

A Characterization of Uniform Distribution

by

Joanna CHACHULSKA

Presented by Stanisław KWAPIEŃ

Summary. Is the Lebesgue measure on $[0, 1]^2$ a unique product measure on $[0, 1]^2$ which is transformed again into a product measure on $[0, 1]^2$ by the mapping $\psi(x, y) = (x, (x + y) \bmod 1)$? Here a somewhat stronger version of this problem in a probabilistic framework is answered. It is shown that for independent and identically distributed random variables X and Y constancy of the conditional expectations of $X + Y - I(X + Y > 1)$ and its square given X identifies uniform distribution either absolutely continuous or discrete. No assumptions are imposed on the supports of the distributions of X and Y .

1. Introduction and the main result. It is well known that the Lebesgue measure on $[0, 1]^2$ is invariant under the mapping $\psi : [0, 1]^2 \rightarrow [0, 1]^2$ defined by $\psi(x, y) = (x, x + y - I(x + y > 1))$. Consequently, ψ transforms a product measure into a product measure. It is natural to ask if the Lebesgue measure is the only measure invariant under ψ . This problem is equivalent to the question if ψ preserves independence of components of a bivariate random vector. Here we are concerned with even weaker assumptions, which, in probabilistic framework, are conveniently expressed in terms of constancy of regressions.

On the other hand, the problem can be viewed as related to the Darmois–Skitovich theorem which states that independence of linear forms in independent random variables implies the normality of these random variables. This celebrated characterization theorem was a source of numerous further investigations including its versions on different algebraic structures. In particular, Stapleton [11] proved its analogue for random variables taking values in a compact Abelian topological group (Γ, \oplus) obtaining a characterization of the Haar measure μ on Γ . That paper was somehow overlooked in later in-

2000 *Mathematics Subject Classification*: 60E05, 62E10, 28A12.

Key words and phrases: uniform distribution, constancy of regression, characterization of probability distribution, probability distribution on algebraic structures.

vestigations of the Darmois–Skitovich characterization on topological groups (see [1], [4]–[10]). For instance Baryshnikov *et al.* [1] considered independent random variables X, Y taking values in the group of real numbers modulo 2π , $([0, 2\pi), \oplus)$, which up to an isomorphism, is the algebraic structure we consider in this paper.

A simplified version of one of the results from [11] reads: Let X and Y be independent random variables (rv's) with a common distribution which is not concentrated on any coset of a proper subgroup of Γ . If X and $U = X \oplus Y$ are independent then μ (the Haar measure) is the common distribution of X and Y .

Let us rewrite the result mentioned above for the group $([0, 1), \oplus)$, where $x \oplus y = x + y - I(x + y > 1)$, $x, y \in [0, 1)$. It states that for two independent identically distributed (i.i.d.) rv's X and Y , defined on a probability space (Ω, \mathcal{F}, P) , which are not of the discrete type, if X and $U = X \oplus Y$ are independent then their distribution is continuous uniform on $[0, 1]$. In this paper the result will be extended in two directions: (1) no assumptions on the support nor type of distributions will be imposed; (2) the condition of independence of U and X will be weakened to constancy of regressions of two functions of U given X (see [6] for a regression version of the Darmois–Skitovich theorem on locally compact Abelian groups). It turns out that under such circumstances not only continuous uniform distribution may appear as the common law for X and Y , but also a family of discrete uniform distributions is admissible.

The regression conditions we are interested in are the following:

$$\begin{aligned} (1) \quad & E[X + Y - I(X + Y > 1) | X] = E(X) \quad \text{a.s.}, \\ (2) \quad & E[(X + Y - I(X + Y > 1))^2 | X] = K \quad \text{a.s.}, \end{aligned}$$

where $K \in \mathbb{R}$ is a constant. The main result of this paper, Theorem 1, states that these two conditions characterize uniform continuous and discrete distributions. Let us stress that no restrictions on the distribution of X and Y are imposed. Also the fact that (1) and (2) hold only almost surely is a source of additional difficulties in the proof.

It is worth mentioning that (1) and (2) imply that there exists $A \in \mathcal{B}$ with $P(X \in A) = 1$ such that for any $a, b \in \mathbb{R}$ with $1 - a, 1 - b \in A$,

$$(3) \quad E(Y | a < Y \leq b) = \frac{1}{2}(a + b) + E(Y) - \frac{1}{2}.$$

Let $\text{supp}(Y)$ denote the support of the distribution of Y . The above property looks quite similar to

$$(4) \quad E(Y | a \leq Y \leq b) = \frac{1}{2}(a + b), \quad \forall a, b \in \text{supp}(Y),$$

which was considered by Herer [3]. As shown there, (4) (also without any

restrictions on distributions) characterizes uniform continuous and discrete laws. Conditions (3) and (4) may seem similar, but there are important differences, besides different right hand sides and different conditionings. Firstly, a and b from (3) may not belong to the support of Y , while a and b in (4) are points of the support of Y . Secondly, condition (3) holds almost surely, while (4) is valid for each point in the support of Y . Much closer to Herer's characterization is the result by Das Gupta, Goswami and Rao [2] who characterized uniform continuous and discrete distributions in terms of order statistics by the condition

$$E(Y_1 | Y_{(1)}, Y_{(n)}) = \frac{1}{2} (Y_{(1)} + Y_{(n)}) \quad \text{a.s.,}$$

where $Y_{(1)}$ and $Y_{(n)}$ are, respectively, the minimum and maximum of i.i.d. rv's Y_1, \dots, Y_n .

Let X be a real random variable with distribution μ (written $X \sim \mu$). We say that X is *uniformly distributed* on $[0, 1]$ if either

- X is continuous with density function $f \equiv 1$, or
- X is discrete and there exist x_1, \dots, x_n and $r \neq 0$ such that $x_i = x_1 + r(i - 1)$, $\mu(\{x_i\}) = 1/n$, $i = 1, \dots, n$, and $S(\mu) = \{x_1, \dots, x_n\}$ (see [3]).

Our main result is the following.

THEOREM 1. *Let X, Y be i.i.d. random variables having distribution μ on \mathbb{R} . If $E|X| < \infty$ and conditions (1), (2) hold, then either*

- $K = 1/3$ and μ is a continuous uniform distribution on $[0, 1]$, or
- there exists n such that $K = K_n = \frac{1}{6n^2}(n + 1)(2n + 1)$ and μ is a discrete uniform distribution with support $\{i/n : i = 1, \dots, n\}$.

The paper is organized as follows. In Section 2 conditions equivalent to (1) and (2) are stated, and we obtain a restriction on the number of non-isolated atoms (Lemma 2.1). The discrete case is considered separately in Section 3. The results of this section are also useful in considering the general case with no restrictions on μ . In Section 4 we show that the support of μ consists of intervals and a finite number of atoms, and we present the proof of the main theorem, which, in a sense, sums up the results of the preceding lemmas.

2. Preliminary facts. Throughout the paper $S(\mu)$ denotes the *support* of a measure μ , that is,

$$(5) \quad S(\mu) = \{x \in \mathbb{R} : \mu(x - \varepsilon, x + \varepsilon) > 0 \text{ for any } \varepsilon > 0\},$$

and $S_d(\mu)$ stands for the set of atoms. Let μ be the distribution of X and let F be its (right continuous) distribution function.

Using well known facts about conditional expectation we introduce conditions that are equivalent to (1) and (2) respectively:

$$(6) \quad F(1 - X) = 1 - X \quad \text{a.s.},$$

$$(7) \quad -X^2 + EY^2 + 2XEY - 2E[YI(X + Y > 1) | X] + X = K \quad \text{a.s.}$$

We say that $x \in \mathbb{R}$ satisfies (6) and (7) if

$$F(1 - x) = 1 - x, \\ -x^2 + EY^2 + 2xEY - 2E[YI(Y > 1 - x)] + x = K.$$

Since (6) and (7) may not hold for some points of $S(\mu)$, we will need the following proposition.

PROPOSITION 2.1. *For every convergent sequence $\{x_n\} \subset S(\mu)$ with $x_n \rightarrow x$, there exists a sequence $\{\tilde{x}_n\} \subset S(\mu)$ such that $\tilde{x}_n \rightarrow x$ and \tilde{x}_n satisfies conditions (6), (7) for $n = 1, 2, \dots$*

Proof. Let $A = S(\mu) \cap \{x \in \mathbb{R} : x \text{ satisfies (6), (7)}\}$. Since $S(\mu)$ is closed, A is a Borel set. The definition of $S(\mu)$ implies that $F(y + \varepsilon) - F(y - \varepsilon) > 0$ for any $y \in S(\mu) \cap A^c$ and $\varepsilon > 0$. Since $P(X \in A) = 1$, for any $\varepsilon > 0$ there exists $y_\varepsilon \in (y - \varepsilon, y + \varepsilon)$ satisfying conditions (6) and (7).

Consider a sequence $\{\varepsilon_n\}$ such that $\varepsilon_n \downarrow 0$. From the above, with an element $x_n \in S(\mu) \cap A^c$ one can associate $\tilde{x}_n \in (x_n - \varepsilon_n, x_n + \varepsilon_n)$ satisfying (6) and (7) for $n = 1, 2, \dots$. Obviously $\tilde{x}_n \rightarrow x$. ■

Note that $A \subset [0, 1]$ and $\mu(A) = 1$. Moreover, every atom in $S(\mu)$ belongs to A . In order to specify the support structure of μ we examine the possible nature of atoms belonging to $S(\mu)$.

LEMMA 2.1. *Let μ be a measure satisfying (6) and (7). There exist only a finite number of atoms that are accumulation points of $S(\mu)$.*

Proof. Let a be an atom. With no loss of generality we can assume that $a \notin \{0, 1\}$. Suppose that there exist two sequences $\{a_n\}, \{c_n\} \subset S(\mu)$ convergent to a such that $a_n \uparrow a$ and $c_n \downarrow a$. Then (6) ensures the existence of sequences $\{1 - a'_n\}, \{1 - c'_n\} \subset S(\mu)$ of elements satisfying (6) such that $a'_n \uparrow a$ and $c'_n \downarrow a$ (Proposition 2.1). Applying (6) to $1 - a'_n$ and $1 - c'_n$ we have $P(X = a) \leq P(X \in (a'_n, c'_n]) = c'_n - a'_n$. Since the right-hand side converges to 0, we obtain $P(X = a) = 0$.

Therefore we shall consider only the cases that a is a left or right accumulation point.

Left accumulation point. Let a be a left accumulation point of $S(\mu)$. Then one can find a sequence $\{a_n\} \subset S(\mu)$ with $a_n \uparrow a$ and $\varepsilon > a$ such that $(a, \varepsilon) \cap S(\mu) = \emptyset$. Hence using (6), we can assert that there exists $\delta > 0$ such that $(\delta, 1 - a) \cap S(\mu) = \emptyset$. Let $\tilde{\varepsilon} = \sup\{\varepsilon \in (a, 1] : (a, \varepsilon) \cap S(\mu) = \emptyset\}$, and $1 - d = \inf\{\delta > 0 : (\delta, 1 - a) \cap S(\mu) = \emptyset\}$. Applying (6) twice we deduce

that there exists a sequence $\{1 - b_n\} \subset S(\mu)$ with $1 - b_n \downarrow 1 - a$. Without loss of generality, one can assume that the elements of this sequence satisfy (6) and (7).

The proof will be divided into three parts:

- (i) $\tilde{\varepsilon} < d$,
- (ii) $\tilde{\varepsilon} > d$,
- (iii) $\tilde{\varepsilon} = d$.

CASE (i). Let us first show that $\tilde{\varepsilon}$ is an atom. Suppose it is not true. Then one can prove that $(1 - d, 1 - a) \cap S(\mu) \neq \emptyset$, which contradicts the definition of $1 - d$. Applying (6) to $\tilde{\varepsilon}$ and a gives $P(X \in (1 - \tilde{\varepsilon}, 1 - a]) = \tilde{\varepsilon} - a$. Since $(1 - \tilde{\varepsilon}, 1 - a) \cap S(\mu) = \emptyset$, it follows that $P(X = 1 - a) = \tilde{\varepsilon} - a > 0$. Hence $1 - a$ is an atom.

Suppose that $1 - d \in S_d(\mu)$. This enables us to apply (6) to $1 - d$ and $1 - a$:

$$(8) \quad P(X \in (a, d]) = d - a.$$

Considering the sequence $\{1 - b_n\}$ and $1 - d$ we obtain

$$(9) \quad P(X \in (b_n, d]) = P(X \in (b_n, a]) + P(X \in (a, d]),$$

which yields $P(X = a) \leq a - b_n$. Since the right-hand side converges to 0 as $n \rightarrow \infty$, $P(X = a) = 0$.

If $1 - d \notin S_d(\mu)$ we proceed in the same manner. We define a sequence $\{1 - d_n\} \subset S(\mu)$ with $1 - d_n \uparrow 1 - d$ and with elements satisfying (6), (7), and instead of $1 - d$, we consider $1 - d_n$. As a result, we get

$$P(X \in (b_n, d_n]) = P(X \in (b_n, a]) + P(X \in (a, d]) + P(X \in (d, d_n]),$$

which yields $P(X = a) \leq d_n - d + a - b_n$. The same argument as above implies $P(X = a) = 0$.

CASE (ii). Using (6), it can be easily checked that $P(X = 1 - d) > 0$. We apply (7) to $1 - b_n$ and $1 - d$ as follows:

$$(10) \quad -(1 - b_n)^2 + EY^2 + 2(1 - b_n)EY - 2E[YI(Y > b_n)] + 1 - b_n = K,$$

$$(11) \quad -(1 - d)^2 + EY^2 + 2(1 - d)EY - 2E[YI(Y > d)] + 1 - d = K.$$

Subtracting (10) from (11) we obtain

$$(d - b_n)(2 - b_n - d) - 2(d - b_n)EY + 2E[YI(b_n < Y \leq d)] - (d - b_n) = 0.$$

It follows that $E[YI(b_n < Y \leq d)] = \frac{1}{2}(d - b_n)(b_n + d - 1 + 2EY)$. Letting $n \rightarrow \infty$, we have

$$(12) \quad E[YI(a \leq Y \leq d)] = \frac{1}{2}(d - a)(a + d - 1 + 2EY).$$

Since there is no atom in $(a, d]$, the above equation is equivalent to

$$(13) \quad aP(X = a) = \frac{1}{2}(d - a)(a + d - 1 + 2EY).$$

On the other hand, applying (6) to $1 - b_n$ and $1 - d$, we have $P(X \in (b_n, d]) = d - b_n$. So taking $n \rightarrow \infty$ we get $P(X \in [a, d]) = d - a$. Since $P(X \in [a, d]) = P(X = a)$, (13) leads to $d = a + 1 - 2EY$. Hence

$$(14) \quad P(X = a) = 1 - 2EY.$$

As a consequence, every atom that is not left-isolated has the same probability mass.

CASE (iii). Assume that $P(X = \tilde{\varepsilon}) = 0$. If this is not true, by the same method as in (i), one can prove that $P(X = a) = 0$.

The definition of $\tilde{\varepsilon}$ implies that there exists a sequence $\{\varepsilon_n\}$ with $\varepsilon_n \downarrow d = \tilde{\varepsilon}$. From (6) we can define a sequence $\{1 - \varepsilon'_n\} \subset S(\mu)$ such that $\varepsilon'_n \downarrow d$ and $1 - \varepsilon'_n$ satisfies (6), (7) (Proposition 2.1). By (7) we get

$$(\varepsilon'_n - b_n)(2 - \varepsilon'_n - b_n) - 2(\varepsilon'_n - b_n)EY + 2E[YI(b_n < Y \leq \varepsilon'_n)] - (\varepsilon'_n - b_n) = 0,$$

which yields

$$E[YI(b_n < Y \leq \varepsilon'_n)] = \frac{1}{2}(\varepsilon'_n - b_n)(2EY - 1 + \varepsilon'_n + b_n).$$

A passage to the limit implies that

$$(15) \quad E[YI(a \leq Y \leq d)] = \frac{1}{2}(d - a)(2EY - 1 + a + d).$$

Since $P(X \in [a, d]) = d - a$ and $(a, d) \cap S(\mu) = \emptyset$ from the above we get (14) again. Note that each atom that is not left-isolated has probability equal to $1 - 2EY$. Thus we have actually proved that there are a finite number of atoms that are not left-isolated.

Right accumulation point. Our next goal is to determine the number of atoms that are not right-isolated. Let a be such an atom. One can define sequences $\{a_n\}, \{1 - b_n\} \subset S(\mu)$ such that $a_n, b_n \downarrow a$. With no loss of generality we can assume that $a_n, 1 - b_n$ satisfy (6) and (7) for $n = 1, 2, \dots$. We can now proceed analogously to the case of the left accumulation point.

CASE (i). Since $\tilde{\varepsilon} < d$, one can verify that $P(X = 1 - d) > 0$. Applying (7) to $1 - d$ and $1 - b_n$ we have

$$(16) \quad E[YI(d < Y \leq b_n)] = \frac{1}{2}(b_n - d)(b_n + d - 1 + 2EY).$$

Letting $n \rightarrow \infty$, we obtain

$$(17) \quad E[YI(d < Y \leq a)] = \frac{1}{2}(a - d)(a + d - 1 + 2EY).$$

From (6) we have $P(X \in (d, a]) = d - a$. Thus (17) gives $P(X = a) = 1 - 2EY$.

CASE (ii). One can show that $P(X = \tilde{\varepsilon}) > 0$. Let us apply (6) to a and $\tilde{\varepsilon}$: $P(X \in (1 - a, 1 - \tilde{\varepsilon}]) = \tilde{\varepsilon} - a$. Thus one can infer that $1 - \tilde{\varepsilon}$ is an atom, a contradiction, since $(1 - a, 1 - d) \cap S(\mu) = \emptyset$.

CASE (iii). Suppose there is an atom at $\tilde{\varepsilon} = d$. By (6) we get

$$(18) \quad P(X \in (1 - a, 1 - d]) = a - d,$$

which gives $P(X = 1 - d) = d - a$. We are now in a position to apply (6) to $1 - d$ and $1 - b_n$. Thus $P(X \in (d, b_n]) = b_n - d$. Letting $n \rightarrow \infty$ we have $P(X \in (d, a]) = a - d$. Since $1 - d$ and $1 - b_n$ satisfy (7), for $n = 1, 2, \dots$ we obtain

$$(19) \quad E[YI(d < Y \leq b_n)] = \frac{1}{2} (b_n - d)(b_n + d - 1 + 2EY).$$

The fact that $b_n \downarrow a$ yields

$$(20) \quad E[YI(d < Y \leq a)] = \frac{1}{2} (a - d)(a + d - 1 + 2EY),$$

and we get $P(X = a) = 1 - 2EY$.

If $P(X = d) = 0$, we proceed in the same manner. Consider a sequence $\{d_n\} \subset S(\mu)$ with elements satisfying (6) and such that $d_n \uparrow d$. Applying (6) to a and d_n we obtain $P(X \in (1 - a, 1 - d_n]) = a - d_n$. Letting $n \rightarrow \infty$, we get $P(X \in (1 - a, 1 - d]) = a - d$. Hence $P(X = 1 - d) > 0$. We can now proceed analogously to the preceding case. Using (7) for $1 - d$ and $1 - b_n$ we obtain $P(X = a) = 1 - 2EY$. By the argument of equal probability mass, we obtain at most a finite number of atoms that are not right-isolated. ■

Let $s = \inf\{x : x \in S(\mu)\}$ and $t = \sup\{x : x \in S(\mu)\}$. As a direct conclusion from the proof of Lemma 2.1 we obtain the following result.

REMARK 2. Let μ be a measure satisfying (6) and (7). Then every non-isolated atom belonging to $S(\mu) \setminus \{s, t\}$ has probability equal to $1 - 2EY$.

3. Discrete distribution. In this section we assume that μ is a discrete measure. Under this assumption we will show that μ is a uniform distribution. First, we prove an auxiliary fact.

PROPOSITION 3.1. *Let μ be a measure satisfying (6) and (7). Then there exists a sequence $\{a_n\}$ of atoms such that either $(a_n, a_{n+1}) \cap S_d(\mu) = \emptyset$ ($a_n < a_{n+1}$) or $(a_{n+1}, a_n) \cap S_d(\mu) = \emptyset$ ($a_{n+1} < a_n$), $n = 1, 2, \dots$*

Proof. We consider only the case of an infinite number of atoms. Since the number of non-isolated atoms is finite, there exists an interval I containing an infinite number of isolated atoms. Suppose that in I there exists

an increasing sequence of atoms (the decreasing case can be treated analogously). Let $a \in S_d(\mu) \cap I$. We construct the sequence $\{a_n\}$ as follows: $a_1 = a$ and

$$(21) \quad a_{n+1} = \sup\{r > a_n : (a_n, r) \cap S_d(\mu) = \emptyset\}, \quad n = 1, 2, \dots$$

Our aim is to show that $a_2 \in S_d(\mu)$. Suppose to the contrary that there is no atom at a_2 . Then we are able to find a sequence $\{z_n\} \subset S_d(\mu)$ with $z_n \downarrow a_2$. This in turn implies the existence of $\{1 - \tilde{z}_n\} \subset S_d(\mu)$ with $1 - \tilde{z}_n \uparrow 1 - a_2$. Applying (6) to z_n and a_1 we have $P(X \in (1 - z_n, 1 - a_1]) = z_n - a_1$. Letting $n \rightarrow \infty$, we obtain $P(X \in [1 - a_2, 1 - a_1]) = a_2 - a_1$. Since atoms in I are isolated, one can prove that they are also isolated in I' (interval symmetric to I with respect to $1/2$), hence $P(X = 1 - a_2) = 0$. Thus $P(X \in (1 - a_2, 1 - a_1]) = a_2 - a_1$, which yields the existence of an atom in $(1 - a_2, 1 - a_1]$, say $1 - b$. From (6) for $1 - \tilde{z}_n$ and $1 - b$ we get $P(X \in (b, \tilde{z}_n]) = \tilde{z}_n - b$, and a passage to the limit indicates that $P(X \in (b, a_2]) = a_2 - b$. According to (21), we obtain $P(X = a_2) > 0$. This contradicts our assumption that a_2 is not an atom. We now proceed by induction to construct the required sequence $\{a_n\}$. ■

The result stated above enables us to give a restriction on the number of atoms in $S_d(\mu)$. The following lemma is crucial for further reasoning.

LEMMA 3.1. *A discrete measure μ satisfying conditions (6) and (7) consists of a finite number of atoms.*

Proof. Suppose that there exist infinitely many atoms in $(0, 1)$. Hence there exists a convergent sequence $\{a_n\} \subset S_d(\mu)$. Let a be its accumulation point. In view of Proposition 3.1, one can choose $\{a_n\}$ in such a way that there are no atoms between consecutive elements of the sequence. We apply (7) to a_n and a_{n+1} :

$$(22) \quad -a_n^2 + EY^2 + 2a_n EY - 2E[YI(Y > 1 - a_n)] + a_n = K,$$

$$(23) \quad -a_{n+1}^2 + EY^2 + 2a_{n+1} EY - 2E[YI(Y > 1 - a_{n+1})] + a_{n+1} = K.$$

Subtracting (23) from (22) we get

$$(24) \quad a_{n+1}^2 - a_n^2 + 2(a_n - a_{n+1})EY + 2E[YI(1 - a_n \geq Y > 1 - a_{n+1})] + a_n - a_{n+1} = 0.$$

On the other hand, by (6) we have

$$(25) \quad P(X \in (1 - a_{n+1}, 1 - a_n]) = a_{n+1} - a_n,$$

which indicates the existence of atoms in $(1 - a_{n+1}, 1 - a_n]$. It can be derived from (6) that $(1 - a_{n+1}, 1 - a_n) \cap S_d(\mu) = \{1 - a_n\}$.

Indeed, suppose that there exist two atoms $x, y \in (1 - a_{n+1}, 1 - a_n]$. For one of them (e.g. x) we have $x \in (1 - a_{n+1}, 1 - a_n)$. Using (6) for x, y we get $P(X \in (1 - y, 1 - x]) = y - x$. Since $(1 - y, 1 - x] \subset (a_n, a_{n+1})$, one can find

an atom in (a_n, a_{n+1}) , which contradicts the definition of the sequence $\{a_n\}$ ($(a_n, a_{n+1}) \cap S_d(\mu) = \emptyset$). Thus there is only one atom in $(1 - a_{n+1}, 1 - a_n]$. We denote it by $1 - b_n$. From (25) we obtain $P(X = 1 - b_n) = a_{n+1} - a_n$ and

$$(26) \quad E[YI(1 - a_n \geq Y > 1 - a_{n+1})] = (1 - b_n)(a_{n+1} - a_n).$$

Using (25) and (26) we get

$$(27) \quad a_{n+1} + a_n - 2EY + 2(1 - b_n) - 1 = 0,$$

which yields

$$(28) \quad EY = \frac{a_{n+1} + a_n - 2b_n + 1}{2}.$$

The sequences $\{a_n\}$ and $\{b_n\}$ converge to a , so the right-hand side of (28) converges to $1/2$. Thus $EY = 1/2$. As a result we have

$$(29) \quad a_{n+1} - b_n = b_n - a_n, \quad n = 1, 2, \dots$$

Applying the same steps to the sequence $\{1 - b_n\}$ we derive the equality

$$(30) \quad b_n - a_n = a_n - b_{n-1}, \quad n = 2, 3, \dots$$

Now from (29) and (30) one gets

$$a_n - b_{n-1} = b_{n-1} - a_{n-1} = a_n - a_{n-1}, \quad n = 2, 3, \dots,$$

which yields

$$a_{n+1} - a_n = b_n - b_{n-1} = a_n - a_{n-1}, \quad n = 2, 3, \dots$$

Hence the points of the sequence $\{a_n\}$ are equispaced. This contradicts the assumption of infinitely many atoms in $S(\mu)$. ■

The next lemma gives the main result for the discrete case.

LEMMA 3.2. *Let X be a random variable having distribution μ . If μ satisfies (6) and (7) then X is uniformly distributed, and there exists n such that μ is supported on the set $\{i/n : i = 1, \dots, n\}$.*

Proof. From Lemma 3.1, there exists n such that $S(\mu) = \{x_1, \dots, x_n\}$, $x_i \in [0, 1]$, $i = 1, \dots, n$. With no loss of generality, we can assume that $0 \leq x_1 < \dots < x_n \leq 1$.

Suppose $x_1 + x_n > 1$. Since $1 - x_n < x_1$, we get $1 - x_n = F(1 - x_n) = 0$, hence $x_n = 1$. From (6) we deduce that $x_{n-k} \in (1 - x_{k+1}, 1 - x_k]$ for $k = 1, \dots, n - 1$, and x_{n-k} is the only atom belonging to $(1 - x_{k+1}, 1 - x_k]$. Applying (7) to x_k and x_{k+1} we obtain

$$(31) \quad x_{k+1}^2 - x_k^2 + 2(x_k - x_{k+1}) + 2E[YI(1 - x_{k+1} < Y \leq 1 - x_k)] + x_k - x_{k+1} = 0,$$

and since $E[YI(1 - x_{k+1} < Y \leq 1 - x_k)] = x_{n-k}(x_{k+1} - x_k)$, we get

$$\begin{aligned} x_k + x_{k+1} - 2EY + 2x_{n-k} - 1 &= 0, & k = 1, \dots, n - 1, \\ x_n &= 1. \end{aligned}$$

It follows that

$$(32) \quad \begin{aligned} x_2 &= 2(x_n - x_{n-1}), \\ x_k - x_{k-2} &= 2(x_{n-k+2} - x_{n-k+1}), & k = 3, \dots, n. \end{aligned}$$

By induction one can show that

$$(33) \quad x_k = \frac{k(k-1)}{2}x_2 - k(k-2)x_1, \quad k = 1, \dots, n.$$

Applying (33) to (32) for $k = 2$, we have $x_2 = 2x_1$, which yields $x_k = k/n$ for $k = 1, \dots, n$.

Let $p_i = P(X = x_i)$, $i = 1, \dots, n$. Condition (6) leads to

$$(34) \quad \begin{aligned} F(1 - x_{n-k}) &= 1 - p_n - p_{n-1} - \dots - p_{k+1} \\ &= 1 - x_{n-k}, & k = 1, \dots, n - 1, \end{aligned}$$

and by an easy computation we get $p_{n-k} = x_{k+1} - x_k$, $k = 1, \dots, n - 1$, $p_n = x_1$. Hence $x_k = k/n$ and $p_{n-k} = x_k - x_{k-1} = 1/n$ for $k = 1, \dots, n$.

Applying the same calculation to the case $x_1 + x_n \leq 1$ we obtain $S(\mu) = \{0, 1/n, \dots, (n - 1)/n\}$. Considering equations analogous to (34), we get

$$\frac{2}{n} = F\left(\frac{1}{n}\right) = F\left(1 - \frac{n-1}{n}\right) = 1 - \frac{n-1}{n} = \frac{1}{n}.$$

This is a contradiction, which completes the proof. ■

4. General case. In this section no additional restriction on μ is given. We consider a measure that can be a mixture of discrete and continuous distributions.

PROPOSITION 4.1. *Let X be a random variable having distribution μ . If μ satisfies (6), (7), and there exists an interval $(a, b) \subset S(\mu)$, then $EX = 1/2$.*

Proof. From (6) we conclude that intervals contained in $S(\mu)$ are symmetric with respect to $1/2$. Since there are a finite number of atoms that are not isolated (Lemma 2.1), and by definition there are no isolated atoms in (a, b) , there exists an interval $(c, d) \subset (a, b)$ such that $(c, d) \cap S_d(\mu) = \emptyset$ and $(1 - d, 1 - c) \cap S_d(\mu) = \emptyset$. For any $1 - z \in (1 - d, 1 - c)$, we define a sequence $\{1 - z_n\} \subset S(\mu)$ such that $1 - z_n \uparrow 1 - z$ and $1 - z_n$ satisfies (6) for $n = 1, 2, \dots$. A passage to the limit as $n \rightarrow \infty$ yields $F(z) = z$ for any $z \in (c, d)$. Considering two points $x_1, x_2 \in (c, d)$ ($x_1 < x_2$) we may construct corresponding sequences $\{x_{1n}\}, \{x_{2n}\} \subset S(\mu)$ with elements satisfying (6) and (7) such that $x_{1n} \rightarrow x_1$ and $x_{2n} \rightarrow x_2$. Applying (7) to $\{x_{1n}\}, \{x_{2n}\}$

and taking the limit as $n \rightarrow \infty$, we obtain the formula

$$(35) \quad x_2^2 - x_1^2 + 2(x_1 - x_2)EY + 2E[YI(1 - x_2 < Y \leq 1 - x_1)] + x_1 - x_2 = 0.$$

Since $F(x) = x$ for $x \in (1 - d, 1 - c)$, we have $EYI(1 - x_2 < Y \leq 1 - x_1) = \frac{1}{2}((1 - x_1)^2 - (1 - x_2)^2)$. Thus from (35) we get $EY = 1/2$. ■

LEMMA 4.1. *Let X be a random variable having distribution μ . If μ satisfies (6), (7) and there exist a, b such that $P(X \in (a, b)) = b - a$ and $(a, b) \cap S_d(\mu) = \emptyset$, then $F(x) = x$ for $x \in (a, b)$.*

Proof. Suppose $a < b$. Then there exists $x \in S(\mu) \setminus S_d(\mu)$ such that $a < x < b$. Define

$$(36) \quad x_* = \sup\{y < x : \exists \varepsilon > 0 (y - \varepsilon, y) \cap S(\mu) = \emptyset\},$$

$$(37) \quad x^* = \inf\{y > x : \exists \varepsilon > 0 (y, y + \varepsilon) \cap S(\mu) = \emptyset\},$$

and $(x_1, x_2) = (x_*, x^*) \cap (a, b)$. From (36) and (37) we conclude that the set $S(\mu) \setminus S_d(\mu)$ is dense in (x_1, x_2) , and consequently in $(1 - x_2, 1 - x_1)$. Let $z \in (x_1, x_2)$. There exists a sequence $\{1 - z_n\}$ of elements satisfying (6) such that $1 - z_n \uparrow 1 - z$. We thus get $F(z_n) = z_n$, $n = 1, 2, \dots$, and $z_n \rightarrow z$. Since F is continuous on $(x_1, x_2) \ni z$, letting $n \rightarrow \infty$, we obtain $F(z) = z$. Therefore X is absolutely continuous on $(x_1, x_2) \cup (1 - x_2, 1 - x_1)$. It remains to prove that $x_1 = a$ and $x_2 = b$. Suppose $x_1 > a$ and $x_2 < b$. We have already noticed that with every point $\tilde{x} \in S(\mu) \setminus S_d(\mu)$ one can associate an interval $I(\tilde{x}) = (\tilde{x}_1, \tilde{x}_2) \ni \tilde{x}$ such that $I(\tilde{x}) \subset S(\mu)$. It follows that $(a, b) \cap S(\mu)$ consists of intervals and $F(x) = x$ on each interval. The assumption $P(X \in (a, b)) = b - a$ implies $x_1 = a$ and $x_2 = b$. Hence $F(x) = x$ for (a, b) . ■

REMARK 3. If F is continuous on $S(\mu)$ and (6), (7) hold, then F is absolutely continuous with density function $f \equiv 1$ on $(0, 1)$.

LEMMA 4.2. *If a measure μ has properties (6) and (7), then there exist only a finite number of atoms in $S(\mu)$.*

Proof. Suppose this is not true. We will define a convergent sequence $\{b_n\} \subset S_d(\mu)$ such that $(b_n, b_{n+1}) \cap S_d(\mu) = \emptyset$. One can find an interval I which I contains an infinite number of atoms and every atom in $I \cup I'$ is an isolated point in $S(\mu)$. (Recall that I' denotes the interval symmetric to I with respect to $1/2$.) Let b be an atom in I . We construct $\{b_n\}$ as follows: $b_1 = b$ and

$$(38) \quad b_{n+1} = \sup\{r > b_n : (b_n, r) \cap S_d(\mu) = \emptyset\}, \quad n = 1, 2, \dots$$

We shall show that $P(X = b_2) > 0$. Suppose that $P(X = b_2) = 0$. Then there exists a sequence $\{z_n\} \subset S(\mu)$ such that $z_n \downarrow b_2$ and z_n satisfies (6), (7) for $n = 1, 2, \dots$. From (6) we have $P(X \in (1 - z_n, 1 - b_1]) = z_n - b_1$. Hence

$P(X \in [1 - b_2, 1 - b_1]) = b_2 - b_1$. Since $1 - b_2$ is not an isolated point in $S(\mu)$, we obtain $P(X = 1 - b_2) = 0$. Then there exists an atom $1 - b \in (1 - b_2, 1 - b_1]$, since in the other case, by Lemma 4.1, $(1 - b_2, 1 - b_1) \subset S(\mu)$, and as a consequence $(b_1, b_2) \subset S(\mu)$, contrary to the assumption that there is an isolated atom at b_1 .

As $P(X = 1 - b_2) = 0$, we define a sequence $\{1 - z'_n\} \subset S(\mu)$ such that $1 - z'_n \uparrow 1 - b_2$ and $1 - z'_n$ satisfies (6) for $n = 1, 2, \dots$. Applying (6) to $1 - b$ and $1 - z'_n$ we obtain $P(X \in (b, z'_n]) = z'_n - b$. A passage to the limit as $n \rightarrow \infty$ gives $P(X \in (b, b_2]) = b_2 - b$. Since $(b, b_2) \cap S_d(\mu) = \emptyset$, Lemma 4.1 implies that $F(x) = x$ for $x \in (b, b_2)$. Consequently, $F(x) = x$ for $x \in (1 - b_2, 1 - b)$. The existence of an atom at $1 - b$ contradicts our assumption that every atom in I' is an isolated point. Hence there is an atom at b_2 . By induction we can prove that there exists a sequence $\{b_n\} \subset S_d(\mu)$ such that $(b_n, b_{n+1}) \cap S_d(\mu) = \emptyset$.

In view of Lemma 3.1, since there are infinitely many atoms, there exist $x \in S(\mu) \setminus S_d(\mu)$ and b_k, b_{k+1} such that $b_k < x < b_{k+1}$ (atoms $b_k, b_{k+1} \in \{b_n\}$); otherwise applying Lemma 3.1 to a subinterval of I such that $\{I \cup I'\} \cap S(\mu) \setminus S_d(\mu) = \emptyset$ leads to contradiction. Define

$$x_1 = \inf\{x \in S(\mu) \setminus S_d(\mu) : x \in (b_k, b_{k+1})\},$$

$$x_2 = \sup\{x \in S(\mu) \setminus S_d(\mu) : x \in (b_k, b_{k+1})\}.$$

From Lemma 4.1 one can deduce that $F(x) = x$ for $x \in (x_1, x_2)$. Since b_k, b_{k+1} are isolated points in $S(\mu)$, we have $(b_k, x_1) \cap S(\mu) = \emptyset$ and $(x_2, b_{k+1}) \cap S(\mu) = \emptyset$. Consider a sequence $\{x_{1n}\} \subset S(\mu)$ such that $x_{1n} \downarrow x_1$ and x_{1n} satisfies (6) for $n = 1, 2, \dots$. Applying (6) to $1 - x_{1n}$ and $1 - b_k$, we get

$$(39) \quad P(X \in [1 - x_1, 1 - b_k]) = x_1 - b_k.$$

One can prove that $[1 - x_1, 1 - b_k] \cap S_d(\mu) = \{1 - b_k\}$ and $P(X = 1 - b_k) = x_1 - b_k$. On the other hand, there exists a sequence $\{x'_{1n}\} \subset S(\mu)$ such that $x'_{1n} \downarrow x_1$ and x'_{1n} satisfies (7) for $n = 1, 2, \dots$. Applying (7) to b_k and x'_{1n} we get

$$(x'_{1n})^2 - b_k^2 + 2(b_k - x'_{1n})EY + 2E[YI(1 - b_k \geq X > 1 - x'_{1n})] + b_k - x'_{1n} = 0.$$

By passing to the limit as $n \rightarrow \infty$, and using (39), we obtain

$$(40) \quad x_1 + b_k - 2EY + 2(1 - b_k) - 1 = 0.$$

From Proposition 4.1 we get $EY = 1/2$, which yields $x_1 = b_k$. Thus there is no atom at b_k . This contradicts our assumption that $S(\mu)$ has an infinite number of atoms. ■

LEMMA 4.3. *The measure μ with properties (6), (7) is either atomic or absolutely continuous.*

Proof. We have actually proved that the support of μ consists of intervals and a finite number of atoms (Lemma 4.2). From the proof of Lemma 4.1 we deduce that intervals contained in $S(\mu)$ are symmetric with respect to $1/2$. Furthermore, one can verify that between two intervals there must be an element belonging to $S(\mu)$. Since the support structure is determined we shall concentrate on the following mutually exclusive cases: (1) atoms exist between intervals in $S(\mu)$, (2) there is only one interval in $S(\mu)$ surrounded by atoms.

The first case, in view of the proof of Lemma 4.2, leads to a contradiction. Let us consider the second case. It suffices to consider an interval $(c, 1 - c)$ and two atoms $P(X = c) > 0$ and $P(X = 1 - c) > 0$. (Using (6) and (7) one can prove that there is no atom in $[0, c) \cup (1 - c, 1]$.) We shall show that $P(X = c) = P(X = 1 - c) = 0$. Applying (6) to c and $1 - c$ we get $P(X \in (c, 1 - c]) = 1 - 2c$. Since $F(x) = x$ for $x \in (c, 1 - c)$ we have $P(X \in (c, 1 - c)) = 1 - 2c$, which yields $P(X = 1 - c) = 0$. Hence using the fact that $EY = 1/2$ we obtain $P(X = c) = 0$. It follows that the existence of an interval in $S(\mu)$ implies absence of atoms. ■

Proof of Theorem 1. From Lemma 4.3 we conclude that μ is purely discrete or absolutely continuous. By Lemma 3.2 the case that μ is discrete leads to a uniform distribution. Supposing that the support of μ consists of n atoms, we obtain $S(\mu) = \{1/n, \dots, 1\}$. The result stated in Remark 2 for continuous distribution implies that μ has density function $f \equiv 1$ on $(0, 1)$. This completes the proof of Theorem 1.

Acknowledgements. I wish to thank Prof. J. Wesolowski for valuable comments and suggestions.

References

- [1] Y. Baryshnikov, B. Eisenberg and W. Stadge, *Independent variables with independent sum and difference: S^1 -case*, J. Multivariate Anal. 45 (1993), 161–170.
- [2] S. Das Gupta, A. Goswami and B. Rao, *On a characterization of uniform distributions*, J. Multivariate Anal. 44 (1993), 102–114.
- [3] W. Herer, *A characterization of uniformly distributed random variable*, Demonstratio Math. 26 (1993), 207–212.
- [4] G. M. Feldman, *On the Skitovich–Darmois theorem on Abelian groups*, Theory Probab. Appl. 37 (1992), 621–631.
- [5] —, *Arithmetic of Probability Distributions and Characterization Problems on Abelian Groups*, Amer. Math. Soc., Providence, RI, 1993.
- [6] —, *A characterization of Gaussian distributions on Abelian groups by constancy of regression*, Theory Probab. Appl. 43 (1999), 477–480.
- [7] —, *A characterization of the Gaussian distribution on Abelian groups*, Probab. Theory Related Fields 126 (2003), 91–102.

- [8] G. M. Feldman and P. Graczyk, *On the Skitovitch–Darmois theorem for compact Abelian groups*, J. Theoret. Probab. 13 (2000), 859–869.
- [9] —, —, *Independent linear statistics on finite Abelian groups*, Ukrainian Math. J. 53 (2001), 499–506.
- [10] D. Neuenschwander and R. Schott, *The Bernstein and Skitovitch–Darmois characterization theorems for Gaussian distributions on groups, symmetric spaces and quantum groups*, Expo. Math. 15 (1997), 289–314.
- [11] J. H. Stapleton, *A characterization of the uniform distribution on a compact topological group*, Ann. Math. Statist. 34 (1963), 319–326.

Joanna Chachulska
Faculty of Mathematics and Information Science
Warsaw University of Technology
Pl. Politechniki 1
00-661 Warszawa, Poland
E-mail: J.Chachulska@mini.pw.edu.pl

Received December 7, 2004;
received in final form July 11, 2005

(7422)