Ann. Inst. Statist. Math.
39 (1987), Part A, 129-140

RELATIONSHIPS BETWEEN TWO EXTENSIONS OF FARLIE-GUMBEL-MORGENSTERN DISTRIBUTION

GWO DONG LIN*

(Received Oct. 26, 1984; revised Nov. 14, 1985)

Summary

In order to increase the dependence between two random variables X and Y obeying the type of Farlie-Gumbel-Morgenstern (FGM) distribution, Johnson and Kotz (1977, *Commun. Statist.*, **6**, 485–496) introduced the $(k-1)$-iteration FGM distribution:

$$H_{1k}=FG+\sum_{j=1}^{k}\alpha_{1j}(FG)^{[j/2]+1}(\bar{F}\bar{G})^{[(j+1)/2]}\ ,$$

where F and G are the respective marginal distributions of X and Y. Recently, Huang and Kotz (1984, *Biometrika*, **71**, 633–636) found the natural parameter space of H_{12} for arbitrary *absolutely continuous* distributions F and G. We extend their result to arbitrary *continuous* distributions F and G and propose another $(k-1)$-iteration FGM distribution:

$$H_{2k}=FG+\sum_{j=1}^{k}\alpha_{2j}(FG)^{[(j+1)/2]}(\bar{F}\bar{G})^{[j/2]+1}\ .$$

For some F and G, the correlation coefficient for H_{2k} is greater than that for H_{1k}.

Further, we find the conditions on F and G under which H_{1k} and H_{2k} have the same natural parameter space. We also find that for arbitrary symmetric distributions F and G with finite means, the covariances between X and Y are the same whatever the joint distribution H_{ik} $(i=1, 2)$ they have. A result of Schucany, Parr and Boyer (1978, *Biometrika*, **65**, 650–653) about the correlation coefficient for FGM distribution is extended to arbitrary distributions F and G. The multivariate case is also discussed.

---

* Work finally completed while visiting Department of Statistics, Stanford University. The author was partially supported by the Chinese National Science Council.

Key words and phrases: FGM distribution, correlation coefficient, absolute continuity, continuity and natural parameter space.

1. Introduction and motivation

A well-known way to construct a bivariate distribution with the given marginal distributions F and G is to consider the Farlie-Gumbel-Morgenstern (FGM) distribution:

$$H(x, y)=F(x)G(y)\{1+\alpha\bar{F}(x)\bar{G}(y)\}\ , \tag{1}$$

where $\bar{F}=1-F$, $\bar{G}=1-G$, and α is a real number such that H is a bivariate distribution. From (1) we also understand that it is impossible to identify the bivariate distribution only by its marginal distributions, since there are many admissible numbers α in general. The set of admissible number α in (1) is called the natural parameter space of H and is denoted by Λ. The usefulness of the FGM distribution H in (1) depends on how many admissible numbers α we have and on what values the correlation coefficient ρ of X and Y may be.

It is trivial that if one of F and G is degenerate, then we have $\Lambda=(-\infty, \infty)$. For nondegenerate distributions F and G, Cambanis [1] showed that H is a bivariate distribution if and only if $\alpha\in\Lambda=[\alpha_{\min}, \alpha_{\max}]$, where

$$\alpha_{\min}=-\min\{(M_F M_G)^{-1}, ((1-m_F)(1-m_G))^{-1}\}\ ,$$

$$\alpha_{\max}=\min\{(M_F(1-m_G))^{-1}, ((1-m_F)M_G)^{-1}\}\ ,$$

and m_F, M_F are the infimum and supremum of the set $\{F(x): -\infty<x<\infty\}-\{0, 1\}$, respectively. For absolutely continuous distributions F and G, we can see that $\alpha_{\min}=-1$, $\alpha_{\max}=1$, and hence $\Lambda=[-1, 1]$, a result of Johnson and Kotz [3]. As to the correlation coefficient ρ of X and Y, ρ may assume the maximal value 1 (minimal value -1, resp.) if we let $F=G$, $\Pr(X=1)=\Pr(X=-1)=1/2$ and $\alpha=4\in\Lambda=[-4, 4]$ ($\alpha=-4$, resp.). However, Schucany, Parr and Boyer [6] showed that $|\rho|\leqq 1/3$ if both F and G are arbitrary absolutely continuous distributions with finite nonzero variances.

In order to increase the dependence between random variables X and Y in (1), Johnson and Kotz [4] proposed the $(k-1)$-iteration FGM distribution:

$$H_{1k}=FG+\sum_{j=1}^{k}\alpha_{1j}(FG)^{[j/2]+1}(\bar{F}\bar{G})^{[(j+1)/2]}\ , \tag{2}$$

where k is any positive integer, $[z]$ denotes the greatest integer less than or equal to z and we have omitted the variables x and y without confusion.

Recently Huang and Kotz [2] considered the one-iteration FGM distribution ($k=2$):

(3) $$H_{12}=FG+\alpha_{11}(FG)(\bar{F}\bar{G})+\alpha_{12}(FG)^2(\bar{F}\bar{G}) .$$

For arbitrary *absolutely continuous* distributions F and G, they found that the natural parameter space Λ_{12} of H_{12} is the set

(4) $$\left\{(\alpha_1, \alpha_2) : |\alpha_1|\leqq 1,\ -\alpha_1-1\leqq\alpha_2\leqq\frac{1}{2}[3-\alpha_1+(9-6\alpha_1-3\alpha_1^2)^{1/2}]\right\}$$

and further

(5) The maximal correlation coefficient ρ corresponding to H_{12} is higher than 1/3 which is the maximal ρ corresponding to $H_{11}=H$, but the former is less than or equal to $((1627/\sqrt{4881})-3)/40=0.5027$;

(6) One single iteration can result in nearly tripling the covariance for certain marginals;

(7) There exist no marginals for which the single iteration will bring about higher negative correlation.

If we exchange the two powers of the third term in (3), namely, if we consider the bivariate distribution

(8) $$H_{22}=FG+\alpha_{21}(FG)(\bar{F}\bar{G})+\alpha_{22}(FG)(\bar{F}\bar{G})^2 ,$$

then we find that for some distributions F and G, the correlation coefficient of X and Y in (8) is greater than that of X and Y in (3) (see Example 2 in Section 5 in detail). This is the motivation to study the other $(k-1)$-iteration *FGM* distribution:

(9) $$H_{2k}=FG+\sum_{j=1}^{k}\alpha_{2j}(FG)^{[(j+1)/2]}(\bar{F}\bar{G})^{[j/2]+1} .$$

In Section 2 we shall prove that for $i=1, 2$ and for arbitrary *continuous* distributions F and G, the natural parameter space Λ_{i2} of H_{i2} is also equal to the set (4). Section 3 will derive the bivariate distribution H_{2k} as Johnson and Kotz [4] did for H_{1k}. In Section 4 we shall study the condition on F and G under which H_{1k} and H_{2k} have the same natural parameter space. Section 5 will prove that for arbitrary symmetric distributions F and G with finite means, the covariances between X and Y are the same whatever the joint distribution H_{ik} $(i=1, 2)$ they have. Based on the results of Section 5, Section 6 will consider the improvements in correlation coefficient of X and Y which obey the *FGM* distribution or one-iteration *FGM* distributions. Finally, the multivariate case is discussed in Section 7.

2. Natural parameter space: the case $k=2$

In this section we shall prove that for $i=1, 2$ and for arbitrary continuous distributions F and G, the natural parameter space Λ_{i2} of H_{i2} is the same as the set (4). Let us recall that the bivariate function H_{12} in (3) is a bivariate distribution if and only if for all $x \leqq x'$ and $y \leqq y'$, we have $\Delta_{x'}\Delta_{y'}H_{12}(x, y) \geqq 0$, where $\Delta_{z'}N(z) \equiv N(z')-N(z)$. For $i=1, 2$, define H_{i2}^{ud} be the bivariate distribution H_{i2} with uniform marginal distributions on $[0, 1]$ and Λ_{i2}^{ud} the natural parameter space of H_{i2}^{ud}. Then we have the following

THEOREM 1. *For arbitrary continuous distributions F and G, the natural parameter space $\Lambda_{i2}=\Lambda_{i2}^{ud}$, $i=1, 2$.*

PROOF. For coefficient $(\alpha_{i1}, \alpha_{i2})$, $H_{i2}^{ud}(u, v)$ is a bivariate distribution on $[0, 1]\times[0, 1]$

$$(10) \quad \Longleftrightarrow \Delta_{u'}\Delta_{v'}H_{i2}^{ud}(u, v) \geqq 0 \qquad \forall\, u \leqq u',\ v \leqq v' \text{ and } u, u', v, v' \in [0, 1]$$

$$(11) \quad \Longleftrightarrow \Delta_{F(x')}\Delta_{G(y')}H_{i2}^{ud}(F(x), G(y)) \geqq 0$$
$$\forall\, x \leqq x',\ y \leqq y' \text{ and } x, x', y, y' \in R=(-\infty, \infty)$$

$$\Longleftrightarrow \Delta_{x'}\Delta_{y'}H_{i2}(x, y) \geqq 0 \qquad \forall\, x \leqq x',\ y \leqq y' \text{ and } x, x', y, y' \in R$$

$$\Longleftrightarrow H_{i2}(x, y) \text{ is a bivariate distribution on } R\times R.$$

Hence $\Lambda_{i2}=\Lambda_{i2}^{ud}$, $i=1, 2$.

Note that Λ_{12}^{ud} is the set (4) due to Huang and Kotz [2], so we have improved their result for arbitrary *continuous* distributions F and G. The continuity condition on F and G is necessary in the direction (11) $\Rightarrow$(10) above. From the proof of Theorem 1, we can understand that for *arbitrary* distributions F and G, $\Lambda_{i2}^{ud} \subset \Lambda_{i2}$ (see, e.g., Example 1 in Section 4 for discrete distributions F and G).

Next, in order to claim that Λ_{22} is the same as the set (4) for arbitrary *continuous* distributions F and G, it suffices to prove the following

LEMMA 1. $\Lambda_{12}^{ud}=\Lambda_{22}^{ud}$.

PROOF. Since $F(x)=x$ and $G(y)=y$ for $x, y \in [0, 1]$, we have

$$H_{12}(x, y)=xy+\alpha_{11}(xy)(1-x)(1-y)+\alpha_{12}(xy)^2(1-x)(1-y)\,, \qquad x, y \in [0, 1]\,,$$

and hence the joint density function of X and Y in this case is

$$h_{12}(x, y)=1+\alpha_{11}(1-2x)(1-2y)+\alpha_{12}(xy)(2-3x)(2-3y)\,, \qquad x, y \in [0, 1]\,.$$

Similarly, the joint density function of X and Y in the other case is

$$h_{22}(x, y) = 1 + \alpha_{21}(1-2x)(1-2y) + \alpha_{22}(1-x)(1-3x)(1-y)(1-3y), \quad x, y \in [0, 1].$$

Taking the transformations $u = 1-x$ and $v = 1-y$ in $h_{12}(x, y)$, we have

$$\begin{aligned} h_{12}^*(u, v) &\equiv h_{12}(1-u, 1-v) \\ &= 1 + \alpha_{11}(1-2u)(1-2v) + \alpha_{12}(1-u)(1-3u)(1-v)(1-3v), \quad u, v \in [0, 1]. \end{aligned}$$

Therefore, for coefficient $\alpha = (\alpha_1, \alpha_2)$, H_{12} is a bivariate distribution

$$\begin{aligned} &\Longleftrightarrow h_{12}(x, y) \geqq 0, \quad \forall x, y \in [0, 1] \\ &\Longleftrightarrow h_{12}^*(u, v) \geqq 0, \quad \forall u, v \in [0, 1] \\ &\Longleftrightarrow h_{22}(x, y) \geqq 0, \quad \forall x, y \in [0, 1] \\ &\Longleftrightarrow H_{22} \text{ is a bivariate distribution} \end{aligned}$$

That is, $\Lambda_{12}^{ud} = \Lambda_{22}^{ud}$.

In Section 4 we shall extend Lemma 1 to a wide class of distributions F and G by another method (Theorem 3). It can be seen that Theorem 1 is still true for the general case, namely, for any fixed $i = 1, 2$ and $k \geqq 2$, the natural parameter spaces Λ_{ik} of H_{ik} are the same for arbitrary continuous distributions F and G.

3. Derivations of H_{1k} and H_{2k}

Johnson and Kotz [4] derived H_{1k} by the following successive $k-1$ steps, so it is named after a $(k-1)$-iteration *FGM* distribution. Substituting $S(x, y) = \Pr(X > x, Y > y)$ for the $\bar{F}\bar{G}$ in (1), and using an equivalent form of the *FGM* distribution,

$$S = \bar{F}\bar{G}\{1 + \beta_1 FG\}, \tag{12}$$

we obtain

$$H_{12} = FG\{1 + \alpha\bar{F}\bar{G}(1 + \beta_1 FG)\}. \tag{13}$$

Then substituting the *FGM* distribution $H = FG(1 + \beta_2\bar{F}\bar{G})$ for the last FG in (13) yields

$$H_{13} = FG\{1 + \alpha\bar{F}\bar{G}[1 + \beta_1(FG(1 + \beta_2\bar{F}\bar{G}))]\}. \tag{14}$$

Continuing this procedure, intersubstituting the forms of (12) and (1), $k-3$ more iterations shall lead to H_{1k}.

Similarly, we can obtain H_{2k} as follows. We first begin with the equivalent form (12) of the FGM distribution. Substituting the FGM distribution $H=FG(1+\alpha_1\bar{F}\bar{G})$ for the FG in (12), we obtain

$$S=\bar{F}\bar{G}\{1+\beta_1 FG(1+\alpha_1\bar{F}\bar{G})\} . \tag{15}$$

Then substituting $S=\bar{F}\bar{G}(1+\beta_2 FG)$, a form of (12), for the last $\bar{F}\bar{G}$ in (15), yields

$$S=\bar{F}\bar{G}\{1+\beta_1 FG[1+\alpha_1\bar{F}\bar{G}(1+\beta_2 FG)]\} .$$

Continuing this procedure, intersubstituting the forms of (1) and (12), $k-3$ more iterations shall lead to

$$S=\bar{F}\bar{G}+\sum_{j=1}^{k}\alpha_{2j}(FG)^{[(j+1)/2]}(\bar{F}\bar{G})^{[j/2]+1} . \tag{16}$$

Recall $S(x, y)=1-F(x)-G(y)+\Pr(X\leqq x, Y\leqq y)$, then we know that (16) and (9) are equivalent.

4. Natural parameter space: the general case

Denote $m_j=[j/2]+1$ and $n_j=[(j+1)/2]$ for convenience. In fact, the results of Sections 4 and 5 remain true for any positive integers m_j and n_j. Huang and Kotz [2] proved that the natural parameter space Λ_{1k} of H_{1k} is convex if both F and G are absolutely continuous. We first assert that their conclusion is also true for arbitrary distributions F and G.

THEOREM 2. *For arbitrary distributions F and G, the natural parameter space Λ_{ik} of H_{ik} is convex, where $i=1, 2$.*

PROOF. We only prove that Λ_{1k} is a convex set since the proof of Λ_{2k} is similar to that of Λ_{1k}. Let $0\leqq p\leqq 1$, $\alpha=(\alpha_1, \alpha_2, \cdots, \alpha_k)\in\Lambda_{1k}$ and $\alpha^*=(\alpha_1^*, \alpha_2^*, \cdots, \alpha_k^*)\in\Lambda_{1k}$. It suffices to prove that $\beta\equiv p\alpha+(1-p)\alpha^*\in\Lambda_{1k}$, that is, to prove that $H_{1k}^{(\beta)}\equiv FG+\sum_{j=1}^{k}(p\alpha_j+(1-p)\alpha_j^*)(FG)^{m_j}(\bar{F}\bar{G})^{n_j}$ is a bivariate distribution, or equivalently, to prove that for all $x\leqq x'$ and $y\leqq y'$, $\Delta_{x'}\Delta_{y'}H_{1k}^{(\beta)}(x, y)\geqq 0$. And the desired result follows from

$$\Delta_{x'}\Delta_{y'}H_{1k}^{(\beta)}(x, y)=p\Delta_{x'}\Delta_{y'}H_{1k}^{(\alpha)}(x, y)+(1-p)\Delta_{x'}\Delta_{y'}H_{1k}^{(\alpha^*)}(x, y)\geqq 0 ,$$

where $H_{1k}^{(\alpha)}$ and $H_{1k}^{(\alpha^*)}$ are bivariate distributions H_{1k} with coefficients α and α^* respectively, and the last inequality is implied by the fact that $\alpha, \alpha^*\in\Lambda_{1k}$.

Theorem 2 is useful in that if we want to find the natural parameter space, then it suffices by this theorem to find its extreme points.

Let us define the new class $\mathscr{D}$ of distributions as follows:

$$\begin{aligned}\mathscr{D} \equiv \{F &: \text{the closure of the range of} \\ &\quad \text{distribution } F \text{ is symmetric about } 1/2\} \\ &= \{F: \forall F(x),\ \exists \{x_n^*\}_{n=1}^{\infty} \ni F(x) = 1 - \lim_{n\to\infty} F(x_n^*)\} .\end{aligned}$$

Notice that $\mathscr{D}$ contains all the continuous distributions and all the symmetric distributions. Then we extend Lemma 1 in Section 2 to the following

THEOREM 3. *For arbitrary distributions* $F, G \in \mathscr{D}$, $\Lambda_{1k} = \Lambda_{2k}$.

PROOF. Recall that for $x \leqq x'$, $y \leqq y'$,

$$\begin{aligned}\Delta_{x'}\Delta_{y'}H_{1k}(x, y) = &(F(x') - F(x))(G(y') - G(y)) \\ &+ \sum_{j=1}^{k} \alpha_{1j}[F(x')^{m_j}(1 - F(x'))^{n_j} - F(x)^{m_j}(1 - F(x))^{n_j}] \\ &\cdot [G(y')^{m_j}(1 - G(y'))^{n_j} - G(y)^{m_j}(1 - G(y))^{n_j}] ,\end{aligned}$$

and that for $x_0 \leqq x_0'$, $y_0 \leqq y_0'$,

$$\begin{aligned}\Delta_{x_0'}\Delta_{y_0'}H_{2k}(x_0, y_0) = &(F(x_0') - F(x_0))(G(y_0') - G(y_0)) \\ &+ \sum_{j=1}^{k} \alpha_{2j}[F(x_0')^{n_j}(1 - F(x_0'))^{m_j} - F(x_0)^{n_j}(1 - F(x_0))^{m_j}] \\ &\cdot [G(y_0')^{n_j}(1 - G(y_0'))^{m_j} - G(y_0)^{n_j}(1 - G(y_0))^{m_j}] .\end{aligned} \tag{17}$$

Suppose $\alpha = (\alpha_{11}, \cdots, \alpha_{1k}) \in \Lambda_{1k}$, that is, for all $x \leqq x'$, $y \leqq y'$, we have $\Delta_{x'}\Delta_{y'} \cdot H_{1k}(x, y) \geqq 0$. Then for any fixed $x_0 \leqq x_0'$, $y_0 \leqq y_0'$, and for any $n = 1, 2, \cdots$, there exist $x_n \leqq x_n'$, $y_n \leqq y_n'$ such that

$$1 - \lim_{n\to\infty} F(x_n) = F(x_0') , \qquad 1 - \lim_{n\to\infty} F(x_n') = F(x_0) ,$$
$$1 - \lim_{n\to\infty} G(y_n) = G(y_0') , \qquad 1 - \lim_{n\to\infty} G(y_n') = G(y_0) ,$$

and hence with coefficient α,

$$\Delta_{x_0'}\Delta_{y_0'}H_{2k}(x_0, y_0) = \lim_{n\to\infty} \Delta_{x_n'}\Delta_{y_n'}H_{1k}(x_n, y_n) \geqq 0 . \tag{18}$$

This means $\alpha \in \Lambda_{2k}$. We have proved $\Lambda_{1k} \subset \Lambda_{2k}$. Similarly, $\Lambda_{2k} \subset \Lambda_{1k}$ and hence $\Lambda_{1k} = \Lambda_{2k}$.

If one of F and G is not a distribution in $\mathscr{D}$, then the result $\Lambda_{1k} = \Lambda_{2k}$ is not always true. See the following example.

Example 1. Let $F = G$, $\Pr(X = -1) = 2/3$ and $\Pr(X = 1) = 1/3$, then $F \notin \mathscr{D}$. It can be seen that

$$H_{12}(-1, -1) = \frac{4}{9} + \frac{4}{81}\alpha_{11} + \frac{16}{729}\alpha_{12} ,$$

$$H_{12}(1,-1)=H_{12}(-1,1)=2/3\ ,\qquad H_{12}(1,1)=1\ ,$$

and that

$$H_{22}(-1,-1)=\frac{4}{9}+\frac{4}{81}\alpha_{21}+\frac{4}{729}\alpha_{22}\ ,$$

$$H_{22}(1,-1)=H_{22}(-1,1)=2/3\ ,\qquad H_{22}(1,1)=1\ .$$

Solving

$$\begin{cases} H_{12}(-1,-1)\geqq 0 \\ H_{12}(1,-1)-H_{12}(-1,-1)\geqq 0 \\ H_{12}(-1,1)-H_{12}(-1,-1)\geqq 0 \\ H_{12}(1,1)-H_{12}(-1,1)-H_{12}(1,-1)+H_{12}(-1,-1)\geqq 0\ , \end{cases}$$

we obtain

$$A_{12}=\left\{(\alpha_1,\alpha_2):\ -\frac{9}{4}\leqq\alpha_1\leqq\frac{9}{2},\ -\frac{81}{16}-\frac{9}{4}\alpha_1\leqq\alpha_2\leqq\frac{81}{8}-\frac{9}{4}\alpha_1\right\}\ .$$

Similarly,

$$A_{22}=\left\{(\alpha_1,\alpha_2):\ -\frac{9}{4}\leqq\alpha_1\leqq\frac{9}{2},\ -\frac{81}{4}-9\alpha_1\leqq\alpha_2\leqq\frac{81}{2}-9\alpha_1\right\}\ .$$

It is clear that $A_{12}\subsetneqq A_{22}$ for this example. On the other hand, if we let $\Pr(X=-1)=1/3$ and $\Pr(X=1)=2/3$, then $A_{12}\supsetneqq A_{22}$. From this example we also understand that it is possible to find A_{ik} ($k>2$) as long as both F and G are finite discrete distributions.

5. Covariance: the general case

Let $X_{k,n}(Y_{k,n})$ denote the k-th smallest order statistic of a sample of size n from arbitrary distribution $F(G)$. (We don't assume the absolute continuity here.) Let $F_{k,n}(G_{k,n})$ be the distribution of $X_{k,n}(Y_{k,n})$. Furthermore, assume that EX and EY exist and are finite, hence implying the finiteness of $E(X_{k,n})\equiv\mu_{k,n}$ and $E(Y_{k,n})\equiv\nu_{k,n}$. Then the identity $F^k\bar{F}^{n-k}=\binom{n}{k}^{-1}(F_{k,n}-F_{k+1,n})$ implies

$$H_{1k}=FG+\sum_{j=1}^{k}\alpha_{1j}\binom{m_j+n_j}{m_j}^{-2}(F_{m_j,m_j+n_j}-F_{m_j+1,m_j+n_j})\cdot(G_{m_j,m_j+n_j}-G_{m_j+1,m_j+n_j})\ ,$$

whose expectation is (see, for example, Royden [5], p. 272)

$$E_{1k}(XY)=EXEY+\sum_{j=1}^{k}\alpha_{1j}\binom{m_j+n_j}{m_j}^{-2}(\mu_{m_j,m_j+n_j}-\mu_{m_j+1,m_j+n_j})\cdot(\nu_{m_j,m_j+n_j}-\nu_{m_j+1,m_j+n_j})\,.$$

Using the triangular identity $(n-k)\mu_{k,n}+k\mu_{k+1,n}=n\mu_{k,n-1}$, we obtain the covariance of X and Y corresponding to H_{1k},

$$\text{(19)}\qquad \begin{aligned}\mathrm{cov}_{1k}(X,Y)&\equiv E_{1k}(XY)-EXEY\\&=\sum_{j=1}^{k}\alpha_{1j}\binom{m_j+n_j-1}{m_j}^{-2}(\mu_{m_j,m_j+n_j-1}-\mu_{m_j+1,m_j+n_j})\\&\quad\cdot(\nu_{m_j,m_j+n_j-1}-\nu_{m_j+1,m_j+n_j})\,.\end{aligned}$$

Similarly, the covariance of X and Y corresponding to H_{2k} is

$$\text{(20)}\qquad \begin{aligned}\mathrm{cov}_{2k}(X,Y)&\equiv\sum_{j=1}^{k}\alpha_{2j}\binom{m_j+n_j-1}{m_j}^{-2}(\mu_{n_j,m_j+n_j}-\mu_{n_j,m_j+n_j-1})\\&\quad\cdot(\nu_{n_j,m_j+n_j}-\nu_{n_j,m_j+n_j-1})\,.\end{aligned}$$

The following theorem states the relationship between cov_{1k} and cov_{2k}.

THEOREM 4. *Let F and G be two arbitrary symmetric distributions with finite expectations, and let $\alpha_{1j}=\alpha_{2j}$, for $j=1,2,\cdots,k$. Then* $\mathrm{cov}_{1k}(X,Y)=\mathrm{cov}_{2k}(X,Y)$.

PROOF. We first assume that $EX=EY=0$. Since F and G are symmetric distributions, we have

$$\mu_{k,n}=-\mu_{n-k+1,n}\qquad\text{and}\qquad \nu_{k,n}=-\nu_{n-k+1,n}\,.$$

Thus, for $j=1,2,\cdots,k$,

$$\begin{aligned}&(\mu_{m_j,m_j+n_j-1}-\mu_{m_j+1,m_j+n_j})(\nu_{m_j,m_j+n_j-1}-\nu_{m_j+1,m_j+n_j})\\&\quad=(\mu_{n_j,m_j+n_j}-\mu_{n_j,m_j+n_j-1})(\nu_{n_j,m_j+n_j}-\nu_{n_j,m_j+n_j-1})\,,\end{aligned}$$

and hence

$$\mathrm{cov}_{1k}(X,Y)=\mathrm{cov}_{2k}(X,Y)\,,$$

for the case $EX=EY=0$. Now, for general case we assume $EX=\mu$, $EY=\nu$, and $X^*=X-\mu$, $Y^*=Y-\nu$. Then applying $EX^*_{k,n}=\mu_{k,n}-\mu$ and $EY^*_{k,n}=\nu_{k,n}-\nu$ to (19) and (20) yields the desired result

$$\mathrm{cov}_{1k}(X,Y)=\mathrm{cov}_{1k}(X^*,Y^*)=\mathrm{cov}_{2k}(X^*,Y^*)=\mathrm{cov}_{2k}(X,Y)\,,$$

in which the second equality follows from the reason $EX^*=EY^*=0$.

If one of F and G is not symmetric, then $\mathrm{cov}_{1k}(X,Y)=\mathrm{cov}_{2k}(X,Y)$ is not always true. See the following example.

Example 2. Let $F=G$ be the triangular distribution $F(x)=x^2$, $x\in$

[0, 1]. Then by the formula

$$\mu_{k,n}=k\binom{n}{k}\int_0^1 F^{-1}(t)t^{k-1}(1-t)^{n-k}dt\ , \tag{21}$$

we can calculate

$$\mu_{1,1}=2/3,\quad \mu_{1,2}=8/15,\quad \mu_{2,2}=4/5,\quad \mu_{1,3}=16/35,\quad \mu_{2,3}=24/35,\quad \mu_{3,3}=6/7\ .$$

Thus for $\alpha_1=\alpha_{11}=\alpha_{21}$ and $\alpha_2=\alpha_{12}=\alpha_{22}$,

$$\text{cov}_{12}(X, Y)=\alpha_1(\mu_{2,2}-\mu_{1,1})^2+\alpha_2(\mu_{3,3}-\mu_{2,2})^2=\alpha_1(2/15)^2+\alpha_2(2/35)^2\ ,$$

and similarly

$$\text{cov}_{22}(X, Y)=\alpha_1(\mu_{1,2}-\mu_{1,1})^2+\alpha_2(\mu_{1,3}-\mu_{1,2})^2=\alpha_1(2/15)^2+\alpha_2(8/105)^2\ .$$

It is clear that $\text{cov}_{12}(X, Y)<\text{cov}_{22}(X, Y)$ if $\alpha_2>0$.

6. Improvements in the correlation coefficient

Based on the results of Section 5, this section will study the correlation coefficient ρ of X and Y which obey the *FGM* distribution or one-iteration *FGM* distributions. As mentioned in Section 1, Schucany et al. [6] proved that $|\rho|\leqq 1/3$ for the *FGM* distribution (1) with *absolutely continuous* marginal distributions F and G. Example 3 below shows that ρ really increases if X and Y obey the one-iteration *FGM* distributions H_{12} or H_{22}. Further, we shall extend in Theorem 5 the result of Schucany et al. [6] to *arbitrary* distributions F and G.

Example 3. (See, also Huang and Kotz [2]). We assume $\alpha_1=\alpha_{11}=\alpha_{21}$ and $\alpha_2=\alpha_{12}=\alpha_{22}$ in this example.

(a) Let $F=G$ be the uniform distribution on [0, 1], then by Theorems 3 and 4 and the result (4), we have

$$\text{cov}_{12}(X, Y)=\text{cov}_{22}(X, Y)=\alpha_1/36+\alpha_2/144\ ,$$

$\rho=\alpha_1/3+\alpha_2/12$, and max $\rho=(\sqrt{13}-1)/6=0.43426$ for both H_{12} and H_{22}.

(b) Let $F=G$ be the standard normal distribution, then similarly,

$$\rho=\text{cov}_{12}(X, Y)=\text{cov}_{22}(X, Y)=\alpha_1/\pi+\alpha_2/(4\pi)\ ,$$

and max $\rho=(\sqrt{13}-1)/(2\pi)=0.41469$ for both H_{12} and H_{22}.

THEOREM 5. *In the FGM distribution, let F and G be arbitrary distributions with finite nonzero variances, then the correlation coefficient ρ satisfies*

$$\frac{1}{3}\alpha_{\min}\leqq\rho\leqq\frac{1}{3}\alpha_{\max}\ ,$$

where $\alpha_{\min}$ and $\alpha_{\max}$ are defined in Section 1.

PROOF. Recall that Formula (21) is also true for arbitrary distribution F if we define the inverse function $F^{-1}(t)\equiv\inf\{x\colon F(x)\geqq t\}$, $t\in(0,1)$. Without loss of generality, we may assume $EX=EY=0$ in the following discussion. By Cauchy-Schwarz inequality we have

$$(\mu_{2,2}-\mu_{1,1})^2=\left(\int_0^1 F^{-1}(t)(2t-1)dt\right)^2\leqq\int_0^1 (F^{-1}(t))^2dt\int_0^1 (2t-1)^2dt=\frac{1}{3}\sigma_X^2\ ,$$

that is,

$$(\mu_{2,2}-\mu_{1,1})/\sigma_X\leqq\frac{1}{\sqrt{3}}\ .$$

Similarly, for distribution G we have

$$(\nu_{2,2}-\nu_{1,1})/\sigma_Y\leqq\frac{1}{\sqrt{3}}\ .$$

Therefore,

$$\rho=\operatorname{cov}(X,Y)/(\sigma_X\sigma_Y)=\alpha[(\mu_{2,2}-\mu_{1,1})/\sigma_X][(\nu_{2,2}-\nu_{1,1})/\sigma_Y]$$
$$\in\left[\frac{1}{3}\alpha_{\min},\ \frac{1}{3}\alpha_{\max}\right]\ .$$

COROLLARY. *In the FGM distribution, let F and G be arbitrary continuous distributions with finite nonzero variances, then* $|\rho|\leqq 1/3$.

Applying Theorem 1 and following the discussions of Huang and Kotz [2], we can understand that the bivariate distributions H_{12} and H_{22} also possess the properties (5), (6) and (7), provided that F and G are two *continuous* distributions. In the case for H_{22} we need the following lemma which can be obtained by replacing X by $-X$ in the lemma of Huang and Kotz [2].

LEMMA 2. *For arbitrary nondegenerate distribution F with finite mean,* $\mu_{2,2}-\mu_{1,1}>\mu_{1,2}-\mu_{1,3}$, *or equivalently,* $\mu_{1,1}-\mu_{1,2}>\mu_{1,2}-\mu_{1,3}$.

7. Multivariate distributions

For bivariate distributions H_{1k} and H_{2k} we have proved in Theorem 3 that $\Lambda_{1k}=\Lambda_{2k}$ if $F, G\in\mathcal{D}$. However, for trivariate distributions,

$$W_{1k}=FGN+\sum_{j=1}^{k}\alpha_{1j}(FGN)^{m_j}(\bar{F}\bar{G}\bar{N})^{n_j}\ ,$$

and

$$W_{2k}=FGN+\sum_{j=1}^{k}\alpha_{2j}(FGN)^{n_j}(\bar{F}\bar{G}\bar{N})^{m_j}\,,$$

we have the following different result.

THEOREM 6. *Let F, G and N be three distributions in $\mathcal{D}$, and let Λ_{ik} denote the natural parameter space of W_{ik} ($i=1, 2$). Then*

$$\Lambda_{1k}=-\Lambda_{2k}\equiv\{-\alpha:\ \alpha\in\Lambda_{2k}\}\ .$$

PROOF. Note that we shall take $\alpha_{2j}=-\alpha_{1j}$ in (17) in order to assure the formula (18) being true in trivariate case.

It is easy to extend the results of Theorems 3 and 6 for any multivariate distributions, that is, $\Lambda_{1k}=\Lambda_{2k}$ or $\Lambda_{1k}=-\Lambda_{2k}$ depends only on the number of variables being even or odd, respectively.

Another type of multivariate *FGM* distributions was discussed by Johnson and Kotz [3] and Shaked [7]. The latter provided some applications to the theory of Bayesian survey sampling and to the reliability theory.

Acknowledgment

The author thanks Professor L. K. Chan and the referees for many valuable suggestions.

INSTITUTE OF STATISTICS, ACADEMIA SINICA, TAIWAN

REFERENCES

[1] Cambanis, S. (1977). Some properties and generalizations of multivariate Eyraud-Gumbel-Morgenstern distributions, *J. Multivar. Anal.*, 7, 551-559.

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

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

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

[5] Royden, H. L. (1972). *Real analysis* (2nd ed.), the Macmillan Company, New York.

[6] Schucany, W., Parr, W. C. and Boyer, J. E. (1978). Correlation structure in Farlie-Gumbel-Morgenstern distributions, *Biometrika*, **65**, 3, 650-653.

[7] Shaked, M. (1975). A note on the exchangeable generalized Farlie-Gumbel-Morgenstern distributions, *Commun. Statist.*, **4**, 711-721.