom{r+k-1}{r-1}\binom{m+n-k-r}{n-r}}{\binom{m+n}{n}}$$

which coincides with Corollary 1 (b) of Wesolowski and Ahsanullah (1998) (see also Katzenbeisser (1985), (1986), Matveychuk and Petunin (1990), Johnson and Kotz (1991), (1994)).

REFERENCES

Bairamov, I.G. (1997) Some distribution free properties of statistics based on record values and characterizations of the distributions trough a record. *J. Appl. Statist. Sci.*, **5**, 1, 17-25.

Bairamov, I.G. and Eryilmaz, S. (2000) Distributional properties of statistics based on minimal spacing and record exceedance statistics. *J. Stat. Plan. Infer.,* **90**, 21-33.

Bairamov, I.G. and Kotz, S. (2001) On distributions of exceedances associated with order statistics and record values for arbitrary distributions. *Statist. Papers.* Vol. **42**, Issue 2, 171-185.

Bairamov, I. G. and Kotz S. (2002) Dependence structure and symmetry of Huang-Kotz FGM distributions and their extensions. *Metrika, **56**, 55-72.*

Cambanis, S. (1993) On Eyraud-Gumbel-Morgenstern random processes in stochastic inequalities, *IMS Lecture Notes-22*, Hayward, California.

Conway, D. (1983) Farlie-Gumbel-Morgenstern distributions. In *Encyclopedia of Statistical Sciences-3* (S.Kotz and N.L.Johnson, eds), 28-31, New York:John Wiley and Sons.

David, H.A.(1981) *Order Statistics.* Wiley. New York.

Eryilmaz, S. (2002) Random threshold models based on multivariate observations. *J. Stat. Plan. Infer.,* in press.

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

Gumbel, E. J. (1960) Bivariate exponential distributions. *Journ. Amer. Statist. Assoc.*, **55**, 698-707.

Gumbel, E.J. and Schelling, H, von (1950) The distribution of the number of exceedances. *Ann. Math. Statist.*, **21**, 247-262.

Huang, J. S. and Kotz, S. (1999) On modifications of the Farlie- Gumbel- Morgenstern distributions. A tough hill to climb. *Metrika,* **49**, 135-145.

Hutchinson, T.R. and Lai, C.D. (1990) *Continuous Bivariate Distributions. Emphasizing Applications*, Adelaide, Australia: Rumsby Scientific Publishing.

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

Johnson, N. L. and Kotz, S. (1977) On some generalized Farlie- Gumbel- Morgenstern distributions- II: Regression, correlation and further generalizations. *Commun. Stat.Theory*, **6**, 485-496.

Johnson, N.L., Kotz, S. and Kemp, A. (1992) *Univariate discrete distributions.* Wiley, New York.

Johnson, N. and Kotz, S. (1991) Some generalizations of Bernoulli and Polya-Eggenberger contagion models. *Statist. Papers*, **32**, 1-17.

Johnson, N. and Kotz, S. (1994) Further comments on Matveychuk and Petunin's generalized Bernoulli model, and nonparametric tests of homogeneity. *J Stat. Plan. Infer.*, **41**, 61-72.

Katzenbeisser, W. (1985) The distribution of two sample location exceedance test statistics under Lehmann alternatives. *Statist. Hefte* (now *Papers*), **26**,131-138.

Katzenbeisser, W. (1986) The exact power of two-sample location tests based on exceedances against shift alternatives. *Math. Op. Forsch. (Statistics),* **20**, 47-54.

Kemp, C.J. (1995) A Farlie-Gumbel-Morgenstern-type bivariate generalized gamma distribution. *Technical Report No 226*, University of the Orange Free State.

Kotz, S. and Seeger J.P. (1991) A new approach to dependence in multivariate distributions. *Advances in Probability Distributions with Given Marginals.* (G. Dall'Aglio et al., eds.) pp. 113-127, Dordrecht, The Netherlalands: Kluwer Academic Publishers.

Kotz, S., Balakrishnan, N.and Johnson, N.L. (2000) *Continuous Multivariate Distributions.* John Wiley&Sons, Inc. New York.

Lai, C. D. and Xie, M. (2000) A new family of positive quadrant dependent bivariate distributions. *Stat. Probabil. Lett.*, **46** , 359-364.

Leadbetter, M.L., Lindgren, G. and Rootzen, H. (1983) *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag, New-York.

Lin, G.D. (1987) Relantionship between two extensions of Farlie-Gumbel-Morgenstern distributions. *Ann. Inst. Statist. Math.*, **39**, 129-140.

Matveychuk, S.A., Petunin Yu.I.(1990/1991) A generalization of the Bernoulli model arising in order statistics. I/II *Ukr. Math. J.*, **42**, 518-528/ **43**, 779-785.

Morgenstern, D. (1956) Einfache Beispiele zweidimensionaler Verteilung, *Mitteislingsblatt für Mathematische Statistik*, **8**, 234-235.

Nelsen, R.B. (1999) *An Introduction to Copulas.* Springer, New York.

Sarkadi, K. (1957) On the distribution of the number of exceedances. *Ann. Math. Stat.*, **28**, 1021-1022.

Wesolowski, J. and Ahsanullah, M. (1998) Distributional properties of exceedance statistics. *Ann. Inst. Statist. Math.*, **50**, 543-565.