Hindawi
Mathematical Problems in Engineering
Volume 2019, Article ID 6963493, 12 pages
https://doi.org/10.1155/2019/6963493

Research Article

A New Generalized Inequality for Covariance in N Dimensions

Mei Liu,[1] Milan Šindelka,[2] and Shi-you Lin [1]

[1] *School of Mathematics and Statistics, Hainan Normal University, Haikou, Hainan, 571158, China*
[2] *Institute of Plasma Physics, Czech Academy of Sciences, Za Slovankou 3, 18200 Prague 8, Czech Republic*

Correspondence should be addressed to Shi-you Lin; linsy1111@foxmail.com

Received 19 August 2018; Revised 15 December 2018; Accepted 24 December 2018; Published 16 January 2019

Academic Editor: Richard I. Avery

Inspired by the work of Zhefei He and Mingjin Wang which was published in the Journal of Inequalities and Applications in 2015, this paper further generalizes some related results to the case of multidimensional random variables. The resulting inequality for covariance is then applied to different multidimensional statistical distributions (multiuniform, multinomial, and multinormal). Coordinate dependence of the inequality is also examined. The obtained formulas could be useful for making estimates in multivariate statistics.

1. Introduction

The concept of covariance appears ubiquitously in probability theory and statistics as the basic measure of correlation between random variables (see, e.g., standard textbooks [1–5]). An intuitive though elusive idea of correlation was transformed into a sound mathematical language in the 19th century by Auguste Bravais and Francis Galton. Readers interested in the fascinating history of this subject are advised to consult a comprehensive work of Stingler [6] or two tutorial articles [7, 8]. An analysis of covariance possesses great practical importance in applied sciences [9, 10], especially in engineering (error analysis, optimum control, probabilistic design, and system identification) [11, 12], in biotechnology [13] and medical sciences [14, 15], or in economy [16].

Naturally, the means/variances/covariances of any ensemble of random variables are uniquely determined by the probability distribution (PD) corresponding to the concrete problem under study. However, both in pure mathematics and in applications, one frequently encounters a case when the pertinent PD is unknown. Under such circumstances, it is often much desirable to provide at least well defined general estimates (constraints) regarding mean/variance/covariance, namely, such estimates which are independent upon the specific PD. These constraints take typically the form of an inequality.

One remarkable result of the abovementioned sort was obtained by Chebyshev as early as in 1882 [1–3, 6]. More explicitly, the so-called Chebyshev inequality enables one to estimate which maximum fraction of values of a given random variable can be located further than a prescribed distance from the mean. Closely related are the subsequently found Ostrowski and Grüss type inequalities [1–3] in all their different variants (as listed in comprehensive monographs [17–19]).

Until nowadays, the works of Chebyshev, Ostrowski, and Grüss have continued to inspire active mathematical research focused on inequalities/estimates of (co)variance. This fact is clearly documented by rich literature dealing with the subject [20–37]. For the purposes of the present article we specifically highlight an inequality for covariance which was derived recently by He and Wang in [20]. Namely, the following theorem has been proven to hold.

Theorem I.

$$\left|\operatorname{cov}\left(f(\xi), g(\xi)\right)\right| \leq 2\left\|f'\right\|_{\infty}\left\|g'\right\|_{\infty} \mathrm{D}\xi. \tag{1}$$

Here $\xi \in (a, b)$ is a single random variable. One assumes that ξ has a finite expectation value $\mathrm{E}\xi$ and a finite variance

$\mathrm{D}\xi$. *Furthermore,* $f, g : [a,b] \longrightarrow \mathbb{R}$ *are any bounded differentiable functions, such that*

$$\|f'\|_\infty = \sup_{x\in(a,b)} |f'(x)| < \infty,$$
$$\|g'\|_\infty = \sup_{x\in(a,b)} |g'(x)| < \infty. \tag{2}$$

The covariance $\mathrm{cov}(f(\xi), g(\xi))$ *is defined explicitly below in Section 2.*

In addition, He and Wang have applied inequality (1) to several concrete probability distributions (uniform, Gamma, Beta, Poisson, and binomial) and derived in this way various other, more specific inequalities.

The purpose of the present article is to generalize the work of He and Wang to the situation when ξ is replaced by an N-dimensional random variable $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_N)$ with $N \geq 1$. More explicitly, we wish to generalize statement (1) and examine subsequently different multidimensional statistical distributions (multiuniform, multinomial, and multinormal).

The paper is structured as follows: Section 2 reviews all the necessary (well known and standard) prerequisites needed for an appropriate multidimensional generalization of statement (1) of Theorem I, such as the concept and basic properties of an expectation, variance, and covariance; the Cauchy-Schwartz inequality; and the Lagrange mean value theorem in N dimensions. Section 3 represents the main core of our work, since it generalizes (1) as mentioned above. In Section 4 we apply our newly derived multidimensional random inequality to concrete multidimensional probability distributions (multiuniform, multinomial, normal, and multinormal). The purpose of Section 5 is to examine coordinate dependence of our generalized inequality (33) stated by Theorem II. Section 6 contains a brief conclusion and prospects of our future research work.

2. Preparatory Considerations

2.1. Multidimensional Random Variables

2.1.1. Discrete Case. Assume that

(i) $\mathscr{M}$ is a countable set, $\mathscr{M} \subset \mathbb{R}^N$

(ii) $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_N)$ is an N-dimensional random variable, $\boldsymbol{\xi} \in \mathscr{M}$

(iii) $\wp(\boldsymbol{\xi}) \geq 0$ $(\forall \boldsymbol{\xi} \in \mathscr{M})$ are the associated probabilities completely determining the statistical distribution of $\boldsymbol{\xi}$

Then

(a) the probabilities $\wp(\boldsymbol{\xi})$ satisfy the normalization condition

$$\sum_{\boldsymbol{\xi}\in\mathscr{M}} \wp(\boldsymbol{\xi}) = 1 \tag{3}$$

(b) an expectation of $\boldsymbol{\xi}$ is given by formula

$$\mathrm{E}\boldsymbol{\xi} = \sum_{\boldsymbol{\xi}\in\mathscr{M}} \boldsymbol{\xi}\wp(\boldsymbol{\xi}). \tag{4}$$

Hereafter we shall assume that $\mathrm{E}\boldsymbol{\xi}$ is well defined and finite

(c) the variance of ξ_j (with $j \in \{1, \ldots, N\}$) is given by formula

$$\mathrm{D}\xi_j = \mathrm{E}\left(\xi_j - \mathrm{E}\xi_j\right)^2. \tag{5}$$

Hereafter we shall assume that $\mathrm{D}\xi_j$ is well defined and finite. Relation (5) can be recast into an equivalent form

$$\mathrm{D}\xi_j = \mathrm{E}\xi_j^2 - \mathrm{E}^2\xi_j \tag{6}$$

Let $f : \mathscr{M} \longrightarrow \mathbb{R}$ be a function which is bounded in $\mathscr{M}$. Boundedness of f means that

$$\|f\|_\infty = \sup_{\boldsymbol{\xi}\in\mathscr{M}} |f(\boldsymbol{\xi})| < \infty. \tag{7}$$

Then

(α) an expectation

$$\mathrm{E}f(\boldsymbol{\xi}) = \sum_{\boldsymbol{\xi}\in\mathscr{M}} f(\boldsymbol{\xi})\wp(\boldsymbol{\xi}) \tag{8}$$

is finite and well defined, since

$$|\mathrm{E}f(\boldsymbol{\xi})| \leq \sum_{\boldsymbol{\xi}\in\mathscr{M}} |f(\boldsymbol{\xi})|\wp(\boldsymbol{\xi}) \leq \|f\|_\infty \sum_{\boldsymbol{\xi}\in\mathscr{M}} \wp(\boldsymbol{\xi}) = \|f\|_\infty < \infty \tag{9}$$

(β) also the variance

$$\mathrm{D}f(\boldsymbol{\xi}) = \mathrm{E}\left(f(\boldsymbol{\xi}) - \mathrm{E}f(\boldsymbol{\xi})\right)^2 = \mathrm{E}f^2(\boldsymbol{\xi}) - \left(\mathrm{E}f(\boldsymbol{\xi})\right)^2 \tag{10}$$

is finite and well defined, since

$$\mathrm{E}f^2(\boldsymbol{\xi}) = \sum_{\boldsymbol{\xi}\in\mathscr{M}} f^2(\boldsymbol{\xi})\wp(\boldsymbol{\xi}) \leq \|f\|_\infty^2 \sum_{\boldsymbol{\xi}\in\mathscr{M}} \wp(\boldsymbol{\xi}) = \|f\|_\infty^2 < \infty \tag{11}$$

2.1.2. Continuous Case. Assume that

(i) Ω is an open set, $\Omega \subseteq \mathbb{R}^N$

(ii) $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_N)$ is an N-dimensional random variable, $\boldsymbol{\xi} \in \Omega$

(iii) $\wp(\boldsymbol{\xi}) \geq 0$ $(\forall \boldsymbol{\xi} \in \Omega)$ are the associated probability densities completely determining the statistical distribution of $\boldsymbol{\xi}$

Then

(a) the probability densities $\wp(\boldsymbol{\xi})$ satisfy the normalization condition

$$\int_\Omega \wp(\boldsymbol{\xi})\,\mathrm{d}^N\xi = 1 \tag{12}$$

(b) an expectation of $\boldsymbol{\xi}$ is given by formula

$$\mathrm{E}\boldsymbol{\xi} = \int_{\Omega} \boldsymbol{\xi}\wp(\boldsymbol{\xi})\,\mathrm{d}^N\xi. \tag{13}$$

Hereafter we shall assume that $\mathrm{E}\boldsymbol{\xi}$ is well defined and finite

(c) the variance of ξ_j (with $j \in \{1,\ldots,N\}$) is given by formula

$$\mathrm{D}\xi_j = \mathrm{E}\left(\xi_j - \mathrm{E}\xi_j\right)^2. \tag{14}$$

Hereafter we shall assume that $\mathrm{D}\xi_j$ is well defined and finite. Similarly as above in (5), also relation (14) can be recast into an equivalent form

$$\mathrm{D}\xi_j = \mathrm{E}\xi_j^2 - \mathrm{E}^2\xi_j \tag{15}$$

Let $f : \Omega \longrightarrow \mathbb{R}$ be a function which is measurable and bounded in Ω. Boundedness of f means that

$$\|f\|_\infty = \sup_{\xi\in\Omega} |f(\boldsymbol{\xi})| < \infty. \tag{16}$$

Then

(α) an expectation

$$\mathrm{E}f(\boldsymbol{\xi}) = \int_{\Omega} f(\boldsymbol{\xi})\wp(\boldsymbol{\xi})\,\mathrm{d}^N\xi \tag{17}$$

is finite and well defined, since

$$\begin{aligned}|\mathrm{E}f(\boldsymbol{\xi})| &\le \int_{\Omega} |f(\boldsymbol{\xi})|\wp(\boldsymbol{\xi})\,\mathrm{d}^N\xi \le \|f\|_\infty \int_{\Omega} \wp(\boldsymbol{\xi})\,\mathrm{d}^N\xi \\ &= \|f\|_\infty < \infty\end{aligned} \tag{18}$$

(β) also the variance

$$\mathrm{D}f(\boldsymbol{\xi}) = \mathrm{E}\left(f(\boldsymbol{\xi}) - \mathrm{E}f(\boldsymbol{\xi})\right)^2 = \mathrm{E}f^2(\boldsymbol{\xi}) - \left(\mathrm{E}f(\boldsymbol{\xi})\right)^2 \tag{19}$$

is finite and well defined, since

$$\begin{aligned}\mathrm{E}f^2(\boldsymbol{\xi}) &= \int_{\Omega} f^2(\boldsymbol{\xi})\wp(\boldsymbol{\xi})\,\mathrm{d}^N\xi \le \|f\|_\infty^2 \int_{\Omega} \wp(\boldsymbol{\xi})\,\mathrm{d}^N\xi \\ &= \|f\|_\infty^2 < \infty\end{aligned} \tag{20}$$

2.2. Covariance. Let $\boldsymbol{\xi} = (\xi_1,\ldots,\xi_N)$ be an N-dimensional random variable. We define the covariance $\mathrm{cov}(\xi_{j_1},\xi_{j_2})$ by prescription

$$\mathrm{cov}\left(\xi_{j_1},\xi_{j_2}\right) = \mathrm{E}\left(\xi_{j_1} - \mathrm{E}\xi_{j_1}\right)\left(\xi_{j_2} - \mathrm{E}\xi_{j_2}\right), \tag{21}$$

with $j_1, j_2 \in \{1,2,\ldots,N\}$. Equivalently one can write

$$\begin{aligned}&\mathrm{cov}\left(\xi_{j_1},\xi_{j_2}\right) \\ &\quad= \mathrm{E}\left(\xi_{j_1}\xi_{j_2} - \xi_{j_1}\mathrm{E}\xi_{j_2} - \xi_{j_2}\mathrm{E}\xi_{j_1} + \mathrm{E}\xi_{j_1}\mathrm{E}\xi_{j_2}\right) \\ &\quad= \mathrm{E}\left(\xi_{j_1}\xi_{j_2}\right) - \mathrm{E}\xi_{j_1}\mathrm{E}\xi_{j_2}.\end{aligned} \tag{22}$$

Recall that $\mathrm{cov}(\xi_{j_1},\xi_{j_1}) = \mathrm{D}\xi_{j_1}$ as follows from (5)-(6).

After taking an absolute value of (21) one gets

$$\left|\mathrm{cov}\left(\xi_{j_1},\xi_{j_2}\right)\right| \le \mathrm{E}\left|\xi_{j_1} - \mathrm{E}\xi_{j_1} \parallel \xi_{j_2} - \mathrm{E}\xi_{j_2}\right|. \tag{23}$$

The well known Cauchy-Schwarz inequality (to be reviewed in the following subsection) implies however

$$\begin{aligned}&\mathrm{E}\left|\xi_{j_1} - \mathrm{E}\xi_{j_1} \parallel \xi_{j_2} - \mathrm{E}\xi_{j_2}\right| \\ &\quad\le \sqrt{\mathrm{E}\left(\xi_{j_1} - \mathrm{E}\xi_{j_1}\right)^2}\sqrt{\mathrm{E}\left(\xi_{j_2} - \mathrm{E}\xi_{j_2}\right)^2}.\end{aligned} \tag{24}$$

Hence one may conclude that

$$\left|\mathrm{cov}\left(\xi_{j_1},\xi_{j_2}\right)\right| \le \sqrt{\mathrm{D}\xi_{j_1}\mathrm{D}\xi_{j_2}}. \tag{25}$$

This means also that $\mathrm{cov}(\xi_{j_1},\xi_{j_2})$ is well defined and finite as long as all the variances $\{\mathrm{D}\xi_j\}_{j=1}^N$ are well defined and finite.

2.3. The Cauchy-Schwartz Inequality. Let $x = (x_1,\ldots,x_N)$ and $y = (y_1,\ldots,y_N)$ be any two vectors in $\mathbb{R}^N$.

$$\left(\sum_{j=1}^N x_j y_j\right)^2 \le \left(\sum_{j=1}^N x_j^2\right)\left(\sum_{j'=1}^N y_{j'}^2\right). \tag{26}$$

The proof is discussed in all standard textbooks of functional analysis, e.g., in [38].

2.4. The Lagrange Mean Value Theorem in N Dimensions. Let f, g be real valued differentiable functions defined in an open convex set $\Omega \subseteq \mathbb{R}^N$. Let $\boldsymbol{\xi} = (\xi_1,\ldots,\xi_N) \in \Omega$, $\boldsymbol{\eta} = (\eta_1,\ldots,\eta_N) \in \Omega$. Convexity of Ω implies that a straight line segment $\{\boldsymbol{\xi} + \gamma(\boldsymbol{\eta} - \boldsymbol{\xi}),\ \gamma \in [0,1]\}$ connecting $\boldsymbol{\xi}$ with $\boldsymbol{\eta}$ is entirely contained within Ω. The Lagrange mean value theorem states that there always exists a number $\gamma(\boldsymbol{\xi},\boldsymbol{\eta}) \in (0,1)$ with the basic property

$$\begin{aligned}&f(\boldsymbol{\eta}) - f(\boldsymbol{\xi}) \\ &\quad= \sum_{j=1}^N \partial_j f\left(\boldsymbol{\xi} + \gamma(\boldsymbol{\xi},\boldsymbol{\eta})(\boldsymbol{\eta} - \boldsymbol{\xi})\right)\left(\eta_j - \xi_j\right).\end{aligned} \tag{27}$$

Because it is not easy to find a proof of statement (27) in available standard textbooks, we prefer to supplement here our own short proof. It is inspired by Theorem 4.2 given on page 378 of [39]: Direct calculation confirms that

$$\begin{aligned}&\int_0^1 \left\{\sum_{j=1}^N \partial_j f\left(\boldsymbol{\xi} + \tilde{\gamma}(\boldsymbol{\eta} - \boldsymbol{\xi})\right)\left(\eta_j - \xi_j\right)\right\}\mathrm{d}\tilde{\gamma} \\ &\quad= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\tilde{\gamma}} f\left(\boldsymbol{\xi} + \tilde{\gamma}(\boldsymbol{\eta} - \boldsymbol{\xi})\right)\mathrm{d}\tilde{\gamma} = f(\boldsymbol{\eta}) - f(\boldsymbol{\xi}).\end{aligned} \tag{28}$$

The Lagrange mean value theorem of the integral calculus (see, e.g., [40]) implies however that

$$\begin{aligned}&\text{l.h.s. of (28)} \\ &\quad= \left.\sum_{j=1}^N \partial_j f\left(\boldsymbol{\xi} + \tilde{\gamma}(\boldsymbol{\eta} - \boldsymbol{\xi})\right)\left(\eta_j - \xi_j\right)\right|_{\tilde{\gamma}=\gamma(\boldsymbol{\xi},\boldsymbol{\eta})},\end{aligned} \tag{29}$$

where $\gamma(\boldsymbol{\xi},\boldsymbol{\eta}) \in (0,1)$ is a generally unknown fixed number (depending of course not only upon $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ but also on the function f). Combination of (28) and (29) yields now immediately the desired claim (27).

3. Multidimensional Generalization of Theorem I

3.1. Preliminaries. Before formulating the above advertised multidimensional generalization of Theorem (1), let us introduce some additional notations and conventions. Recall that $\boldsymbol{\xi}$ is an N-dimensional random variable with a finite expectation and a finite variance (see Section 2.1 for details). We define an auxiliary quantity

$$\mathrm{D}\xi = \sum_{j=1}^{N} \mathrm{D}\xi_j. \tag{30}$$

Symbol Ω will hereafter stand for an open convex subset of $\mathbb{R}^N$ (case $\Omega = \mathbb{R}^N$ is also allowed to occur). If $\boldsymbol{\xi}$ is a discrete random vector, then we tacitly assume $\mathscr{M} \subset \Omega$.

Assume now that two functions $f, g : \Omega \longrightarrow \mathbb{R}$ are continuous and differentiable in Ω. Assume also that all the partial derivatives $\{\partial_j f\}_{j=1}^N$, $\{\partial_j g\}_{j=1}^N$ are bounded in Ω. Then one may define auxiliary symbols

$$\begin{aligned} \left\|f_j\right\|_\infty &= \sup_{\zeta\in\Omega} \left|\partial_j f(\boldsymbol{\zeta})\right| < \infty, \\ \left\|g_j\right\|_\infty &= \sup_{\zeta\in\Omega} \left|\partial_j g(\boldsymbol{\zeta})\right| < \infty. \end{aligned} \tag{31}$$

$$(1 \le j \le N).$$

Subsequently one may introduce additional notations $\mathfrak{f}_\infty$ and $\mathfrak{g}_\infty$ through the formulas

$$\begin{aligned} \mathfrak{f}_\infty &= \sqrt{\sum_{j=1}^{N} \left\|f_j\right\|_\infty^2}, \\ \mathfrak{g}_\infty &= \sqrt{\sum_{j=1}^{N} \left\|g_j\right\|_\infty^2}. \end{aligned} \tag{32}$$

3.2. Our Basic Theorem. Now we are ready to state our own multidimensional generalization of Theorem I.

Theorem II. *Assuming the above specified notations and conventions, one has*

$$\left|\operatorname{cov}\left(f(\boldsymbol{\xi}), g(\boldsymbol{\xi})\right)\right| \le 2\mathfrak{f}_\infty \mathfrak{g}_\infty \mathrm{D}\xi. \tag{33}$$

Proof. Since the proof is a bit lengthy, we shall conveniently divide it into parts:

(i) let $\boldsymbol{x} \in \Omega$ be a fixed parameter. Then $f(\boldsymbol{x})$ is just a fixed number, and

$$f(\boldsymbol{x}) - \mathrm{E}f(\boldsymbol{\xi}) = \mathrm{E}\left(f(\boldsymbol{x}) - f(\boldsymbol{\xi})\right). \tag{34}$$

Define a quantity

$$Q(\boldsymbol{x}) = \left(f(\boldsymbol{x}) - \mathrm{E}f(\boldsymbol{\xi})\right)^2 = \mathrm{E}^2\left(f(\boldsymbol{x}) - f(\boldsymbol{\xi})\right). \tag{35}$$

The Lagrange mean value theorem (Section 2.4) states that

$$\begin{aligned} &f(\boldsymbol{x}) - f(\boldsymbol{\xi}) \\ &\quad = \sum_{j=1}^{N} \partial_j f\left(\boldsymbol{\xi} + \gamma(\boldsymbol{\xi},\boldsymbol{x})(\boldsymbol{x}-\boldsymbol{\xi})\right)\left(x_j - \xi_j\right), \end{aligned} \tag{36}$$

where $0 < \gamma(\boldsymbol{\xi},\boldsymbol{x}) < 1$. Hence

$$\begin{aligned} Q(\boldsymbol{x}) &= \mathrm{E}^2\left\{\sum_{j=1}^{N} \partial_j f\left(\boldsymbol{\xi} + \gamma(\boldsymbol{\xi},\boldsymbol{x})(\boldsymbol{x}-\boldsymbol{\xi})\right)\left(x_j - \xi_j\right)\right\} \\ &= \left|\mathrm{E}\left\{\sum_{j=1}^{N} \partial_j f\left(\boldsymbol{\xi} + \gamma(\boldsymbol{\xi},\boldsymbol{x})(\boldsymbol{x}-\boldsymbol{\xi})\right)\left(x_j - \xi_j\right)\right\}\right|^2 \\ &= \left|\sum_{j=1}^{N} \mathrm{E}\left\{\partial_j f\left(\boldsymbol{\xi} + \gamma(\boldsymbol{\xi},\boldsymbol{x})(\boldsymbol{x}-\boldsymbol{\xi})\right)\left(x_j - \xi_j\right)\right\}\right|^2 \\ &\le \left(\sum_{j=1}^{N} \left|\mathrm{E}\left\{\partial_j f\left(\boldsymbol{\xi} + \gamma(\boldsymbol{\xi},\boldsymbol{x})(\boldsymbol{x}-\boldsymbol{\xi})\right)\left(x_j - \xi_j\right)\right\}\right|\right)^2. \end{aligned} \tag{37}$$

In the last line of (37) we have used the triangle inequality

(ii) by definition (17), we have

$$\begin{aligned} &\mathrm{E}\left\{\partial_j f\left(\boldsymbol{\xi} + \gamma(\boldsymbol{\xi},\boldsymbol{x})(\boldsymbol{x}-\boldsymbol{\xi})\right)\left(x_j - \xi_j\right)\right\} \\ &\quad = \int_\Omega \partial_j f\left(\boldsymbol{\xi} + \gamma(\boldsymbol{\xi},\boldsymbol{x})(\boldsymbol{x}-\boldsymbol{\xi})\right)\left(x_j - \xi_j\right) \wp(\boldsymbol{\xi})\, \mathrm{d}^N\xi. \end{aligned} \tag{38}$$

But the involved integral satisfies an inequality

$$\begin{aligned} &\left|\int_\Omega \partial_j f\left(\boldsymbol{\xi} + \gamma(\boldsymbol{\xi},\boldsymbol{x})(\boldsymbol{x}-\boldsymbol{\xi})\right)\left(x_j - \xi_j\right) \wp(\boldsymbol{\xi})\, \mathrm{d}^N\xi\right| \\ &\quad \le \int_\Omega \left|\partial_j f\left(\boldsymbol{\xi} + \gamma(\boldsymbol{\xi},\boldsymbol{x})(\boldsymbol{x}-\boldsymbol{\xi})\right)\right| \left|x_j - \xi_j\right| \wp(\boldsymbol{\xi})\, \mathrm{d}^N\xi \\ &\quad \le \left\|f_j\right\|_\infty \int_{\mathbb{R}^N} \left|x_j - \xi_j\right| \wp(\boldsymbol{\xi})\, \mathrm{d}^N\xi \\ &\quad = \left\|f_j\right\|_\infty \mathrm{E}\left|x_j - \xi_j\right|. \end{aligned} \tag{39}$$

Relations (38)-(39) correspond of course to continuous statistical distributions. Analogical formulas apply also in the case of discrete random variables (one merely replaces $\int_\Omega \mathrm{d}^N\xi$ by $\sum_{j\in\mathscr{M}}$). We leave the details to the reader

(iii) after plugging (39) into (37) one finds that

$$Q(\boldsymbol{x}) \le \left(\sum_{j=1}^{N} \left\|f_j\right\|_\infty \mathrm{E}\left|x_j - \xi_j\right|\right)^2. \tag{40}$$

Cauchy-Schwarz inequality (26) implies however

$$\left(\sum_{j=1}^{N}\left\|f_j\right\|_\infty \mathrm{E}\left|x_j-\xi_j\right|\right)^2 \le \sum_{j=1}^{N}\left\|f_j\right\|_\infty^2 \sum_{j=1}^{N}\mathrm{E}^2\left|x_j-\xi_j\right|. \quad (41)$$

Hence also

$$Q(\boldsymbol{x}) \le \mathfrak{f}_\infty^2 \sum_{j=1}^{N}\mathrm{E}^2\left|x_j-\xi_j\right|, \quad (42)$$

where

$$\mathfrak{f}_\infty = \sqrt{\sum_{j=1}^{N}\left\|f_j\right\|_\infty^2}, \quad (43)$$

consonantly with (32)

(iv) proceeding further, the variance relation

$$\mathrm{E}\left|x_j-\xi_j\right|^2 - \mathrm{E}^2\left|x_j-\xi_j\right| = \mathrm{D}\left|x_j-\xi_j\right| \ge 0 \quad (44)$$

implies $\mathrm{E}^2|x_j-\xi_j| \le \mathrm{E}(x_j-\xi_j)^2$, giving in turn

$$Q(\boldsymbol{x}) \le \mathfrak{f}_\infty^2 \sum_{j=1}^{N}\mathrm{E}\left(x_j-\xi_j\right)^2. \quad (45)$$

Yet the variance formula

$$\mathrm{E}\left(x_j-\xi_j\right)^2 - \mathrm{E}^2\left(x_j-\xi_j\right) = \mathrm{D}\left(x_j-\xi_j\right) \quad (46)$$

yields $\mathrm{E}(x_j-\xi_j)^2 = \mathrm{E}^2(x_j-\xi_j) + \mathrm{D}(x_j-\xi_j) = (x_j - \mathrm{E}\xi_j)^2 + \mathrm{D}\xi_j$.

This is valid since $\mathrm{D}x_j = 0$ and $\mathrm{D}(-\xi_j) = \mathrm{D}\xi_j$. We may thus conclude that

$$Q(\boldsymbol{x}) \le \mathfrak{f}_\infty^2 \sum_{j=1}^{N}\left[\left(x_j-\mathrm{E}\xi_j\right)^2 + \mathrm{D}\xi_j\right] \quad (47)$$

(v) comparison of (35) and (47) provides an inequality

$$\left(f(\boldsymbol{x}) - \mathrm{E}f(\boldsymbol{\xi})\right)^2 \le \mathfrak{f}_\infty^2 \sum_{j=1}^{N}\left[\left(x_j-\mathrm{E}\xi_j\right)^2 + \mathrm{D}\xi_j\right]. \quad (48)$$

So far, an entity $\boldsymbol{x}$ was treated as a fixed parameter

(vi) now we shall set $\boldsymbol{x}$ to be a random variable equivalent to $\boldsymbol{\xi}$. Consequently we take the expectation over (48). This results in

$$\mathrm{E}\left(f(\boldsymbol{\xi}) - \mathrm{E}f(\boldsymbol{\xi})\right)^2 \le \mathfrak{f}_\infty^2 \sum_{j=1}^{N}\left[\mathrm{E}\left(\xi_j-\mathrm{E}\xi_j\right)^2 + \mathrm{D}\xi_j\right]. \quad (49)$$

Yet, by the definition of the variance (19), one can convert (49) into a simple outcome

$$\mathrm{D}f(\boldsymbol{\xi}) \le \mathfrak{f}_\infty^2 2\mathrm{D}\xi, \quad (50)$$

where of course

$$\mathrm{D}\xi = \sum_{j=1}^{N}\mathrm{D}\xi_j, \quad (51)$$

consonantly with (30)

(vii) completely analogical sequence of considerations can be applied also to the case of $g(\boldsymbol{\xi})$. One would arrive towards an inequality

$$\mathrm{D}g(\boldsymbol{\xi}) \le \mathfrak{g}_\infty^2 2\mathrm{D}\xi, \quad (52)$$

where

$$\mathfrak{g}_\infty = \sqrt{\sum_{j=1}^{N}\left\|g_j\right\|_\infty^2}. \quad (53)$$

Combination of (25), (50), and (52) provides now the desired final claim

$$\left|\mathrm{cov}\left(f(\boldsymbol{\xi}), g(\boldsymbol{\xi})\right)\right| \le \sqrt{\mathrm{D}f(\boldsymbol{\xi})\,\mathrm{D}g(\boldsymbol{\xi})} \le 2\mathfrak{f}_\infty\mathfrak{g}_\infty\mathrm{D}\xi. \quad (54)$$

This is exactly as stated above in (33). Thus, our Theorem II is proven

□

4. An Application of Theorem II on Different Probability Distributions

In the present section, we shall derive some new inequalities by applying our basic Theorem II to three different types of probability distributions of an N-dimensional random variable. The definitions of the distributions discussed below can be found in [4, 41].

4.1. Multiuniform Distribution

4.1.1. Definition. The probability density function $\wp(\boldsymbol{\xi})$ is given by formula

$$\wp(\boldsymbol{\xi}) = \begin{cases} \beta & \text{if } \boldsymbol{\xi} \in \Omega, \\ 0 & \text{if otherwise.} \end{cases} \quad (55)$$

Notice that

(a) the probability densities $\wp(\boldsymbol{\xi})$ satisfy the normalization condition

$$\int_\Omega \wp(\boldsymbol{\xi})\,\mathrm{d}^N\xi = 1. \quad (56)$$

Therefore

$$\beta\int_\Omega \mathrm{d}^N\xi = 1. \quad (57)$$

Define

$$V_\Omega = \int_\Omega \mathrm{d}^N\xi; \tag{58}$$

this is the volume of region Ω. Then

$$\beta = \frac{1}{V_\Omega} \tag{59}$$

(b) the expectation of ξ_j (with $j \in \{1, \ldots, N\}$) is given by formula combining (55), (59), and (13),

$$\mathrm{E}\xi_j = \frac{1}{V_\Omega}\int_\Omega \xi_j \mathrm{d}^N\xi \tag{60}$$

(c) the variance of ξ_j (with $j \in \{1, \ldots, N\}$) is then given by formula combining (15) and (60),

$$\begin{aligned}\mathrm{D}\xi_j &= \mathrm{E}\xi_j^2 - \mathrm{E}^2\xi_j \\ &= \frac{1}{V_\Omega}\int_\Omega {\xi_j}^2 \mathrm{d}^N\xi - \left(\frac{1}{V_\Omega}\int_\Omega \xi_j \mathrm{d}^N\xi\right)^2.\end{aligned} \tag{61}$$

After recalling (30) one can write

$$\begin{aligned}\mathrm{D}\xi &= \sum_{j=1}^N \mathrm{D}\xi_j \\ &= \frac{1}{V_\Omega}\int_\Omega |\boldsymbol{\xi}|^2 \mathrm{d}^N\xi - \sum_{i=1}^N\left(\frac{1}{V_\Omega}\int_\Omega \xi_j \mathrm{d}^N\xi\right)^2\end{aligned} \tag{62}$$

4.1.2. An Application of Theorem II

Theorem II-1. *Assume that two functions $f, g : \Omega \longrightarrow \mathbb{R}$ are continuous and differentiable in Ω. Assume also that all the partial derivatives $\{\partial_j f\}_{j=1}^N$, $\{\partial_j g\}_{j=1}^N$ are bounded in Ω. Recalling notations (31) and (32), one has*

$$\begin{aligned}&|\mathfrak{T}(f,g)| \\ &\le 2\mathfrak{f}_\infty \mathfrak{g}_\infty \left[\frac{1}{V_\Omega}\int_\Omega |\boldsymbol{\xi}|^2 \mathrm{d}^N\xi - \sum_{i=1}^N\left(\frac{1}{V_\Omega}\int_\Omega \xi_j \mathrm{d}^N\xi\right)^2\right],\end{aligned} \tag{63}$$

where

$$\begin{aligned}\mathfrak{T}(f,g) &= \frac{1}{V_\Omega}\int_\Omega f(\boldsymbol{\xi})\, g(\boldsymbol{\xi})\, \mathrm{d}^N\xi - \frac{1}{V_\Omega}\int_\Omega f(\boldsymbol{\xi})\, \mathrm{d}^N\xi \\ &\quad \cdot \frac{1}{V_\Omega}\int_\Omega g(\boldsymbol{\xi})\, \mathrm{d}^N\xi.\end{aligned} \tag{64}$$

Proof. Let $\boldsymbol{\xi}$ be an N-dimensional random variable which possesses the multiuniform distribution. According to (55), (59), and (17) we have

$$\mathrm{E}f(\boldsymbol{\xi}) = \frac{1}{V_\Omega}\int_\Omega f(\boldsymbol{\xi})\, \mathrm{d}^N\xi, \tag{65}$$

and therefore

$$\begin{aligned}&\mathrm{E}f(\boldsymbol{\xi})\, g(\boldsymbol{\xi}) - \mathrm{E}f(\boldsymbol{\xi})\, \mathrm{E}g(\boldsymbol{\xi}) \\ &\quad = \frac{1}{V_\Omega}\int_\Omega f(\boldsymbol{\xi})\, g(\boldsymbol{\xi})\, \mathrm{d}^N\xi - \frac{1}{V_\Omega}\int_\Omega f(\boldsymbol{\xi})\, \mathrm{d}^N\xi \\ &\quad \cdot \frac{1}{V_\Omega}\int_\Omega g(\boldsymbol{\xi})\, \mathrm{d}^N\xi = \mathfrak{T}(f,g).\end{aligned} \tag{66}$$

Then

$$\begin{aligned}\mathfrak{T}(f,g) &= \left|\mathrm{E}f(\boldsymbol{\xi})\, g(\boldsymbol{\xi}) - \mathrm{E}f(\boldsymbol{\xi})\, \mathrm{E}g(\boldsymbol{\xi})\right| \\ &= \left|\mathrm{cov}\left(f(\boldsymbol{\xi}), g(\boldsymbol{\xi})\right)\right| \le 2\mathfrak{f}_\infty \mathfrak{g}_\infty \mathrm{D}\xi,\end{aligned} \tag{67}$$

due to Theorem II, (33). After inserting (62) one gets

$$\begin{aligned}&\mathfrak{T}(f,g) \\ &\le 2\mathfrak{f}_\infty \mathfrak{g}_\infty \left[\frac{1}{V_\Omega}\int_\Omega |\boldsymbol{\xi}|^2 \mathrm{d}^N\xi - \sum_{i=1}^N\left(\frac{1}{V_\Omega}\int_\Omega \xi_j \mathrm{d}^N\xi\right)^2\right].\end{aligned} \tag{68}$$

Thus the proof is completed. □

4.1.3. Special Case: $\wp(\boldsymbol{\xi})$ Defined within an N-Dimensional Rectangular Box. Assume that Ω corresponds to a rectangular box. Stated mathematically,

$$\Omega = \times_{j=1}^N \left(a_j, b_j\right), \tag{69}$$

where $a_j, b_j \in \mathbb{R}$ and $a_j < b_j$. Then,

(a) according to (58),

$$V_\Omega = \Pi_{j=1}^N \left(b_j - a_j\right) \tag{70}$$

(b) the expectation of ξ_j is equal to

$$\mathrm{E}\xi_j = \frac{1}{b_j - a_j}\int_{a_j}^{b_j} \xi_j \mathrm{d}\xi_j = \frac{b_j + a_j}{2}, \tag{71}$$

whereas the expectation of ${\xi_j}^2$ is

$$\mathrm{E}{\xi_j}^2 = \frac{1}{b_j - a_j}\int_{a_j}^{b_j} {\xi_j}^2 \mathrm{d}\xi_j = \frac{{b_j}^2 + {a_j}^2 + b_j a_j}{3} \tag{72}$$

(c) the variance of ξ_j can be expressed as

$$\begin{aligned}\mathrm{D}\xi_j &= \mathrm{E}\xi_j^2 - \mathrm{E}^2\xi_j = \frac{{b_j}^2 + {a_j}^2 + b_j a_j}{3} - \left(\frac{b_j + a_j}{2}\right)^2 \\ &= \frac{\left(b_j - a_j\right)^2}{12}.\end{aligned} \tag{73}$$

Subsequently we have

$$\mathrm{D}\xi = \sum_{j=1}^{N} \mathrm{D}\xi_j = \frac{1}{12}\sum_{i=1}^{N}\left(b_j - a_j\right)^2 \tag{74}$$

Now we are ready to make the following statement.

Theorem II-2. *One has*

$$\left|\widetilde{\mathfrak{T}}(f,g)\right| \le \frac{1}{6}\sum_{i=1}^{n}(b_i - a_i)^2\, \mathfrak{f}_\infty \mathfrak{g}_\infty, \tag{75}$$

where

$$\begin{aligned}\widetilde{\mathfrak{T}}(f,g) &= \frac{1}{\Pi_{j=1}^{N}\left(b_j - a_j\right)}\int_{a_N}^{b_N}\cdots\int_{a_1}^{b_1} f(\boldsymbol{\xi})\, g(\boldsymbol{\xi})\, \mathrm{d}\xi_1 \ldots \mathrm{d}\xi_N \\ &\quad - \frac{1}{\Pi_{j=1}^{N}\left(b_j - a_j\right)}\int_{a_N}^{b_N}\cdots\int_{a_1}^{b_1} f(\boldsymbol{\xi})\, \mathrm{d}\xi_1 \ldots \mathrm{d}\xi_N \\ &\quad \cdot \frac{1}{\Pi_{j=1}^{N}\left(b_j - a_j\right)}\int_{a_N}^{b_N}\cdots\int_{a_1}^{b_1} g(\boldsymbol{\xi})\, \mathrm{d}\xi_1 \ldots \mathrm{d}\xi_N.\end{aligned} \tag{76}$$

Proof. The proof follows immediately after substituting (69), (70), and (74) into (63) and (64). □

4.2. Multinomial Distribution

4.2.1. Definition. Assume that

(i) $\mathscr{M} = \{\boldsymbol{\xi} \in (\mathbb{N} \cup \{0\})^N \mid \xi_1 + \ldots + \xi_N = n,\ n \in \mathbb{N}\}$, where

(a) n is a given fixed number of independent trials

(b) N is the number of possible outcomes in each trial,

and here the outcomes can be labeled as $1, 2, \ldots, j, \ldots, N$

(c) ξ_j is the number of occurrences of the outcome j during the whole n trials

(ii) p_j is the prescribed probability of the outcome j for a single trial, and one has

$$\sum_{j=1}^{N} p_j = 1 \tag{77}$$

Then

(a) the probability distribution function $\wp(\boldsymbol{\xi})$ is given by formula

$$\wp(\boldsymbol{\xi}) = \frac{n!}{\xi_1!\xi_2!\ldots\xi_N!}p_1^{\xi_1}p_2^{\xi_2}\ldots p_N^{\xi_N} \tag{78}$$

(b) expectation is equal to

$$\mathrm{E}\xi_j = np_j \tag{79}$$

(c) variance is equal to

$$\mathrm{D}\xi = \sum_{j=1}^{N}\mathrm{D}\xi_j = \sum_{j=1}^{N} np_j\left(1 - p_j\right) \tag{80}$$

4.2.2. An Application of Theorem II

Theorem II-3. *Assume that two functions $f, g : \mathscr{M} \longrightarrow \mathbb{R}$ are continuous and differentiable in Ω where $\mathscr{M} \subset \Omega \subseteq \mathbb{R}^N$. Assume also that all the partial derivatives $\{\partial_j f\}_{j=1}^{N}$, $\{\partial_j g\}_{j=1}^{N}$ are bounded in Ω. Recalling notations (31) and (32), one has*

$$\begin{aligned}&\left|\sum_{\xi_1=0}^{n}\sum_{\xi_2=0}^{n-\xi_1}\cdots\sum_{\xi_N=0}^{n-\xi_1-\ldots-\xi_{N-1}} f(\boldsymbol{\xi})\, g(\boldsymbol{\xi})\, \frac{n!}{\xi_1!\xi_2!\ldots\xi_N!}p_1^{\xi_1}p_2^{\xi_2}\ldots p_N^{\xi_N}\right. \\ &\quad - \sum_{\xi_1=0}^{n}\sum_{\xi_2=0}^{n-\xi_1}\cdots\sum_{\xi_N=0}^{n-\xi_1-\ldots-\xi_{N-1}} f(\boldsymbol{\xi})\, \frac{n!}{\xi_1!\xi_2!\ldots\xi_N!}p_1^{\xi_1}p_2^{\xi_2}\ldots p_N^{\xi_N} \\ &\quad \left.\times \sum_{\xi_1=0}^{n}\sum_{\xi_2=0}^{n-\xi_1}\cdots\sum_{\xi_N=0}^{n-\xi_1-\ldots-\xi_{N-1}} g(\boldsymbol{\xi})\, \frac{n!}{\xi_1!\xi_2!\ldots\xi_N!}p_1^{\xi_1}p_2^{\xi_2}\ldots p_N^{\xi_N}\right| \\ &\quad \le 2\mathfrak{f}_\infty\mathfrak{g}_\infty n\sum_{j=1}^{N}p_j\left(1 - p_j\right).\end{aligned} \tag{81}$$

Proof. Let $\boldsymbol{\xi}$ be an N-dimensional random variable which possesses the multinomial distribution. According to (78) and (8) we have

$$\begin{aligned}\mathrm{E}f(\boldsymbol{\xi}) &= \sum_{\xi_1=0}^{n}\sum_{\xi_2=0}^{n-\xi_1}\cdots\sum_{\xi_N=0}^{n-\xi_1-\ldots-\xi_{N-1}} f(\boldsymbol{\xi}) \\ &\quad \cdot \frac{n!}{\xi_1!\xi_2!\ldots\xi_N!}p_1^{\xi_1}p_2^{\xi_2}\ldots p_N^{\xi_N};\end{aligned} \tag{82}$$

therefore

$$\begin{aligned}&\mathrm{E}f(\boldsymbol{\xi})\, g(\boldsymbol{\xi}) - \mathrm{E}f(\boldsymbol{\xi})\,\mathrm{E}g(\boldsymbol{\xi}) \\ &\quad = \sum_{\xi_1=0}^{n}\sum_{\xi_2=0}^{n-\xi_1}\cdots\sum_{\xi_N=0}^{n-\xi_1-\ldots-\xi_{N-1}} f(\boldsymbol{\xi})\, g(\boldsymbol{\xi}) \\ &\quad \cdot \frac{n!}{\xi_1!\xi_2!\ldots\xi_N!}p_1^{\xi_1}p_2^{\xi_2}\ldots p_N^{\xi_N} \\ &\quad - \sum_{\xi_1=0}^{n}\sum_{\xi_2=0}^{n-\xi_1}\cdots\sum_{\xi_N=0}^{n-\xi_1-\ldots-\xi_{N-1}} f(\boldsymbol{\xi})\, \frac{n!}{\xi_1!\xi_2!\ldots\xi_N!}p_1^{\xi_1}p_2^{\xi_2} \\ &\quad \ldots p_N^{\xi_N} \times \sum_{\xi_1=0}^{n}\sum_{\xi_2=0}^{n-\xi_1}\cdots\sum_{\xi_N=0}^{n-\xi_1-\ldots-\xi_{N-1}} g(\boldsymbol{\xi}) \\ &\quad \cdot \frac{n!}{\xi_1!\xi_2!\ldots\xi_N!}p_1^{\xi_1}p_2^{\xi_2}\ldots p_N^{\xi_N}.\end{aligned} \tag{83}$$

Yet (33) of Theorem II implies

$$\begin{aligned}&\left|\mathrm{E}f(\boldsymbol{\xi})\, g(\boldsymbol{\xi}) - \mathrm{E}f(\boldsymbol{\xi})\,\mathrm{E}g(\boldsymbol{\xi})\right| = \left|\mathrm{cov}\left(f(\boldsymbol{\xi}), g(\boldsymbol{\xi})\right)\right| \\ &\quad \le 2\mathfrak{f}_\infty\mathfrak{g}_\infty\mathrm{D}\xi.\end{aligned} \tag{84}$$

Combination of (83), (84), and (80) yields desired inequality (81); thus the proof is completed. □

4.3. One-Dimensional Normal Distribution. Since the normal distribution was not discussed in paper [20] of He and Wang, we shall discuss first the case $N = 1$ of a single random variable. Later in the subsequent subsection we shall extend our result to the case of two random variables ($N = 2$).

4.3.1. Definition. Assume here

(i) $\Omega = (-\infty, +\infty)$

(ii) $\xi \in \Omega$ is a one-dimensional random variable

(iii) $\mu, \sigma > 0$ are given prescribed parameters

Then

(a) the probability density function $\wp(\xi)$ is given by formula

$$\wp(\xi) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(\xi-\mu)^2/2\sigma^2} \tag{85}$$

(b) the expectation is equal to

$$\mathrm{E}\xi = \mu \tag{86}$$

(c) the variance is equal to

$$\mathrm{D}\xi = \sigma^2 \tag{87}$$

4.3.2. An Application of Theorem I

Theorem I-1. *Assume that two functions $f, g : \Omega \longrightarrow \mathbb{R}$ are continuous and differentiable in Ω. Assume also that the derivatives f', g' are bounded in Ω. Then*

$$\begin{aligned} &\left| \int_{-\infty}^{+\infty} f(\xi)\, g(\xi) \frac{1}{\sigma\sqrt{2\pi}} e^{-(\xi-\mu)^2/2\sigma^2} \mathrm{d}\xi \right. \\ &\quad - \int_{-\infty}^{+\infty} f(\xi) \frac{1}{\sigma\sqrt{2\pi}} e^{-(\xi-\mu)^2/2\sigma^2} \mathrm{d}\xi \\ &\quad \left. \cdot \int_{-\infty}^{+\infty} g(\xi) \frac{1}{\sigma\sqrt{2\pi}} e^{-(\xi-\mu)^2/2\sigma^2} \mathrm{d}\xi \right| \le 2 \left\| f' \right\|_\infty \left\| g' \right\|_\infty \\ &\quad \cdot \sigma^2. \end{aligned} \tag{88}$$

Proof. Let ξ be a random variable which possesses the normal distribution. According to (85) and (17) we have

$$\mathrm{E}f(\xi) = \int_{-\infty}^{+\infty} f(\xi) \frac{1}{\sigma\sqrt{2\pi}} e^{(\xi-\mu)^2/2\sigma^2} \mathrm{d}\xi; \tag{89}$$

therefore

$$\begin{aligned} &\mathrm{E}f(\xi)\, g(\xi) - \mathrm{E}f(\xi)\, \mathrm{E}g(\xi) \\ &\quad = \int_{-\infty}^{+\infty} f(\xi)\, g(\xi) \frac{1}{\sigma\sqrt{2\pi}} e^{-(\xi-\mu)^2/2\sigma^2} \mathrm{d}\xi \\ &\quad - \int_{-\infty}^{+\infty} f(\xi) \frac{1}{\sigma\sqrt{2\pi}} e^{-(\xi-\mu)^2/2\sigma^2} \mathrm{d}\xi \\ &\quad \cdot \int_{-\infty}^{+\infty} g(\xi) \frac{1}{\sigma\sqrt{2\pi}} e^{-(\xi-\mu)^2/2\sigma^2} \mathrm{d}\xi. \end{aligned} \tag{90}$$

Yet (1) of Theorem I implies

$$\begin{aligned} &\left| \mathrm{E}f(\xi)\, g(\xi) - \mathrm{E}f(\xi)\, \mathrm{E}g(\xi) \right| = \left| \mathrm{cov}\left(f(\xi), g(\xi)\right) \right| \\ &\quad \le 2\mathfrak{f}_\infty \mathfrak{g}_\infty \mathrm{D}\xi. \end{aligned} \tag{91}$$

Combination of (90), (91), and (87) yields desired inequality (88); thus the proof is completed. □

4.3.3. Standard Normal Distribution. The standard normal distribution corresponds to choosing $\mu = 0$ and $\sigma = 1$. Accordingly we have

$$\wp(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2}. \tag{92}$$

After substituting (92) into (88) one arrives to the following outcome.

Theorem I-2.

$$\begin{aligned} &\left| \int_{-\infty}^{+\infty} f(\xi)\, g(\xi) \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \mathrm{d}\xi \right. \\ &\quad \left. - \int_{-\infty}^{+\infty} f(\xi) \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \mathrm{d}\xi \int_{-\infty}^{+\infty} g(\xi) \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \mathrm{d}\xi \right| \\ &\quad \le 2 \left\| f' \right\|_\infty \left\| g' \right\|_\infty. \end{aligned} \tag{93}$$

4.4. Standard Bivariate Normal Distribution

4.4.1. Definition. Assume that

(i) $\Omega = \mathbb{R}^2$

(ii) $\boldsymbol{\xi} = (\xi_1, \xi_2)$ with ξ_1 and ξ_2 independent

(iii) $\phi(\xi_j)$ denotes the standard normal probability density function of ξ_j, with $j \in \{1, 2\}$:

$$\phi\left(\xi_j\right) = \frac{1}{\sqrt{2\pi}} e^{-\xi_j^2/2} \tag{94}$$

Then

(a) the probability density function $\wp(\boldsymbol{\xi})$ is given by formula

$$\wp(\boldsymbol{\xi}) = \phi\left(\xi_1\right) \phi\left(\xi_2\right) = \frac{1}{2\pi} e^{-(\xi_1^2+\xi_2^2)/2} \tag{95}$$

(b) the expectation is equal to

$$\mathrm{E}\boldsymbol{\xi} = \mathbf{0} \tag{96}$$

(c) the variance is equal to

$$\mathrm{D}\xi = \mathrm{D}\xi_1 + \mathrm{D}\xi_2 = 2 \tag{97}$$

4.4.2. An Application of Theorem II

Theorem II-4. *Assume that two functions $f, g : \Omega \longrightarrow \mathbb{R}$ are continuous and differentiable in Ω. Assume also that all the partial derivatives $\{\partial_j f\}_{j=1}^N$, $\{\partial_j g\}_{j=1}^N$ are bounded in Ω. Recalling (31) and (32), one has*

$$\begin{aligned}\Bigg|\iint f(\boldsymbol{\xi})\, g(\boldsymbol{\xi})\, \frac{1}{2\pi} e^{(\xi_1^2+\xi_2^2)/2} d\xi_1 d\xi_2 \\ - \iint f(\boldsymbol{\xi})\, \frac{1}{2\pi} e^{(\xi_1^2+\xi_2^2)/2} d\xi_1 d\xi_2 \\ \cdot \iint g(\boldsymbol{\xi})\, \frac{1}{2\pi} e^{(\xi_1^2+\xi_2^2)/2} d\xi_1 d\xi_2 \Bigg| \le 4\mathfrak{f}_\infty \mathfrak{g}_\infty.\end{aligned} \tag{98}$$

Proof. Let $\boldsymbol{\xi}$ be a 2-dimensional random variable which possesses the standard bivariate normal distribution. According to (95) and (17), one has

$$\mathrm{E} f(\boldsymbol{\xi}) = \iint f(\boldsymbol{\xi})\, \frac{1}{2\pi} e^{(\xi_1^2+\xi_2^2)/2} d\xi_1 d\xi_2; \tag{99}$$

therefore

$$\begin{aligned}&\mathrm{E} f(\boldsymbol{\xi})\, g(\boldsymbol{\xi}) - \mathrm{E} f(\boldsymbol{\xi})\, \mathrm{E} g(\boldsymbol{\xi}) \\ &\quad = \iint f(\boldsymbol{\xi})\, g(\boldsymbol{\xi})\, \frac{1}{2\pi} e^{(\xi_1^2+\xi_2^2)/2} d\xi_1 d\xi_2 \\ &\qquad - \iint f(\boldsymbol{\xi})\, \frac{1}{2\pi} e^{(\xi_1^2+\xi_2^2)/2} d\xi_1 d\xi_2 \\ &\qquad \cdot \iint g(\boldsymbol{\xi})\, \frac{1}{2\pi} e^{(\xi_1^2+\xi_2^2)/2} d\xi_1 d\xi_2.\end{aligned} \tag{100}$$

Yet (33) of Theorem II implies

$$\begin{aligned}&\left|\mathrm{E} f(\boldsymbol{\xi})\, g(\boldsymbol{\xi}) - \mathrm{E} f(\boldsymbol{\xi})\, \mathrm{E} g(\boldsymbol{\xi})\right| = \left|\mathrm{cov}\left(f(\boldsymbol{\xi}), g(\boldsymbol{\xi})\right)\right| \\ &\quad \le 2\mathfrak{f}_\infty \mathfrak{g}_\infty \mathrm{D}\xi.\end{aligned} \tag{101}$$

Combination of (100), (101), and (97) yields desired inequality (98); thus the proof is completed. □

5. Coordinate Dependence of Theorem II

Our basic relation (33) of Theorem II is formulated in terms of an N-dimensional random variable $\boldsymbol{\xi} \in \Omega \subseteq \mathbb{R}^N$. Clearly, for a given N-dimensional statistical problem there are many different (mutually equivalent) choices of N independent random variables (coordinates) $\boldsymbol{\xi}$ in terms of which the probability distribution $\wp(\boldsymbol{\xi})$ can be expressed. It is not a priori clear whether or not our basic random inequality (33) is coordinate dependent. As we show below on a concrete example, an answer is affirmative. This means also that an inequality (33) can be optimized (made stronger) via carrying out a suitable coordinate transformation.

5.1. Correlated Bivariate Normal Distribution. As an illustrative example we shall take the correlated bivariate normal distribution described in [4]. Correspondingly, we have $N = 2$, $\Omega = \mathbb{R}^2$, and $\boldsymbol{\xi} = (x, y)$. The probability distribution of $\boldsymbol{\xi}$ is characterized by formula

$$\wp(\boldsymbol{\xi}) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-(x^2-2\rho xy+y^2)/2(1-\rho^2)}, \tag{102}$$

where $\rho \in [0, 1]$ is a fixed correlation parameter.

Importantly, the above introduced correlated bivariate normal distribution can be converted into an uncorrelated normal distribution via performing a bijective coordinate transformation $\boldsymbol{\xi} = (x, y) \longleftrightarrow \boldsymbol{\eta} = (x, z)$. Here, according to [4],

$$z = \frac{y - \rho x}{\sqrt{1-\rho^2}}; \tag{103}$$

$$y = \rho x + \sqrt{1-\rho^2}\, z. \tag{104}$$

The associated probability density $\mathscr{P}(\boldsymbol{\eta})$ takes the form already discussed above in (95), i.e.,

$$\mathscr{P}(\boldsymbol{\eta}) = \frac{1}{2\pi} e^{-(x^2+z^2)/2}. \tag{105}$$

Direct calculation yields $\mathrm{D}x = \mathrm{D}z = 1$; thus $\mathrm{D}\eta = \mathrm{D}x + \mathrm{D}z = 1 + 1 = 2$. Hence also

$$\begin{aligned}\mathrm{D}\xi = \mathrm{D}x + \mathrm{D}y &= \mathrm{D}x + \mathrm{D}\left\{\rho x + \sqrt{1-\rho^2}\, z\right\} \\ &= \mathrm{D}x + \rho^2 \mathrm{D}x + \left(1-\rho^2\right)\mathrm{D}z = 2.\end{aligned} \tag{106}$$

5.2. An Application of Theorem II: Correlated Case. For the sake of maximum simplicity, we shall apply Theorem II on a concrete case of two elementary functions

$$f(x, y) = x + y; \tag{107}$$

$$g(x, y) = x - y. \tag{108}$$

Direct calculation yields

$$\begin{aligned}\|f_x\| &= \sup_{(x,y)\in\mathbb{R}^2} \left|\frac{\partial f}{\partial x}\right| = 1, \\ \|g_x\| &= \sup_{(x,y)\in\mathbb{R}^2} \left|\frac{\partial g}{\partial x}\right| = 1; \\ \|f_y\| &= \sup_{(x,y)\in\mathbb{R}^2} \left|\frac{\partial f}{\partial y}\right| = 1, \\ \|g_y\| &= \sup_{(x,y)\in\mathbb{R}^2} \left|\frac{\partial g}{\partial y}\right| = 1; \\ \mathfrak{f}_\infty &= \sqrt{\|f_x\|^2 + \|f_y\|^2} = \sqrt{2}, \\ \mathfrak{g}_\infty &= \sqrt{\|g_x\|^2 + \|g_y\|^2} = \sqrt{2}.\end{aligned} \tag{109}$$

Relation (33) of Theorem II boils then down to

$$\left|\operatorname{cov}\left(f\left(\boldsymbol{\xi}\right), g\left(\boldsymbol{\xi}\right)\right)\right| \leq 8, \tag{110}$$

where

$$\begin{aligned}\operatorname{cov}\left(f\left(\boldsymbol{\xi}\right), g\left(\boldsymbol{\xi}\right)\right) &= \mathrm{E}f\left(\boldsymbol{\xi}\right) g\left(\boldsymbol{\xi}\right) - \mathrm{E}f\left(\boldsymbol{\xi}\right) \mathrm{E}g\left(\boldsymbol{\xi}\right) \\ &= \iint f\left(x, y\right) g\left(x, y\right) \\ &\quad \cdot \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-(x^2-2\rho xy+y^2)/2(1-\rho^2)} \mathrm{d}x\,\mathrm{d}y \\ &\quad - \iint f\left(x, y\right) \\ &\quad \cdot \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-(x^2-2\rho xy+y^2)/2(1-\rho^2)} \mathrm{d}x\,\mathrm{d}y \iint g\left(x, y\right) \\ &\quad \cdot \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-(x^2-2\rho xy+y^2)/2(1-\rho^2)} \mathrm{d}x\,\mathrm{d}y.\end{aligned} \tag{111}$$

5.3. An Application of Theorem II: Uncorrelated Case. In the coordinates $\boldsymbol{\eta} = (x, z)$ introduced above by (104) we have

$$\begin{aligned}F\left(x, z\right) &= f\left(x, y\left(x, z\right)\right) = x + \left(\rho x + \sqrt{1-\rho^2} z\right) \\ &= \left(1+\rho\right) x + \sqrt{1-\rho^2} z;\end{aligned} \tag{112}$$

$$\begin{aligned}G\left(x, y\right) &= g\left(x, y\left(x, z\right)\right) = x - \left(\rho x + \sqrt{1-\rho^2} z\right) \\ &= \left(1-\rho\right) x - \sqrt{1-\rho^2} z.\end{aligned} \tag{113}$$

Direct calculation yields

$$\begin{aligned}\left\|F_x\right\| &= \sup_{(x,z)\in\mathbb{R}^2} \left|\frac{\partial F}{\partial x}\right| = 1+\rho, \\ \left\|G_x\right\| &= \sup_{(x,z)\in\mathbb{R}^2} \left|\frac{\partial G}{\partial x}\right| = 1-\rho; \\ \left\|F_z\right\| &= \sup_{(x,z)\in\mathbb{R}^2} \left|\frac{\partial F}{\partial z}\right| = \sqrt{1-\rho^2}, \\ \left\|G_z\right\| &= \sup_{(x,z)\in\mathbb{R}^2} \left|\frac{\partial G}{\partial z}\right| = \sqrt{1-\rho^2}; \\ \mathfrak{F}_\infty &= \sqrt{\left\|F_x\right\|^2 + \left\|F_z\right\|^2} = \sqrt{2}\sqrt{1+\rho}, \\ \mathfrak{G}_\infty &= \sqrt{\left\|G_x\right\|^2 + \left\|G_z\right\|^2} = \sqrt{2}\sqrt{1-\rho}.\end{aligned} \tag{114}$$

According to our above derived statement (98) of Theorem II-4, relation (33) of Theorem II boils down to

$$\left|\operatorname{cov}\left(F\left(\boldsymbol{\eta}\right), G\left(\boldsymbol{\eta}\right)\right)\right| \leq 8\sqrt{1-\rho^2}, \tag{115}$$

where

$$\begin{aligned}\operatorname{cov}\left(F\left(\boldsymbol{\eta}\right), G\left(\boldsymbol{\eta}\right)\right) &= \mathrm{E}f\left(\boldsymbol{\xi}\right) g\left(\boldsymbol{\xi}\right) - \mathrm{E}f\left(\boldsymbol{\xi}\right) \mathrm{E}g\left(\boldsymbol{\xi}\right) \\ &= \iint F\left(x, z\right) G\left(x, z\right) \frac{1}{2\pi} e^{(x^2+z^2)/2} \mathrm{d}x\,\mathrm{d}z \\ &\quad - \iint F\left(x, z\right) \frac{1}{2\pi} e^{(x^2+z^2)/2} \mathrm{d}x\,\mathrm{d}z \\ &\quad \cdot \iint G\left(x, z\right) \frac{1}{2\pi} e^{(x^2+z^2)/2} \mathrm{d}x\,\mathrm{d}z.\end{aligned} \tag{116}$$

5.4. Analysis. Functions $f(x, y)$ and $F(x, z)$ represent the same random variable, just expressed in different coordinates. This is similar for $g(x, y)$ and $G(x, z)$. If so, then covariances (111) and (116) must be equal. This can easily be verified explicitly as a consistency check. One starts with (111) and performs a substitution $(x, y) \longrightarrow (x, z)$ following (104). The corresponding Jacobian is equal to

$$J\left(x, z\right) = \begin{vmatrix} \dfrac{\partial x}{\partial x} & \dfrac{\partial y}{\partial x} \\ \dfrac{\partial x}{\partial z} & \dfrac{\partial y}{\partial z} \end{vmatrix} = \begin{vmatrix} 1 & \rho \\ 0 & \sqrt{1-\rho^2} \end{vmatrix} = \sqrt{1-\rho^2}. \tag{117}$$

Straightforward manipulations confirm then that

$$\operatorname{cov}\left(f\left(\boldsymbol{\xi}\right), g\left(\boldsymbol{\xi}\right)\right) = \operatorname{cov}\left(F\left(\boldsymbol{\eta}\right), G\left(\boldsymbol{\eta}\right)\right), \tag{118}$$

exactly as being claimed.

According to (118), the l.h.s. of (110) and (115) are equal. However, the r.h.s. of (110) differs from the r.h.s. of (115), the latter being ρ-dependent. This means in turn that statement (33) of our basic Theorem II does generally depend upon the coordinates chosen. Coordinate transformations can be thus exploited to optimize (strengthen) inequality (33).

6. Conclusion and Prospects

In summary, our basic Theorem II stated above generalizes a recently derived random inequality of [20] to the case of multidimensional random variables. Six subsequent additional results (Theorems I-1~I-2, Theorems II-1~II-4) apply then Theorem II to different frequently encountered statistical distributions (multiuniform, multinomial, normal, and multinormal). Furthermore, we show in Section 5 that basic inequality (33) of Theorem II is coordinate dependent (and thus optimizable via carrying out suitable coordinate transformations). The just mentioned formulas and insights could be useful for making estimates in multivariate statistics.

Finally, we find it useful to list below three open questions which may be worthy of examining in the future. Namely,

(i) we have assumed above that Ω is a convex open set. Any nonconvex path connected open set $\widetilde{\Omega}$ can actually be made convex via a suitable coordinate transformation. So Theorem II turns out to be even more general than stated in Section 3. It is not a priori clear, however, what would happen in the case when Ω is a disconnected open set. This remains to be seen

(ii) an optimization of inequality (33) via coordinate transformations represents a very promising direction of further research

(iii) it might be desirable to supplement some real life application which would reflect the practical value of our inequality (33). For example, one may think about applications in physics or economy

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Natural Science Foundation of Hainan Province (Grant no. 2018CXTD338; contribution rate: 30%), the National Natural Science Foundation of China (Grant no. 11761027; contribution rate: 50%), the Scientific Research Foundation of Hainan Province Education Bureau (Grant no. Hnky2016-14; contribution rate: 10%), and the Educational Reform Foundation of Hainan Province Education Bureau (Grant no. Hnjg2017ZD-13; contribution rate: 10%).

References

[1] S. Ross, *A First Course in Probability*, Pearson, London, UK, 8th edition, 2010.

[2] J. T. McClave, *Statistics*, Pearson, London, UK, 2017.

[3] W. Mendenhall, R. J. Beaver, and B. M. Beaver, *Introduction to Probability and Statistics*, Brooks/Cole Publishing Co., Pacific Grove, Calif, USA, 14th edition, 2013.

[4] J. Pitman, Probability, Springer, New York, NY, USA, 2011.

[5] M. Bilodeau and D. Brenner, *Theory of Multivariate Statistics*, Springer, New York, NY, USA, 1999.

[6] S. M. Stigler, *The History of Statistics: The Measurement of Uncertainty before 1900*, Harvard University Press, Cambridge, Mass, USA, 1990.

[7] L. Debnath and K. Basu, "A short history of probability theory and its applications," *International Journal of Mathematical Education in Science and Technology*, vol. 46, no. 1, pp. 13–39, 2015.

[8] S. E. Fienberg, "A brief history of statistics in three and one-half chapters: A review essay," *Statistical Science*, vol. 7, no. 2, pp. 208–225, 1992.

[9] S. M. Ross, *Introduction to Probability and Statistics for Engineers and Scientists*, Academic Press, San Diego, Calif, USA, 5th edition, 2014.

[10] T. T. Soong, *Fundamentals of Probability and Statistics for Engineers*, Wiley, England, UK, 2004.

[11] T. Z. Fahidy, "Basic applications of the analysis of variance and covariance in electrochemical science and engineering," in *Modern Aspects of Electrochemistry*, vol. 40, pp. 37–74, Springer, New York, NY, USA, 2007.

[12] G. G. Vining and S. Kowalski, *Statistical Methods for Engineers*, Centage Learning, Boston, Mass, USA, 3rd edition, 2011.

[13] T. Coffey and H. Yang, *Statistics for Biotechnology Process Development*, Chapman and Hall/CRC, Boca Raton, Fla, USA, 2018.

[14] J. C. Bailar and D. C. Hoaglin, *Medical Uses of Statistics*, Wiley, Hoboken, NJ, USA, 3rd edition, 2009.

[15] T. Park and Y. J. Lee, "Covariance models for nested repeated measures data: analysis of ovarian steroid secretion data," *Statistics in Medicine*, vol. 21, no. 1, pp. 143–164, 2002.

[16] N. Balakrishnan, *Methods and Applications of Statistics in Business, Finance, and Management Science*, Wiley, Hoboken, NJ, USA, 2010.

[17] M. J. Cloud, B. C. Drachman, and L. Lebedev, *Inequalities: With Applications to Engineering*, Springer, New York, NY, USA, 2014.

[18] D. S. Mitrinovic, J. Pecaric, and A. M. Fink, *Classical and New Inequalities in Analysis*, Springer, New York, NY, USA, 1993.

[19] G. A. Anastassiou, *Advanced Inequalities*, World Scientific Publishing, Singapore, 2011.

[20] Z. He and M. Wang, "An inequality for covariance with applications," *Journal of Inequalities and Applications*, vol. 413, 2015.

[21] G. Anastassiou, "On Grüss type multivariate integral inequalities," *Mathematica Balkanica*, vol. 17, no. 1-2, 2003.

[22] G. A. Anastassiou, "Multivariate Chebyshev-Grüss and comparison of integral means type inequalities via a multivariate Euler type identity," *Demonstratio Mathematica*, vol. 40, no. 3, pp. 537–558, 2007.

[23] M. W. Alomari, "New Grüss type inequalities for double integrals," *Applied Mathematics and Computation*, vol. 228, pp. 102–107, 2014.

[24] A. Florea and C. P. Niculescu, "A note on Ostrowskis inequality," *Journal of Inequalities and Applications*, vol. 2005, no. 5, Article ID 459, 2005.

[25] N. S. Barnett and S. S. Dragomir, "An inequality of Ostrowski's type for cumulative distribution functions," *Kyungpook Mathematical Journal*, vol. 39, no. 2, pp. 303–311, 1999.

[26] N. S. Barnett, P. Cerone, and S. S. Dragomir, "Inequalities for Random Variables Over a Finite Interval," in *RGMIA Monographs*, Victoria University, Melbourne, Australia, 2004.

[27] B. G. Pachpatte, "On Grüss type inequalities for double integrals," *Journal of Mathematical Analysis and Applications*, vol. 267, no. 2, Article ID 454, 2002.

[28] B. G. Pachpatte, "On multidimensional Grüss type inequalities," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 3, no. 2, 2002.

[29] X. Li, R. N. Mohapatra, and R. S. Rodriguez, "Grüss-type inequalities," *Journal of Mathematical Analysis and Applications*, vol. 267, no. 2, pp. 434–443, 2002.

[30] R. P. Agarwal, N. S. Barnett, P. Cerone, and S. S. Dragomir, "A survey on some inequalities for expectation and variance," *Computers & Mathematics with Applications*, vol. 49, no. 2-3, pp. 429–480, 2005.

[31] F. Ahmad, N. S. Barnett, and S. S. Dragomir, "New weighted Ostrowski and Chebyshev type inequalities," *Nonlinear Analysis*, vol. 71, no. 12, pp. e1408–e1412, 2009.

[32] P. Cerone and S. S. Dragomir, "Bounding the Chebyshev functional for the Riemann-Stieltjes integral via a Beesack

inequality and applications," *Computers & Mathematics with Applications*, vol. 58, no. 6, pp. 1247–1252, 2009.

[33] G. Farid, "Straightforward Proofs of Ostrowski Inequality and Some Related Results," *International Journal of Analysis*, vol. 2016, Article ID 3918483, 5 pages, 2016.

[34] W. Liu, A. Tuna, and Y. Jiang, "On weighted Ostrowski type, Trapezoid type, Grüss type and Ostrowski-Grüss like inequalities on time scales," *Applicable Analysis: An International Journal*, vol. 93, no. 3, pp. 551–571, 2014.

[35] A. Tuna and D. Daghan, "Generalization of Ostrowski and Ostrowski-Grüss type inequalities on time scales," *Computers & Mathematics with Applications*, vol. 60, no. 3, pp. 803–811, 2010.

[36] A. Tuna, Y. Jiang, and W. Liu, "Weighted Ostrowski, Ostrowski-Grüss and Ostrowski-Chebyshev type inequalities on time scales," *Publicationes Mathematicae Debrecen*, vol. 81, no. 1-2, pp. 81–102, 2012.

[37] B. Zheng and Q. Feng, "Generalized n-dimensional Ostrowski type and Grüss type inequalities on time scales," *Journal of Applied Mathematics*, vol. 2014, Article ID 434958, 11 pages, 2014.

[38] J. Conway, *A Course in Point Set Topology*, Springer International Publishing, Switzerland, 2014.

[39] S. Lang, *Undergraduate Analysis*, Springer-Verlag, New York, NY, USA, 1983.

[40] M. Comenetz, *Calculus: The elements*, World Scientific, Singapore, 2002.

[41] K. Chung and F. AitSahlia, *Elementary Probability Theory*, Springer, New York, NY, USA, 2011.