Statistics & Probability Letters 48 (2000) 105–119

Explicit bounds on Levy–Prohorov distance for a class of multidimensional distribution functions

Isha Dewan *, B.L.S. Prakasa Rao

Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi 110016, India

Received April 1999; received in revised form July 1999

Abstract

Let $F(x_1,\dots,x_k)$ and $G(x_1,\dots,x_k)=F_{X_1}(x_1)\dots F_{X_k}(x_k)$, where $F_{X_i}(x_i)$, $1\leqslant i\leqslant k$, are the one-dimensional marginal distributions of F, be two distribution functions on $\mathbb{R}^k$. Here, we obtain explicit bounds for the Levy–Prohorov distance between F and G using some general results due to Yurinskii (1975, Theory Probab. Appl. 20, 1–10).

Keywords: Multidimensional distribution functions; Levy–Prohorov distance; Cumulants

1. Introduction

It is known that if F and G are two distribution functions on the real line, then

$$\sup_{-\infty<x<\infty}|F(x)-G(x)|\leqslant \sup_{B\in\mathscr{C}}|P(B)-Q(B)|\leqslant 2\sup_{-\infty<x<\infty}|F(x)-G(x)|, \tag{1.1}$$

where P and Q denote the probability measures corresponding to the distribution functions F and G, respectively (cf. Prohorov and Rozanov, 1969, pp. 160). Here $\mathscr{C}$ is the class of all convex subsets of the real line. This result does not hold if class $\mathscr{C}$ is replaced by class $\mathscr{B}$, the class of all Borel subsets of the real line. See the counterexample given below due to Babu (1998).

Example 1.1 (Babu, 1998). Let F be the standard normal distribution and G the discrete distribution which puts mass $\frac{1}{4}$ at each of the points $z_{1/4}$, 0, $z_{3/4}$, and 3 where $F(z_a)=a$ for any $0<a<1$. Then

$$\sup_{B\in\mathscr{B}}|P(B)-Q(B)|=1$$

* Corresponding author.
E-mail address: isha@isid.ac.in (I. Dewan)

as P and Q are mutually singular. But

$$\sup_{-\infty<x<\infty} |F(x) - G(x)| \leqslant 0.25.$$

So

$$\sup_{B\in\mathscr{B}} |P(B) - Q(B)| \leqslant 2 \sup_{-\infty<x<\infty} |F(x) - G(x)|$$

does not hold.

The question is whether there is a result analogous to (1.1) in higher dimensions connecting the difference between two distribution functions and the total variation of the difference between the probability measures generated by them. The problem arose in estimating the quantity:

$$\left| \int_{\mathbb{R}^k} g\,\mathrm{d}F(x_1,\dots,x_k) - \int_{\mathbb{R}^k} g\,\mathrm{d}G(x_1,\dots,x_j)\,\mathrm{d}H(x_{j+1},\dots,x_k) \right|,$$

where F, G and H denote the distribution functions of $(X_1,\dots,X_k)$, $(X_1,\dots,X_j)$, and $(X_{j+1},\dots,X_k)$, respectively.

Remark 1.2. The relation between $\int_{\mathbb{R}^k} |f(\boldsymbol{x})-g(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}$ and $\sup_{B\in\mathscr{B}} |P(B)-Q(B)|$, where f and g are densities of F and G with respect to the Lebesgue measure on $\mathbb{R}^k$ and $\mathscr{B}$ is the σ-algebra of Borel subsets of $\mathbb{R}^k$, is well known. Here F and G could be distribution functions on any finite-dimensional space $\mathbb{R}^k$. It is known that (cf. Strasser, 1985, p. 7)

$$\sup_{B\in\mathscr{B}} |P(B) - Q(B)| = \frac{1}{2}\int_{\mathbb{R}^k} |f(\boldsymbol{x}) - g(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}.$$

Result (1.1) quoted at the beginning on the supremum over convex sets on the absolute difference of probability measures generated by distribution functions on the real line does not hold even for the class of convex sets in $\mathbb{R}^2$. The following example due to Babu (1998) demonstrates the point.

Example 1.3 (Babu, 1998). Let F denote the distribution function corresponding to the uniform measure μ on the unit square. Suppose v denotes the measure that puts mass 0.1 at the upper right vertex of the unit square, and distributes the rest of the mass 0.9 uniformly on the remaining part of the diagonal. Let G denote the distribution function corresponding to v. Clearly,

$$F(x,y) = \begin{cases} 0 & \text{if } \min(x,y)\leqslant 0, \\ xy & \text{if } 0<x,y<1, \\ 1 & \text{if } \min(x,y)\geqslant 1 \end{cases}$$

and

$$G(x,y) = \begin{cases} 0 & \text{if } \min(x,y)\leqslant 0, \\ 0.9\min(x,y) & \text{if } 0<\min(x,y)<1, \\ 1 & \text{if } \min(x,y)\geqslant 1. \end{cases}$$

Hence,

$$\Delta = \sup_{x,y} |F(x,y) - G(x,y)| = (0.45)^2.$$

On the other hand, if A denotes the open triangle below the diagonal in the unit square (i.e. $A=\{(x,y)\colon 0<y<x<1\}$), then A is a convex set, $v(A)=0$ and $\mu(A)=0.5$. Consequently, $2\Delta < 0.5 \leqslant \sup |\mu(B) - v(B)|$, where the supremum is taken over all convex sets. Hence the statement that $\sup\{|\mu(B)-\nu(B)|\colon B \text{ convex}\} \leqslant 2\Delta$ is false.

However, it should be noted that in both the examples discussed above the two distributions are mutually singular.

Our aim in this paper is to obtain bounds on the Levy–Prohorov distance between two probability measures generated by a random vector $\boldsymbol{X}=(X_1,\dots,X_k)$ and another random vector $\boldsymbol{Y}=(Y_1,\dots,Y_k)$ where the component Y_i has the same distribution as that of X_i for $1\leqslant i\leqslant k$ but the components Y_i, $1\leqslant i\leqslant k$ are stochastically independent. We will compute bounds in terms of the moments related to the joint distribution of $\boldsymbol{X}$. Our results are based on general results of Yurinskii (1975).

2. Preliminaries

2.1. Cumulants of functions of random vectors

We extend some results on cumulants of functions of random vectors along the same lines as that of Block and Fang (1988). They are used later to prove the main results.

Consider a random vector $(X_1,\dots,X_r)$, where $E|X_i|^r < \infty$, $i=1,\dots,r$.

Definition 2.1 (Block and Fang, 1988). The rth-order joint cumulant of $(X_1,\dots,X_r)$, denoted by $\mathrm{cum}(X_1,\dots,X_r)$, is defined by

$$\mathrm{cum}(X_1,\dots,X_r)=\sum(-1)^{p-1}(p-1)!\left(E\prod_{j\in v_1}X_j\right)\dots\left(E\prod_{j\in v_p}X_j\right), \tag{2.1}$$

where summation extends over all partitions $(v_1,\dots,v_p)$, $p=1,2\dots,r$, of $(1,\dots,r)$.

For real-valued functions f_i, $i=1,\dots,r$, assume that $E|f_i(X_i)|^r<\infty$. The proof of the following lemma is along the same lines as the proof of Lemma 1 of Block and Fang (1988).

Lemma 2.2. *If $E|f_i(X_i)|^m<\infty$, then*

$$E[f_1(X_1)\dots f_m(X_m)]-\prod_{i=1}^{m}E[f_i(X_i)]=\sum \mathrm{cum}(f_k(X_k),\ k\in v_1)\dots\mathrm{cum}(f_k(X_k),\ k\in v_p), \tag{2.2}$$

where $\sum$ extends over all partitions $(v_1,\dots,v_p)$, $p=1,\dots,m-1$, of $\{1,\dots,m\}$.

In particular, for $m=3$, we have

$$\begin{aligned} E[f_1(X_1)f_2(X_2)f_3(X_3)]-\prod_{i=1}^{3}E[f_i(X_i)] &=\mathrm{cum}(f_1(X_1),f_2(X_2),f_3(X_3))+E[f_1(X_1)]\mathrm{cum}(f_3(X_3),f_2(X_2))\\ &\quad+E[f_2(X_2)]\mathrm{cum}(f_1(X_1),f_3(X_3))+E[f_3(X_3)]\mathrm{cum}(f_1(X_1),f_2(X_2)). \end{aligned} \tag{2.3}$$

Note that

$$\mathrm{cum}(f_1(X_1),f_2(X_2))=\mathrm{Cov}(f_1(X_1),f_2(X_2)), \tag{2.4}$$

and if f_1 is differentiable, then

$$f_1(X_1) - f_1(0) = \int_{-\infty}^{\infty} f_1'(x_1)[\varepsilon(x_1) - I_{(-\infty,x_1]}(X_1)]\,\mathrm{d}x_1, \tag{2.5}$$

where

$$\varepsilon(x) = \begin{cases} 1 & \text{if } x \geqslant 0, \\ 0 & \text{otherwise} \end{cases}$$

and $f'(x)$ is the derivative of $f(x)$.

Therefore,

$$E[f_1(X_1)] - f_1(0) = \int_{-\infty}^{\infty} f_1'(x_1)[\varepsilon(x_1) - F_{X_1}(x_1)]\,\mathrm{d}x_1, \tag{2.6}$$

where $F_{X_i}(x_i)$ is the distribution function of X_i.

Then, from Fubini's theorem, we get

$$\begin{aligned} E[(f_1(X_1) - f_1(0))\dots(f_r(X_r) - f_r(0))] &= E\left[\prod_{i=1}^{r}\int_{-\infty}^{\infty} f_i'(x_i)[\varepsilon(x_i) - I_{(-\infty,x_i]}(X_i)]\mathrm{d}x_i\right] \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} E\left(\prod_{i=1}^{r} f_i'(x_i)[\varepsilon(x_i) - I_{(-\infty,x_i]}(X_i)]\,\mathrm{d}x_i\right) \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\prod_{i=1}^{r} f_i'(x_i)\left[\prod_{i=1}^{r}\varepsilon(x_i) - \sum_{j=1}^{r}\prod_{k\neq j}\varepsilon(x_k)F(\boldsymbol{x}^{(j)})\right. \\ &\quad \left. + \sum_{i<j}\prod_{k\neq j}\varepsilon(x_k)F(\boldsymbol{x}^{(i,j)}) + \cdots + (-1)^r F(\boldsymbol{x})\right]\mathrm{d}x_1\dots\mathrm{d}x_r. \end{aligned} \tag{2.7}$$

Here $\boldsymbol{x}^{(i_1,\dots,i_k)}$ represents $(x_1,\dots, x_{i_1-1},x_{i_1+1},\dots, x_{i_k-1},x_{i_k+1},\dots,x_r)$ and $F(\boldsymbol{x}^{(i_1,\dots,i_k)})$ is the distribution function of $\boldsymbol{X}^{(i_1,\dots,i_k)}$.

Using the above results we can prove the following theorem.

Theorem 2.3. *If $E|f_i(X_i)|^r < \infty$ and f_i is differentiable for $i = 1,\dots,r$, then*

$$\mathrm{cum}(f_1(X_1),\dots,f_r(X_r)), = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\prod_{i=1}^{r} f_i'(x_i)\mathrm{cum}(\chi_{X_1}(x_1),\dots,\chi_{X_r}(x_r))\,\mathrm{d}x_1\dots\mathrm{d}x_r, \tag{2.8}$$

where

$$\chi_{X_i}(x_i) = \begin{cases} 1 & \text{if } X_i \geqslant x_i, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. The proof follows from (2.7) and the fact that

$$\mathrm{cum}(f_1(X_1),\dots,f_r(X_r)) = \mathrm{cum}((f_1(X_1) - f_1(0)),\dots,(f_r(X_r) - f_r(0))).$$

Remark 2.4. Using properties (i) and (iv) in Block and Fang (1988), we can extend the above results to complex-valued functions f_i.

2.2. Yurinskii's bound

We first discuss some general results on Levy–Prohorov distance due to Yurinskii (1975).

Let F, G and H be probability distributions on $\mathbb{R}^k$ and let $|.|$ be some norm in $\mathbb{R}^k$. Let $L(F,G)$ be the Levy–Prohorov distance between F and G corresponding to $|.|$, that is, the lower bound of all positive numbers ε such that for any closed set $A \in \mathbb{R}^k$ and for its ε-neighbourhood A^ε in the sense of the norm $|.|$,

$$\mu_F(A) < \mu_G(A^\varepsilon) + \varepsilon, \quad \mu_G(A) < \mu_F(A^\varepsilon) + \varepsilon, \tag{2.9}$$

where μ_F denotes the probability measure corresponding to F.

Suppose that G has a density $g(\boldsymbol{x})$ satisfying the condition

$$\int_{\mathbb{R}^k} |g(\boldsymbol{x}+\boldsymbol{h}) - g(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \leqslant \Gamma|\boldsymbol{h}|, \quad \boldsymbol{h} \in \mathbb{R}^k \tag{2.10}$$

for some constant $\Gamma > 0$. Yurinskii (1975) proved that

$$L(F,G) \leqslant c_1(1+\Gamma)\int_{\mathbb{R}^k} |\boldsymbol{x}| H(\mathrm{d}\boldsymbol{x}) + c_2\, L(F*H, G*H), \tag{2.11}$$

where c_1 and c_2 are absolute constants and $*$ denotes the convolution operation.

Suppose F and G are distribution functions such that

$$\int_{\mathbb{R}^k} |\boldsymbol{x}|^\ell F(\mathrm{d}\boldsymbol{x}) < \infty, \quad \int_{\mathbb{R}^k} |\boldsymbol{x}|^\ell G(\mathrm{d}\boldsymbol{x}) < \infty, \quad \ell = \left[\frac{k}{2}\right] + 1. \tag{2.12}$$

Further suppose that H is a fixed distribution function on $\mathbb{R}^k$ with density $h(\boldsymbol{x})$ such that

$$\int_{\mathbb{R}^k} |\boldsymbol{x}|^\ell h(\boldsymbol{x})\,\mathrm{d}x < \infty \tag{2.13}$$

and the characteristic function

$$\eta(\boldsymbol{t}) = \int_{\mathbb{R}^k} \exp(\mathrm{i}(\boldsymbol{t},\boldsymbol{x}))h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \tag{2.14}$$

vanishes for $|\boldsymbol{t}| \geqslant 1$. Then it follows from Yurinskii (1975) that there exists an absolute constant C, possibly depending on the choice of H but not on F or G, such that

$$L(F,G) \leqslant C\left\{\frac{1+\Gamma}{T} + \left(\int_{|\boldsymbol{t}|\leqslant T} [|\varphi(\boldsymbol{t}) - \gamma(\boldsymbol{t})|^2 + \sum_{i=1}^{\ell} |\mathscr{D}^i(\varphi(\boldsymbol{t}) - \gamma(\boldsymbol{t}))|^2]\,\mathrm{d}\boldsymbol{t}\right)^{1/2}\right\}, \tag{2.15}$$

where

$$\mathscr{D}^i w(\boldsymbol{t}) = \left(\sum_{\|\alpha\|=i} |D^\alpha w(\boldsymbol{t})|^2\right)^{1/2} \tag{2.16}$$

and

$$D^\alpha w(\boldsymbol{t}) = \frac{\partial^{\|\alpha\|} w(\boldsymbol{t})}{\partial^{\alpha_1} t_1 \ldots \partial^{\alpha_k} t_k}, \tag{2.17}$$

where φ and γ are the characteristic functions of F and G, respectively, $\boldsymbol{t} = (t_1, \ldots, t_k)$ and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)$. Here $\|\boldsymbol{\alpha}\| = \alpha_1 + \cdots + \alpha_k$. Throughout the following discussion, the absolute constant may depend on the choice of H but not on F or G.

3. Bound in the bivariate case

Suppose F is a bivariate distribution function and it has the density f with marginal distributions F_X and F_Y and densities f_X and f_Y, respectively. Let $G(x,y)=F_X(x)F_Y(y)$.

It is easy to see that

$$\gamma(t_1,t_2)=\varphi(t_1,0)\varphi(0,t_2)$$

and hence

$$\frac{\partial\gamma(t_1,t_2)}{\partial t_1}=\frac{\partial\varphi(t_1,0)}{\partial t_1}\varphi(0,t_2),$$

$$\frac{\partial\gamma(t_1,t_2)}{\partial t_2}=\varphi(t_1,0)\frac{\partial\varphi(0,t_2)}{\partial t_2}$$

and

$$\frac{\partial^2\gamma(t_1,t_2)}{\partial t_1\partial t_2}=\frac{\partial\varphi(t_1,0)}{\partial t_1}\frac{\partial\varphi(0,t_2)}{\partial t_2}$$

whenever they exist.

In particular, there exists an absolute constant C such that

$$L(F,G)\leqslant C\left\{\frac{1+\Gamma}{T}+\left(\int_{|t|\leqslant T}\left[|\varphi(t_1,t_2)-\varphi(t_1,0)\varphi(0,t_2)|^2+\left|\frac{\partial\varphi(t_1,t_2)}{\partial t_1}-\frac{\partial\varphi(t_1,0)}{\partial t_1}\varphi(0,t_2)\right|^2\right.\right.\right.$$
$$\left.\left.\left.+\left|\frac{\partial\varphi(t_1,t_2)}{\partial t_2}-\varphi(t_1,0)\frac{\partial\varphi(0,t_2)}{\partial t_2}\right|^2+\left|\frac{\partial^2\varphi(t_1,t_2)}{\partial t_1\partial t_2}-\frac{\partial\varphi(t_1,0)}{\partial t_1}\frac{\partial\varphi(0,t_2)}{\partial t_2}\right|^2\right]\mathrm{d}t\right)^{1/2}\right\}. \tag{3.1}$$

Note that

$$\varphi(t_1,t_2)-\varphi(t_1,0)\varphi(0,t_2)=E[\mathrm{e}^{\mathrm{i}t_1X+\mathrm{i}t_2Y}]-E[\mathrm{e}^{\mathrm{i}t_1X}]E[\mathrm{e}^{\mathrm{i}t_2Y}]=\mathrm{Cov}(\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y}),$$

$$\frac{\partial\varphi(t_1,t_2)}{\partial t_1}-\frac{\partial\varphi(t_1,0)}{\partial t_1}\varphi(0,t_2)=E[\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X+\mathrm{i}t_2Y}]-E[\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X}]E[\mathrm{e}^{\mathrm{i}t_2Y}]=\mathrm{Cov}(\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y}),$$

$$\frac{\partial\varphi(t_1,t_2)}{\partial t_2}-\varphi(t_1,0)\frac{\partial\varphi(0,t_2)}{\partial t_2}=\mathrm{Cov}(\mathrm{e}^{\mathrm{i}t_1X},\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y})$$

and

$$\frac{\partial^2\varphi(t_1,t_2)}{\partial t_1\partial t_2}-\frac{\partial\varphi(t_1,0)}{\partial t_1}\frac{\partial\varphi(0,t_2)}{\partial t_2}=\mathrm{Cov}(\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y})$$

under some moment conditions. Let $\xi(t_1,t_2)$ be the integrand under the integral sign on the right-hand side of inequality (2.15). Note that

$$\xi(t_1,t_2)=|\mathrm{Cov}(\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y})|^2+|\mathrm{Cov}(\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y})|^2+|\mathrm{Cov}(\mathrm{e}^{\mathrm{i}t_1X},\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y})|^2$$
$$+|\mathrm{Cov}(\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y})|^2. \tag{3.2}$$

It is known that if h_1 and h_2 are real-valued differentiable functions such that $\mathrm{Cov}(h_1(X),h_2(Y))$ exists, then

$$\mathrm{Cov}(h_1(X),h_2(Y))=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h_1'(x)h_2'(y)H_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

where

$$H_{X,Y}(x,y) = P(X > x, Y > y) - P(X > x)P(Y > y)$$
$$= P(X \leqslant x,\ Y \leqslant y) - P(X \leqslant x)P(Y \leqslant y)$$

(cf. Newman, 1980; Prakasa Rao, 1993). It is easy to see that the above result extends to complex-valued functions $h_1(x)$ and $h_2(y)$ provided that the real and imaginary parts are differentiable. Let

$$h_1(x) = (\mathrm{i}x)^r \mathrm{e}^{\mathrm{i}t_1x} \quad \text{and} \quad h_2(y) = (\mathrm{i}y)^s \mathrm{e}^{\mathrm{i}t_2y},$$

where $r \geqslant 0$ and $s \geqslant 0$. Then

$$h_1'(x) = (\mathrm{i}x)^r \mathrm{i}t_1 \mathrm{e}^{\mathrm{i}t_1x} + r(\mathrm{i}x)^{r-1}\mathrm{i}\mathrm{e}^{\mathrm{i}t_1x},$$
$$h_2'(y) = (\mathrm{i}y)^s \mathrm{i}t_2 \mathrm{e}^{\mathrm{i}t_2y} + s(\mathrm{i}y)^{s-1}\mathrm{i}\mathrm{e}^{\mathrm{i}t_2y},$$

where we interpret the second term on the right-hand side of the above equations as zero whenever $r = 0$ or $s = 0$. Hence,

$$\mathrm{Cov}(h_1(X), h_2(Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}t_1x+\mathrm{i}t_2y} H_{X,Y}(x,y)[\mathrm{i}^{r+s+2}x^r y^s t_1 t_2 + rs\mathrm{i}^{r+s}x^{r-1}y^{s-1} + \mathrm{i}^{r+s+1}x^r y^{s-1}t_1 s$$
$$+ \mathrm{i}^{r+s+1}x^{s-1}y^s t_2 r]\,\mathrm{d}x\,\mathrm{d}y. \tag{3.3}$$

In particular,

$$|\mathrm{Cov}(h_1(X), h_2(Y))| \leqslant \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [|x^r y^s t_1 t_2| + |x^{r-1}y^{s-1}rs| + |x^r y^{s-1}t_1 s| + |x^{r-1}y^s t_2 r|]|H_{X,Y}(x,y)|\,\mathrm{d}x\,\mathrm{d}y. \tag{3.4}$$

Hence, we have

$$J_1 \equiv |\mathrm{Cov}(\mathrm{e}^{\mathrm{i}t_1X}, \mathrm{e}^{\mathrm{i}t_2Y})| \leqslant |t_1t_2| \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |H_{X,Y}(x,y)|\,\mathrm{d}x\,\mathrm{d}y,$$

$$J_2 \equiv |\mathrm{Cov}(\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X}, \mathrm{e}^{\mathrm{i}t_2Y})| \leqslant \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{|xt_1t_2| + |t_2|\}|H_{X,Y}(x,y)|\,\mathrm{d}x\,\mathrm{d}y,$$

$$J_3 \equiv |\mathrm{Cov}(\mathrm{e}^{\mathrm{i}t_1X}, \mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y})| \leqslant \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{|yt_1t_2| + |t_1|\}|H_{X,Y}(x,y)|\,\mathrm{d}x\,\mathrm{d}y \tag{3.5}$$

and

$$J_4 \equiv |\mathrm{Cov}(\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X}, \mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y})| \leqslant \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [|xyt_1t_2| + |xt_1| + |yt_2| + 1]|H_{X,Y}(x,y)|\,\mathrm{d}x\,\mathrm{d}y. \tag{3.6}$$

Let

$$A_{X,Y}^{rs} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x^r y^s||H_{X,Y}(x,y)|\,\mathrm{d}x\,\mathrm{d}y. \tag{3.7}$$

Then

$$J_1 \leqslant |t_1t_2|A_{X,Y}^{00},$$
$$J_2 \leqslant |t_1t_2|A_{X,Y}^{10} + |t_2|A_{X,Y}^{00},$$
$$J_3 \leqslant |t_1t_2|A_{X,Y}^{01} + |t_1|A_{X,Y}^{00}$$

and

$$J_4 \leqslant |t_1 t_2| A^{11}_{X,Y} + |t_1| A^{10}_{X,Y} + |t_2| A^{01}_{X,Y} + A^{00}_{X,Y}. \tag{3.8}$$

In particular, it follows that

$$\xi(t_1,t_2) \leqslant \{|t_1t_2|^2 + (|t_2| + |t_1t_2|)^2 + (|t_1| + |t_1t_2|)^2 + (|t_1t_2| + |t_1| + |t_2| + 1)^2\} \left(\max_{\substack{i=0,1\\ j=0,1}} A^{ij}_{X,Y} \right)^2 \tag{3.9}$$

Suppose the norm $|\boldsymbol{t}|$ is the Euclidean norm $|\boldsymbol{t}| = (t_1^2 + t_2^2)^{1/2}$. Note that

$$\frac{t_1^2 + t_2^2}{2} \geqslant t_1 t_2$$

and hence $T^2/2 \geqslant t_1 t_2$ if $|\boldsymbol{t}| \leqslant T$. Similarly, for $T \geqslant 1$,

$$(|t_2| + |t_1t_2|)^2 \leqslant 2(t_2^2 + t_1^2 t_2^2) \leqslant 2\left(T^2 + \frac{T^4}{4}\right) \leqslant \frac{5}{2} T^4,$$

$$(|t_1| + |t_1t_2|)^2 \leqslant \tfrac{5}{2} T^4$$

and

$$\begin{aligned}(|t_1t_2| + |t_1| + |t_2| + 1)^2 &\leqslant 4(t_1^2 t_2^2 + t_1^2 + t_2^2 + 1)\\ &\leqslant 4\left(\frac{T^4}{4} + T^2 + T^2 + T^2\right)\\ &\leqslant 13 T^4.\end{aligned}$$

Therefore,

$$\xi(t_1,t_2) \leqslant C_0 T^4 \left(\max_{\substack{i=0,1\\ j=0,1}} A^{ij}_{X,Y} \right)^2$$

for $T \geqslant 1$, where C_0 is an absolute constant and hence, for $T \geqslant 1$,

$$\begin{aligned}\int_{|\boldsymbol{t}| \leqslant T} \xi(t_1,t_2)\,\mathrm{d}\boldsymbol{t} &\leqslant \left(C_0 T^4 \left(\max_{\substack{i=0,1\\ j=0,1}} A^{ij}_{X,Y} \right)^2 \right) C_1 T^2\\ &= C_2 T^6 \left(\max_{\substack{i=0,1\\ j=0,1}} A^{ij}_{X,Y} \right)^2,\end{aligned} \tag{3.10}$$

where C_2 is an absolute constant. Relations (3.1), (3.5) and (3.6) show that, for every $T \geqslant 1$,

$$L(F,G) \leqslant C_3 \left\{ \frac{1+\Gamma}{T} + T^3 \max_{\substack{i=0,1\\ j=0,1}} A^{ij}_{X,Y} \right\}, \tag{3.11}$$

where C_3 is an absolute constant. It is clear that (3.11) holds trivially for $0 < T < 1$. Suppose T is chosen so that

$$\frac{1+\Gamma}{T} = T^3 \max_{\substack{i=0,1\\j=0,1}} A^{ij}_{X,Y}.$$

Then it follows that

$$T = \left(\frac{1+\Gamma}{\max_{\substack{i=0,1\\j=0,1}} A^{ij}_{X,Y}} \right)^{1/4}$$

and

$$L(F,G) \leqslant C_4 \left\{ (1+\Gamma)^{3/4} \left[\max_{\substack{i=0,1\\j=0,1}} A^{ij}_{X,Y} \right]^{1/4} \right\}, \tag{3.12}$$

where C_4 is an absolute constant.

Hence the following theorem holds.

Theorem 3.1. *Suppose F and G are distribution functions on $\mathbb{R}^2$ with $G(x,y) = F_X(x)F_Y(y)$ where F_X and F_Y are the marginal distributions of F. Further suppose that G has a density function satisfying* (2.10) *and*

$$\int_{\mathbb{R}^2} |\boldsymbol{x}|^2 F(\mathrm{d}\boldsymbol{x}) < \infty, \quad \int_{\mathbb{R}^2} |\boldsymbol{x}|^2 G(\mathrm{d}\boldsymbol{x}) < \infty.$$

Let $L(F,G)$ be the Levy–Prohorov distance between F and G. Then,

$$L(F,G) \leqslant C \left\{ (1+\Gamma)^{3/4} \left[\max_{\substack{i=0,1\\j=0,1}} A^{ij}_{X,Y} \right]^{1/4} \right\},$$

where C is an absolute constant.

Remark 3.2. If the random variables X and Y are associated then, it is easy to check that there exists an absolute constant C such that

$$L(F,G) \leqslant C \left\{ (1+\Gamma)^{3/4} \left[\max_{\substack{i=0,1\\j=0,1}} A^{*}_{ij} \right]^{1/4} \right\}, \tag{3.13}$$

where

$$A^{*}_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x|^i |y|^j H_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Note that

$$A_{00}^* = \int_{-\infty}^{\infty}\int_{\infty}^{\infty} H_{X,Y}(x,y)\,dx\,dy = \mathrm{Cov}(X,Y) \geqslant 0.$$

It is known that bound (3.13) ensures the fact that

$$\sup_{A\in\mathscr{A}} |\mu_F(A) - \mu_G(A)| \leqslant C\left\{(1+\Gamma)^{3/4}\left[\max_{\substack{i=0,1\\ j=0,1}} A_{ij}^*\right]^{1/4}\right\}, \tag{3.14}$$

where $\mathscr{A}$ is the class of Lipschitz sets (cf. Yurinskii, 1975) with respect to F or G. Recall that $G(x,y) = F_X(x)F_Y(y)$ where F_X and F_Y are the marginal distributions of F. It is plausible that the bound in (3.14) cannot be obtained from the bound given in (3.15) below due to Bagai and Prakasa Rao (1991), from the examples discussed at the beginning of Section 1.

Theorem 3.3. *Let X and Y be associated random variables with bounded continuous density function f_X and f_Y, respectively. Then there exists a constant C depending on f_X and f_Y such that*

$$\sup_{x,y} |P(X \leqslant x, Y \leqslant y) - P(X \leqslant x)P(Y \leqslant y)| \leqslant C\,\mathrm{Cov}^{1/3}(X,Y). \tag{3.15}$$

4. Bound in the trivariate case

Suppose F is a trivariate distribution function and it has the density f with marginal distribution F_X, F_Y and F_Z and marginal densities f_X, f_Y, and f_Z, respectively. Let $G(x,y,z) = F_X(x)F_Y(y)F_Z(z)$. It is easy to see that

$$\gamma(t_1,t_2,t_3) = \varphi(t_1,0,0)\varphi(0,t_2,0)\varphi(0,0,t_3)$$

and hence

$$\frac{\partial\gamma(t_1,t_2,t_3)}{\partial t_1} = \frac{\partial\varphi(t_1,0,0)}{\partial t_1}\varphi(0,t_2,0)\varphi(0,0,t_3)$$

and

$$\frac{\partial^2\gamma(t_1,t_2,t_3)}{\partial t_1\partial t_2} = \frac{\partial\varphi(t_1,0,0)}{\partial t_1}\frac{\partial\varphi(0,t_2,0)}{\partial t_2}\varphi(0,0,t_3)$$

whenever they exist.

Similarly, we have $\partial\gamma(t_1,t_2,t_3)/\partial t_2$, $\partial\gamma(t_1,t_2,t_3)/\partial t_3$, $\partial^2\gamma(t_1,t_2,t_3)/\partial t_3\partial t_2$ and $\partial^2\gamma(t_1,t_2,t_3)/\partial t_1\partial t_3$.

Relation (2.15) implies that

$$L(F,G) \leqslant C\left\{\frac{1+\Gamma}{T} + \left(\int_{|t|\leqslant T}[|\varphi(t_1,t_2,t_3) - \varphi(t_1,0,0)\varphi(0,t_2,0)\varphi(0,0,t_3)|^2\right.\right.$$

$$+\left|\frac{\partial\varphi(t_1,t_2,t_3)}{\partial t_1} - \frac{\partial\varphi(t_1,0,0)}{\partial t_1}\varphi(0,t_2,0)\varphi(0,0,t_3)\right|^2$$

$$+\left|\frac{\partial\varphi(t_1,t_2,t_3)}{\partial t_2} - \varphi(t_1,0,0)\frac{\partial\varphi(0,t_2,0)}{\partial t_2}\varphi(0,0,t_3)\right|^2$$

$$+\left|\frac{\partial\varphi(t_1,t_2,t_3)}{\partial t_3}-\varphi(t_1,0,0)\varphi(0,t_2,0)\frac{\partial\varphi(0,0,t_3)}{\partial t_3}\right|^2$$
$$+\left|\frac{\partial^2\varphi(t_1,t_2,t_3)}{\partial t_1\partial t_2}-\frac{\partial\varphi(t_1,0,0)}{\partial t_1}\frac{\partial\varphi(0,t_2,0)}{\partial t_2}\varphi(0,0,t_3)\right|^2$$
$$+\left|\frac{\partial^2\varphi(t_1,t_2,t_3)}{\partial t_1\partial t_3}-\frac{\partial\varphi(t_1,0,0)}{\partial t_1}\varphi(0,t_2,0)\frac{\partial\varphi(0,0,t_3)}{\partial t_3}\right|^2$$
$$+\left|\frac{\partial^2\varphi(t_1,t_2,t_3)}{\partial t_2\partial t_3}-\varphi(t_1,0,0)\frac{\partial\varphi(0,t_2,0)}{\partial t_2}\frac{\partial\varphi(0,0,t_3)}{\partial t_3}\right|^2\mathrm{d}\boldsymbol{t}\Bigg)^{1/2}\Bigg\}. \tag{4.1}$$

Note that, by Lemma 2.2,

$$\begin{aligned}&\varphi(t_1,t_2,t_3)-\varphi(t_1,0,0)\varphi(0,t_2,0)\varphi(0,0,t_3)\\&\quad=E[\mathrm{e}^{\mathrm{i}t_1X+\mathrm{i}t_2Y+\mathrm{i}t_3Z}]-E[\mathrm{e}^{\mathrm{i}t_1X}]E[\mathrm{e}^{\mathrm{i}t_2Y}]E[\mathrm{e}^{\mathrm{i}t_3Z}]\\&\quad=\mathrm{Cum}(\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{e}^{\mathrm{i}t_3Z})+E[\mathrm{e}^{\mathrm{i}t_1X}]\mathrm{Cov}[\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{e}^{\mathrm{i}t_3Z}]\\&\qquad+E[\mathrm{e}^{\mathrm{i}t_2Y}]\mathrm{Cov}[\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_3Z}]+E[\mathrm{e}^{\mathrm{i}t_3Z}]\mathrm{Cov}[\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y}],\end{aligned} \tag{4.2}$$

$$\begin{aligned}&\frac{\partial\varphi(t_1,t_2,t_3)}{\partial t_1}-\frac{\partial\varphi(t_1,0,0)}{\partial t_1}\varphi(0,t_2,0)\varphi(0,0,t_3)\\&\quad=\mathrm{Cum}(\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{e}^{\mathrm{i}t_3Z})+E[\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X}]\mathrm{Cov}[\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{e}^{\mathrm{i}t_3Z}]\\&\qquad+E[\mathrm{e}^{\mathrm{i}t_2Y}]\mathrm{Cov}[\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_3Z}]+E[\mathrm{e}^{\mathrm{i}t_3Z}]\mathrm{Cov}[\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y}],\end{aligned} \tag{4.3}$$

$$\begin{aligned}&\frac{\partial\varphi(t_1,t_2,t_3)}{\partial t_2}-\varphi(t_1,0,0)\frac{\partial\varphi(0,t_2,0)}{\partial t_2}\varphi(0,0,t_3)\\&\quad=\mathrm{Cum}(\mathrm{e}^{\mathrm{i}t_1X},\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{e}^{\mathrm{i}t_3Z})+E[\mathrm{e}^{\mathrm{i}t_1X}]\mathrm{Cov}[\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{e}^{\mathrm{i}t_3Z}]\\&\qquad+E[\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y}]\mathrm{Cov}[\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_3Z}]+E[\mathrm{e}^{\mathrm{i}t_3Z}]\mathrm{Cov}[\mathrm{e}^{\mathrm{i}t_1X},\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y}],\end{aligned} \tag{4.4}$$

$$\begin{aligned}&\frac{\partial\varphi(t_1,t_2,t_3)}{\partial t_3}-\varphi(t_1,0,0)\varphi(0,t_2,0)\frac{\partial\varphi(0,0,t_3)}{\partial t_3}\\&\quad=\mathrm{Cum}(\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{i}Z\mathrm{e}^{\mathrm{i}t_3Z})+E[\mathrm{e}^{\mathrm{i}t_1X}]\mathrm{Cov}[\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{i}Z\mathrm{e}^{\mathrm{i}t_3Z}]\\&\qquad+E[\mathrm{e}^{\mathrm{i}t_2Y}]\mathrm{Cov}[\mathrm{e}^{\mathrm{i}t_1X},\mathrm{i}Z\mathrm{e}^{\mathrm{i}t_3Z}]+E[\mathrm{i}Z\mathrm{e}^{\mathrm{i}t_3Z}]\mathrm{Cov}[\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y}],\end{aligned} \tag{4.5}$$

$$\begin{aligned}&\frac{\partial^2\varphi(t_1,t_2,t_3)}{\partial t_1\partial t_2}-\frac{\partial\varphi(t_1,0,0)}{\partial t_1}\frac{\partial\varphi(0,t_2,0)}{\partial t_2}\varphi(0,0,t_3)\\&\quad=\mathrm{Cum}(\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{e}^{\mathrm{i}t_3Z})+E[\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X}]\mathrm{Cov}[\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{e}^{\mathrm{i}t_3Z}]\\&\qquad+E[\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y}]\mathrm{Cov}[\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_3Z}]+E[\mathrm{e}^{\mathrm{i}t_3Z}]\mathrm{Cov}[\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{i}Y\mathrm{e}^{\mathrm{i}t_2Y}],\end{aligned} \tag{4.6}$$

$$\begin{aligned}&\frac{\partial^2\varphi(t_1,t_2,t_3)}{\partial t_1\partial t_3}-\frac{\partial\varphi(t_1,0,0)}{\partial t_1}\varphi(0,t_2,0)\frac{\partial\varphi(0,0,t_3)}{\partial t_3}\\&\quad=\mathrm{Cum}(\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{i}Z\mathrm{e}^{\mathrm{i}t_3Z})+E[\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X}]\mathrm{Cov}[\mathrm{e}^{\mathrm{i}t_2Y},\mathrm{i}Z\mathrm{e}^{\mathrm{i}t_3Z}]\\&\qquad+E[\mathrm{e}^{\mathrm{i}t_2Y}]\mathrm{Cov}[\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{i}Z\mathrm{e}^{\mathrm{i}t_3Z}]+E[\mathrm{i}Z\mathrm{e}^{\mathrm{i}t_3Z}]\mathrm{Cov}[\mathrm{i}X\mathrm{e}^{\mathrm{i}t_1X},\mathrm{e}^{\mathrm{i}t_2Y}]\end{aligned} \tag{4.7}$$