

An extension of the Darmois–Skitovitch theorem to a class of dependent random variables

Abram Kagan^{a,1}, Jacek Wesołowski^{b,*}

^a*Department of Mathematics, University of Maryland, College Park, MD 20742, USA*

^b*Mathematical Institute, Warsaw University of Technology, 00-661 Warsaw, Plac Politechniki 1, Poland*

Received September 1997; received in revised form May 1999

Abstract

Linear transformation of a factorizable distribution is a product measure iff it is Gaussian under some natural assumptions. The result extends the classical Darmois–Skitovitch theorem. © 2000 Elsevier Science B.V. All rights reserved

Keywords: Multivariate Gaussian distribution; Factorizable distribution; Characteristic function; Darmois–Skitovitch theorem; Characterizations of probability distributions

1. Introduction

The following remarkable result was proved independently in Darmois (1953) and Skitovitch (1953) and is now known as the Darmois–Skitovitch theorem (DST).

Theorem 1. *Let X_1, \dots, X_n be independent random variables. If linear forms $L_1 = a_1X_1 + \dots + a_nX_n$ and $L_2 = b_1X_1 + \dots + b_nX_n$ are independent then for each index i ($i = 1, \dots, n$) for which $a_ib_i \neq 0$, X_i is normal.*

This theorem stimulated much research aimed at extending the DST in different directions such as: (i) considering more abstract structures for the X 's while preserving independence relations (see, e.g., Ghurye and Olkin, 1962; Krakowiak, 1985; Feldman, 1988); (ii) replacing independence of linear forms by constancy of regression conditions (see, e.g., Lukacs and King, 1954; Laha, 1956; Kagan and Zinger, 1985); (iii) analytical weakening of the independence condition on the linear forms (Kagan, 1987, 1988a,b); (iv) replacing X_1, \dots, X_n with a stochastic process (see Plucińska and Wesołowski, 1995).

* Corresponding author.

¹ The main idea of this paper arose while the first author was visiting Warsaw University of Technology in October 1995.

In this paper, while preserving independence of linear forms, a special form of dependence for X_1, X_2, \dots, X_n is considered. It is based on factorization properties of the characteristic function (ch. f.) of a random vector which was introduced in Kagan (1988b):

Definition 1. A random vector $\mathbf{X}=(X_1, \dots, X_n)$ (or its distribution) belongs to the class $\mathcal{D}_{n,k}$ (is k -factorizable) iff its ch. f. ϕ has the form

$$\phi(t_1, \dots, t_n) = \prod_{1 \leq i_1 < \dots < i_k \leq n} R_{i_1 \dots i_k}(t_{i_1}, \dots, t_{i_k}) \quad (1)$$

for any $(t_1, \dots, t_n) \in \mathbb{R}^n$, where $R_{i_1 \dots i_k}$ is a continuous complex-valued function such that $R_{i_1 \dots i_k}(0, \dots, 0) = 1$ for any $1 \leq i_1 < \dots < i_k \leq n$.

The random vector \mathbf{X} (or its distribution) belongs to the class $\mathcal{D}_{n,k}(\text{loc})$ (is locally k -factorizable) if the representation (1) holds in some neighborhood of the origin.

Properties and examples of $\mathcal{D}_{n,k}$ measures were thoroughly studied in Section 2 of Kagan (1988b). (Observe that 1-factorizability means independence of components and in this sense the above definition extends the concept of independence.) The main result of that paper was the following version of the DST:

Theorem 2. Let (L_1, \dots, L_m) denote the vector of linear forms in independent random variables X_1, \dots, X_n . That is,

$$L_j = a_{j1}X_1 + \dots + a_{jn}X_n, \quad j = 1, \dots, m.$$

If the joint distribution of (L_1, \dots, L_m) belongs to the class $\mathcal{D}_{m,m-1}(\text{loc})$ then for each index k ($k = 1, \dots, n$) for which $a_{1k} \dots a_{mk} \neq 0$, (i.e., X_k enters each of the forms) X_k has a normal distribution.

A group theoretic version of this theorem has been given by Lisyanoy (1995). The study of factorizable measures has been continued in Wesolowski (1991a,b,1994,1997) involving, for example, the ch. f. representation, relationships to Gaussian conditional structure, the central limit problem, and decomposition questions.

Here a kind of a dual version of Theorem 2 is considered. While keeping the linear forms independent, the joint distribution of the X 's is assumed to be factorizable. As a consequence, normality of the linear forms but not of the parent X 's is obtained. This is partially due to the fact that the Cramér decomposition theorem may not hold for $\mathcal{D}_{n,k}$ classes in general.

2. The DST for factorizable measures

Consider a k -factorizable random vector \mathbf{X} , and denote by \mathbf{L} a vector of independent linear forms in its components. It turns out that under rather general assumptions the linear forms themselves are necessarily Gaussian. Unfortunately, it does not allow one to conclude that even the univariate marginals of the \mathbf{X} are Gaussian except for the case of $k = 1$, i.e., independence.

The main result is contained in:

Theorem 3. Let $n > 1$. Assume that $\mathbf{X} \in \mathcal{D}_{n,k}$ for some k , $1 \leq k \leq n$, and $\mathbf{L} = (L_1, \dots, L_m) = \mathbf{A}\mathbf{X} \in \mathcal{D}_{m,1}$, where $\mathbf{A} = [a_{ij}]$ is an $m \times n$ real matrix, for some m , $k + 1 \leq m \leq n$.

Then L_j is a normal random variable if $\mathbf{e}_j \in \text{span} \{\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_k}\}$ for any $1 \leq j_1 < \dots < j_k \leq n$, where $\mathbf{a}_l = (a_{1l}, \dots, a_{ml})$ is the l th column of the matrix \mathbf{A} , $l = 1, \dots, n$, and $\mathbf{e}_j = (0, \dots, 0, 1, 0, \dots, 0)$ is the j th element of the canonical basis in \mathbb{R}^m , $j = 1, \dots, m$.

Before we give the proof of this result, let us emphasize that normality of linear forms in k -factorizable random variables does not imply that the X_k 's are also normal. This follows from the more general considerations in Wesolowski (1997). An explicit and easy example is the following: consider Gaussian random vector (Y_1, Y_2, Y_3) and an independent non-Gaussian random vector (U_1, U_2, U_3) with independent components. Define $X_1 = Y_1 - U_2 + U_3$, $X_2 = U_1 + Y_2 - U_3$, $X_3 = -U_1 + U_2 + Y_3$. Then it can be easily checked that $\mathbf{X} = (X_1, X_2, X_3) \in \mathcal{D}_{3,2}$ and is non-Gaussian, while $L = X_1 + X_2 + X_3 = Y_1 + Y_2 + Y_3$ is Gaussian.

Proof of Theorem 3. Denote by ϕ_j the ch. f. of L_j , $j = 1, \dots, m$. Then by the independence assumption one has, for any $\mathbf{t} = (t_1, \dots, t_m) \in \mathbb{R}^m$,

$$\prod_{j=1}^m \phi_j(t_j) = E \exp(i(\mathbf{t}, \mathbf{L})) = E \exp(i(\mathbf{t}, A\mathbf{X})) = E \exp(i(A^T \mathbf{t}, \mathbf{X})) = \phi_X(A^T \mathbf{t}),$$

where ϕ_X is the ch. f. of \mathbf{X} . Consequently, the factorization property (1) of \mathbf{X} implies

$$\prod_{j=1}^m \phi_j(t_j) = \prod_{1 \leq j_1 < \dots < j_k \leq n} R_{j_1, \dots, j_k}((\mathbf{a}_{j_1}, \mathbf{t}), \dots, (\mathbf{a}_{j_k}, \mathbf{t}))$$

for any $\mathbf{t} \in \mathbb{R}^m$. Next, we take logarithms of the both sides of the above equation. Then, in a neighborhood of the origin,

$$\sum_{j=1}^m \psi_j(t_j) = \sum_{1 \leq j_1 < \dots < j_k \leq n} r_{j_1, \dots, j_k}((\mathbf{a}_{j_1}, \mathbf{t}), \dots, (\mathbf{a}_{j_k}, \mathbf{t})), \tag{2}$$

where the ψ 's and r 's are logarithms of the respective ϕ 's and R 's.

Suppose the multiindex $\mathbf{j} = (j_1, \dots, j_k)$, $1 \leq j_1 < \dots < j_k \leq n$, takes J different values. Let us enumerate them by numbers from 1 through J in an arbitrary order. Once the order is chosen, the notation $\mathbf{j} = k$, $k \in \{1, \dots, J\}$ means that the multiindex \mathbf{j} takes the value number k in our enumeration.

Denote by $A(\mathbf{j})$ the matrix with rows $\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_k}$. Then (2) can be rewritten as

$$\sum_{j=1}^m \psi_j(t_j) = \sum_{j=1}^J r_j(A(\mathbf{j})\mathbf{t}) \tag{3}$$

where \mathbf{t} takes values in a neighborhood of the origin in \mathbb{R}^m .

Now take any vector \mathbf{u}_1 (with sufficiently small norm) orthogonal to all rows of $A(\mathbf{1})$ and consider $\mathbf{u}_1 + \mathbf{t}$ instead of \mathbf{t} in (3). Subtracting (3) from such a new equation one easily sees that the first term of the sum on the right hand side cancels. Hence

$$\sum_{j=1}^m \psi_{j,1}(t_j) = \sum_{j=2}^J r_{j,1}(A(\mathbf{j})\mathbf{t}),$$

where

$$\psi_{j,1}(a) = \psi_j(a) - \psi_j(a - b), \quad j = 1, \dots, m,$$

for any $|a|, |b|$ sufficiently small and

$$r_{j,1}(A(\mathbf{j})\mathbf{t}) = r_j(A(\mathbf{j})[\mathbf{t} + \mathbf{u}_1]) - r_j(A(\mathbf{j})\mathbf{t}).$$

After having performed the l th step ($1 < l \leq J$) one has the equation

$$\sum_{j=1}^m \psi_{j,l}(t_j) = \sum_{j=l+1}^J r_{j,l}(A(\mathbf{j})\mathbf{t}),$$

where

$$\psi_{j,l}(a) = \psi_{j,l-1}(a) - \psi_{j,l-1}(a - b), \quad j = 1, \dots, m,$$

for any $|a|, |b|$ sufficiently small and

$$r_{j,l}(A(\mathbf{j})\mathbf{t}) = r_{j,l-1}((A\mathbf{j})[\mathbf{t} + \mathbf{u}_l]) - r_{j,l-1}(A(\mathbf{j})\mathbf{t})$$

for \mathbf{t} from a neighborhood of the origin and with any \mathbf{u}_l (with sufficiently small norm) orthogonal to all rows of $A(l)$. Finally, after J steps, the right-hand side vanishes, so that

$$\sum_{j=1}^m \psi_{j,J}(t_j) = 0,$$

where the left-hand side contains only summands $\psi_{j,J}(t_j)$ such that e_j is not contained in the hyperplanes spanned by rows of $A(\mathbf{j})$, $\mathbf{j} = 1, \dots, J$. Let B be the set of such j 's. Then for any $j \in B$ and small $|t|$ it follows that $\psi_{j,J}(t) = 0$. Upon returning to the original ψ_j 's and following the known results (see, e.g., Kagan et al., 1973, Chapter 1.5.) we obtain $\psi_j(t) = Q_j(t)$, $j \in B$, where t takes values in a neighbourhood of the origin and Q_j is a polynomial of degree at most $J - 1$.

According to the classical theorem due to Marcinkiewicz (1938), if in some neighborhood of the origin a ch. f. $\psi(t) = \exp(P(t))$, where P is a polynomial, then ψ is the ch. f. of some normal distribution. Application of this fact to ϕ_j concludes the proof. \square

Remarks. 1. Theorem 3 is a straightforward extension of the original DST. It follows from the fact that in the case of independence, i.e. when $k = 1$, independence of a pair of linear forms based on a common set of independent X 's can be always reduced to the situation when all the X 's are present in each of the linear forms by simply deleting those which are not present in both. This is possible since in the resulting equation for the ch. f., cancellation can be done in a neighborhood of the origin. Hence the DST follows from Theorem 3 via the Cramér decomposition theorem.

2. In general, the Cramér theorem does not work for k -factorizable X 's. The fact that the m -dimensional linear transformation $L(X)$ has a Gaussian distribution does not imply that X is Gaussian too – see the example after Theorem 3. However, for $m \geq k$ the additional assumption that all $(k - 1)$ -variate marginals of the X are Gaussian implies joint normality (see Theorem 2 in Wesołowski, 1997). Hence it follows that if it is assumed additionally in Theorem 3 that the $(k - 1)$ -dimensional marginals of X are normal, then X is n -variate Gaussian.

3. Observe that if $k = m - 1$, $n = m$ and A is non-singular, then L_j is normal if only $b_{1j} \dots b_{mj} \neq 0$, where $B = A^{-1}$. This is a trivial consequence of the main result of Kagan (1988b).

4. From the method used in the proof of Theorem 3 it follows that, if L_1 and (L_2, \dots, L_m) are independent (under the obvious assumption on coefficients of linear forms) then L_1 is Gaussian.

5. The assumption of Theorem 3 that the hyperplanes $\text{span}\{\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_k}\}$ do not contain respective vectors from the canonical basis cannot simply be omitted. (Similarly, the analogous condition cannot be omitted from the hypotheses of the DST.)

In fact, take $\mathbf{X} = (X_1, X_2, X_3)$ such that X_1 is independent of (X_2, X_3) . Then, plainly $\mathbf{X} \in \mathcal{D}_{3,2}$. If now $L_1 = X_1 + a_{12}X_2 + a_{13}X_3$, $L_2 = a_{22}X_2 + a_{23}X_3$, $L_3 = a_{32}X_2 + a_{33}X_3$, then L_1, L_2, L_3 are independent whenever the linear form $a_{12}X_2 + a_{13}X_3$, L_2, L_3 are independent (i.e. (X_2, X_3) is Gaussian).

Therefore, in principle, the normality of L_1 cannot be deduced from independence of L_1, L_2, L_3 since the distribution of X_1 can be arbitrary.

Acknowledgements

Thanks are due to Paul J. Smith, an anonymous referee, and the Editor whose suggestions improved the presentation of the material.

References

- Darmois, G., 1953. Analyse générale des liaisons stochastiques. *Revue Institute Internationale Statistique* 21, 2–8.
- Feldman, G.M., 1988. Characterization of the Gaussian distribution on groups by independence of linear statistics. *Doklady Akademii Nauk SSSR* 301, 558–560 (in Russian).
- Ghurye, S.G., Olkin, I., 1962. A characterization of the multivariate normal distribution. *Ann. Math. Statist.* 33, 533–541.
- Kagan, A.M., 1987. A class of two-dimensional distributions arising in connection with the Cramér and Darmois–Skitovitch theorems. *Theory Probab. Appl.* 32, 349–351 (in Russian).
- Kagan, A.M., 1988a. A multivariate analog of the Cramér theorem on components of the Gaussian distribution. In: V.M. Zolotarev, V.V. Kalashnikov (Eds.), *Stability Problems of Stochastic Models, Lecture Notes in Mathematics*, vol. 1412. Springer, New York, pp. 68–77.
- Kagan, A.M., 1988b. New classes of dependent random variables and a generalization of the Darmois–Skitovitch theorem to the case of several linear forms. *Theory Probab. Appl.* 33, 305–314 (in Russian).
- Kagan, A.M., Linnik, Yu.V., Rao, C.R., 1973. *Characterization Problems of Mathematical Statistics*. Wiley, New York.
- Kagan, A.M., Zinger, A.A., 1985. A refinement of the Darmois–Skitovitch and Heyde theorems. In: V.M. Zolotarev, V.V. Kalashnikov, (Eds.), *Stability Problems of Stochastic Models, Lecture Notes in Mathematics*, vol. 1155. Springer, New York, pp. 81–94.
- Laha, R.G., 1956. On a characterization of the normal distribution from properties of suitable linear statistics. *Ann. Math. Statist.* 27, 126–139.
- Lisyanoy, S.V., 1995. On a group analogue of the Kagan theorem. *Theory Probab. Appl.* 40, 180–183 (in Russian).
- Lukacs, E., King, E.P., 1954. A property of the normal distribution. *Ann. Math. Statist.* 25, 389–394.
- Marcinkiewicz, J., 1938. Sur une propriété de la loi de Gauss. *Mathemat. Zeitschrift* 44, 622–638.
- Plucińska, A., Wesółowski, J., 1995. Gaussian processes via independence of linear forms. In: A.V. Skorokhod, Yu.V. Borovskikh (Eds.), *Exploring Stochastic Laws*. VSP International Publishers, Zeist (The Netherlands), pp. 383–393.
- Skitovitch, W.P., 1953. On a property of the normal distribution. *Doklady Akademii Nauk SSSR* 89, 217–219 (in Russian).
- Wesółowski, J., 1991a. On determination of multivariate distribution by its marginals for the Kagan class. *J. Multivar. Anal.* 36, 314–319.
- Wesółowski, J., 1991b. Gaussian conditional structure of the second order and the Kagan classification of multivariate distributions. *J. Multivar. Anal.* 39, 79–86.
- Wesółowski, J., 1994. The Lyapounov central limit theorem for factorizable arrays. *Proc. American Math. Soc.* 122, 165–174.
- Wesółowski, J., 1997. On some distributional and limit properties of factorizable distributions. *Statist. Probab. Lett.* 32, 329–337.