Rev.R.Acad.Cienc.Exact.Fís.Nat. (Esp)
Vol. 93, N.º 2, pp 231-239, 1999
Matemáticas

ESTIMACION DE LA FUNCION CUANTIL ASOCIADA A UNA COPULA *C(x, y)* MEDIANTE LA FUNCION CUANTIL EMPIRICA

(cópula/contorno de distribución/función cuantil, función cuantil empírica)

L. Imlahi, M. Ezzerg and A. Chakak

Facultad de Ciencias C.P. 2121. Tetuán-Marruecos

Presentado por F. J. Girón el 25 de noviembre de 1998. Aceptado el 14 de abril de 1999.

SUMMARY

The objective of this work is to generalize some results concerning the unidimensional quantile function to the multidimensional case.

RESUMEN

El objetivo de este trabajo es generalizar ciertos resultados concernientes a la función cuantil unidimensional al caso multidimensional.

INTRODUCCIÓN

Sean X e Y dos variables aleatorias de funciones de distribución marginales $F(x)$ y $G(y)$ respectivamente, y de función de distribución conjunta $H(x, y)$.

Podemos siempre escribir $H(x, y) = C(F(x), G(y))$ siempre que $F(x)$ y $G(y)$ sean continuas. Alternativamente, se tiene que: $C(u, v) = H(F^{-1}(u), G^{-1}(v))$, $0 \leq u, v \leq 1$, donde $F^{-1}(u) = \inf \{x : F(x) \geq u\}$ y $G^{-1}(v) = \inf \{y : G(y) \geq v\}$.

La función $C(u, v)$, llamada frecuentemente cópula, es entonces una función de distribución de un vector aleatorio cuyas marginales son uniformes sobre [0, 1]. Para w fijado en (0, 1], el conjunto de soluciones de la ecuación $C(u, v) = w$ es llamado contorno de distribución [2].

En [1], hemos definido una función $\psi(u, w)$ que describe este contorno de distribución, llamada función cuantil asociada a la cópula $C(u, v)$, por

$$\psi(u, w) = \inf \{v : C(u, v) \geq w\}, \quad 0 < w \leq u \leq 1.$$

Sea $C_n(u, v)$ una sucesión de cópulas y $\psi_n(u, w)$ sus funciones cuantiles. Una de las propiedades estudiadas en [1] es:

$$C_n(u,v) \xrightarrow[n\to\infty]{} C(u,v) \Leftrightarrow \psi_n(u,w) \xrightarrow[n\to\infty]{} \psi(u,w)$$

donde C es una cópula y $\psi(u, w)$ su función cuantil.

En este artículo vamos a estudiar la convergencia de la función cuantil empírica $\psi_n(x, u) = \inf \{y : F_n(x, y) \geq u\}$, $0 < u \leq x \leq 1$, hacia $\psi(x, u)$; donde

$F_n(x,y) = \frac{1}{n}\sum_{i=1}^{n} 1_{[X_i \leq x, Y_i \leq y]}$. Precisamente demostramos que:

a) $\psi_n(x,u) \xrightarrow[n\to\infty]{} \psi(x,u)$ c.s.

b) $\sqrt{n}\left(\psi_n(x,u) - \psi(x,u)\right)$ es asintóticamente normal. Lo que permite dar un intervalo de confianza para $\psi(x, u)$, cuyos límites dependen de $\psi'(u) = \frac{\partial}{\partial u}\psi(x,u)$, generalmente desconocida.

Después, damos dos descomposiciones de Bahadur, la primera utilizando $\psi_n(x, u)$ y la segunda utilizando los estadísticos de orden $Y_{(i)}$. Estas dos descomposiciones de Bahadur permiten construir un intervalo de confianza para $\psi(x, u)$, cuyos límites no dependen de $\psi'(u)$, y la diferencia entre los límites inferiores y aquélla entre los límites superiores de dos intervalos de confianza es igual a $\theta(n^{-1/2})$ casi seguramente. Además, sus coeficientes de confianza tienen el mismo límite.

Convergencia de $\psi_n(x, u)$ hacia $\psi(x, u)$

Sea (X, Y) un vector aleatorio de cópula $C(x, y)$ y cuyas variables aleatorias marginales son uniformes sobre [0, 1].

Consideramos una muestra $(X_1, Y_1), \ldots, (X_n, Y_n)$ extraída de (X, Y), y $F_n(x,y) = \frac{1}{n}\sum_{i=1}^{n} 1_{[X_i \leq x, Y_i \leq y]}$. Es conocido, según la ley de los grandes números, que la sucesión de v.a. $F_n(x, y)$ converge casi seguramente hacia $C(x, y)$, y que la distribución de $F_n(x, y)$ es:

$$P\left[F_n(x,y) \le t\right] = \sum_{k=0}^{[nt]} \binom{n}{k} C^k(x,y) \cdot \left(1 - C(x,y)\right)^{n-k}$$

donde $[nt]$ = parte entera de nt.

Definición 1. Para x, u fijos, $0 < u \le x < 1$, se designa por $\psi_n(x, u)$ la función cuantil asociada a $F_n(x, y)$, definida por:

$$\psi_n(x, u) = \inf \{y : F_n(x, y) \ge u\}$$

más exactamente, para $w \in \Omega$:

$$\psi_n(x, u)(\omega) = \inf \{y : F_n(x, y)(\omega) \ge u\}$$

Proposición 1. La distribución de la v.a. $\psi_n(x, u)$ es:

$$G_n(t) = P\left[\psi_n(x,u) \le t\right] = \sum_{k=m}^{n} \binom{n}{k} C^k(x,t)\left(1 - C(x,t)\right)^{n-k}.$$

donde

$$m = \begin{cases} nu & \text{si } nu \text{ es entero} \\ [nu]+1 & \text{sino} \end{cases}$$

y su densidad $g_n(t)$ es:

$$g_n(t) = n\binom{n-1}{m-1}\left(C(x,t)\right)^{m-1}\left(1 - C(x,t)\right)^{n-m} \frac{d}{dt} C(x,t).$$

Demostración:

$$P\left[\omega : \psi_n(x,u)(\omega) \le t\right] = P\left[\omega : u \le F_n(x,t)(\omega)\right]$$
$$= P\left[nu \le \sum_{i=1}^{n} 1_{[X_i \le x, Y_i \le t]}\right]$$
$$= \begin{cases} 1 - P\left[\sum_{i=1}^{n} 1_{[X_i \le x, Y_i \le t]} < nu\right] & \text{si } nu \text{ es entero} \\ 1 - P\left[\sum_{i=1}^{n} 1_{[X_i \le x, Y_i \le t]} \le nu\right] & \text{sino.} \end{cases}$$

de donde $G_n(t) = \sum_{k=m}^{n} \binom{n}{k} C^k(x,t)\left(1 - C(x,t)\right)^{n-k}$, con

$$m = \begin{cases} nu & \text{si } nu \text{ es entero} \\ [nu]+1 & \text{sino} \end{cases}$$

y la densidad (cuando existe, es decir, cuando existe $\frac{d}{dt} C(x,t)$) es:

$$g_n(t) = \frac{d}{dt} G_n(t).$$

Proposición 2. Sea C una cópula estrictamente creciente en y, y $\psi_n(x, u)$ la función asociada a $F_n(x, y)$, entonces $\psi_n(x,u) \underset{n\to\infty}{\overset{c.s.}{\to}} \psi(x,u)$.

Demostración. Sea $\varepsilon > 0$, se sabe que $F_n(x, \psi(x, u) - \varepsilon)$ converge casi seguramente hacia $C(x, \psi(x, u) - \varepsilon)$ y $F_n(x, \psi(x, u) + \varepsilon) \overset{c.s.}{\to} C(x, \psi(x, u) + \varepsilon)$, por consiguiente, para todo ω fijo, $\exists N_1 : \forall n \ge N_1$ se tiene que:

$F_n(x, \psi(x, u) - \varepsilon) < u$; porque por definición,

$C(x, \psi(x, u) - \varepsilon)$ es estrictamente inferior a u.

De igual modo existe N_2, tal que $\forall n \ge N_2$ se tiene que $u < F_n(x, \psi(x, u) + \varepsilon)$, porque

$C(x, \psi(x, u) + \varepsilon) > u$ ya que por hipótesis, C es estrictamente creciente en y, así $\forall n \ge N = \max \{N_1, N_2\}$, se tiene:

$$F_n(x, \psi(x, u) - \varepsilon) < u < F_n(x, \psi(x, u) + \varepsilon)$$

Por tanto,

$$\Omega \overset{p.s.}{=} \bigcup_m \bigcap_{n \ge m} \left\{F_n\left(x, \psi(x,u) - \varepsilon\right) < u < F_n\left(x, \psi(x,u) + \varepsilon\right)\right\}$$ y se tiene:

$$1 = \lim_{m\to\infty} P\left[F_n\left(x, \psi(x,u) - \varepsilon\right) < u < F_n\left(x, \psi(x,u) + \varepsilon\right), n \ge m\right]$$

Es decir:

$$1 = \lim_{m\to\infty} P\left[\psi(x,u) - \varepsilon \le \psi_n(x,u) \le \psi(x,u) + \varepsilon,\ n \ge m\right]$$

o también

$$P\left[\sup_{n \ge m} \left|\psi_n(x,u) - \psi(x,u)\right| \le \varepsilon\right] \underset{m\to\infty}{\to} 1.$$

Proposición 3. Bajo las hipótesis de la proposición 2, se tiene:

$$P\left[\left|\psi_n(x,u) - \psi(x,u)\right| > \varepsilon\right] \le 2 \exp -2n\delta_\varepsilon^2$$

donde

$$\delta_\varepsilon = \min \{C(x, \psi(x, u) + \varepsilon) - u, u - C(x, \psi(x, u) - \varepsilon)\}$$

Demostración:

$$P\left[\left|\psi_n(x,u) - \psi(x,u)\right| > \varepsilon\right] = P\left[\psi_n(x,u) > \psi(x,u) + \varepsilon\right] +$$
$$+ P\left[\psi_n(x,u) < \psi(x,u) - \varepsilon\right]$$

$$P[\psi_n(x,u) > \psi(x,u)+\varepsilon] = P[u > F_n(\psi(x,u)+\varepsilon)]$$
$$= P\left[nu > \sum_{i=1}^{n} 1_{[X_i \le x, Y_i \le \psi(x,u)+\varepsilon]}\right]$$
$$= P\left[n(u - C(x,\psi(x,u)+\varepsilon)) > \sum_{i=1}^{n}(V_i - EV_i)\right]$$

donde $V_i = 1_{[X_i \le x, Y_i \le \psi(x,u)+\varepsilon]}$

El lema de Hoeffding ([3] p. 75) nos da:

$$P[\psi_n(x,u) > \psi(x,u)+\varepsilon] \le \exp\ -2n\ \delta_1^2$$

donde $\delta_1 = C(x,\ \psi(x,\ u) + \varepsilon) - u$.

De la misma manera se demuestra que:

$$P[\psi_n(x,u) < \psi(x,u)-\varepsilon] \le \exp\ -2n\ \delta_2^2$$

donde $\delta_2 = u - C(x,\ \psi(x,\ u) - \varepsilon)$, de donde se tiene el resultado:

$$P[|\psi_n(x,u) - \psi(x,u)| > \varepsilon] \le 2\ \exp\ -2n\ \delta_\varepsilon^2$$

con $\delta_\varepsilon = \min\ \{C(x,\ \psi(x,\ u) + \varepsilon) - u,\ u - C(x,\ \psi(x,\ u) - \varepsilon)\}$

Comportamiento asintótico de $\psi_n(x,\ u)$

Enunciamos un teorema análogo a aquello que describe el comportamiento asintótico de la sucesión de funciones cuantiles asociadas a $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{[X_i \le x]}$ en el caso de una v.a. unidimensional ([3] p. 77).

Teorema 1. Sea $C(x,\ y)$ una cópula estrictamente creciente en y, y $\psi(x,\ u)$ su función cuantil.

Si $\frac{\partial}{\partial u}\psi(x,u)$ existe en u_0, designado por $\psi'(u_0) > 0$, entonces

$$\frac{n^{1/2}[\psi_n(x,u_0) - \psi(x,u_0)]}{\sqrt{u_0(1-u_0)}\ \psi'(u_0)} \to \mathcal{N}(0,1).$$

Demostración:

$$G_n(t) = P\left[\frac{n^{1/2}[\psi_n(x,u) - \psi(x,u)]}{\sqrt{u(1-u)}\psi'(u)} \le t\right]$$
$$= P\left[\psi_n(x,u) \le n^{-1/2}At + \psi(x,u)\right]$$

donde $A = \sqrt{u(1-u)}\psi'(u)$

$$G_n(t) = P\left[nu \le \sum_{i=1}^{n} 1_{[X_i \le x, Y_i \le n^{-1/2}At + \psi(x,u)]}\right]$$
$$= P[nu \le Z_n(\Delta_{nt})]$$

donde $\Delta_{nt} = C(x,\ n^{-1/2}\ At\ +\ \psi(x,\ u))$ y $Z_n(\Delta_{nt})$ sigue la distribución binomial de parámetros n, Δ_{nt}; $Z_n(\Delta_{nt}) \rightsquigarrow B(n,\ \Delta_{nt})$, de donde $G_n(t) = P[-C_{nt} \le Z_n^*(\Delta_{nt})]$ donde

$$C_{nt} = \frac{n(\Delta_{nt} - u)}{\sqrt{n\Delta_{nt}(1-\Delta_{nt})}}.$$

Denotamos por $\phi(t)$ la función de distribución de la normal $\mathcal{N}(0,\ 1)$.

$$\phi(t) - G_n(t) = \phi(t) - 1 + P[Z_n^*(\Delta_{nt}) < -C_{nt}]$$
$$= \phi(t) - \phi(C_{nt}) + 1 - \phi(-C_{nt}) - 1 +$$
$$+ P[Z_n^*(\Delta_{nt}) < -C_{nt}]$$
$$= \phi(t) - \phi(C_{nt}) + P[Z_n^*(\Delta_{nt}) < -C_{nt}] - \phi(-C_{nt})$$

ahora bien, el teorema de Berry-Esséen ([3] p. 33) nos da:

$$\left|P[Z_n^*(\Delta_{nt}) < -C_{nt}] - \phi(-C_{nt})\right| \le \frac{33}{4}\ \frac{(1-\Delta_{nt})^2 + \Delta_{nt}^2}{n^{1/2}\sqrt{\Delta_{nt}(1-\Delta_{nt})}} \to 0$$

porque $\Delta_{nt} \underset{n\to\infty}{\longrightarrow} u$.

Y $\left|\phi(t) - \phi(C_{n,t})\right| \underset{n\to\infty}{\longrightarrow} 0$ si $C_{nt} \underset{n\to\infty}{\longrightarrow} t$; ahora bien,

$$C_{nt} = \frac{n(\Delta_{nt} - u)}{\sqrt{n\Delta_{nt}(1-\Delta_{nt})}}$$
$$= \frac{n^{1/2}\left[\psi_x^{-1}(n^{-1/2}At + \psi(x,u)) - u\right]}{\sqrt{\Delta_{nt}(1-\Delta_{nt})}}$$
$$= \frac{tA}{\sqrt{\Delta_{nt}(1-\Delta_{nt})}} \cdot \frac{\left[\psi_x^{-1}(n^{-1/2}At + \psi(x,u)) - u\right]}{n^{-1/2}At}$$

y

$$\lim_{n\to\infty} \frac{tA}{\sqrt{\Delta_{nt}(1-\Delta_{nt})}} = \frac{tA}{\sqrt{u(1-u)}}$$
$$\lim_{n\to\infty} \frac{\psi_x^{-1}(n^{-1/2}At + \psi(x,u)) - u}{n^{-1/2}At} = (\psi_x^{-1})'(\psi(x,u))$$
$$= \frac{1}{\psi'(u)}$$

de donde $\lim_{n\to\infty} C_{nt} = \frac{tA}{\sqrt{u(1-u)}\psi'(u)} = t.$

Observación 1:

$P\left[n^{1/2}\left[\psi_n(x,u)-\psi(x,u)\right]\le 0\right]\underset{n\to\infty}{\longrightarrow} 1/2$ sin que esté satisfecha la hipótesis de existencia de $\psi'(u)$, porque $C_{nt} = 0$ para $t = 0$.

Corolario 1. Sea $C(x, y)$ una cópula verificando las hipótesis del teorema 1.

Si además $\frac{\partial}{\partial u}\psi(x,u)>0$ para todo $x > u$, entonces:

$$\int_u^1 P\left[\frac{n^{1/2}\left[\psi_n(x,u)-\psi(x,u)\right]}{\sqrt{u(1-u)}\psi'(u)}\le\alpha\right]\frac{dx}{1-u}\underset{n\to\infty}{\longrightarrow}\phi(\alpha).$$

Demostración. Se demuestra aplicando directamente la convergencia dominada.

Descomposición de Bahadur utilizando $\psi_n(x, u)$

En el caso de una v.a. unidimensional, el siguiente teorema de Bahadur propone una representación del cuantil empírico F_n^{-1} en función del cuantil F^{-1} y de la función empírica F_n.

Teorema 2 (de Bahadur) ([3] p. 91).

Si $F''(F^{-1}(u))$ existe y $F'(F^{-1}(u)) > 0$, entonces:

$$F_n^{-1}(u)=F^{-1}(u)+\frac{u-F_n\left(F^{-1}(u)\right)}{F'\left(F^{-1}(u)\right)}+R_n$$

con

$$R_n=O\left(\left(\frac{\log n}{n}\right)^{3/4}\right)\ \text{c.s., } n \text{ suficientemente grande.}$$

Definición 2:

$R_n = O(g(n))$ *c.s.* $\Leftrightarrow \exists\Omega_0 \subset \Omega$ tal que $P(\Omega_0) = 1$, y $\forall\omega \in \Omega_0$, $\exists B(\omega)$ verificando $|R_n(\omega)| \le B(\omega).g(n)$, $\forall n$ suficientemente grande.

Este teorema está basado sobre los tres lemas siguientes:

Lema 1. Si $F'(F^{-1}(u)) > 0$, entonces:

$$\left|F_n^{-1}(u)-F^{-1}(u)\right|\le\frac{2}{F'\left(F^{-1}(u)\right)}\cdot\left(\frac{\log n}{n}\right)^{1/2}\ c.s.$$

Observación 2. Si además $F''(F^{-1}(u))$ existe,

$$\left|F_n^{-1}(u)-F^{-1}(u)\right|\le\frac{1}{F'\left(F^{-1}(u)\right)}\left(\frac{\log(\log n)}{n}\right)^{1/2}\ c.s.$$

(ver [3] p. 96).

Lema 2. Si $F''(F^{-1}(u))$ existe y si $T_n \overset{c.s.}{\longrightarrow} F^{-1}(u)$ entonces:

$F(T_n) - F(F^{-1}(u)) = (T_n - F^{-1}(u))\, F'(F^{-1}(u)) + O((T_n - F^{-1}(u))^2)$.

Lema 3 (de Bahadur).

Supongamos que $F''(F^{-1}(u))$ existe y $F'(F^{-1}(u)) > 0$. Sea (a_n) una sucesión de números reales positivos tales que: $a_n \sim c_0\, n^{-1/2} (\log n)^q$, n suficientemente grande con $c_0 > 0$ y $q \ge 1/2$ y si designamos por:

$$H_{un}=\sup_{|x|\le a_n}\left\|\left[F_n\left(F^{-1}(u)+x\right)-F_n\left(F^{-1}(u)\right)\right]-\left[F\left(F^{-1}(u)+x\right)-F\left(F^{-1}(u)\right)\right]\right\|$$

entonces,

$$H_{un}=O\left(n^{-3/4}.(\log n)^{(1+q)/2}\right)c.s.$$

Aplicamos estos resultados a las cópulas absolutamente continuas del vector (X, Y), con $X \rightsquigarrow U[0, 1]$ e $Y \rightsquigarrow U[0, 1]$.

Para x fijo, denotamos por $F(y) = C(x, y)$, $F_n^{-1}(u) = \psi_n(x, u)$ y $F^{-1}(u) = \psi(x, u)$.

Lema 4. Si $\frac{\partial}{\partial y}C\left(x,\psi(x,u)\right)=F'\left(\psi(x,u)\right)$ existe y es estrictamente positiva, entonces:

$$\left|\psi_n(x,u)-\psi(x,u)\right|\le\frac{2}{F'(\psi)}\left(\frac{\log n}{n}\right)^{1/2}\ c.s.$$

Demostración. Se hace lo mismo que en el lema 1. En efecto, ya se ha probado:

$P\left[\left|\psi_n-\psi\right|>\varepsilon\right]\le 2\exp-2n\,\delta_\varepsilon^2$, donde $\delta_\varepsilon = \min\{F(\psi + \varepsilon) - u,\ u - F(\psi - \varepsilon)\}$.

Ahora bien,

$$F(\psi + \varepsilon) - u = F'(\psi).\varepsilon + \theta(\varepsilon)$$

y para $\varepsilon=\varepsilon_n=\frac{2}{F'(\psi)}\left(\frac{\log n}{n}\right)^{1/2}$ se tiene:

$$F(\psi+\varepsilon_n)-u=2\left(\frac{\log n}{n}\right)^{1/2}+\theta(\varepsilon_n)\ge\left(\frac{\log n}{n}\right)^{1/2}$$

para n suficientemente grande.

Así como, $u - F(\psi - \varepsilon_n) \ge \left(\frac{\log n}{n}\right)^{1/2}$, entonces $2n\,\delta_{\varepsilon_n}^2\ge 2\log n$, por consiguiente $P\left[\left|\psi_n-\psi\right|>\varepsilon_n\right]\le\frac{2}{n^2}$. Se sigue que:

$$\sum_{n=1}^{n} P\left[|\psi_n - \psi| > \varepsilon_n\right] < +\infty$$

y el lema de Borel-Cantelli implica que hay solamente un número finito de sucesos $\{|\psi_n - \psi| > \varepsilon_n\}$ que están realizados, es decir, para n suficientemente grande, $|\psi_n - \psi| \leq \varepsilon_n$ *c.s.*

Lema 5. Si $F''(\psi(x, u))$ existe, entonces

$$F(\psi_n) - F(\psi) = (\psi_n - \psi)\, F'(\psi) + O((\psi_n - \psi)^2)$$

Demostración. En un entorno abierto de $\psi(x, u)$, se tiene:

$$F(\psi + s) - F(\psi) = sF'(\psi) + \frac{s^2}{2} F''(\psi) + \theta(s^2)$$

y por definición de $\theta(s^2)$, se tiene:

$$\forall \varepsilon > 0,\ \exists \eta > 0 : |s| < \eta \quad \left|\theta(s^2)\right| \leq \varepsilon s^2$$

pues para n suficientemente grande:

$$\left|F(\psi_n) - F(\psi) - (\psi_n - \psi)F'(\psi)\right| \leq \left(\varepsilon + \frac{|F''(\psi)|}{2}\right)(\psi_n - \psi)^2$$

es decir, $F(\psi_n) - F(\psi) = (\psi_n - \psi)\, F'(\psi) + O((\psi_n - \psi)^2)$.

Lema 6. Si $F''(\psi)$ existe y $F'(\psi) > 0$ y si

$$a_n = \frac{2}{F'(\psi)} n^{-1/2} (\log n)^q,\ q \geq 1/2$$

entonces

$$H_{un} = O\left(n^{-3/4} (\log n)^{(1+q)/2}\right)$$

donde

$$H_{un} = \sup_{|s| \leq a_n} \left|\left[F_n(\psi + s) - F_n(\psi)\right] - \left[F(\psi + s) - F(\psi)\right]\right|.$$

Demostración. Ponemos $F(y) = C(x, y)$, x fijo.

Sea (b_n) una sucesión de enteros positivos, tales que

$$b_n \sim \frac{2}{F'(\psi)} n^{1/4} (\log n)^q \underset{n \to \infty}{\longrightarrow} +\infty$$

entonces $[-a_n, a_n] = \bigcup_{l=-b_n}^{b_n - 1} \left[\frac{la_n}{b_n}, \frac{(l+1)a_n}{b_n}\right]$.

Denotamos por:

$$G_{l,n} = \left[F_n\left(\psi + \frac{la_n}{b_n}\right) - F_n(\psi)\right] - \left[F\left(\psi + \frac{la_n}{b_n}\right) - F(\psi)\right]$$

y

$$\alpha_{l,n} = F\left(\psi + \frac{(l+1)a_n}{b_n}\right) - F\left(\psi + \frac{la_n}{b_n}\right)$$

Para todo $s \in \left[\frac{la_n}{b_n}, \frac{(l+1)a_n}{b_n}\right]$, se tiene:

$$F_n\left(\psi + \frac{la_n}{b_n}\right) \leq F_n(\psi + s) \leq F_n\left(\psi + \frac{(l+1)a_n}{b_n}\right)$$

y

$$F\left(\psi + \frac{la_n}{b_n}\right) \leq F(\psi + s) \leq F\left(\psi + \frac{(l+1)a_n}{b_n}\right)$$

porque F_n y F son crecientes.

Estas dos últimas desigualdades implican:

$G_{l,n} - \alpha_{l,n} \leq [F_n(\psi + s) - F_n(\psi)] - [F(\psi + s) - F(\psi)] \leq G_{l+1,n} + \alpha_{l,n}$

de esto se deduce: $H_{u,n} \leq K_n + B_n$ donde

$K_n = \max\{|G_{l,n}|,\ l = -b_n, ..., b_n\}$

$B_n = \max\{\alpha_{l,n},\ l = -b_n, ..., b_n - 1\}$

- Probemos que $B_n = O(n^{-3/4})$

$$\alpha_{l,n} = F\left(\psi + \frac{a_n(l+1)}{b_n}\right) - F\left(\psi + \frac{a_n l}{b_n}\right)$$

$$= \int_{\psi + \frac{a_n l}{b_n}}^{\psi + \frac{a_n(l+1)}{b_n}} F'(s)ds \leq \frac{a_n}{b_n}.\sup F'(s)$$

$$\leq \frac{a_n}{b_n} \sup_{|s| \leq a_n} F'(\psi + s) \leq \frac{a_n}{b_n}$$

porque $F'(\psi + s) = \int_0^x c(t, \psi + s)dt \leq 1$, por consiguiente

$$B_n = max\{\alpha_{l,n}, l = -b_n, ..., b_n - 1\} \leq \frac{a_n}{b_n} = n^{-3/4}.$$

- Probemos que $K_n = O\left(n^{-3/4} (\log n)^{\frac{q+1}{2}}\right)$. Sabemos que:

$$|G_{l,n}| = \left|\left[F_n\left(\psi + \frac{a_n l}{b_n}\right) - F_n(\psi)\right] - \left[F\left(\psi + \frac{a_n l}{b_n}\right) - F(\psi)\right]\right|$$

Por tanto,

$$n|G_{l,n}| = \left| \sum_{i=1}^{n} \left(1_{\left[X_i \le x, Y_i \le \psi + \frac{a_n l}{b_n}\right]} - 1_{[X_i \le x, Y_i \le \psi]} \right) - \right.$$

$$\left. - n \left[F\left(\psi + \frac{a_n l}{b_n} \right) - F(\psi) \right] \right|$$

• Si $l > 0$, se pone:

$$V_i = 1_{\left[X_i \le x, Y_i \le \psi + \frac{a_n l}{b_n}\right]} - 1_{[X_i \le x, Y_i \le \psi]}$$

y

$$n|G_{l,n}| = \left| \sum_{i=1}^{n} (V_i - E(V_i)) \right|$$

• Si $l < 0$, se pone:

$$V_i = 1_{[X_i \le x, Y_i \le \psi]} - 1_{\left[X_i \le x, Y_i \le \psi + \frac{a_n l}{b_n}\right]}$$

y

$$n|G_{l,n}| = \left| \sum_{i=1}^{n} (V_i - E(V_i)) \right|.$$

En los dos casos, las variables aleatorias V_i son independientes de Bernoulli de parámetro P_l donde $P_l = F\left(\psi + \frac{la_n}{b_n}\right) - F(\psi)$ si $l > 0$ y por tanto $P_l \le F(\psi + a_n) - F(\psi)$. Y $P_l = F(\psi) - F\left(\psi + \frac{la_n}{b_n}.\right)$ si $l < 0$ y por consiguiente $P_l \le F(\psi) - F(\psi - a_n)$. En los dos casos $l > 0$ o $l < 0$, se tiene:

$$P\left[|G_{l,n}| > t\right] = P\left[n|G_{l,n}| \ge nt\right] = P\left[\left| \sum_{i=1}^{n} (V_i - E(V_i)) \right| \ge nt\right],$$

según el lema de Bernstein ([3] p. 95),

$$P\left[|G_{l,n}| > t\right] \le 2 \exp - \frac{nt^2}{2(P_l + t)}$$

Ponemos ahora $t = t_n = C_1 n^{-3/4} (\log n)^{\frac{q+1}{2}}$ y puesto que:

$P_l \le F(\psi + a_n) - F(\psi)$ si $l > 0$ entonces, $P_l < 2a_n F'(\psi)$ para n suficientemente grande, porque:

$$\frac{F(\psi + a_n) - F(\psi)}{a_n} \to F'(\psi) < 2F'(\psi).$$

Así como para $l < 0$, $P_l < 2a_n F'(\psi)$ para n suficientemente grande.

Por consiguiente,

$$\frac{nt_n^2}{2(P_l + t_n)} \ge \frac{n^{-1/2} C_1^2 (\log n)^{q+1}}{2\left[2a_n F'(\psi) + C_1 n^{-3/4} (\log n)^{\frac{1+q}{2}}\right]}$$

$$= \frac{n^{-1/2} C_1^2 (\log n)^{q+1}}{2\left[4n^{-1/2} (\log n)^q + C_1 n^{-3/4} (\log n)^{\frac{1+q}{2}}\right]}$$

$$= \frac{C_1^2 \log n}{2\left[4 + C_1 n^{-1/4} (\log n)^{\frac{1-q}{2}}\right]}$$

ahora bien, $q \ge 1/2$; así pues $n^{-1/4} (\log n)^{\frac{1-q}{2}} \underset{n \to \infty}{\longrightarrow} 0$. Por consiguiente, para n suficientemente grande

$$2\left[4 + C_1 n^{-1/4} (\log n)^{\frac{1-q}{2}}\right] < 9$$

es decir:

$$\frac{nt_n^2}{2(P_l + t_n)} > \frac{C_1^2 \log n}{9}.$$

Ya que C_1 es una constante positiva arbitraria se elige de manera que $\frac{C_1^2}{9} \ge 2$ es decir: $C_1 \ge \sqrt{18}$, así $P\left[|G_{l,n}| \ge t_n\right] \le 2 \exp(-2 \log n) = \frac{2}{n^2}$.

Por consiguiente:

$$\sum_{n=1}^{\infty} P\left[|G_{l,n}| \ge t_n\right] < +\infty$$

y el lema de Borel-Cantelli nos permite escribir

$$|G_{l,n}| \le \sqrt{18} \cdot n^{-3/4} (\log n)^{\frac{1+q}{2}} \ c.s.$$

de donde:

$$K_n = max\left\{|G_{l,n}|, l = -b_n, \ldots, b_n\right\} \le \sqrt{18} n^{-3/4} (\log n)^{\frac{1+q}{2}} \ c.s.$$

es decir,

$$K_n = O\left(n^{-3/4} (\log n)^{\frac{1+q}{2}}\right) c.s.$$

Con los resultados de los lemas 4, 5 y 6, tenemos el resultado siguiente:

Teorema 3. Si $F''(\psi(x, u))$ existe y $F'(\psi(x, u)) > 0$, entonces

$$\psi_n(x,u) = \psi(x,u) + \frac{u - F_n(\psi(x,u))}{F'(\psi(x,u))} + R_n$$

donde $R_n = O\left(\left(\frac{\log n}{n}\right)^{3/4}\right)$ *c.s.* para n suficientemente grande y $F(y) = C(x, y)$.

Demostración. Se sabe que $\psi_n(x, u) \overset{c.s.}{\to} \psi(x, u)$ y puesto que $F''(\psi)$ existe, el lema 5 nos da: $F(\psi_n) - F(\psi) = (\psi_n - \psi) F'(\psi) + \Delta_n$ donde $\Delta_n = O((\psi_n - \psi)^2)$.

Se puede igualmente escribir:

$F(\psi_n) - F(\psi) = F_n(\psi_n) - F_n(\psi) - F_n(\psi_n) + F_n(\psi) + F(\psi_n) - F(\psi)$

lo que implica:

$F_n(\psi_n) - F_n(\psi) = (\psi_n - \psi) F'(\psi) + \Delta_n + \Delta'_n$

donde $\Delta'_n = [F_n(\psi_n) - F_n(\psi)] - [F(\psi_n) - F(\psi)]$.

Ponemos $F_n(\psi_n) = u + \Delta''_n$, pues $u - F_n(\psi) = (\psi_n - \psi) F'(\psi) + \Delta_n + \Delta'_n + \Delta''_n$ es decir:

$$\psi_n = \psi + \frac{u - F_n(\psi)}{F'(\psi)} - (\Delta_n + \Delta'_n + \Delta''_n)/F'(\psi)$$

Ponemos $R_n = -(\Delta_n + \Delta'_n + \Delta''_n)/F'(\psi)$.

- Se tiene $|\Delta_n| \leq C(\omega)(\psi_n - \psi)^2$ y el lema 4 implica $|\psi_n - \psi|^2 \leq B(\omega)n^{-1} \log n$, pues $|\Delta_n| \leq k(\omega)n^{-1} \log n \leq k(\omega)\, n^{-3/4} (\log n)^{3/4}$, es decir, $\Delta_n = O(n^{-3/4}(\log n)^{3/4})$.

- $|\Delta'_n| \leq H_{un}$ del lema 6 con $q = \frac{1}{2}$, así pues $\Delta'_n = O\left(n^{-3/4}(\log n)^{3/4}\right)$.

- $|\Delta''_n| = |F_n(\psi_n) - u| \leq \frac{1}{n} \leq n^{-3/4}(\log n)^{3/4}$, es decir, $\Delta''_n = O(n^{-3/4}(\log n)^{3/4})$, así $R_n = O\left(\left(\frac{\log n}{n}\right)^{3/4}\right)$.

Algoritmo para el cálculo de $\psi_n(x, u)$

He aquí un algoritmo para el cálculo de $\psi_n(x, u)$ para x y u fijos, $0 < u \leq x \leq 1$.

Sea $(X_1, Y_1), ..., (X_n, Y_n)$ una muestra procedente de (X, Y).

Ponemos $A_x = \{(X_i, Y_i)$, tal que $X_i \leq x$, $i = 1, ..., n\}$ y $m_n = card\ A_x$.

Observación 2. $\lim_{n\to\infty} m_n = +\infty$, sino, es decir, si $\lim_{n\to\infty} m_n$ es finito, entonces

$$o = \lim_{n\to\infty} \frac{m_n}{n} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{[X_i \leq x]} \overset{c.s.}{=} x > 0.$$

Sea $Y_{(j)}$, $j = 1, ..., m_n$ el estadístico de orden de las proyecciones de A_x sobre Y.

- Si $u \geq \frac{m_n}{n}$, se toma $\psi_n(x, u) = 1$

- si $u < \frac{m_n}{n}$, entonces $\exists k$ entero, $k \leq m_n - 1$ tal que $\frac{k}{n} < u \leq \frac{k+1}{n}$. Y en este caso, se tiene:

$$F_n\left(x, Y_{(k)}\right) = \frac{k}{n} < u \leq \frac{k+1}{n} = F_n\left(x, Y_{(k+1)}\right)$$

y puesto que la función $y \to F_n(x, y)$ es escalonada de salto $\frac{1}{n}$, se tiene pues

$$\psi_n(x, y) = Y_{(k+1)}.$$

Resumen y algoritmo:

$$\psi_n(x,u) = \begin{cases} 1 & si\ u \geq \frac{m_n}{n} \\ Y_{nu} & si\ nu \text{ es entero y } u < \frac{m_n}{n} \\ Y_{[nu]+1} & si\ nu \text{ no es entero y } u < \frac{m_n}{n}. \end{cases}$$

Descomposición de Bahadur utilizando los estadísticos de orden $Y_{(i)}$

Para x fijo, consideramos los estadísticos de orden $Y_{(i)}$ de las proyecciones de A_x sobre Y.

Como en la descomposición de Bahadur para $\psi_n(x, u)$, se necesita el lema siguiente:

Lema 7. Sea (k_n) una sucesión de enteros tales que:

$$\frac{k_n}{n} = u + \theta\left(\left(\frac{\log n}{n}\right)^{1/2}\right)$$

entonces $\left|Y_{k_n} - \psi(x,u)\right| \leq \frac{2}{F'(\psi)} \cdot \left(\frac{\log n}{n}\right)$ *c.s.* donde $F(y) = C(x, y)$ verifica las hipótesis del lema 4.

Demostración. Se utiliza la misma técnica que para:

$$|\psi_n(x,u)-\psi(x,u)| \le \frac{2}{F'(\psi)}\left(\frac{\log n}{n}\right)^{1/2} \ c.s.$$

En efecto:

$$P\left[\left|Y_{(k_n)}-\psi\right|>\varepsilon_n\right]=P\left[Y_{(k_n)}>\psi+\varepsilon_n\right]+P\left[Y_{k_n}<\psi-\varepsilon_n\right]$$

$$y \ \ P\left[Y_{k_n}>\psi+\varepsilon_n\right]=P\left[\frac{k_n}{n}>F_n(x,\psi+\varepsilon_n)\right]=$$

$$= P\left[n\left(F(\psi+\varepsilon_n)-\frac{k_n}{n}\right)<\sum_{i=1}^{n}(V_i-EV_i)\right]$$

donde $V_i=-1_{[X_i\le x, Y_i\le \psi+\varepsilon_n]}$.

Así pues, el lema de Hoeffding ([3] p. 75) nos da:

$$P\left[Y_{k_n}>\psi+\varepsilon_n\right]\le \exp\ -2n\delta_1^2$$

donde $\delta_1 = F(\psi + \varepsilon_n) - \frac{k_n}{n}$. Lo mismo para:

$$P\left[Y_{k_n}<\psi-\varepsilon_n\right]\le \exp\ -2n\delta_2^2$$

donde $\delta_2 = \frac{k_n}{n} - F(\psi + \varepsilon_n)$, de donde se tiene el resultado:

$$P\ [|Y_{(k_n)} - \psi| > \varepsilon_n] \le 2\ \exp\ - 2n\delta^2$$

con $\delta = \min(\delta_1, \delta_2)$ y para $\varepsilon_n=\frac{2}{F'(\psi)}\left(\frac{\log n}{n}\right)^{1/2}$ se tiene $2n\ \delta^2 \ge 2\log n$. En efecto:

$$F(\psi + \varepsilon_n) = u + \varepsilon_n\, F'(\psi) + \theta(\varepsilon_n)$$

según la fórmula de Taylor-Young, por consiguiente:

$$F(\psi+\varepsilon_n)-\frac{k_n}{n}=\varepsilon_n F'(\psi)+\theta(\varepsilon_n)-\theta\left(\left(\frac{\log n}{n}\right)^{1/2}\right)$$

$$=2\left(\frac{\log n}{n}\right)^{1/2}+\theta(\varepsilon_n)-\theta\left(\left(\frac{\log n}{n}\right)^{1/2}\right)$$

de donde: $\left(\frac{n}{\log n}\right)^{1/2}\left[F(\psi+\varepsilon_n)-\frac{k_n}{n}\right]\underset{n\to\infty}{\to} 2$

así pues, para n suficientemente grande

$$\left(\frac{n}{\log n}\right)\delta_1^2>1 \Leftrightarrow 2n\delta_1^2>2\log n$$

Se hace lo mismo para probar que

$$2n\delta_2^2>2\log n$$

lo que implica $2n\ \delta^2 > 2\log n$.

Entonces:

$P\left[\left|Y_{(k_n)}-\psi\right|>\varepsilon_n\right]\le\frac{2}{n^2}$ y por el lema de Borel-Cantelli se llega a la conclusión $\left|Y_{(k_n)}-\psi\right|\le\varepsilon_n\ c.s.$

Teorema 4. Bajo las hipótesis del teorema 1 sobre $C(x, y)$, sea

$\frac{k_n}{n}=u+\theta\left(\left(\frac{\log n}{n}\right)^{1/2}\right)$, entonces

$$Y_{(k_n)}=\psi(x,u)+\left[\frac{k_n}{n}-F_n(x,\psi(x,u))\right]/F'(\psi)+\tilde{R}_n$$

donde $\tilde{R}_n=\mathrm{O}\left(\left(\frac{\log n}{n}\right)^{3/4}\right)$.

Demostración. Se hace lo mismo que en el teorema 1, tomando $Y_{(k_n)}$ en lugar de ψ_n.

Corolario 2. Sea (k_n) una sucesión de enteros tales que:

$$\frac{k_n}{n}=u+\frac{k}{n^{1/2}}+\theta\left(n^{-1/2}\right), k \text{ real}$$

entonces

$$a)\ \ \sqrt{n}\left(Y_{(k_n)}-\psi_n\right)\overset{p.s.}{\to}\frac{k}{F'(\psi)}$$

$$b)\ \ \sqrt{n}\left(Y_{(k_n)}-\psi\right)\overset{L}{\to}\mathcal{N}\left(\frac{k}{F'(\psi)},\frac{u(1-u)}{(F'(\psi))^2}\right).$$

Demostración:

a) Directamente a partir de las dos descomposiciones de Bahadur.

b) Según el teorema de Slutsky, ya que

$$\sqrt{n}\left(Y_{(k_n)}-\psi_n\right)\overset{p.s.}{\to}\frac{k}{F'(\omega)}$$

y

$$\sqrt{n}(\psi_n-\psi)\overset{L}{\to}\mathcal{N}\left(0,\frac{u(1-u)}{(F'(\psi))^2}\right),$$

$$F'(\psi)=\frac{1}{\psi'(u)}$$

INTERVALOS DE CONFIANZA PARA $\psi(x, u)$

Hemos visto que

$$\frac{\sqrt{n}(\psi_n(x,u)-\psi(x,u))}{\sqrt{u(1-u)}\cdot\psi'(u)} \xrightarrow{L} \mathcal{N}(0, 1)$$

Esto nos permite dar un intervalo de confianza para $\psi(x. u)$.

Para $0 < \alpha < 1$, sea K_α tal que $P\ [Z \le K_\alpha] = 1 - \alpha$ donde $Z \to \mathcal{N}(0, 1)$.

Proposición 4. Se pone

$$I_{Q_n} = \left[\psi_n - \frac{K_\alpha\psi'(u)\sqrt{u(1-u)}}{n^{1/2}},\ \psi_n + \frac{K_\alpha\psi'(u)\sqrt{u(1-u)}}{n^{1/2}}\right]$$

• El coeficiente de confianza de I_{Q_n} tiende hacia $1 - 2\alpha$ cuando n tiende hacia $+\infty$.

• Long $I_{Q_n} = \dfrac{2K_\alpha\psi'(u)\sqrt{u(1-u)}}{n^{1/2}} \underset{n\to\infty}{\to} 0.$

Pero los límites de I_{Q_n} dependen de $\psi'(u)$ que es en general desconocida. Así pues, tenemos interés en dar un intervalo de confianza con los límites independientes de $\psi'(u)$.

Sean (k_{1n}) y (k_{2n}) dos sucesiones de enteros tales que:

$$\frac{k_{1n}}{n} \sim u - \frac{K_\alpha\sqrt{u(1-u)}}{n^{1/2}};\quad \frac{k_{2n}}{n} \sim u + \frac{K_\alpha\sqrt{u(1-u)}}{n^{1/2}}$$

Se pone

$$I_{S_n} = \left[Y_{(k_{1n})}, Y_{(k_{2n})}\right]$$

Proposición 5.

a) $P\left[\psi(x,u) \in I_{S_n}\right] \underset{n\to\infty}{\to} 1 - 2\alpha$

b) $Long\ I_{S_n} \sim \dfrac{2K_\alpha\psi'(u)\sqrt{u(1-u)}}{n^{1/2}}$ *c.s.* para n suficientemente grande.

Demostración.

b) La diferencia entre los límites inferiores de I_{Q_n} y I_{S_n} es:

$$\psi_n(x,u) - \frac{K_\alpha\psi'(u)\sqrt{u(1-u)}}{n^{1/2}} - Y_{(k_{1n})} =$$

$$= n^{-1/2}\left[\sqrt{n}\left(\psi_n - Y_{(k_{1n})}\right) - K_\alpha\psi'(u)\sqrt{u(1-u)}\right] = \theta\left(n^{-1/2}\right)\ c.s.$$

según el corolario 2.

Se hace lo mismo para la diferencia entre los límites superiores de I_{Q_n} y I_{S_n}.

$$Y_{(k_{2n})} - \psi_n - \frac{k_\alpha\psi'(u)\sqrt{u(1-u)}}{n^{1/2}} =$$

$$= n^{-1/2}\left[\sqrt{n}\left(Y_{(k_{2n})} - \psi_n\right) - K_\alpha\psi'(u)\sqrt{u(1-u)}\right] = \theta\left(n^{-1/2}\right).\ c.s.$$

Por consiguiente los límites superiores, así como los límites inferiores de I_{Q_n} y I_{S_n} coinciden con una diferencia igual casi seguramente a $\theta\ (n^{-1/2})$, es decir,

$Long\ I_{S_n} \sim \dfrac{2K_\alpha\psi'(u)\sqrt{u(1-u)}}{n^{1/2}}$ *c.s.* para n suficientemente grande.

a) $P\left[Y_{(k_{1n})} > \psi\right] = P\left[\sqrt{n}\left(Y_{(k_{1n})} - \psi\right) > 0\right] =$

$$= P\left[\frac{\sqrt{n}\left(Y_{(k_{1n})} - \psi\right) + K_\alpha\psi'(u)\sqrt{u(1-u)}}{\psi'(u)\sqrt{u(1-u)}} > K_\alpha\right] \underset{n\to\infty}{\to} P[Z > K_\alpha] = \alpha$$

(corolario 2).

Se hace lo mismo para $P\left[Y_{(k_{2n})} < \psi\right] \underset{n\to\infty}{\to} \alpha$, así pues

$$P\left[\psi \in I_{S_n}\right] \underset{n\to\infty}{\to} 1 - 2\alpha$$

de donde se tiene el resultado.

Observación 3. Todos estos resultados son válidos para el caso multidimensional, definiendo la función cuantil asociada a $C(x_1, ..., x_m)$ por:

$\psi\ (x_1, ..., x_{m-1}, u) = \inf\ \{x_m : C\ (x_1, ..., x_m) \ge u\}$ con $0 < u \le C\ (x_1, ..., x_{m-1}, 1)$

AGRADECIMIENTOS

Este trabajo ha sido realizado mediante un proyecto marroquí-andaluz financiado por la Junta de Andalucía.

REFERENCIAS

1. Chakak, A. & Ezzerg, M. (1996) Quantile fonctions for bivariate distributions with given margins, *Journal de Mathématiques du Maroc* **4**, 39-50.

2. Conway, D.A. (1979) Multivariate distributions with specified marginals, Stanford University technical report. Stanford, CA: Stanford University.

3. Serfling, R.J. (1980) Approximation theorems of Mathematical statistics. John Wiley & Sons, New York.