THE INSTITUTE OF STATISTICS

THE CONSOLIDATED UNIVERSITY OF NORTH CAROLINA

Dependence Properties of Iterated Generalized Bivariate Farlie-Gumbel-Morgenstern Distributions

by

S. Kotz
Temple University
Philadelphia, Pennsylvania

and

N. L. Johnson
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina

Institute of Statistics Mimeo Series #1106
March, 1977

DEPARTMENT OF STATISTICS
Chapel Hill, North Carolina

Dependence Properties of Iterated Generalized Bivariate Farlie-Gumbel-Morgenstern Distributions*

by

S. Kotz*

Temple University
Philadelphia, Pennsylvania

and

N. L. Johnson**

University of North Carolina
Chapel Hill, North Carolina

Institute of Statistics Mimeo Series #1106

March, 1977

*This research was supported by the Air Force Office of Scientific Research, Grant No. 75-2837.

**This research was supported by the U.S. Army Research Office under Grant DAA G29-74-C-0030.

1. Introduction

The bivariate Farlie-Gumbel-Morgenstern (FGM) distribution has joint cumulative distribution function (cdf)

$$F_{12}(x_1,x_2) = F_1(x_1)F_2(x_2)\{1 + \alpha S_1(x_1)S_2(x_2)\} \tag{1}$$

(with $|\alpha| \le 1$)

where $F_j(x_j)$ is the cdf of X_j (j=1,2) and $S_j(x_j) = 1 - F_j(x_j)$ is the survival distribution function (sdf) of X_j .

In Johnson and Kotz (1976) we discussed certain generalizations of (1). Motivated by some remarks of Farlie (1960), we suggested replacing the product $S_1(x_1)S_2(x_2)$ (which would be the joint sdf of X_1 and X_2 if they were independent) in (1) by the joint sdf of another bivariate FGM distribution, namely

$$S_{12}^{(1)}(x_1,x_2) = S_1(x_1)S_2(x_2)\{1 + \alpha_1 F_1(x_1)F_2(x_2)\} \tag{2}$$

with $|\alpha_1| \le 1$.

This procedure can be iterated in various ways. In particular, we can replace $F_1(x_1)F_2(x_2)$ in (2) by the FGM type distribution

$$F_{12}^{(2)}(x_1,x_2) = F_1(x_1)F_2(x_2)\{1 + \alpha_2 S_1(x_1)S_2(x_2)\}. \tag{3}$$

(Of course $F_{12}^{(2)}$ might itself be of the form of a first stage iterate FGM distribution, like (2), but here we will restrict ourselves to the relatively simple form (3).)

One purpose of introducing these iterated generalizations of bivariate FGM distributions was to check if one could remedy a defect of bivariate

FGM distributions — namely, the relatively low correlation coefficients that can be attained with such distributions.

For FGM distributions with normal marginals the correlation coefficient is $\alpha\pi^{-1}$, and, since $|\alpha| \leq 1$, this means that the absolute magnitude of the correlation coefficient cannot exceed π^{-1} (= 0.32 approx.)

In the present note we find that iterations of the kind described above cannot increase the maximum possible dependence (as measured by an index suggested by Schweizer and Wolff (1976)), between the variables by more than about 0.12, no matter how many iterations are performed. We therefore investigate a few other modifications of bivariate FGM distributions which might increase dependence more substantially, but we find that necessary restrictions on the values of the α's are such as to seriously limit the possibilities.

2. Schweizer & Wolff's Measure of Dependence

The measure of dependence introduced by Schweizer & Wolff (1976) is essentially (for continuous joint distributions), the expected value of

$$12|F_{12}(X_1,X_2) - F_1(X_1)F_2(X_2)| \qquad (4)$$

(or, equivalently, of

$$12|S_{12}(X_1,X_2) - S_1(X_1)S_2(X_2)|). \qquad (4')$$

Denoting this by the symbol $(\sigma\omega)_{X_1,X_2}$ it is clear that it is unchanged by any monotonic transformation is applied to either (or both) of the variables X_1,X_2. In particular we can suppose that such transformations have been applied to produce variables X_1, X_2 which are each uniformly distributed between 0 and 1. Then

$$(\sigma\omega)_{X_1,X_2} = 12 \int_0^1 \int_0^1 |F_{12}(x_1,x_2) - x_1x_2| dx_1 dx_2 \ . \qquad (5)$$

If X_1 and X_2 are mutually independent, $(\sigma\omega)_{X_1,X_2} = 0$. The maximum possible value of $(\sigma\omega)_{X_1,X_2}$ is 1, and this is attained when, with probability 1, X_1 and X_2 are monotonic functions of each other.

(When X_1 and X_2 are each uniformly distributed over (0,1), if X_1 and X_2 are each monotonic functions of the other then we must have either $X_1 = X_2$ or $X_1 = 1 - X_2$.)

We will use this measure of dependence in our studies, rather than the better-known mean square contingency. This is quite closely related to the Schweizer-Wolff index,(see Section 5)— but the comparisons using it are not so clearcut as using the latter measure.

3. Dependence in Generalized FGM Distributions

If the marginal distributions are each uniform over (0,1) then the FGM cdf (1) is

$$x_1x_2\{1 + \alpha(1-x_1)(1-x_2)\}$$

and

$$(\sigma\omega)_{X_1X_2} = 12|\alpha| \int_0^1 \int_0^1 x_1(1-x_1)x_2(1-x_2)dx_1dx_2$$

$$= 12|\alpha|(\tfrac{1}{2} - \tfrac{1}{3})^2 = \tfrac{1}{3}|\alpha| \ . \qquad (6)$$

This is of course the value of $(\sigma\omega)_{X_1X_2}$ for <u>any</u> continuous FGM distribution with cdf (1).

For the once iterated FGM distribution with uniform (0,1) marginal distributions we have

$$F_{12}(x_1,x_2) = x_1x_2[1 + \alpha(1-x_1)(1-x_2)\{1 + \alpha_1x_1x_2\}] \tag{7}$$

whence

$$(\sigma\omega)_{X_1X_2} = 12\int_0^1\int_0^1 |\alpha\pi_1(1-x_1)x_2(1-x_2) + \alpha\alpha_1x_1^2(1-x_1)x_2^2(1-x_1)|dx_1dx_2 .$$

For given numerical values of α, α_1 we will get the largest value for $(\sigma\omega)_{X_1X_2}$ by having α, α_1 of the same sign. Taking $\alpha>0$ and $\alpha_1 > 0$ we obtain

$$(\sigma\omega)_{X_1X_2} = 12\{\alpha(\tfrac{1}{2} - \tfrac{1}{3})^2 + \alpha\alpha_1(\tfrac{1}{3} - \tfrac{1}{4})^2\} = \tfrac{1}{3}\alpha + \tfrac{1}{12}\alpha\alpha_1 = \tfrac{1}{3}(1 + \tfrac{1}{4}\alpha_1)\alpha. \tag{8}$$

Similarly, we find for the twice iterated FGM distribution, with $\alpha \geq 0$, $\alpha_1 \geq 0$, $\alpha_2 \geq 0$

$$(\sigma\omega)_{X_1X_2} = \tfrac{1}{3}\alpha + \tfrac{1}{12}\alpha\alpha_1 + \tfrac{1}{75}\alpha\alpha_1\alpha_2 . \tag{9}$$

Generally for the r-th iterated FGM distribution with no negative α's

$$(\sigma\omega)_{X_1X_2} = 12\sum_{j=1}^{r} a_j\alpha'_j$$

where $\alpha'_1 = \alpha$

$$\alpha'_j = \alpha \prod_{i=1}^{j-1} \alpha_i \qquad (j\geq2)$$

and

$$a_j = \begin{cases}\{B(\frac{j+3}{2}, \frac{j+3}{2})\}^2 & \text{for } j \text{ odd}\\ \{B(\frac{j+4}{2}, \frac{j+2}{2})\}^2 & \text{for } j \text{ even}\end{cases} . \tag{11}$$

where B(i,j) is the Beta function.

Since $0 \le \alpha'_r \le \alpha'_{r-1} \le \ldots \le \alpha'_1 = \alpha$, it follows that

$$(\sigma\omega)_{X_1X_2} \le 12\alpha \sum_{j=1}^{r} a_j \le 12\alpha \sum_{j=1}^{\infty} a_j$$

and so

$$(\sigma\omega)_{X_1X_2} \le 12\alpha \sum_{h=2}^{\infty} [\{B(h,h)\}^2 + \{B(h+1,\ h)\}^2]$$

$$= 15\alpha \sum_{h=2}^{\infty} \{B(h,h)\}^2 \qquad (12)$$

(since $B(h+1,\ h) = \tfrac{1}{2}B(h,h)$).

Now $$\sum_{h=2}^{\infty} \{B(h,h)\}^2 = 6^{-2} + 30^{-2} + 140^{-2} + 630^{-2} + \ldots$$

$$< 6^{-2} + 30^{-2} + 140^{-2} + 630^{-2}(1 + 4^{-2} + 4^{-4} + \ldots)$$

$$< 0.029 \qquad (13)$$

and so

$$(\sigma\omega)_{X_1X_2} < 15 \times 0.029\alpha = 0.435\alpha \qquad (14)$$

however many iterations are used.

This shows that iterations of the kind described in Section 1 cannot increase the measure of dependence $(\sigma\omega)_{X_1,X_2}$ above 0.435.

The reason for the low limit on the measure of dependence is in the rapid decrease of $B(h,h)$ as h increases. A natural way to avoid this would be to use distributions of form

$$F_{12}(x_1,x_2) = F_1(x_1)F_2(x_2)[1 + \alpha\{S_1(x_1)\}^{\phi_1}\{S_2(x_2)\}^{\phi_2}]$$

with $0 < \phi_j < 1$ $(j = 1,2)$. Unfortunately this is not a proper distribution function for any absolutely continuous marginals $F_1(x_1)$, $F_2(x_2)$. We have

$$\frac{\partial^2 F_{12}}{\partial x_1 \partial x_2} = f_1(x_1)f_2(x_2)[1 + \prod_{j=1}^{2} \{S_j(x_j)\}^{-(1-\phi_j)}\{1 - (1+\phi_j)F_j(x_j)\}].$$

By choosing x_1 so that $(1+\phi_1)F(x_1) > 1$ and x_2 so that $S_2(x_2)$ is sufficiently small we make the expression in square brackets negative. (Since $F_1(x_1) < 1$ and $S_2(x_2) > 0$, we can always find x_1, x_2 so that $f_1(x_1) > 0$ and $f_2(x_2) > 0$.)

Some other modifications of bivariate FGM distributions are discussed in the next section, with special reference to possible increases in dependence.

4. Some Other Generalized FGM Distributions

4.1 The following example shows directly how the restrictions on the values of the α's militates against realization of substantial increases in $(\sigma\omega)_{X_1,X_2}$.

We consider a family of distributions obtained by modifying (1) with replacement of

$$F_1(x_1)F_2(x_2) \quad \text{by} \quad F_1(x_1)F_2(x_2)\{1+\alpha_1 S_1(x_1)S_2(x_2)\}$$

and

$$S_1(x_1)S_2(x_2) \quad \text{by} \quad S_1(x_1)S_2(x_2)\{1+\alpha_2 F_1(x_1)F_2(x_2)\}.$$

Transforming so that each marginal distribution is uniform over $(0,1)$ we have

$$F_{12}(x_1,x_2) = x_1x_2\{1+\alpha_1(1-x_1)(1-x_2)\}\{1+\alpha(1-x_1)(1-x_2)(1+\alpha_2x_1x_2)\}$$

$$(0 < x_j < 1,\ j=1,2). \quad (15)$$

Hence

$$(\sigma\omega)_{X_1,X_2} = 12\int_0^1\int_0^1 |(\alpha+\alpha_1)x_1x_2(1-x_1)(1-x_2) + \alpha\alpha_1x_1x_2(1-x_1)^2(1-x_2)^2$$

$$+ \alpha\alpha_2x_1^2x_2^2(1-x_1)(1-x_2) + \alpha\alpha_1\alpha_2x_1^2x_2^2(1-x_1)^2(1-x_2)^2|dx_1dx_2 .$$

If all α's are nonnegative then

$$(\sigma\omega)_{X_1,X_2} = \frac{1}{3}(\alpha + \alpha_1) + \frac{1}{12}\alpha(\alpha_1 + \alpha_2) + \frac{1}{75}\alpha\alpha_1\alpha_2 . \tag{16}$$

If we only needed the conditions $|\alpha| \le 1$, $|\alpha_1| \le 1$, $|\alpha_2| \le 1$ then it would appear that we could get values of $(\sigma\omega)_{X_1,X_2}$ as near to $\frac{2}{3} + \frac{1}{6} + \frac{1}{75} = 0.846$ as desired.

However, the α's must be such that

$$\frac{\partial^2 F}{\partial x_1 \partial x_2} = 1 + (\alpha+\alpha_1)(1-2x_1)(1-2x_2) + \alpha\alpha_1(1-x_1)(1-3x_1)(1-x_2)(1-3x_2)$$
$$+ \alpha\alpha_2 x_1(2-3x_1)(2-3x_2) + 4\alpha\alpha_1\alpha_2 x_1(1-x_1)(1-2x_1)x_2(1-x_2)(1-2x_2) \ge 0 \tag{17}$$

for x_1, x_2 with $0 \le x_1 \le 1$, $0 \le x_2 \le 1$.

Putting $x_1 = 0$, $x_2 = 1$ gives

$$\alpha + \alpha_1 \le 1.$$

This implies that

$$(\sigma\omega)_{X_1X_2} \le \frac{1}{3} + \frac{1}{12}\alpha(1 + \alpha_2 - \alpha) + \frac{1}{75}\alpha(1-\alpha)\alpha_2$$
$$\le \frac{1}{3} + \frac{1}{12}\alpha(2-\alpha) + \frac{1}{75}\alpha(1-\alpha)$$
$$\le 0.417 \quad \text{for } 0 \le \alpha \le 1. \tag{18}$$

4.2 The modified FGM family defined by

$$F_{12}(x_1,x_2) = F_1(x_1)F_2(x_2)[1 + \alpha \min(S_1(x_1), S_2(x_2))] \tag{19}$$

or equivalently

$$F_{12}(x_1,x_2) = F_1(x_1)F_2(x_2)[1 + \alpha\{1 - \max(F_1(x_1),F_2(x_2))\}] \qquad (19)'$$

has a somewhat larger value of $(\sigma\omega)_{X_1X_2}$. We have

$$\begin{aligned}(\sigma\omega)_{X_1X_2} &= 12\,\alpha \int_0^1 \int_0^1 x_1x_2\{1 - \max(x_1,x_2)\}dx_1dx_2 \\ &= 12\,\alpha \times 2 \int_0^1 \int_0^{x_2} x_1x_2(1-x_2)dx_1dx_2 \\ &= 12\,\alpha \int_0^1 x_2^3(1-x_2)dx_2 \\ &= 0.6\alpha. \end{aligned} \qquad (20)$$

This is considerably larger (if $\alpha>0$) than the values so far obtained, but it still cannot exceed 0.6.

5. Comparison with Mean Contingency

The mean contingency (as defined by Rényi (1970, p. 282)) for an absolutely continuous distribution with standard uniform marginal distribution is

$$\phi^2_{X_1,X_2} = \int_0^1 \int_0^1 [f_{12}(x_1,x_2)]^2 dx_1dx_2 - 1 \qquad (21)$$

where $f_{12} = \dfrac{\partial^2 F_{12}}{\partial x_1 \partial x_2}$ is the joint probability density function of X_1 and X_2 .

Table 1 summarizes values of $(\sigma\omega)_{X_1X_2}$ and the corresponding values of $\phi^2_{X_1X_2}$, for the distributions discussed in this note.

TABLE 1 Comparison of $(\sigma\omega)_{X_1X_2}$ and $\phi^2_{X_1X_2}$

from eq.	Value of $(\sigma\omega)_{X_1X_2}$ value	Corresponding value of $\phi^2_{X_1X_2}$
(6)	$\frac{1}{3}\lvert\alpha\rvert$	$\frac{1}{9}\alpha^2$
(8)	$\frac{1}{3}(1+\frac{1}{4}\alpha_1)\alpha$	$\frac{1}{9}(1+\frac{1}{2}\alpha_1+\frac{4}{25}\alpha_1^2)\alpha^2$
(9)	$\frac{1}{3}\alpha+\frac{1}{12}\alpha\alpha_1+\frac{1}{75}\alpha\alpha_1\alpha_2$	$\frac{1}{9}(1+\frac{1}{2}\alpha_1+\frac{4}{25}\alpha_1^2+\frac{2}{25}\alpha_1\alpha_2+\frac{2}{25}\alpha_1^2\alpha_2+\frac{4}{1225}\alpha_1^2\alpha_2^2)\alpha^2$
(16)	$\frac{1}{3}(\alpha+\alpha_1)+\frac{1}{12}\alpha(\alpha_1+\alpha_2)+\frac{1}{75}\alpha\alpha_1\alpha_2$	$\frac{1}{9}\{(\alpha+\alpha_1)^2+\frac{1}{2}\alpha(\alpha+\alpha_1)(\alpha_1+\alpha_2)+\frac{4}{25}\alpha^2(\alpha_1^2+\alpha_2^2)$ $+\frac{2}{25}(\alpha+\alpha_1)\alpha\alpha_1\alpha_2+\frac{1}{50}\alpha^2\alpha_1\alpha_2(\alpha_1+\alpha_2+1)$ $+\frac{4}{1225}\alpha^2\alpha_1^2\alpha_2^2\}$

We note that for (6)

$$(\sigma\omega)^2_{X_1X_2} = \phi^2_{X_1X_2}$$

and for (8)

$$(\sigma\omega)^2_{X_1X_2} = \phi^2_{X_1X_2} - \frac{13}{1200}\alpha^2\alpha_1^2 ,$$

and for (9), approximately

$$(\sigma\omega)^2_{X_1X_2} \doteq \phi^2_{X_1X_2} - \frac{13}{1200}\left[(1+\frac{4\alpha_2}{13})^2 + (\frac{16}{35}\alpha_2)^2\right]\alpha^2\alpha_1^2 .$$

REFERENCES

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

JOHNSON, N.L. and KOTZ, S.(1976). On some generalized Farlie-Gumbel-Morgenstern distributions - II. Regression, correlation and further generalizations. Mimeo Series No. 1080, Institute of Statistics, University of North Carolina. (A slightly different version will appear in *Communications in Statistics*.)

RÉNYI, A. (1970). *Probability Theory*. Amsterdam: North Holland Publishing Co.

SCHWEIZER, B. & WOLFF, E. F. (1976). Sur une mésure de dépendance pour les variables aléatoires, *C.R. Acad. Sci., Paris, Sér. A*, 283, 659-661.