Theory Probab. Appl.
Vol. 42, No. 1

VARIANCE INEQUALITIES FOR COVARIANCE KERNELS AND APPLICATIONS TO CENTRAL LIMIT THEOREMS*

T. CACOULLOS†, N. PAPADATOS†, and V. PAPATHANASIOU†

Abstract. A simple estimate for the error in the CLT, valid for a wide class of absolutely continuous r.v.'s, is derived without Fourier techniques. This is achieved by using a simple convolution inequality for the variance of *covariance kernels* or w-functions in conjunction with bounds for the total variation distance. The results are extended to the multivariate case. Finally, a simple proof of the classical Darmois–Skitovich characterization of normality is obtained.

Key words. convolution inequality, covariance kernels, CLT, rate of convergence, characterization of normality

PII. S0040585X97976039

1. Introduction and summary. Let X be a normalized (absolutely) continuous random variable (r.v.) with a distribution function (d.f.) F and density f, and let Z be a standard normal r.v. with a d.f. Φ. Recently, Cacoullos, Papathanasiou, and Utev (henceforth, CPU) [8] obtained (inter alia) the following bound for the total variation distance (TVD) $\rho(F,\Phi)=\sup\{|F(A)-\Phi(A)|, A \text{ Borel}\}$ between F and Φ (or X and Z), namely,

$$\rho(F,\,\Phi)\leqq 2\mathbf{E}\big|w(X)-1\big|\leqq 2\sqrt{\operatorname{Var}[w(X)]}, \tag{1.1}$$

where the (characterizing) covariance kernel $w(\cdot)$ is defined for every x in the interval support of X by the relation

$$w(x)f(x)=-\int_{-\infty}^{x} tf(t)\,dt=\int_{x}^{\infty} t\,f(t)\,dt. \tag{1.2}$$

Note that the second inequality in (1.1) follows from $(\mathbf{E}|w-1|)^2\leqq \mathbf{E}(w-1)^2$ and the fact that, for any r.v. X with finite second moment, $\mathbf{E}[w(X)]=1$, as easily verified from (1.2) or the basic *covariance identity* (see [5])

$$\operatorname{cov}\big[X,\,g(X)\big]=\mathbf{E}Xg(X)=\mathbf{E}\big[w(X)\,g'(X)\big] \tag{1.3}$$

for any absolutely continuous g. Furthermore, by using (1.1) and the law of large numbers, CPU proved a strong version (L^1 convergence of the densities) of the central limit theorem (CLT).

More recently, Papathanasiou [14] extended the above results for a standardized continuous random vector $\mathbf{X}=(X_1,\dots,X_p)'$ with d.f. F, by showing that under appropriate conditions

$$\rho(F,\,\Phi_p)\leqq 2\sum_{i=1}^{p}\mathbf{E}\big|w^i(\mathbf{X})-1\big|\leqq 2\sum_{i=1}^{p}\sqrt{\operatorname{Var}[w^i(\mathbf{X})]}, \tag{1.4}$$

*Received by the editors March 15, 1996. This work was partially supported by the Ministry of Industry, Energy and Technology of Greece under grant 1369.

http://www.siam.org/journals/tvp/42-1/97603.html

†University of Athens, Department of Mathematics, Panepistemiopolis, 157 84 Athens, Greece.

where $\Phi_p(x_1,\dots,x_p) = \Phi(x_1)\cdots\Phi(x_p)$ is the d.f. of p independent standard normal r.v.'s and $\mathbf{w} = (w^1,\dots,w^p)'$: $\mathbf{R}^p \longrightarrow \mathbf{R}^p$ is the p-dimensional covariance kernel associated with the random vector $\mathbf{X}$ (for the definition, see [6]). The strengthened multivariate CLT was also proved under the assumption that $\mathbf{E}[w^i(\mathbf{X})]^2 < \infty$ for $i = 1,\dots,p$.

The key role to the derivation of the above limit theorems is based on (1.1), (1.4), and the following theorem.

THEOREM 1.1 [7]. *Let $X_1, X_2, \dots, X_n, \dots$ be independent identically distributed (i.i.d.) absolutely continuous r.v.'s with*

$$\mathbf{E}[X_1] = 0, \quad \mathbf{E}[X_1^2] = 1 \quad \textit{and} \quad \mathbf{E}\big[w^2(X_1)\big] < \infty,$$

where w is the covariance kernel of X_1. Then,

$$\operatorname{Var}\big[w_n(S_n)\big] \longrightarrow 0 \quad \textit{as } n, \tag{1.5}$$

where $S_n = (X_1 + \cdots + X_n)/\sqrt{n}$ and w_n is the w-function of S_n (for a proof of the multivariate analogue of (1.5), see [6]).

In the present paper, exploiting the bound (1.1) of TVD, we derive an appropriate convolution inequality (Lemma 2.1) for the variance of the w-function, perhaps of independent interest. This is used to provide not only a very simple proof of the CLT (i.e., Theorem 1.1), but also to establish the corresponding rate of convergence (Theorem 2.1), of order (at least) $n^{-1/2}$ for the i.i.d. case. It should be noted that such rates of convergence were independently given by Sirazhdinov and Mamatov [17] under the assumption of finiteness of the third moment (see also [15]). Also, for Δ_n, defined by

$$\Delta_n = \sup_x \big|F_n(x) - \Phi(x)\big|, \tag{1.6}$$

the Berry–Esseen type bounds are, of course, well known; so are those concerning D_n, where

$$D_n = \sup\big\{\big|F_n(C) - \Phi(C)\big|,\ C \text{ convex}\big\} \tag{1.7}$$

(see, for example, [2], [3], [4], [9], [13], [18], and references therein). As regards the convergence of ρ_n to zero, where

$$\rho_n = \rho(F_n, \Phi) = \sup\big\{\big|F_n(A) - \Phi(A)\big|,\ A \text{ Borel}\big\} \tag{1.8}$$

(here and everywhere in this paper, F_n is the d.f. of the standardized sum S_n of n independent r.v.'s and $F_X(A) = \mathbf{P}\{X \in A\}$), the known results are those of Prokhorov [16], Barron [1], Mayer-Wolf [12], and CPU [8], without establishing rates of convergence. Note also the complications involved in establishing Sirazhdinov and Mamatov's results through Fourier techniques, and also Barron's and Mayer-Wolf's results through entropy and information inequalities, respectively, in contrast to the present simplifying approach.

Furthermore, the results are easily extended to the case of nonidentically distributed r.v.'s (Theorem 2.2), as well as in the multivariate case (Theorems 3.1, 3.2). Moreover, an application of the above convolution inequality to the classical Darmois–Skitovich characterization of normality, via the independence of two different linear forms of independent r.v.'s, is given in section 4 (Theorem 4.1). Another use of the present TVD bound (2.1) is in proving CLT's for random sums $S_N = X_1 + \cdots + X_N$, in the usual case where N is independent of $X_1, X_2, \ldots$; this, however, is beyond the scope of the present investigation and will be the object of a separate paper.

However, the preceding simplifications and/or improvements are "paid" by the assumption that the variance of $w(X)$ is finite, which for the Pearson family entails the finiteness of the fourth moment of X. It is perhaps worth investigating the possibility of relaxing this condition.

2. Rate of convergence in the CLT. In the present paper we concentrate on the TVD ρ_n defined by (1.8), so that every result holding for this distance obviously continues to hold for D_n, as well as Δ_n. However, the results are applicable to the family $\mathbf{C}$ of r.v.'s X satisfying the following conditions:

(i) $\mathbf{E}[X] = 0$, $\mathbf{E}[X^2] = 1$ (without any loss of generality);

(ii) X is absolutely continuous with density f and its support is an interval (not necessarily finite);

(iii) the w-function of X (see (1.2)) has finite second moment:

$$\mathbf{E}\big[w^2(X)\big] < \infty.$$

Note that $\mathbf{C}$ is wide enough within the family of the standardized absolutely continuous r.v.'s, including, for example, the Pearson system of distributions with finite fourth moments, as it follows from Korwar's characterization [11] of the Pearson family by a quadratic w-function. Regarding the restriction (ii) of interval support, this does not essentially affect the validity of the results, since for large n (depending on the gaps where $f(x) = 0$), S_n attains an interval support.

The main result of this section is stated in the following theorem.

THEOREM 2.1. *Suppose $X_1, \ldots, X_n, \ldots$ are i.i.d. r.v.'s with $X_1 \in \mathbf{C}$. Then,*

$$\rho_n \leqq \frac{c}{\sqrt{n}}, \tag{2.1}$$

where the constant c can be taken as

$$c = 2\sqrt{\operatorname{Var}[w_1(X_1)]} \tag{2.2}$$

and w_1 is the w-function associated with X_1.

The proof of Theorem 2.1 is based on the convolution inequality given by the following lemma.

LEMMA 2.1. *Let X, Y be independent r.v.'s with $X \in \mathbf{C}$, $Y \in \mathbf{C}$ and consider the r.v. $S = aX + bY$, where a, b are real constants such that $a^2 + b^2 = 1$. Then, $S \in \mathbf{C}$ and*

$$\operatorname{Var}\big[w_S(S)\big] \leqq a^4 \operatorname{Var}\big[w_X(X)\big] + b^4 \operatorname{Var}\big[w_Y(Y)\big], \tag{2.3}$$

where w_X, w_Y, w_S are the w-functions of X, Y, S, respectively.

Proof. First observe that S has a density with interval support, mean zero, and variance one, so that the function w_S is well defined by (1.2). From (1.3), for any absolutely continuous function g, the following covariance identity holds (c.f. Stein's identity for a normal r.v.):

$$\mathbf{E}\big[Sg(S)\big] = \mathbf{E}\big[w_S(S)\, g'(S)\big], \tag{2.4}$$

where g' is the derivative of g.

On the other hand,

$$\begin{aligned}\mathbf{E}\big[Sg(S)\big] &= \mathbf{E}\big[(aX+bY)g(aX+bY)\big]\\ &= a\mathbf{E}\Big\{\mathbf{E}\big[Xg(aX+bY)\mid Y\big]\Big\} + b\mathbf{E}\Big\{\mathbf{E}\big[Yg(aX+bY)\mid X\big]\Big\}\\ &= a^2\mathbf{E}\big[w_X(X)\,g'(S)\big] + b^2\mathbf{E}\big[w_Y(Y)\,g'(S)\big],\end{aligned}$$

so that we have

$$\mathbf{E}\big[w_S(S)\,g'(S)\big] = \mathbf{E}\big[\big(a^2w_X(X)+b^2w_Y(Y)\big)\,g'(S)\big]. \tag{2.5}$$

Applying (2.5) to $g'(x) = w_S(x)\,I\{w_S(x) \leqq N\}$ (a bounded function), we have

$$\begin{aligned}\Big(\mathbf{E}\big[w_S^2(S)\,I\{\,w_S(S)\leqq N\}\big]\Big)^2 &= \Big(\mathbf{E}\big[\big(a^2w_X(X)+b^2w_Y(Y)\big)\,w_S(S)\,I\{\,w_S(S)\leqq N\}\big]\Big)^2\\ &\leqq \mathbf{E}\big[\big(a^2w_X(X)+b^2w_Y(Y)\big)^2\big]\,\mathbf{E}\big[w_S^2(S)\,I\{\,w_S(S)\leqq N\}\big]\end{aligned}$$

by the Cauchy–Schwarz inequality. Thus,

$$\mathbf{E}\big[w_S^2(S)\,I\{\,w_S(S)\leqq N\}\big] \leqq \mathbf{E}\big[\big(a^2w_X(X)+b^2w_Y(Y)\big)^2\big], \tag{2.6}$$

which, taking the limit as $N\to\infty$, implies that $\mathbf{E}[w_S^2(S)] < \infty$, and hence $S \in \mathbf{C}$. Now, (2.6) leads to

$$\mathbf{E}\big[w_S^2(S)\big] \leqq \mathbf{E}\big[\big(a^2w_X(X)+b^2w_Y(Y)\big)^2\big], \tag{2.7}$$

which is equivalent to (2.3), since $\mathbf{E}[w_X(X)] = 1$ for any r.v. X.

Remark 2.1. Obviously Lemma 2.1 also holds when $S = \sum_{i=1}^n a_iX_i$ with $\sum_{i=1}^n a_i^2 = 1$:

$$\operatorname{Var}\big[w_S(S)\big] \leqq \sum_{i=1}^n a_i^4\operatorname{Var}\big[w_{X_i}(X_i)\big].$$

In the following lemma we give a slight variation of the above inequality which in turn characterizes normality within the class $\mathbf{C}$.

LEMMA 2.2. *Under the conditions of Lemma* 2.1,

$$\mathbf{E}\big[w_S^2(S)\big] \leqq a^2\mathbf{E}\big[w_X^2(X)\big] + b^2\mathbf{E}\big[w_Y^2(Y)\big], \tag{2.8}$$

with equality if and only if X, Y, *and* S *are standard normal*, *provided* $ab \neq 0$.

Proof. It is easily seen that

$$\begin{aligned}&a^2\mathbf{E}\big[w_X^2(X)\big] + b^2\mathbf{E}\big[w_Y^2(Y)\big] - \mathbf{E}\big[\big(a^2w_X(X)+b^2w_Y(Y)\big)^2\big]\\ &\quad= a^2b^2\big\{\mathbf{E}\big[w_X^2(X)\big]+\mathbf{E}\big[w_Y^2(Y)\big]-2\big\} = a^2b^2\big\{\operatorname{Var}\big[w_X(X)\big]+\operatorname{Var}\big[w_Y(Y)\big]\big\} \geqq 0\end{aligned}$$

and the assertion follows from (2.7).

Obviously the equality in (2.8) holds when $ab = 0$ or when X and Y are both normal; as for the "only if" part, observe that the equality in (2.8) implies the equality in (2.7). Hence, if $ab \neq 0$ we must have $\operatorname{Var}[w_X(X)] = \operatorname{Var}[w_Y(Y)] = 0$ by the above argument, which in turn implies that both X, Y, (and S) are normal (see [7, Characterization 4]).

Remark 2.2. From the above proof it is clear that (2.3) (or (2.7)) is much stronger than (2.8), since the crucial factors a^4 and b^4 become a^2 and b^2, respectively. (This is clearly seen in the proof of Corollary 2.1 below, where the induction argument could not apply if the former would be replaced by the latter.) It should be noted, however, that the characterization result is lost in (2.3), since the equality is also attained if, e.g., both X, Y are i.i.d. *exponential* random variables and $a = b = \frac{\sqrt{2}}{2}$.

COROLLARY 2.1. *If $X_1, \ldots, X_n, \ldots$ are i.i.d. r.v.'s such that $X_1 \in \mathbf{C}$, then $S_n = (X_1 + \cdots + X_n)/\sqrt{n} \in \mathbf{C}$. Furthermore, for the sequence $\sigma_n = \mathrm{Var}[w_n(S_n)]$ we have*

$$\sigma_n \leqq \frac{\sigma_1}{n}, \tag{2.9}$$

where w_n is the w-function of S_n.

Proof. The assertion follows immediately either from Remark 2.1 with $a_1 = \cdots = a_n = 1/\sqrt{n}$ or from (2.3) with $X = S_{n-1}$, $Y = X_n$, $S = S_n$, $a = \sqrt{(n-1)/n}$, $b = 1/\sqrt{n}$, using induction on n, since

$$S_n = \sqrt{\frac{n-1}{n}}\, S_{n-1} + \sqrt{\frac{1}{n}}\, X_n, \quad n = 1, 2, \ldots .$$

Remark 2.3. Corollary 2.1 gives a direct proof of Theorem 1.1 above (c.f. [7, Theorem 4]).

Proof of Theorem 2.1. From (1.1) we have $\rho_n \leqq 2\sqrt{\sigma_n}$, while from Corollary 2.1 $\sigma_n \leqq \sigma_1/n$, which completes the proof.

The above results can be easily extended to the general case where the distributions of the X's are not identical, as shown by the following theorem.

THEOREM 2.2. *Suppose that the independent r.v.'s $X_1, \ldots, X_n, \ldots$ are in $\mathbf{C}$. Then,*

$$\rho_n \leqq 2\,\frac{C_n}{\sqrt{n}},$$

where C_n can be taken as

$$C_n = \Big(\max_{1 \;\; k \;\; n} \big\{ \mathrm{Var}\big[w_{X_k}(X_k)\big] \big\} \Big)^{1/2}.$$

Proof. From Remark 2.1 it follows that

$$\mathrm{Var}\big[w_{S_n}(S_n)\big] \leqq \frac{1}{n^2} \sum_{k=1}^{n} \mathrm{Var}\big[w_{X_k}(X_k)\big] \leqq \frac{1}{n}\, C_n^2$$

and the assertion follows from (1.1).

Therefore, $C_n = o(\sqrt{n})$ as $n \to \infty$ provides a sufficient condition for the convergence (in total variation) of S_n to the standard normal, while the assumption that C_n remains bounded leads to a rate of convergence of order $n^{-1/2}$.

3. Multivariate extensions. The results can be easily extended in $\mathbf{R}^p$ in an obvious manner. Here, the subfamily $\mathbf{C}$ consists of the continuous random vectors $\mathbf{X}$ with mean zero and dispersion matrix I_p (the $p \times p$ identity matrix), such that $\mathbf{E}[w^i(\mathbf{X})]^2 < \infty$, $i = 1, \ldots, p$.

By using similar arguments and the covariance identities of [6] in combination with the multivariate result of [14] (see (1.4) above), one can easily establish the multivariate analogues of Lemma 2.1, Theorems 2.1, 2.2, and Corollary 2.1 as well. We, therefore, formulate without proof the following results.

LEMMA 3.1. *Let* $\mathbf{X}$, $\mathbf{Y}$ *be independent random vectors in* $\mathbf{C}$ *and consider the random vector* $\mathbf{S} = a\mathbf{X} + b\mathbf{Y}$, *where* a, b *are real scalars such that* $a^2 + b^2 = 1$. *Then,* $\mathbf{S} \in \mathbf{C}$ *and*

$$\mathrm{Var}\big[w^i_{\mathbf{S}}(\mathbf{S})\big] \leqq a^4 \mathrm{Var}\big[w^i_{\mathbf{X}}(\mathbf{X})\big] + b^4 \mathrm{Var}\big[w^i_{\mathbf{Y}}(\mathbf{Y})\big], \quad i = 1, \ldots, p,$$

where $w^i_{\mathbf{X}}$, $w^i_{\mathbf{Y}}$, $w^i_{\mathbf{S}}$ *are the ith components of the* $\mathbf{w}$*-functions of* $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{S}$, *respectively.*

COROLLARY 3.1. *If* $\mathbf{X}_1, \ldots, \mathbf{X}_n, \ldots$ *are i.i.d. random vectors such that* $\mathbf{X}_1 \in \mathbf{C}$, *then* $\mathbf{S}_n = (\mathbf{X}_1 + \cdots + \mathbf{X}_n)/\sqrt{n} \in \mathbf{C}$. *Moreover, for the sequences* $\sigma^i_n = \mathrm{Var}[w^i_n(\mathbf{S}_n)]$ *we have*

$$\sigma^i_n \leqq \frac{\sigma^i_1}{n}, \quad i = 1, \ldots, p,$$

where $\mathbf{w}_n = (w^1_n, \ldots, w^p_n)'$ *is the* $\mathbf{w}$*-function of* $\mathbf{S}_n$.

THEOREM 3.1. *Under the conditions of Corollary* 3.1 *we have*

$$\rho_n \leqq \frac{c}{\sqrt{n}},$$

where the constant c *can be taken as*

$$c = 2 \sum_{i=1}^{p} \sqrt{\mathrm{Var}[w^i_1(\mathbf{X}_1)]}$$

and $\mathbf{w}_1 = (w^1_1, \ldots, w^p_1)'$ *is the* $\mathbf{w}$*-function associated with* $\mathbf{X}_1$.

Finally, the above theorem can be easily extended to the following theorem.

THEOREM 3.2. *Suppose that the independent random vectors* $\mathbf{X}_1, \ldots, \mathbf{X}_n, \ldots$ *are in* $\mathbf{C}$. *Then,*

$$\rho_n \leqq 2\,\frac{C_n}{\sqrt{n}},$$

where C_n *can be taken as*

$$C_n = \sum_{i=1}^{p} \Big(\max_{1 \le k \le n} \big\{ \mathrm{Var}\big[w^i_{\mathbf{X}_k}(\mathbf{X}_k)\big] \big\} \Big)^{1/2}.$$

4. A simple proof of a classical characterization for the normal distribution. Several, more or less complicated, proofs have been given for the well-known Darmois–Skitovich theorem, characterizing the normal distribution by the independence of two linear forms in n independent r.v.'s (see [10] and references therein). Itoh gave a complete proof for $n = 2$ via the convolution inequality for the Fisher information. Here, by using the convolution inequality (2.3), a relatively simple proof of the same result is given under stronger conditions, namely for the family $\mathbf{C}$.

THEOREM 4.1. *Let the independent r.v.'s* X_1, $X_2 \in \mathbf{C}$ *and suppose that the r.v.'s* Y_1, Y_2 *are independent, where*

$$Y_1 = X_1 \cos\theta + X_2 \sin\theta, \qquad Y_2 = -X_1 \sin\theta + X_2 \cos\theta,$$

for some θ *which is not a multiple of* $\pi/2$. *Then, the r.v.'s* X_1, X_2, Y_1, Y_2 *are standard normal.*

Proof. It follows from Lemma 2.1 that Y_1, $Y_2 \in \mathbf{C}$. Hence, from (2.3) we conclude that

$$\mathrm{Var}\big[w_{Y_1}(Y_1)\big] \leqq \cos^4\theta\,\mathrm{Var}\big[w_{X_1}(X_1)\big] + \sin^4\theta\,\mathrm{Var}\big[w_{X_2}(X_2)\big],$$
$$\mathrm{Var}\big[w_{Y_2}(Y_2)\big] \leqq \sin^4\theta\,\mathrm{Var}\big[w_{X_1}(X_1)\big] + \cos^4\theta\,\mathrm{Var}\big[w_{X_2}(X_2)\big];$$

thus,

$$\operatorname{Var}\left[w_{Y_1}(Y_1)+w_{Y_2}(Y_2)\right] \leqq (\cos^4\theta+\sin^4\theta)\operatorname{Var}\left[w_{X_1}(X_1)+w_{X_2}(X_2)\right]. \tag{4.1}$$

On the other hand, since

$$X_1 = Y_1\cos\theta - Y_2\sin\theta, \qquad X_2 = Y_1\sin\theta + Y_2\cos\theta,$$

we conclude, using similar arguments, that

$$\operatorname{Var}\left[w_{X_1}(X_1)+w_{X_2}(X_2)\right] \leqq (\cos^4\theta+\sin^4\theta)\operatorname{Var}\left[w_{Y_1}(Y_1)+w_{Y_2}(Y_2)\right]. \tag{4.2}$$

Therefore,

$$\operatorname{Var}\left[w_{X_1}(X_1)\right] = \operatorname{Var}\left[w_{X_2}(X_2)\right] = \operatorname{Var}\left[w_{Y_1}(Y_1)\right] = \operatorname{Var}\left[w_{Y_2}(Y_2)\right] = 0,$$

and the desired result follows from Characterization 4 in [7].

Note that one could equally apply Lemma 2.2 in the above proof, using exactly the arguments of Itoh.

REFERENCES

[1] A. Barron, *Entropy and the central limit theorem*, Ann. Probab., 14 (1986), pp. 336–342.

[2] R. N. Bhattacharya, *On the errors of normal approximation*, Ann. Probab., 3 (1975), pp. 815–828.

[3] R. N. Bhattacharya and R. Ranga Rao, *Normal Approximation and Asymptotic Expansion*, Krieger, Melbourne, FL, 1986.

[4] E. Bolthausen, *An estimate of the remainder in a combinatorial central limit theorem*, Z. Wahrscheinlichkeitstheorie verw. Gebiete, 66 (1984), pp. 379–386.

[5] T. Cacoullos and V. Papathanasiou, *Characterizations of distributions by variance bounds*, Statist. Probab. Lett., 7 (1989), pp. 351–356.

[6] T. Cacoullos and V. Papathanasiou, *Lower variance bounds and a new proof of the central limit theorem*, J. Multivariate Anal., 43 (1992), pp. 173–184.

[7] T. Cacoullos, V. Papathanasiou, and S. Utev, *Another characterization of the normal law and a proof of the central limit theorem connected with it*, Theory Probab. Appl., 37 (1992), pp. 581–588.

[8] T. Cacoullos, V. Papathanasiou, and S. Utev, *Variational inequalities with examples and an application to the central limit theorem*, Ann. Probab., 22 (1994), pp. 1607–1618.

[9] F. Götze, *On the rate of convergence in the multivariate CLT*, Ann. Probab., 19 (1991), pp. 724–739.

[10] Y. Itoh, *An application of the convolution inequality for the Fisher information*, Ann. Inst. Statist. Math., 41 (1989), pp. 9–12.

[11] R. M. Korwar, *On characterizations of distributions by mean absolute deviation and variance bounds*, Ann. Inst. Statist. Math., 43 (1991), pp. 287–295.

[12] E. Mayer-Wolf, *The Cramer–Rao functional and limiting laws*, Ann. Probab., 18 (1990), pp. 840–850.

[13] R. Michel, *Generalization and application of a result of C. C. Heyde*, Ann. Probab., 10 (1982), pp. 1066–1068.

[14] V. Papathanasiou, *Multivariate variational inequalities and the central limit theorem*, to appear.

[15] V. V. Petrov, *Sums of Independent Random Variables*, Springer-Verlag, Berlin, New York, 1975.

[16] Yu. V. Prokhorov, *On a local limit theorem for densities*, Dokl. Akad. Nauk SSSR, 83 (1952), pp. 797–800 (in Russian).

[17] S. Kh. Sirazhdinov and M. Mamatov, *On convergence in the mean for densities*, Theory Probab. Appl., 7 (1962), pp. 424–429.

[18] T. J. Sweeting, *Speed of convergence for the multidimensional central limit theorem*, Ann. Probab., 5 (1977), pp. 28–41.