Qüestiió - V. 10, n.º 3 (setembre 1986) pp. 133-141

CONSTRUCCIÓN DE DISTRIBUCIONES CON MARGINALES MULTIVARIANTES DADAS

PEDRO SÁNCHEZ ALGARRA
UNIVERSIDAD DE BARCELONA

Este trabajo discute ciertos aspectos acerca de la construcción de distribuciones multivariantes con las marginales multivariantes fijas. Se prueba un resultado general que permite la construcción de pseudodistribuciones, que sin embargo no es adecuado para construir distribuciones mediante familias paramétricas conocidas. Entonces se propone una nueva familia, dependiendo sobre una función real contínua y un parámetro. Algunas propiedades de este sistema son estudiadas y relacionadas con otras familias.

Keywords: MULTIVARIATE DISTRIBUTIONS; MULTIVARIATE FIXED MARGINALS; STOCHASTIC DEPENDENCE; FARLIE-GUMBEL-MORGENSTERN DISTRIBUTIONS.

1. INTRODUCCION.

Sean $F(x)$, $G(y)$ dos funciones de distribución de probabilidad. La clase de distribuciones bivariantes $\mathcal{F}$ cuyas marginales son $F(x)$, $G(y)$, llamada clase de Fréchet, ha sido bien estudiada por diferentes autores -- (Fréchet, Hoeffding, Dall'Aglio, Kimeldorf, Sampson, Ruiz-Rivas, etc.). Véase /12/ para una amplia exposición de este tema.

En este trabajo se estudia la siguiente generalización: sean $F(x_1,\dots,x_m)$, $G(y_1,\dots,y_n)$ dos distribuciones multivariantes m y n dimensionales respectivamente; nos interesa encontrar las distribuciones $H(x_1,\dots,x_m,y_1,\dots,y_n)$ cuyas marginales son F y G, es decir,

$$\begin{aligned} H(x_1,\dots,x_m,\infty,\dots,\infty) &= F(x_1,\dots,x_m) \\ H(\infty,\dots,\infty,y_1,\dots,y_n) &= G(y_1,\dots,y_n) \end{aligned} \quad (1)$$

Sin embargo, este problema no es una mera generalización del caso $m = n = 1$. En efecto, resulta que para $m + n > 2$, la generalización de las cotas de Fréchet, la generalización del sistema FGM (Farlie-Gumbel-Morgenstern), etc., da lugar a pseudodistribuciones de probabilidad, es decir, a funciones $S:R^k \to R$ tales que:

$$\text{a)} \lim_{x_i \to -\infty} S(x_1,\dots,x_k) = 0 \qquad i = 1,\dots,n$$

$$\text{b)} \ S \text{ es creciente en cada variable.} \qquad (2)$$

$$\text{c)} \ F_i(x_i) = \lim_{\substack{x_j \to \infty \\ \forall j \neq i}} S(x_1,\dots,x_k)$$

es una función de distribución de probabilidad.

De hecho, en los casos citados, la sustitución de las funciones univariantes por funciones multivariantes, nos permite obtener pseudodistribuciones que cumplen la condición más fuerte.

$$\text{c')} \lim_{\forall y_i \to \infty} S(x_1,\dots,x_m,y_1,\dots,y_n) = F(x_1,\dots,x_m)$$

$$\lim_{\forall x_i \to \infty} S(x_1,\dots,x_m,y_1,\dots,y_n) = G(y_1,\dots,y_n)$$

Esta condición es imprescindible para el propósito de este trabajo.

Toda función de distribución de probabilidad $F(x_1,\dots,x_k)$ debe cumplir /8/ que dado el intervalo de R^k, $(a_1,a_1+h_1]\times\dots\times(a_k+h_k]$, la siguiente cantidad

- Pedro Sánchez Algarra - Universitat de Barcelona - Facultat de Biologia - Departament de Bioestadística - Av. Diagonal, 637 - 08028 Barcelona.

- Article rebut el febrer de 1986.

$$\Delta F = F(a_1+h_1,\ldots,a_k+h_k) - F(a_1,a_2+h_2,\ldots,a_k+h_k)-\ldots - F(a_1+h_1,\ldots,a_{k-1}+h_{k-1},a_k)+\ldots+ +(-1)^k . F(a_1,\ldots,a_k)$$

verifique

$$\Delta F \geq 0$$

En el caso absolutamente contínuo esta condición es equivalente a que:

$$\frac{\partial^n F}{\partial x_1 \partial x_2 . \partial x_n} \geq 0$$

Una pseudodistribución S verifica las condiciones (2), pero no necesariamente la (3) -- (correspondería, por tanto, a una medida con signo). Se pueden construir distribuciones y pseudodistribuciones a partir de un teorema que enunciamos seguidamente.

Teorema 1: Sea $U(u_1,\ldots,u_k)$ una función de distribución contínua con marginales uniformes en $[0,1]$, y sean $F_1,\ldots,F_k$ funciones de distribución de probabilidad contínuas. Sea H la función construida mediante el cambio $u_i = F_i(x_i)$ en U, es decir,

$$H = Uo(F_1,\ldots,F_k) \tag{4}$$

Se verifica:

1) Si cada F_i es distribución unidimensional, entonces H es una distribución k-dimensional.

2) Si cada F_i es distribución n_i-dimensional, entonces H es una pseudodistribución N-dimensional, siendo $N=n_1+\ldots+n_k$.

Demostración:

Como

$$\lim_{u_i \to 0} U(k_1,\ldots,k_k) = 0$$

$$\lim_{\substack{u_j \to 1 \\ j \neq i}} U(u_1,\ldots,u_k) = u_i$$

$$U(u_1,\ldots,u_i+h,\ldots,u_k)-U(k_1,\ldots,u_i,\ldots u_k) \geq 0$$

deducimos fácilmente que H verifica las condiciones (2). Por otra parte, si hacemos

$$F_i(a_i) = u_i \qquad F_i(a_i+h_i) = u_i+d_i$$

Como U verifica la condición (3), H también la verifica. Luego queda probada la primera parte.

Supongamos ahora que cada F_i es n_i-dimensional. Entonces, como

$$\lim_{x_{p_i} \to -\infty} F_p(x_{p_1},\ldots,x_{p_{n_p}}) = 0$$

tenemos que

$$\lim_{x_{p_i} \to -\infty} H = U(F_1,\ldots,0,\ldots,F_k) = 0$$

Por otra parte

$$\lim_{\substack{x_{p_j} \to \infty \\ \forall j \neq i}} F_p(x_{p_1},\ldots,x_{p_{n_p}}) = F_{p_i}(x_{p_i})$$

$$\lim_{\substack{x_{q_j} \to \infty \\ \forall j}} F_p(c_{q_1},\ldots,x_{q_{n_q}}) = 1$$

Tenemos también que

$$\lim_{\substack{x_j \to \infty \\ \forall j \neq p_i}} H=U(1,\ldots,F_{p_i}(x_{p_i}),\ldots,1) = F_{p_i}(x_{p_i})$$

Se comprueba también sin dificultad que H es creciente en cada variable. Sin embargo, la condición (3) en general no se cumple, salvo que U sea lineal. Veamos un contraejemplo.

Consideremos la distribución con marginales uniformes

$$U(u_1,u_2)=u_1u_2(1+1/2(1-u_1)(1-u_2))$$

y las funciones de distribución

$$F(x,y) = xy \qquad 0 \leq x,\ y \leq 1$$

$$G(z) = z \qquad 0 \leq z \leq 1$$

Entonces:

$$H(x,y,z)=xyz(1+1/2(1-xy)(1-z))$$

$$\frac{\partial^3 H}{\partial x \partial y \partial z} = 1+1/2(1-xy-2z+3xyz$$

que es negativa para ciertos valores x,y,z (por ejemplo, para $x = y = 0.98$, $z = 0.1$). Luego H no es una distribución.

CLASE DE FRECHET DE PSEUDODISTRIBUCIONES.

Sean $F(x_1,\ldots,x_m)$, $G(y_1,\ldots,y_n)$ dos distribuciones. Llamaremos clase de Fréchet $\mathcal{S}$ al conjunto de pseudodistribuciones $S(x_1,\ldots,x_m,y_1,\ldots,y_n)$ cuyas marginales sean F, G, es decir, que verifiquen (1).

Dados dos sucesos A, B, de probabilidades $P(A)$, $P(B)$, la bien conocida desigualdad

$$\max\{P(A)+P(B)-1,0\}\le P(A\ B) \le \min\{P(A),P(B)\}$$

nos lleva, como en el caso bivariante, a -- considerar las siguientes funciones

$$S^- = \max\{F+G-1,\ 0\}$$
$$S^+ = \min\{F,G\}$$

llamadas cotas de Fréchet. Para toda $S\in\mathcal{S}$ se verifica

$$S^- \le S \le S^+ \qquad (6)$$

Sin embargo, salvo que $m = n = 1$, S^- y S^+ no son distribuciones sino tan solo pseudodistribuciones. En efecto, supongamos que para ciertos números reales $x_1,x_2,y,a>0,b>0,c>0$ se verifica

$$G(y)\le G(y+c)\le F(x_1+a,x_2)$$
$$G(y)\le F(x_1,x_2)<F(x_1,x_2+b)\le G(y+c)\le F(x_1+a,x_2+b)$$

Tales desigualdades son compatibles entre sí. Tomando entonces la cota S^+ tenemos

$$\Delta S^+=G(y+c)-G(y)-G(y+c)-F(x_1,x_2+b)+G(y)+$$
$$+G(y)+F(x_1,x_2)-G(y)=F(x_1,x_2)-$$
$$-F(x_1,x_2+b)<0$$

Análogamente se comprueba que S^- no es en general distribución de probabilidad.

Otro elemento destacado de $\mathcal{S}$ es

$$H^O = F.G$$

que corresponde al caso de independencia estocástica entre $(X_1,\ldots,X_m)$ y $(Y_1,\ldots,Y_n)$. Evidentemente en este caso H^O da lugar a -- una verdadera distribución de probabilidad.

Respecto a las cotas que son distribuciones, podemos seguir dos caminos para obtenerlas.

En primer lugar, como a partir de $F(x_1,\ldots,x_m)$ obtenemos, por paso al límite, sus correspondientes distribuciones marginales $F_1,\ldots,F_m$ y análogamente obtenemos $G_1,\ldots,G_n$, entonces podemos construir la cota superior

$$H^+ =\{\min F_1(x_1),\ldots,F_m(x_m),\ G_1(y_1),\ldots,G_n(y_n)$$

y los $(2^{m+n-1}-1)$ elementos minimales H^- (véase /12/), verificándose entonces que para toda distribución H cuyas marginales sean $F_1,\ldots,F_m$, $G_1,\ldots,G_n$, (en particular aque-- llas cuyas marginales m y n variantes sean F, G), verifique (casi seguramente)

$$\begin{matrix} H \le H^+ & \text{(cota máxima)} \\ H \le H^- \Rightarrow H = H^- & \text{(elemento minimal)} \end{matrix} \qquad (7)$$

Sin embargo, no todas las distribuciones verificando (7) tienen marginales F, G. Por ejemplo, basta considerar H^+ y cualquier H^-.

En segundo lugar podemos caracterizar los elementos máximo y mínimo de la clase bivariante de Fréchet con marginales dadas continuas, como la distribución H^+ definida sobre (X, Y) si y sólo si $F(X) = G(Y)$ (c.s.) y la distribución H^- sí y sólo si $F(X)=1-G(Y)$ (c.s.). La generalización de esta idea nos lleva al estudio de las distribuciones singulares H^+ y H^- (si existen) correspondientes a las relaciones funcionales

$$F(X_1,\ldots,X_m) = G(Y_1,\ldots,Y_n) \quad \text{(c.s.)}$$
$$F(X_1,\ldots,X_m) = 1-G(Y_1,\ldots,Y_n) \quad \text{(c.s.)}$$

respectivamente.

Sin embargo, en el caso no bivariante, estas relaciones podrían no ser posibles. Por ejemplo:

Sean X_1, X_2 v.a. independientes y uniformes en $(0,1)$, Y también uniforme en $(0,1)$. Entonces es imposible la relación funcional

$$F(X_1,X_2) = X_1.X_2 = Y = G(Y)$$

pues la función de densidad de la variable $Z = X_1.X_2$ es $f(z) = -\ln(z)$, $0\le z\le 1$, luego la distribución de Z no es uniforme y no puede coincidir con la de Y.

3. ALGUNAS FAMILIAS PARAMETRICAS DE PSEUDO-DISTRIBUCIONES.

Dado que la clase de Fréchet es muy amplia, como en el caso bivariante, nos interesará hallar familias paramétricas $\mathcal{S}_\theta \subset \mathcal{S}$, dependientes de un cierto parámetro θ y que sean funcionales en F, G. Un camino natural para construir tales familias consiste en adaptar a nuestro caso el sistema FGM, el sistema Cuadras-Augé, la familia trasladada de la normal, etc. Sin embargo, veremos que sólo llegamos a obtener pseudodistribuciones. En la sección siguiente proponemos una nueva familia, que es capaz de proporcionar verdaderas distribuciones de probabilidad con marginales multivariantes.

3.1. SISTEMA FGM

La representación uniforme del sistema FGM (Farlie-Gumbel-Morgenstern) es /12/

$$U(u,v) = uv[1+\alpha(1-u).(1-v)] \qquad -1\leq\alpha\leq 1 \tag{8}$$

Utilizando (4) obtenemos la pseudodistribución

$$H = FG[1+\alpha(1-F)(1-G)] \tag{9}$$

Para comprobar que H no es en general una distribución, tomemos $m = 2$, $n = 1$, supongamos F, G absolutamente contínuas y sean f, g las funciones de densidad, F_{x_1}, F_{x_2} las derivadas parciales respecto x_1, x_2

Entonces

$$\frac{\partial^3 H}{\partial x_1 \partial x_2 \partial y} = fg[1+\alpha(1-2F)(1-2G)]-2\alpha.F_{x_1}F_{x_2} - 2\alpha.F_{x_1}F_{x_2}g(1-2G) \tag{10}$$

La primera parte de (10) es análoga al caso $m = n = 1$. Sin embargo, ahora aparece un término negativo que impide que se cumpla (3).

3.2. SISTEMA DE CUADRAS-AUGÉ

La traslación uniforme de este sistema es /6/.

$$U(u,v) = \begin{cases} u^{1-\theta}v & \text{si } u\geq v \qquad u,v\in[0,1] \\ u.v^{1-\theta} & \text{si } u<v \qquad 0\leq\theta\leq 1 \end{cases} \tag{11}$$

Utilizando (4) obtenemos la pseudodistribución

$$H = \begin{cases} F^{1-\theta}G & \text{si } F(x_1,\ldots,x_m)\geq G(y_1,\ldots,y_n) \\ F\,G^{1-\theta} & \text{si } F(x_1,\ldots,x_m)< G(y_1,\ldots,y_n) \end{cases} \tag{12}$$

Sea $m = 2$, $n = 1$, y tomemos $\theta = 1/2$. Entonces

$$\frac{\partial^2 H}{\partial x_1 \partial x_2} = \begin{cases} \dfrac{2f\,F - Fx_1Fx_2}{4F^{3/2}}.G & \text{si } F\geq G \\ f.G^{1/2} & \text{si } F< G \end{cases}$$

Estudiemos el salto de esta función en la región tal que $F = G$.

$$\lim_{\substack{F\to G\\ F<G}} fG^{1/2} - \lim_{\substack{F\to G\\ F>G}} \frac{2fF-Fx_1Fx_2}{4F^{3/2}}.G = F^{1/2} - \frac{2fF-Fx_1Fx_2}{F^{1/2}} = A$$

La derivada de H en el sentido de Schwartz /13/ es

$$\frac{\partial^3 H}{\partial x_1 \partial x_2 \partial y} = \begin{Bmatrix} \dfrac{2fF-Fx_1Fx_2}{4F^{3/2}}.g & \text{si } F>G \\ 1/2fgG^{-1/2} & \text{si } F<G \end{Bmatrix} + A\delta_{\{F=G\}} \tag{13}$$

donde $\delta_{\{F=G\}}$ es la distribución de Dirac en la región de R^{m+n} tal que F=G. Como (13) puede ser negativa, concluimos que (12) no es, en general, una distribución.

3.3. OTROS SISTEMAS DE PSEUDODISTRIBUCIONES.

Por idéntico camino al seguido en los apartados anteriores, podemos construir otras pseudodistribuciones. Por ejemplo, el método de traslación de Mardia proporcionaría la familia

$$H\rho = N(B^{-1}F,\ B^{-1}G) \tag{14}$$

donde N es la función de distribución de la normal bivariante $N(0,0;\ 1,1;\ \rho)$, ρ es el coeficiente de correlación, B es la función de distribución N(0,1) /10/.

La familia bivariante de Ali, Mikhail y Hag /1/ nos permite definir el sistema

$$H_\theta = \frac{FG}{1-\theta(1-F)(1-G)} \qquad (15)$$

Finalmente, la familia propuesta por Ruiz-Rivas /10/ da lugar a

$$\begin{aligned} H_\theta &= \max\{F^{-\theta}+G^{-\theta}-1,\ 0\}^{-1/\theta} && \text{si } -1\le\theta\le 0 \\ &= F.G. && \text{si } \theta = 0 \qquad (16) \\ &= F-(\max\{F^\theta+(1-G)^\theta-1,0\})^{1/\theta} && \text{si } 0\le\theta\le 1 \end{aligned}$$

donde, en todos los casos, F y G son distribuciones en R^m y R^n. (para construir otras familias, consúltese /4/, /10/).

Sin embargo, (14), (15), (16) no son distribuciones, y por lo tanto no sirven para construir distribuciones con marginales multivariantes, aunque, de acuerdo con el Teorema 1 dan lugar a pseudodistribuciones.

4. UNA FAMILIA DE DISTRIBUCIONES CON MARGINALES MULTIVARIANTES.

En esta sección construimos una familia bivariante que depende de un parámetro α y una cierta función Q . Podremos construir distribuciones con marginales multivariantes utilizando el Teorema 1, salvo que con una adecuada elección de Q conseguiremos evitar los inconvenientes expuestos en la sección anterior.

Sea Q(x) una función contínua sobre el intervalo [-1/2, 1/2] , positiva sobre (-1/2, 1/2), tal que:

a) $Q(-1/2) = Q(1/2) = 0$ (17)

b) Q(x) es $K = \max\{m,n\}$ veces derivable; su primera derivada, $q(x) = Q'(x)$, es una función impar y monótona en el intervalo [-1/2, 1/2] (18)

Ejemplos interesantes son:

$$1) \quad Q_1(x) = \begin{cases} 2x+1 & -1/2\le x<0 \\ -2x+1 & 0\le x<1/2 \end{cases} \qquad (19)$$

que en realidad no es derivable en x = 0, pero esto no será ningún problema para lo que haremos después.

$$2) \quad Q_2(x) = 1-4x^2 \qquad -1/2\le x\le 1/2 \qquad (20)$$

$$3) \quad Q_3(x) = \cos(x\pi) \qquad -1/2\le x\le 1/2 \qquad (21)$$

4.1. CASO BIVARIANTE.

Definamos la siguiente familia bivariante con marginales uniformes

$$U(u_1,u_2) = u_1u_2+\alpha^2\, Q(u_1-1/2)\, Q(u_2-1/2) \qquad (22)$$

La función de densidad es

$$u(u_1,u_2) = 1+\alpha^2 q(u_1-1/2)\, q(u_2-1/2) \qquad (23)$$

La condición (18) implica que el mínimo de $q(x).q(y)$ es $q(-1/2)\ q(1/2) = -q^2(-1/2)$. Luego, a fin de que (23) sea positiva es necesario que α verifique

$$-q(-1/2)^{-1}\le\alpha\le q(-1/2)^{-1} \qquad (24)$$

Sean ahora F(x), G(y) dos distribuciones univariantes. (22) nos proporciona entonces la familia bivariante

$$H_\alpha(x,y) = F(x)G(y)+\alpha^2 Q(F(x)-1/2)Q(G(y)-1/2) \qquad (25)$$

donde α verifica (24). La correspondiente función de densidad, en el caso de que F, G sean absolutamente continuas, es

$$h(x,y) = f(x)g(y)[1+\alpha^2 q(F(x)-1/2)q(G(y)-1/2)]$$

Es fácil comprobar que (25) es una familia paramétrica de distribuciones cuyas marginales son F, G.

4.2. CASO MULTIVARIANTE.

Sean $F(x_1,\dots,x_m)$, $G(y_1,\dots,y_n)$ dos distribuciones multivariantes absolutamente contínuas. Aplicando (4) pero tomando (22) obtenemos la familia uniparamétrica

$$\begin{aligned} H(x_1,\dots,x_m, y_1,\dots,y_n) &= F(x_1,\dots,x_m)\ G(y_1,\dots,y_n) \\ &+\alpha^2 Q[F(x_1,\dots,x_m)-1/2]\ Q\,[G(y_1,\dots,y_n)-1/2] \qquad (26) \end{aligned}$$

De entrada, (26) es una pseudodistribución que, para ciertas formas de Q dará lugar a una --

distribución con marginales F, G. Observamos, previamente, que si n=3, y F es derivable, las derivadas de Q(F) son

$$\frac{\partial Q(F)}{\partial x_1} = Q'(F).Fx_1 \qquad \frac{\partial^2 Q(F)}{\partial x_1 \partial x_2} = Q''(F)Fx_1Fx_2+Q'(F)Fx_1x_2$$

$$\frac{\partial^3 Q(F)}{\partial x_1 \partial x_2 \partial x_3} = Q'''(F)Fx_1Fx_2Fx_3+Q''(F)Fx_1Fx_2x_3 +$$

$$+Q''(F)Fx_2Fx_1x_3+Q''(F)Fx_1x_2Fx_3+Q'(F)Fx_1x_2x_3$$

Hemos visto que los sistemas anteriores presentaban dificultades, precisamente porque las derivadas compuestas de orden superior generan términos negativos en la función de densidad. Estudiemos pues la familia (26).

En primer lugar tomemos $Q = Q_1$ (véase (19)). Entonces $Q' = q_1$, $Q'' = Q''' = \ldots = 0$, con lo cual obtenemos la familia de densidades

$$\begin{aligned} h(x,y) &= f(x)g(y)+4\alpha^2 f(x)g(y) \text{ si } F(x),\ G(y)>1/2 \\ &\qquad\qquad \text{ó } F(x),\ G(y)<1/2 \\ &= f(x)g(y)-4\alpha^2 f(x)g(y) \text{ en caso contrario} \end{aligned}$$

siendo $x = (x_1,\ldots,x_m)$, $y = (y_1,\ldots,y_n)$. Según (24), α debe cumplir

$$-1/2 \leq \alpha \leq +1/2$$

Tomemos ahora $Q = Q_2$ (ver(20)). Obtenemos la familia de distribuciones

$$H(x,y)=F(x)G(y)+16\alpha^2 F(x)[1-F(x)]G(y)[1-G(y)] \tag{28}$$

$$x \in R^m \quad y \in R^n \quad -a\leq\alpha\leq a$$

donde (-a, a) es un intervalo que en general dependerá de las distribuciones marginales F, G. Obsérvese que (28) es análogo al sistema FGM (ver (9)) salvo los límites de variación del parámetro α. Por otra parte, $Q_2' = q_2$, $Q_2'' = -8$, $Q_2''' = 0$, luego

$$\frac{\partial^m Q_2(F-1/2)}{\partial x_1 \partial x_2 \ldots \partial x_m} = -8[Fx_1Fx_2\ldots x_m+Fx_1x_2Fx_3\ldots x_m+\ldots$$

$$Fx_1\ldots x_{m-1}Fx_m \ldots + Fx_1\ldots x_{m-1}Fx_m+(F-1/2)f]$$

Derivando (28) obtenemos la función de densidad

$$h = f.g+64\alpha^2[K(F)+(F-1/2)f][L(G)+(G-1/2)g] \tag{29}$$

$$x \in R^m \quad y \in R^n \quad -a\leq\alpha\leq+a$$

donde K(F), L(G) son funcionales positivos en F, G, puesto que las derivadas parciales $F_{x_{i_1}\ldots x_{i_k}}$ son no negativas y además no decrecientes en cada variable.

A continuación vamos a construir un procedimiento para calcular a. Sin pérdida de generalidad podemos considerar las traslaciones uniformes /12/ $u_i = F_i(x_i)$, $v_j = G_j(y_j)$, que nos define sendas distribuciones u, v sobre $[0,1]^m$, $[0,1]^n$ respectivamente. Entonces es $f = g = 1$.

Calculemos los extremos de (K(U) + (U-1/2)), etc.,

$$\delta_1 = \sup_{x\in R^m}\{K(U)+(U-1/2)\}=(K(U)+1/2)_{u_i=1}>0$$

$$\delta_2 = \inf_{y\in R^n}\{L(V)+(V-1/2)\}=(L(V)-1/2)_{v_i=0}<0$$

por ser U,K(U), V,L(V) funciones crecientes en cada variable. A fin de que (29) sea positiva deberá cumplirse

$$1+64\alpha^2\delta_1\delta_2>0$$

es decir

$$|\alpha|<(8\sqrt{-\delta_1.\delta_2})^{-1} = c$$

Análogamente, si calculamos

$$\delta'_1= \inf_{x\in R^m}\{K(U)+(U-1/2)\}=(K(U)-1/2)_{u_i=0}<0$$

$$\delta'_2= \sup_{y\in R^n}\{L(V)+(V-1/2)\}=(L(V)+1/2)_{v_i=1}>0$$

obtenemos la desigualdad

$$|\alpha|<(8\sqrt{-\delta'_1.\delta'_2})^{-1} = d$$

Así, el intervalo de variación de es [-a,+a] siendo

$$a = \min\{c,d\}$$

que puede calcularse en cada caso concreto.

Por ejemplo, dadas las distribuciones $F(x_1,x_2) = F_1(x_1)F_2(x_2)$, $G(y)$ (m=2,n=1), es fácil ver que $\delta_1 = 3/2$, $\delta_2 = -1/2$, $\delta'_1 = -1/2$, $\delta'_2 = 1/2$, luego $a = (4\sqrt{3})^{-1}$. La densidad de una distribución trivariante con marginales $F = F_1.F_2$, G será entonces

$$h(x,y) = f(x)g(y) +$$
$$+64\alpha^2[F(x)+(F(x)-1/2)f(x)][(G(y)-1/2)g(y)]$$
$$x\epsilon R^2,\ y\epsilon R,\quad -(4\sqrt{3})^{-1}\leq\alpha\leq(4\sqrt{3})^{-1}$$

Sin embargo, cambiando las marginales, puede variar el valor de a.

Finalmente, tomemos $Q_3(x) = \cos(x\Pi)$. Vamos a considerar tres casos:

a) $m = n = 1$

$$h=f.g\{1+\alpha^2\Pi^2 \mathrm{sen}[\Pi(F-1/2)]\mathrm{sen}[\Pi(G-1/2)]\} \qquad (30)$$

$$-\frac{1}{\Pi} \leq \alpha \leq \frac{1}{\Pi}$$

b) $m = n = 2$

$$h=f.g+\{\Pi \mathrm{sen}[\Pi(F-1/2)]f+\Pi^2\cos[\Pi(F-1/2)]F_{x_1}$$
$$\{\Pi \mathrm{sen}[\Pi(G-1/2)]g+\Pi^2\cos[\Pi(G-1/2)]G_{y_1}G_{y_2}$$

$$-\frac{1}{\Pi} \leq \alpha \leq \frac{1}{\Pi}$$

c) Para $m, n > 2$ aparecen expresiones negativas que podrían impedir que (26) fuera una distribución.

5. OTRAS FAMILIAS DE DISTRIBUCIONES CON MARGINALES MULTIVARIANTES.

Finch y Groblicki /17/ introducen un método de construcción de distribuciones bivariantes que puede ser generalizado al caso de marginales multivariantes. Sea $r(u,v)$ una función definida sobre el cuadrado unidad - tal que

$$\int_0^1 r(u,v)\,dv = 0 \qquad \int_0^1 r(u,v)\,du = 0 \qquad (32)$$

$$r(u,v) \geq -1 \qquad (33)$$

Entonces es evidente que

$$H(x,y) = f(x)g(y)[1+r(F(x),G(y))] \qquad (34)$$

da lugar a una familia bivariante con marginales F, G.

Una densidad de la forma (33) fue propuesta previamente por Cohen y Zaparovanny /3/ tomando

$$r(u,v) = c[h(u,v)-h_1(u)-h_2(v)+1] \qquad (35)$$

donde $h(u,v)$ es una densidad sobre el cuadrado unidad con marginales h_1, h_2. La constante c se calcula de modo que se cumpla (33).

Cohen /2/ generaliza la familia (34), donde $r(u,v)$ es (35), al caso de marginales multivariantes como sigue:

Sean $u = (u_1,\ldots,u_m)$, $v = (v_1,\ldots,v_n)$, $h(u,v)$ una densidad sobre el hipercubo unidad de dimensión $m + n$ con marginales $h_1(u)$, $h_2(v)$. Consideremos entonces la función

$$\rho(u,v) = h - h_1 - h_2 + 1$$

y los cambios de variables

$$\begin{aligned} u_i &= u_i(x_1,\ldots,x_m) \qquad i = 1,\ldots,m \\ v_i &= v_i(y_1,\ldots,y_n) \qquad i = 1,\ldots,n \end{aligned} \qquad (36)$$

Indicando $x = (x_1,\ldots,x_m)$, $y = (y_1,\ldots,y_n)$, $dx = dx_1,\ldots,dx_m$ $dv = dv_1,\ldots,dv_n$, etc., y simbolizando (36) con la notación $u = u(x)$, $v = v(y)$, supongamos que los jacobianos de las transformaciones (36) son

$$J_1 = f \qquad J_2 = g \qquad (37)$$

donde $f(x)$, $g(y)$ son densidades marginales. Entonces

$$h_1 = \int_0^1 h(u,v)\,dv = \int_{-\infty}^{+\infty} h(u,v(y))J_2 dy =$$
$$= \int_{-\infty}^{+\infty} g(y)h(u,v(y))\,dy$$
$$1 = \int_0^1 h_2(v)\,dv = \int_{-\infty}^{+\infty} h_2(v(y))J_2 dy =$$
$$\int_{-\infty}^{+\infty} g(y)h_2(v(y))\,dy$$

Además es fácil ver que

$$\int_{-\infty}^{+\infty} f(x)g(y)\rho(u(x),\ v(y))\,dx dy = 0$$

relación que nos permite definir la familia

$$h(x,y) = f(x)g(y)[1+c\rho(u(x),v(y))]$$
$$x\epsilon R^m,\quad y\epsilon R^n \qquad (38)$$

cuyas densidades marginales son $f(x)$, $g(y)$.

Por otra parte, la generalización de (34) po dría consistir en definir una función r(u,v) sobre $[0,1]^{m+n}$ verificando (con el cambio oportuno de notaciones) las relaciones (32), (33). Eligiendo una transformación de variables (36) que verifique (37), tendremos

$$0 = \int_0^1 r(u,v)\,du\,dv =$$

$$= \int_{-\infty}^{+\infty} r(u(x), v(y))\, f(x)\, g(y)\, dx\,dy \qquad (39)$$

Luego podremos definir la familia

$$h(x,y) = f(x)\,g(y).[1+r(u(x),v(y))]$$

$$x \in R^m, \quad y \in R^n \qquad (40)$$

que comprende (37) como caso particular.

La familia propuesta en (26) es similar a la (40), pero (26) tiene la ventaja de venir expresada en términos de las funciones de distribución. Además, a través de la función Q, podemos obtener diferentes soluciones concretas.

6. DEPENDENCIA Y CORRELACION.

La familia (26) goza de algunas propiedades relativas a la dependencia estocástica y a la correlación que vamos a estudiar.

6.1. DEPENDENCIA F-ORTANTE.

Supongamos que F, G son F-ortantes dependientes, es decir, /9/

$$F(x_1,\ldots,x_m) \geq F_1(x_1)\ldots\ldots F_m(x_m)$$

donde $F_1,\ldots,F_m$ son las distribuciones marginales de F. Análogamente para G. Entonces, es evidente que, dado que Q es positiva,

$$H(x_1,\ldots,x_m, y_1,\ldots,y_n) \geq$$

$$\geq F_1(x_1)\ldots F_m(x_m)\,G_1(y_1)\ldots G_n(y_n)$$

Luego la familia (26) conserva la dependencia F-ortante.

6.2. COVARIANZA Y CORRELACION.

En el caso m = n = 1, tengamos en cuenta la siguiente fórmula para la covarianza /9/

$$\sigma_{xy} = \int_{R^2} (H(x,y) - F(x)G(y))\,dx\,dy \qquad (41)$$

Luego

$$\sigma_{xy} = \alpha^2 \int_{R^2} Q(F(x)-1/2)\,Q(G(y)-1/2)\,dx\,dy \qquad (42)$$

y por lo tanto la covarianza es proporcional a α^2.

Si las marginales son uniformes y tomamos $Q = Q_1$, obtenemos $\sigma xy = \alpha^2/4$ y el coeficiente de correlación es

$$\rho_1 = 3\alpha^2 \qquad -\tfrac{1}{2} \leq \alpha \leq +\tfrac{1}{2}$$

Finalmente para $Q = Q_3$ es $\sigma_{xy} = 4\alpha^2/\Pi^2$

$$\rho_3 = \frac{48\alpha^2}{\Pi^2} \qquad -\frac{1}{\Pi} \leq \alpha \leq \frac{1}{\Pi}$$

6.3. OTRA MEDIDA DE DEPENDENCIA.

Centrándonos también en el caso bivariante, es interesante estudiar la medida de dependencia estocástica

$$\theta = \int_{R^2} [H(x,y) - F(x)G(y)]\,dF(x)\,dG(y) \qquad (43)$$

Sustituyendo

$$\theta = \alpha^2 \int_{R^2} Q(F(x)-1/2)\,Q(G(y)-1/2)\,df(x)\,dG(y) \qquad (44)$$

Como θ verifica (ver /5/)

$$\theta = P([X<X'] \cap [Y<Y']) - 1/4$$

donde (X,Y) son variables aleatorias independientes de (X',Y'), ambas con distribución H(x,y), es fácil ver que θ es invariante por transformaciones monótonas de las variables. Luego bastará calcular (44) en el caso de marginales uniformes. Pero entonces θ coincide con el coeficiente de correlación. Luego, tomando Q_1, Q_2, Q_3 obtenemos respectivamente

$$\theta_1 = 8\alpha^2 \qquad \theta_2 = 12\alpha^2 \qquad \theta_3 = 48\alpha^2/\Pi^2$$

7.BIBLIOGRAFIA.

/1/ ALI, M.M.; MIKHAIL, N.N.; HAG, M.S.: "A class of bivariate distributions including the bivariate logistic". J. of Multivariate Analysis, 8, 405-412 (1978).

/2/ COHEN, L.: "Probability distributions with given multivariate marginales". J. of Math. Phys., 25(8), 2402-2403 (1984).

/3/ COHEN, L.; ZAPAROVANNY, Y.I.: "Positive quantum joint distributions". J. of Math. Phys., 21(4), 794-796 (1980).

/4/ CUADRAS, C.M.: "Sobre medidas de dependencia estocástica invariantes por --- transformaciones monótonas". Homenatge a F. d'A.Sales. Contribuciones Científiques. Fac. de Matematiques, U. de Barcelona, 28-47 (1985).

/5/ CUADRAS, C.M.: "Problemas de Probabilidades y Estadística". 6ª.Ed. Vol. I. Probabilidades. Prom. Pub. Univ., Barcelona (1985).

/6/ CUADRAS, C.M., AUGE, J.: "A continuous general multivariate distribution and its properties". Commun. Stat., Theor. Meth., A10 (4), 339-353 (1981).

/7/ FINCH, P.D.; GROBLICKI, R.: "Bivariate probability densities with given marginals". Found. of Physics, 14(6), 549-552 (1984).

/8/ FISZ, M.: "Probability Theory and Mathematical Statistics", 3ª ed. Wiley, New York (1963).

/9/ LEHMANN, E.L.: "Some concepts of dependence". Ann Math. Statistic., 37, 1137-1153 (1966).

/10/ MARDIA, K.V.: "Families of Bivariate Distributions". Griffing Stat. Monog. and Courses, Griffing, London (1970).

/11/ RUIZ-RIVAS, C.: "Un nuevo sistema bivariante y sus propiedades". Estadística Española, 87, 47-54 (1981).

/12/ SANCHEZ ALGARRA, P.: "Algunos aspectos acerca de la construcción de distribuciones multivariantes con marginales dadas". QUESTIIO, V.9(3), 149-161 (1985).

/13/ SCHWARTZ, L.: "Théorie des Distributions". Hermann, Paris (1957).