mathematics

Article

On the Condition of Independence of Linear Forms with a Random Number of Summands

Abram M. Kagan [1] and Lev B. Klebanov [2,*]

[1] Department of Mathematics, University of Maryland, College Park, MD 20742, USA; akagan@umd.edu
[2] Department of Probability and Mathematical Statistics, Charles University, 18675 Prague, Czech Republic
* Correspondence: klebanov@karlin.mff.cuni.cz

Abstract: The property of independence of two random forms with a non-degenerate random number of summands contradicts the Gaussianity of the summands.

Keywords: independent linear forms; random number of summands; Gaussian distribution

check for updates

Citation: Kagan, A.M.; Klebanov, L.B. On the Condition of Independence of Linear Forms with a Random Number of Summands. *Mathematics* **2021**, *9*, 1516. https://doi.org/10.3390/math9131516

Academic Editor: Panagiotis-Christos Vassiliou

Received: 20 April 2021
Accepted: 23 June 2021
Published: 29 June 2021

1. Introduction

The interest in the characterization of probability distributions appeared during the 1920s. The original document was an article by G. Polya [1], which highlights an interesting characteristic of Gaussian distribution due to the property of the identity distribution of a random variable and a special linear form. The first result on the characterization of the distribution by the independence of two linear forms of two random variables was achieved by S.N. Bernstein [2], which again appears to be a characterization of the Gaussian distribution. The result obtained by S.N. Bernstein was essentially generalized by Skitovich [3] and Darmois [4]. In their works, the independence of two linear forms from an arbitrary number of random variables was considered. However, the result appeared to be the same. Independence took place for normally distributed variables only. In 1972, the first monograph on the characterization problems in statistics was published, written by A.M. Kagan, Yu.V. Linnik and C.R. Rao [5]. The monograph contained the results on the characterization of probability distributions from many fields of statistics, probability and their applications. Among these were, of course, the results related to independence properties, generalizing those noted above. Nevertheless, all results on the independence of linear forms led to the characterizations of normal distribution. Note that previously unsolved problems were formulated in the monograph [5], the solutions of which could contribute to a better understanding of the role and significance of the characterizations of probability distributions. During the period since the publication of [5], some of these problems have been solved, which led to the further development of the corresponding theory. One of the unsolved problems remained the following (see [5], pp. 460–461): let $\{X_i\}$ be a sequence of independent (and to start with, identically distributed) r.v.'s, let $\{a_i\}$ and $\{b_i\}$ be two sequences of real numbers, and let τ be a Markovian stopping time (cf. Chapter 12 also): then, construct the 'linear forms' with a random number of summands:

$$L_1 = \sum_{i=1}^{\tau} a_i X_i \quad \text{and} \quad L_2 = \sum_{i=1}^{\tau} b_i X_i.$$

Investigate the conditions for the independence of L_1 and L_2. Under what conditions on τ and the sequences $\{a_i\}$ and $\{b_i\}$ would the independence of L_1 and L_2 imply the normality of X_i?

Here, we consider the case:

$$L_1 = \sum_{i=1}^{\tau_1} a_i X_i \quad \text{and} \quad L_2 = \sum_{i=1}^{\tau_2} b_i X_i.$$

when τ_1 and τ_2 are identically distributed random variables independent of the sequence $\{X_i\}$ and from each other. The variant with $\tau_1 = \tau_2 = \tau$ may almost surely be considered in the same way which would lead to the same result.

2. Main Result

Our main result is given by the following Theorem.

Theorem 1. *Suppose that $\{X_i\}$ is a sequence of independent non-degenerate random variables and τ is a positive integer-valued random variable independent on $\{X_i\}$. Linear forms:*

$$L_1 = \sum_{i=1}^{\tau_1} a_i X_i \quad \text{and} \quad L_2 = \sum_{i=1}^{\tau_2} b_i X_i$$

are independent for normally distributed $\{X_i\}$ if and only if $\tau_1 = \tau_2 = n$ with probability 1 (n is a positive integer constant) and $\sum_{i=1}^{n} a_i b_i \sigma_j^2 = 0$ where σ_j^2 is a variance of X_j.

Proof. Suppose the opposite. This means that there are normally distributed X_j with the parameters $m_j \in \mathbb{R}^1$, $\sigma_j \in \mathbb{R}^1_+$ $(j = 1, 2, \ldots)$ and a positive integer-valued non-degenerated random variable τ independent of the sequence $\{X_j\}$ and such that L_1 and L_2 are independent. The independence of L_1 and L_2 may be written in terms of characteristic functions, which has the form:

$$\mathbb{E}\exp\{isL_1 + itL_2\} = \mathbb{E}\exp\{isL_1\}\mathbb{E}\exp\{itL_2\},$$

or, in a more detailed form:

$$\sum_{k=1}^{\infty}\prod_{j=1}^{k} f_j(a_j s + b_j t) p_k = \sum_{k=1}^{\infty}\prod_{j=1}^{k} f_j(a_j s) p_k \sum_{n=1}^{\infty}\prod_{j=1}^{n} f_j(b_j t) p_n, \tag{1}$$

where $p_k = \mathbb{P}\{\tau = k\}$, and $f_j(u) = \exp\{im_j t - \sigma_j^2 u^2/2\}$ is a characteristic function of the normal distribution with the parameters m_j and σ_j $(j = 1, 2, \ldots)$. The relation (1) may be written in more detail as

$$\begin{gathered}\textstyle\sum_{k=1}^{\infty} p_k \exp\{is\sum_{j=1}^{k} a_j m_j + it\sum_{j=1}^{k} b_j m_j - \frac{s^2}{2}A_k - stC_k - \frac{t^2}{2}B_k\} = \\ \textstyle= \sum_{k=1}^{\infty} p_k \exp\{is\sum_{j=1}^{k} a_j m_j - \frac{s^2}{2}A_k\}\cdot\sum_{n=1}^{\infty} p_n \exp\{it\sum_{j=1}^{n} b_j m_j - \frac{t^2}{2}B_n\},\end{gathered} \tag{2}$$

where:

$$A_k = \sum_{j=1}^{k} a_j^2\sigma_j^2, \quad B_k = \sum_{j=1}^{k} b_j^2\sigma_j^2, \quad C_k = \sum_{j=1}^{k} a_j b_j \sigma_j^2. \tag{3}$$

Let us note that A_k and B_k are non-decreasing functions of k and strictly increasing on the set of all k for which $p_k \neq 0$.

Denote by k_o the minimal value of k for which $p_k \neq 0$. Let us fix arbitrary t in (2) and let $s \to \infty$. It is clear that the right-hand-side of (2) tends toward zero and:

$$\begin{gathered}\textstyle\sum_{k=1}^{\infty} p_k \exp\{is\sum_{j=1}^{k} a_j m_j - \frac{s^2}{2}A_k\}\cdot\sum_{n=1}^{\infty} p_n \exp\{it\sum_{j=1}^{n} b_j m_j - \frac{t^2}{2}B_n\} = \\ \textstyle p_{k_o}\exp\{is\sum_{j=1}^{k_o} a_j m_j - A_{k_o}s^2/2\}\sum_{n=1}^{\infty} p_n \exp\{it\sum_{j=1}^{n} b_j m_j - \frac{t^2}{2}B_n\}\big(1 + o(1)\big)\end{gathered} \tag{4}$$

as $s \to \infty$.

However, for the left-hand-side of (2) we have:

$$\sum_{k=1}^{\infty} p_k \exp\{is\sum_{j=1}^{k} a_j m_j + it\sum_{j=1}^{k} b_j m_j - \frac{s^2}{2}A_k - stC_k - \frac{t^2}{2}B_k\} =$$
$$p_{k_o} \exp\{is\sum_{j=1}^{k_o} a_j m_j + it\sum_{j=1}^{k_o} b_j m_j - \frac{s^2}{2}A_{k_o} - stC_{k_o} - \frac{t^2}{2}B_{k_o}\}\Big(1+o(1)\Big) \tag{5}$$

as $s \to \infty$. From (4) and (5), it follows that:

$$p_{k_o} \exp\{is\sum_{j=1}^{k_o} a_j m_j - A_{k_o}s^2/2\} \sum_{n=1}^{\infty} p_n \exp\{it\sum_{j=1}^{n} b_j m_j - \frac{t^2}{2}B_n\} =$$

$$p_{k_o} \exp\{is\sum_{j=1}^{k_o} a_j m_j + it\sum_{j=1}^{k_o} b_j m_j - \frac{s^2}{2}A_{k_o} - stC_{k_o} - \frac{t^2}{2}B_{k_o}\}\Big(1+o(1)\Big)$$

or

$$\sum_{n=k_o}^{\infty} p_n \exp\{it\sum_{j=k_o}^{n} b_j m_j - \frac{t^2}{2}B_n\} = \exp\{it\sum_{j=1}^{k_o} b_j m_j - stC_{k_o} - \frac{t^2}{2}B_{k_o}\}\Big(1+o(1)\Big) \tag{6}$$

as $s \to \infty$ for arbitrary fixed $t \in \mathbb{R}^1$. The left-hand-side of (6) does not depend on s. Therefore, $C_{k_o} = \sum_{j=1}^{k_o} a_j b_j \sigma_j^2 = 0$ and:

$$\sum_{n=k_o}^{\infty} p_n \exp\{it\sum_{j=k_o}^{n} b_j m_j - \frac{t^2}{2}B_n\} = \exp\{it\sum_{j=1}^{k_o} b_j m_j - \frac{t^2}{2}B_{k_o}\}\Big(1+o(1)\Big). \tag{7}$$

However, $B_n > B_{k_o}$ for any $n > k_o$ and t is arbitrary. Therefore, $\exp\{it\sum_{j=k_o}^{n} b_j m_j - \frac{t^2}{2}B_n\}/\exp\{-\frac{t^2}{2}B_{k_o}\} = o(1)$ as $t \to 0$. Taking this into account and passing to absolute values in both sides of (7), we can see that:

$$p_o \exp\{-\frac{t^2}{2}B_{k_o}\} = \exp\{-\frac{t^2}{2}B_{k_o}\}\Big(1+o(1)\Big). \tag{8}$$

The relation (8) implies $p_o = 1$. □

3. Conclusions

As mentioned in the introduction, the independence of two linear forms usually leads to the normality of the corresponding random variables. Our Theorem 1 shows that this is not the case in the situation under consideration. The reason for this dependence of the forms for non-degenerate distributed, τ_1 and τ_2, lies in the random number of random variables used. Similar facts may be found in the publications [6,7] for the cases of some different problems. One can see a big difference between problems with a fixed or a random number of variables involved.

Author Contributions: Conceptualization, A.M.K. and L.B.K.; methodology, A.M.K. and L.B.K.; formal analysis, A.M.K. and L.B.K.; investigation, A.M.K. and L.B.K.; writing—original draft preparation, A.M.K. and L.B.K.; writing—review and editing, A.M.K. and L.B.K.; funding acquisition, L.B.K. Both authors have read and agreed to the published version of the manuscript.

Funding: The study was partially supported by grant GAČR 19-04412S (Lev Klebanov).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Polya, G. Herleitung des Gausssćhen Fehrlergesetzes aus einer Funktionalgleichung. *Math. Z.* **1923**, *18*, 96–108. (In German) [CrossRef]
2. Bernstein, S.N. On a property characterizing Gaussian law. *Proc. Leningr. Polytech. Inst.* **1941**, *3*, 21–22. (In Russian)
3. Skitovich, V.P. On a property of the normal distribution. *Rep. Acad. Sci. USSR* **1953**, *89*, 217–219. (In Russian)
4. Darmois, G. Analyse générale des liaisons stochastiques. *Rev. Inst. Intern. Stat.* **1953**, *30*, 80–101. (In France)
5. Kagan, A.M.; Linnik, Y.V.; Rao, C.R. *Characterization Problems in Mathematical Statistics*; John Wiley & Sons Inc.: Hoboken, NJ, USA, 1973.
6. Kakosyan, A.V.; Klebanov, L.B.; Melamed, J.A. *Characterization of Distributions by the Method of Intensively Monotone Operators*; Springer: Berlin/Heidelberg, Germany, 1984.
7. Klebanov, L.B. Linear Statistics with Random Coefficients and Characterization of Hyperbolic Secant Distribution. *arXiv* **2019**, arXiv:1905.09910v1.