Probab. Th. Rel. Fields 73, 43–49 (1986)

**Probability Theory and Related Fields**



# Characterizing the Distributions of the Random Variables $X_1$, $X_2$, $X_3$ by the Distribution of $(X_1-X_3, X_2-X_3)$

Zoltán Sasvári

Technische Universität, Sektion Mathematik, DDR-8027 Dresden, Mommsenstr. 13, German Democratic Republic

I. Kotlarski [2] showed that, for three independent random variables $X_1$, $X_2$, $X_3$ the distribution of the random vector

$$Y=(X_1-X_3, X_2-X_3) \tag{1}$$

determines the distributions of $X_1$, $X_2$ and $X_3$ up to a change of the location if the characteristic function $\psi$ of $Y$ does not vanish. In [5], P.G. Miller proved that the statement remains true if $X_1$, $X_2$, $X_3$ are $n$-dimensional random vectors and the requirement that $\psi$ does not vanish is replaced by the requirement that $X_1$, $X_2$ and $X_3$ have analytic characteristic functions.

For results concerning random variables with values in linear topological spaces or in locally compact Abelian groups see [3] and [6], respectively.

Paper [9] contains generalizations of the results of I. Kotlarski and P.G. Miller in the case, where the random variables $X_1$, $X_2$ and $X_3$ have the same distribution. Moreover, examples are given, where the distribution of $Y$ does not determine the distributions of $X_1$, $X_2$ and $X_3$.

In this paper we generalize some results of [9] without the assumption that $X_1$, $X_2$, $X_3$ are identically distributed. In Sect. 3 we give new examples, where the distribution of $Y$ does not determine the distributions of $X_1$, $X_2$ and $X_3$. The word "support" of a function is used for the set of points where the function is not zero.

## 1. Preliminaries

Throughout this note $S_f$ will denote the support of a function $f$ on $R^n$. In the case $n=1$ we set

$$S_f^+ = S_f \cap (0, \infty) \quad \text{and} \quad n_f(t) = \frac{\lambda(S_f^+ \cap (0, t))}{t}, \qquad t>0,$$

where $\lambda$ denotes the Lebesgue measure on $R$. For the algebraic sum of the sets $A$, $B \subset R^n$ we shall use the notation $A+B$.

In the proof of Theorem 1 we need the following simple fact.

**Lemma 1.** *Let $A$ and $B$ be Lebesgue measurable subsets of the interval $(0, t)$, $t>0$. If $\lambda(A)+\lambda(B)>t$, then there exist $a \in A$ and $b \in B$ such that*

$$t=a+b.$$

*Proof.* Since $t-A \subset (0, t)$ and $\lambda(t-A)=\lambda(A)$, we have $\lambda(t-A)+\lambda(B)>t=\lambda(0, t)$. Therefore $(t-A) \cap B \neq \emptyset$, that is, there exist $a \in A$ and $b \in B$ such that $t-a=b$ and thus $t=a+b$. Q.E.D.

## 2. The Case where $X_1$, $X_2$ and $X_3$ are not Supposed to be Identically Distributed

**Theorem 1.** *Let $X_1$, $X_2$, $X_3$ be independent random variables with characteristic functions $f$, $g$ and $h$, respectively. If the inequalities*

$$\begin{aligned} n_f(t)+n_g(t)&>1, \\ n_f(t)+n_h(t)&>1, \\ n_g(t)+n_h(t)&>1 \end{aligned} \tag{2}$$

*hold for every $t>0$, then the distribution of* (1) *determines the distributions of $X_1$, $X_2$ and $X_3$ up to a change of the location.*

*Proof.* Denote the characteristic function of $Y=(X_1-X_3, X_2-X_3)$ by $\psi(t_1, t_2)$. We have

$$\begin{aligned} \psi(t_1, t_2) &= E\exp(i(t_1(X_1-X_3)+t_2(X_2-X_3))) \\ &= E(\exp(it_1X_1)\exp(it_2X_2)\exp(i(-t_1-t_2)X_3)) \\ &= E\exp(it_1X_1)E\exp(it_2X_2)E\exp(i(-t_1-t_2)X_3) \\ &= f(t_1)g(t_2)h(-t_1-t_2). \end{aligned}$$

Let $X'_1$, $X'_2$, $X'_3$ another three independent random variables, having characteristic functions $f'$, $g'$, $h'$, respectively. The random vectors $Y=(X_1-X_3, X_2-X_3)$ and $Y'=(X'_1-X'_3, X'_2-X'_3)$ have the same distribution if and only if the functional equation

$$f(t_1)g(t_2)h(-t_1-t_2)=f'(t_1)g'(t_2)h'(-t_1-t_2) \tag{3}$$

holds for every $t_1$, $t_2 \in R$.

We shall show that (2) and (3) imply

$$f'(t)=\exp(iat)f(t), \quad g'(t)=\exp(iat)g(t), \quad h'(t)=\exp(iat)h(t) \tag{4}$$

for all $t \in R$, where $a$ is a real number.

Using (3) and a well-known result of the theory of functional equations, it can be shown [5] that (4) holds in some neighborhood of 0.

Let

$$K_f = \sup\{T \in R : f'(t) = \exp(iat) f(t) \text{ for all } t \in [-T, T]\},$$

and define $K_g$ and $K_h$ similarly.

Suppose first that $K_h \leqq K_g$ and $K_h \leqq K_f$. We want to show that $K_h = K_g = K_h = \infty$. If this were not true, we would have for every integer $n \geqq 1$ a number $t_n \in R$ such that

$$K_h < t_n < K_h + \frac{1}{n} \quad \text{and} \quad h'(t_n) \neq \exp(iat_n) h(t_n).$$

Setting in (2) $t = K_h$ and using Lemma 1 we see that there exist real numbers $t_f \in S_f^+$ and $t_g \in S_g^+$ such that

$$K_h = t_f + t_g.$$

From this we conclude that $K_h$ is in the open set

$$O = [(0, K_h) \cap S_f^+] + [(0, K_h) \cap S_g^+]. \tag{5}$$

For a sufficiently large $n$ we have $t_n \in O$. Applying (5) we obtain

$$t_n = \bar{t}_f + \bar{t}_g$$

where $\bar{t}_f \in (0, K_h) \cap S_f^+$ and $\bar{t}_g \in (0, K_h) \cap S_g^+$.

Since $\bar{t}_f < K_f$ and $\bar{t}_g < K_g$, by substituing $t_1 = \bar{t}_f$ and $t_2 = \bar{t}_g$ in (3) we obtain

$$f(\bar{t}_f) g(\bar{t}_g) h(-t_n) = \exp(ia\bar{t}_f) f(\bar{t}_f) \exp(ia\bar{t}_g) g(\bar{t}_g) h'(-t_n).$$

Using the fact that $f(\bar{t}_f) \neq 0$ and $g(\bar{t}_g) \neq 0$, we have

$$h(-t_n) = \exp(i(a\bar{t}_f + a\bar{t}_g)) h'(-t_n) = \exp(iat_n) h'(-t_n)$$

and so

$$h'(t_n) = \exp(iat_n) h(t_n).$$

This contradiction establishes that $K_h = \infty$. Hence we have also $K_f = K_g = \infty$.

Let us now suppose that $K_f \leqq K_g$ and $K_f \leqq K_h$. Substituing $-t_1 - t_2 = x$ and $t_2 = y$ in (3) we obtain

$$f(-x-y) g(y) h(x) = f'(-x-y) g'(y) h'(x) \qquad \text{for every } x, y \in R.$$

Therefore, we can repeat the above argument with the roles of $f$ and $h$ interchanged. In the case $K_g \leqq K_f$, $K_g \leqq K_h$ we substitute $-t_1 - t_2 = x$ and $t_1 = y$. Q.E.D.

From Theorem 1 we obtain immediately the following generalization of the above mentioned results of I. Kotlarski and P.G. Miller.

**Corollary 1.** *If two of the characteristic functions $f$, $g$ and $h$ are analytic or have no zeros, then the distribution of* (1) *determines the distributions of $X_1$, $X_2$ and $X_3$ up to a change of the location.*

*Remark 1.* In view of condition (2) it is natural to ask, how might the support or the set of zeros of a characteristic function look. A.I. Iljinskij gave in [1] the following characterization.

A closed set $Z \subset R$ is the set of zeros of a characteristic function if and only if $Z = -Z$ and $0 \notin Z$.

Let $X$ be a random variable and $f$ its characteristic function. In [8] the set of zeros of $f$ is characterised in the following three cases:

a) all moments of $f$ exist,

b) $X$ is a lattice random variable,

c) $X$ is discret.

*Remark 2.* If the supports $S_f$, $S_g$, $S_h$ are bounded, for example, $S_f, S_g, S_h \subset (-1, 1)$, then (2) is not satisfied and the distribution of (1) does not determine the distributions of $X_1$, $X_2$ and $X_3$ up to a change of the location.

In fact, let $h'$ be a characteristic function having the properties $h'(t) = h(t)$ for $|t| \leqq 2$ and $h' \neq h$. For the construction of such characteristic functions see [7]. Setting $f' = f$ and $g' = g$ we have

$$f(t_1)\, g(t_2)\, h(-t_1 - t_2) = f'(t_1)\, g'(t_2)\, h'(-t_1 - t_2),$$

but there is no $a \in R$ for which $h'(t) = \exp(iat)\, h(t)$.

*Remark 3.* If we only require that two of the inequalities in (2) are satisfied, then the statement of Theorem 1 becomes false. To see this, let $f$, $g$, $f'$, $g'$ be as in Remark 2. We choose the characteristic functions $h$ and $h'$ so that $h > 0$, $h \neq h'$ and $h(t) = h'(t)$ for $|t| \leqq 2$. (We can find simple examples of $h$ and $h'$ if we choose convex characteristic functions.) Then we have (3) and two inequalities in (2) are satisfied but there is no $a \in R$ for which $h'(t) = \exp(iat)\, h(t)$.

*Remark 4.* As the example in [9] shows, "$>1$" cannot be replaced by "$\geqq 1$" in (2).

*Remark 5.* Suppose that we know the characteristic function $\psi$ of the random variable $Y = (X_1 - X_3, X_2 - X_3)$. Generally $\psi$ does not determine the supports $S_f$, $S_g$, $S_h$ uniquely. However, if we set

$$S_1^\psi = \{t_1 \in R \colon \psi(t_1, t_2) \neq 0 \text{ for some } t_2 \in R\},$$

$$S_2^\psi = \{t_2 \in R \colon \psi(t_1, t_2) \neq 0 \text{ for some } t_1 \in R\},$$

$$S_3^\psi = \{t \in R \colon \psi(t_1, t_2) \neq 0 \text{ for some } t_1,\ t_2 \in R \text{ with } t_1 + t_2 = t\},$$

then we have $S_1^\psi \subset S_f$, $S_2^\psi \subset S_g$, $S_3^\psi \subset S_h$. Therefore, setting

$$n_k^\psi(t) = \frac{\lambda(S_k^\psi \cap (0, t))}{t}$$

for $k = 1, 2, 3$ and $t > 0$, we get

$$n_1^\psi(t) \leqq n_f(t), \qquad n_2^\psi(t) \leqq n_g(t), \qquad n_3^\psi(t) \leqq n_h(t).$$

Thus we have

**Theorem 1′.** *Let $\psi$ be the characteristic function of the random vector* (1) *and suppose that the inequalities*

$$\begin{aligned} n_1^{\psi}(t)+n_2^{\psi}(t)&>1,\\ n_1^{\psi}(t)+n_3^{\psi}(t)&>1,\\ n_2^{\psi}(t)+n_3^{\psi}(t)&>1 \end{aligned} \tag{2'}$$

*hold for every $t>0$. Then the distribution of* (1) *determines the distributions of $X_1$, $X_2$, $X_3$ up to a change of the location.*

We now give another generalization of the results of I. Kotlarski and P.G. Miller.

**Theorem 2.** *Let $\psi$ be the characteristic function of the random vector* (1). *If*

$$\bar{S}_{\psi}=R^2 \tag{6}$$

*then the distribution of* (1) *determines the distributions of $X_1$, $X_2$, $X_3$ up to a change of the location.*
(By $\bar{A}$ we denote the closure of the set $A$.)

*Proof.* Denote by $f$, $g$, $h$ the characteristic functions of $X_1$, $X_2$ and $X_3$, respectively. From (6) if follows that

$$\bar{S}_f=R, \quad \bar{S}_g=R \quad \text{and} \quad \bar{S}_h=R. \tag{7}$$

Let $t>0$ and define

$$A=(0,t)\cap S_f, \qquad B=(0,t)\cap S_g \quad \text{and} \quad C=(0,t)\cap S_h.$$

The set $t-A$ is open and we have $t-A\subset(0,t)$. From $\bar{B}=[0,t]$ it follows that $(t-A)\cap B\neq\emptyset$. Thus there exist $a\in A$ and $b\in B$ such that $t-A=b$, and so,

$$t=a+b. \tag{8}$$

The same argument can be used to prove that there exist $a'\in A$, $b''\in B$, $c'$, $c''\in C$ such that

$$t=a'+c' \quad \text{and} \quad t=b''+c''. \tag{9}$$

We can now repeat the proof of Theorem 1, if we use (8) and (9) instead of Lemma 1. Q.E.D.

*Remark 6.* Theorem 2 can be easily generalized in the case where $X_1$, $X_2$ and $X_3$ are $n$-dimensional random vectors.

*Remark 7.* It is easy to see that (8) and (9) remain true if we only require that two of the equalities in (7) are satisfied. Thus we obtain the following generalization of Corollary 1.

**Corollary 2.** *If two of the sets $S_f$, $S_g$, $S_h$ are dense in $R$ then the distribution of* (1) *determines the distributions of $X_1$, $X_2$ and $X_3$ up to a change of the location.*

## 3. The Identically Distributed Case

Let $C_p$, $p>0$, denote the family of characteristic functions $f$ on $R$ satisfying the following conditions:

a) $f(t)\neq 0$ for $|t|<p$,

b) $f(t)=0$ for $|t|\geqq p$.

By $C_{p,q}$, $0<p\leqq q$, we denote the family of functions $g$ satisfying

a) $g$ is periodic of period $2q$,

b) there exists $f\in C_p$ such that $g(t)=f(t)$ for $|t|\leqq q$.

P. Lévy [4] has shown that $g\in C_{p,q}$ implies that $g$ is a characteristic function. We give now a generalization of Theorem 1 in [9].

**Theorem 3.** *Let $X_1$, $X_2$, $X_3$ be independent identically distributed random variables having the common characteristic function $f\in C_{p,q}$. The distribution of* (1) *determine the distribution of $X_1$, $X_2$ and $X_3$ up to a change of the location if and only if*

$$p>\tfrac{2}{3}q.$$

*Proof.* Since $X_1$, $X_2$ and $X_3$ are identically distributed we have the functional equation

$$f(t_1)f(t_2)f(-t_1-t_2)=f'(t_1)f'(t_2)f'(-t_1-t_2), \quad t_1,t_2\in R \tag{10}$$

instead of (3).

In [9] it was proved that (10) has nontrivial solutions if and only if $p\leqq\frac{2}{3}q$. In the case $p\leqq\frac{2}{3}q$ the solution of (10) was to be shown in the form

$$f'(t)=c^k\exp(iat)f(t), \quad t\in((2k-1)q, \quad (2k+1)q), \quad k=0,\pm 1,\pm 2,\ldots,$$

where $a\in R$ and $c$ is a complex number with $|c|=1$. We have only to show that $f'$ is a characteristic function. Now, this fact follows from Remark 8 in [7]. Q.E.D.

*Acknowledgements.* The author wishes to thank W. Wolf for his help and encouragement.

## References

1. Iljinskij, A.: On the zeros and the argument of characteristic functions. Teor. Verojatnost. i Primenen (Russian) **20**, 421–427 (1975)
2. Kotlarski, I.: On characterizing the gamma and the normal distribution. Pacific J. Math. **20**, 69–76 (1967)
3. Kotlarski, I.: A theorem on joint probability distributions in stochastic locally convex linear topological spaces. Colloquium Mathematicum **19**, 175–177 (1968)
4. Lévy, P.: Quelques problèmes non résolus de la théorie des fonctions caractéristiques. Annali Mat. Pura e Appl. (**4**), 315–332 (1961)
5. Miller, P.G.: Characterizing the distributions of three independent $n$-dimensional random variables $X_1$, $X_2$, $X_3$, having analytic characteristic functions by the joint distribution of $(X_1+X_3, X_2+X_3)$. Pacific J. Math. **35**, 487–491 (1970)

6. Rao, B.L.S.P.: On a characterization of probability distributions on locally compact Abelian groups. Z. Wahrscheinlichkeitstheor. Verw. Geb. **9**, 98–100 (1968)
7. Sasvári, Z.: Einige Bemerkungen über die Fortsetzung positiv definiter Funktionen. Z. Anal. Anwend. **3**(5), 435–440 (1984)
8. Sasvári, Z.: Über die Nullstellenmenge von charakteristischen Funktionen. Math. Nachrichten **121**, 33–40 (1985)
9. Sasvári, Z., Wolf, W.: Über die Wiederherstellung der Ausgangsverteilung durch die Statistik $(X_1-X_3, X_2-X_3)$. Math. Operationsforsch. Statist. Ser. Statist **16**(2), 233–241 (1985)

Received February 10, 1984; in revised form January 22, 1986