# ON THE THEORY OF DISPERSIONS FOR PROBABILITY DISTRIBUTIONS ON COMPACT GROUPS

1970

SovietRxiv

## Abstract

## Full Text

UDC 519

V. M. MAKSIMOV

# ON THE THEORY OF DISPERSIONS FOR PROBABILITY DISTRIBUTIONS ON COMPACT GROUPS

*(Presented by Academician A. N. Kolmogorov on 21 XI 1969)*

The role of the concept of dispersion in classical probability theory is well known. In particular, it is indispensable in the study of questions connected with the summation of independent random variables. When considering analogous problems on compact groups, expressing any convergence properties in terms of numerical characteristics of distributions is also attractive, especially if one takes into account the nontriviality of multiplication in noncommutative groups in comparison with the usual addition of real numbers.

Let $G$ be an arbitrary compact group. Denote by $\mathfrak{M}$ the set of all Borel measures on $G$. Measures concentrated at elements of $G$ will be denoted by the same elements and often called shifts. As usual, $e_1$ is the identity in all groups under consideration. The invariant measure on a subgroup $g$, $g \subseteq G$, is denoted by $n_g$. Regarding the operation of composition of measures as multiplication in $\mathfrak{M}$, we obtain that $\mathfrak{M}$ is a compact associative semigroup.

We now single out the minimum necessary properties that a dispersion must possess in order to obtain, in general form, conditions for convergence of compositions of measures, and in order that, starting from these properties, one can describe a general construction of dispersions for the semigroup of measures $\mathfrak{M}$. For convenience we adopt the multiplicative form of dispersion (one can, obviously, pass to the customary additive form by taking logarithms). The description of dispersions for measures on finite groups is given in ([1]).

**Definition.** Let $E$ be a closed semigroup of measures containing the left and right shifts, $E \subseteq \mathfrak{M}$. Any real function $D$ on $E$ is called a **dispersion for distributions from** $E$ if the following conditions are fulfilled:

1°. $0 \leq D(\mu) < \infty$ for all $\mu \in E$ and $D \neq \text{const}$.

2°. $D$ is continuous, i.e., $D(\mu_n) \to D(\mu)$ if $\mu_n \to \mu$ weakly, $\mu_n, \mu \in E$.

3°. $D(\nu\mu) = D(\nu)D(\mu)$ for all $\mu, \nu \in E$.

4°. $D(n_g) = 0$ for all invariant measures from $E$, where $g \neq e_1$.

The following important properties of dispersions follow from 1°–4°.

a) $D(e_1) = 1$. Indeed, by 1°, $D \neq 0$. Therefore there exists a measure $\mu$, $\mu \in E$, such that $D(\mu) \neq 0$. But then from 3° we obtain:

$$D(\mu) = D(\mu e_1) = D(\mu)D(e_1) \neq 0,$$

whence $D(e_1) = 1$.

b) $D(a) = 1$ for any $a \in G$. Consider the set $\overline{\{a^n\}}$, $n = 1, \infty$. This will be a commutative subgroup of $G$. Therefore there exists a sequence $n_i$ such that $a^{n_i} \to e_1$ as $n_i \to \infty$. By 2°–3° and a) we have

$$D(a)^{n_i} = D(a^{n_i}) \to D(e_1) = 1.$$

Consequently, $D(a) = 1$.

c) $D(\mu) \leq 1$ for all $\mu \in E$. If $D(\mu) = 1$, then the measure $\mu$ is concentrated at one element of $G$.

Indeed, by the main result of ($^2$) it follows that there exist elements $a_n$ of the group $G$ such that

$$\mu^n a_n \to n_g.$$

If $\mu$ is not concentrated at one element of $G$, then $g \neq e_1$. Then by 2°–4° and b) we have:

$$D(\mu)^n = D(\mu^n) = D(\mu^n \cdot a_n) \to D(n_g) = 0,$$

i.e., $D(\mu) < 1$. If $\mu$ is concentrated at one element of $G$, then by b) $D(\mu) = 1$.

If $\mu \in E$ is the distribution of a random variable $\xi$, then we put $D(\xi) = D(\mu)$. The properties of $D(\xi)$, evidently, follow from $1^0$–$4^0$.

We now give an application of the notion of dispersion to the characterization of a certain property of a sequence of measures $\{\mu_n\}$. We shall say of the sequence $\{\mu_n\}$ that it is of type $e_1$ if every sequence $\mu_{n_i} \cdots \mu_{n_i+m_i}$, $m_i \geq 0$, with $n_i \to \infty$, has as limit points only shifts. A sequence of independent random variables is called of type $e_1$ if the corresponding sequence of measures is of type $e_1$. If the independent variables on $G$, $\{\xi_n\}$, are of type $e_1$, then it can be shown that there exist elements $a_n$ of $G$ such that, for the sequence $\{\xi_n'\}$, where $\xi_n' = a_n^{-1}\xi_n a_{n+1}$, the product $\xi_i' \cdots \xi_n'$ will converge almost everywhere as $n \to \infty$ for every $i$. Thus, in order to clarify almost everywhere convergence, it is necessary to know the type of the sequence.

**Proposition 1.** *Let the measures $\{\mu_n\}$ belong to $E$. In order that the sequence $\{\mu_n\}$ be of type $e_1$, it is necessary and sufficient that the series $\sum(1 - D(\mu_n))$ converge. (For sufficiency it is assumed that the series converges for at least one dispersion.)*

For arbitrary noncommutative groups it is easy to indicate semigroups of measures in $\mathfrak{M}$ for which a dispersion can be defined. However, this can no longer be done for the whole of $\mathfrak{M}$, defined on an arbitrary compact group $G$.

**Proposition 2.** *If a compact group $G$ is infinite-dimensional or zero-dimensional but contains an infinite number of elements, then no dispersion exists on the semigroup of measures $\mathfrak{M}$.*

**Proof.** The groups indicated in the proposition contain, in every neighborhood of the identity, subgroups (3). Therefore there exists a sequence of subgroups $g_i$, $g_i \neq e_1$, contracting to $e_1$. Consequently, $n_{g_i} \to e_1$ weakly. Suppose now that some dispersion $D$ is defined in $\mathfrak{M}$. Then, by property $4^0$, $D(n_{g_i}) = 0$, while by the continuity property of $D$,

$$\lim_{i\to\infty} D(n_{g_i}) = D(e_1) = 1,$$

which contradicts a).

It follows from Proposition 2 that compact groups admitting a generalization of the notion of dispersion are either finite groups or finite-dimensional groups, which are Lie groups (3). Since in (1) the dispersions of measures on finite groups are described, we shall give the general form and properties of dispersions of all distributions on compact Lie groups.

For what follows, the notion of a weak dispersion will be useful. By this we mean a function $D$ on $\mathfrak{M}$ satisfying only the first three conditions in the definition of dispersion. It will be shown that all dispersions are constructed from weak dispersions.

**Proposition 3.** *Let $G$ be an arbitrary compact group. If $D$ is a weak dispersion on $\mathfrak{M}$ of the group $G$, then there exists a normal divisor $N$ of the group $G$ for which $D(n_N) = 1$, and every subgroup $g$ for which $D(n_g) = 1$ is contained in $N$.*

For the proof of this proposition one can use the same scheme as in the proof of the analogous assertion in (1). For this it is only necessary to apply Zorn' s lemma.

We shall call the subgroup $N$ the kernel of the weak dispersion. We denote a weak dispersion with kernel $N$ by $D_N$. Thus a dispersion can be regarded as a weak dispersion with kernel $e_1$.

**Lemma 1.** *If the support of some measure $\nu$ is contained in $N$, then*

$$D_N(\nu) = 1.$$

Indeed, since the support of $\nu$ is contained in $N$, we have $\nu n_N = n_N$. Therefore, taking Proposition 3 into account,

$$1 = D_N(n_N) = D_N(\nu n_N) = D_N(\nu) \times D_N(n_N) = D_N(\nu).$$

The function $D(\nu) = D_{N_1}(\nu) D_{N_2}(\nu)$ is a weak dispersion.

**Lemma 2.** *The kernel of a weak dispersion $D$ equal to the product of weak dispersions with kernels $N_1$ and $N_2$ is equal to $N_1 \cap N_2$.*

Indeed, suppose that for some subgroup $g$ the value $D_{N_1}(n_g)D_{N_2}(n_g)$ is equal to $D(n_g) = 1$. Then $D_{N_1}(n_g) = D_{N_2}(n_g) = 1$. By Proposition 3, the subgroup $g$ must be contained both in $N_1$ and in $N_2$, i.e. $g \subseteq N_1 \cap N_2$. On the other hand, by Lemma 1, $D_{N_1}(n_{N_1 \cap N_2}) = D_{N_2}(n_{N_1 \cap N_2}) = 1$, and therefore $D(n_{N_1 \cap N_2}) = 1$. Consequently, $N_1 \cap N_2$ is the kernel of the weak dispersion $D = D_{N_1} D_{N_2}$.

**Corollary 1.** *If $N_1 \cap N_2 = e_1$, then the product $D_{N_1} D_{N_2}$ will be a dispersion.*

In particular, the product of dispersions by any weak dispersion again gives a dispersion. Thus, Corollary 1 shows that, in the formation of dispersions, weak dispersions play a large role. As will be seen from Proposition 4, this is not accidental. The totality of all weak dispersions with kernel $N$ forms a semigroup under multiplication. We shall denote this semigroup by $\mathcal{D}_N(G)$.

**Proposition 4.** $\mathcal{D}_N(G) \sim \mathcal{D}_{e_1}(G/N)$.

**Proof.** Let $\varphi$ be the natural mapping $G \to G/N$. If $\mu$ is a measure on $G$, then denote by $\varphi(\mu)$ the naturally induced measure on $G/N$. It is clear that

$$\varphi(\mu_1\mu_2) = \varphi(\mu_1)\varphi(\mu_2).$$

Suppose that for measures $\mu_1$ and $\mu_2$ on $G$ we have $\varphi(\mu_1) = \varphi(\mu_2)$. Then $\mu_1 n_N = \mu_2 n_N$, whence

$$D_N(\mu_1) = D_N(\mu_1 n_N) = D_N(\mu_2 n_N) = D_N(\mu_2).$$

That is, if we put $D(\varphi(\mu)) = D_N(\mu)$, then $D$ will be a single-valued function on all measures of the group $G/N$ and will satisfy items 1°–3°. However, $D$ also satisfies item 4°. To see this, note that in the contrary case there would be a subgroup $g$ in $G/N$ for which

$$D(n_g) = 1 = D_N(\varphi^{-1}(n_g)) = D_N(n_{gN}).$$

But the equality $D_N(n_{gN}) = 1$ for $g \neq e_1$ contradicts Proposition 3. Conversely, if $D_{e_1}$ is a dispersion on $G/N$, then for any measure $\mu$ on $G$ put

$$D_N(\mu) = D_{e_1}(\varphi(\mu)).$$

It is clear that $D_N$ satisfies items 1°–3°. The proposition is proved.

**Corollary 2.** *The group $G/N$, where $N$ is the kernel of some weak dispersion, is either finite or a Lie group.*

Indeed, this follows from Propositions 2 and 4.

Thus, despite the great generality in the definition of weak dispersions, weak dispersions in principle do not go beyond the set of all dispersions on finite groups and Lie groups. Therefore, by virtue of Corollary 1, in order to obtain the general form of dispersions it suffices to describe all weak dispersions on compact Lie groups.

A compact group $G$ has a countable number of irreducible representations. Therefore they can be arranged in pairs $\{Q_i, \overline{Q_i}\}$, $i = 1, \infty$, so that each pair (up to equivalence) occurs only once. Let $R_n$ be the linear space over the field of real numbers generated by the imaginary and real parts of the elements of the matrix $Q_n$. By the orthogonality relations for irreducible representations (3), the spaces $R_m$ and $R_n$ are orthogonal for $m \neq n$. If

$$Q_n(x) = \|g_{ij}(x)\|,$$

then

$$\left\|\int g_{ij}(x)\mu(dx)\right\| = Q_n(\mu)$$

is the Fourier coefficient of the measure $\mu$. We have

$$Q_n(\nu\mu) = Q_n(\nu)Q_n(\mu)$$

for $\nu, \mu \in \mathfrak{M}$. Therefore

$$|\det Q_n(\mu)| = \Gamma_n$$

is a weak dispersion on $\mathfrak{M}$. It can be shown that the kernel $\Gamma_n$ is equal to the kernel of the representation $Q_n$. Let $R$ be the algebraic sum of the rings $R_n$, $n = 1, \infty$, and let $\mathfrak{M}_0$ be the semigroup of measures on $G$ having a finite number of nonzero Fourier coefficients. Then by (5), $R$ is dense in the space $C$ of all continuous real-valued functions on $G$, and $\mathfrak{M}_0$ is dense in $\mathfrak{M}$. Define multiplication in $C$ as convolution with respect to the invariant measure on $G$. The spaces $R_n$ are closed with respect to this multiplication and therefore are finite-dimensional rings. It follows from (5) that the rings $R_n$ are semisimple.

**Proposition 5.** *The ring $R_n$ is simple.*

**Proof.** If $Q_n \nsim \overline{Q_n}$, then the functions

$$f = \sum(\alpha_{kl}u_{kl} + \beta_{kl}v_{kl}),$$

where $u_{kl}, v_{kl}$ are the real and imaginary parts of $g_{kl}$, fill $R_n$ and are determined by the coefficients uniquely. The correspondence

$$f \to \|\alpha_{kl}+$$

$+i\beta_{kl}\|$ will be isomorphic. Therefore the ring $R_n$ is simple. If $Q_n \sim \overline{Q}_n$ and $R_n$ is not simple, then in $R_n$ there is a two-sided ideal $R_0$, distinct from $R_n$ and 0. By virtue of the orthogonality of $R_m, R_n$ and the density of $R$ in $C$, $R_0$ will be an ideal in $C$. Then from (5) $R_0$ contains the left and right translates of a function $f \in R_0$ by arbitrary elements of $G$. Consequently, if $\{f_i(x)\}$, $i = 1, l$, is a basis of $R_0$, then $f_i(ax) = \sum \alpha_{ij}(a)f_j(x)$, or $f(ax) = \alpha(a)f(x)$ in vector form. Taking into account the linear independence of the $f_i$, we obtain $f(aa'x) = \alpha(a)f(a'x) = \alpha(a)\alpha(a')f(x) = \alpha(aa')f(x)$, and therefore $\alpha(aa') = \alpha(a)\alpha(a')$. From the linear independence of the $f_j$ there exist elements $x_\nu$ such that det $\|f_j(x_\nu)\| \neq 0$. Consequently, the $a_{ij}(a)$—the elements of the matrix $\alpha(a)$ —are expressed linearly in terms of the continuous functions $f_i(ax_\nu)$. Thus $\alpha(a)$ is a linear representation of $G$. Since $f_i(ax_\nu) \in R_0$, the $a_{ij}(a)$—the elements of the representation $\alpha(a)$—also belong to $R_0$. But $\alpha(a)$, as a representation of $G$, decomposes into a direct sum of irreducible representations. Therefore, taking into account the orthogonality of $R_m$ and $R_n$, we obtain $R_0 = R_n$. The proposition is proved.

As a consequence of the simplicity of $R_n$ and the density of $R$ in $C$, we obtain

**Proposition 6.** If $\Gamma$ is a continuous homomorphism, under multiplication, of the ring $C$ into the nonnegative numbers, not identically equal to zero, then $\Gamma = \Gamma_{n_1}^{\alpha_1} \cdots \Gamma_{n_s}^{\alpha_s}$, $\alpha_i > 0$. This representation is unique.

Now we can formulate the main result.

**Theorem.** Every dispersion $D$ on $\mathfrak{M}$ is representable in a unique way in the form $D = \Gamma_{n_1}^{\alpha_1} \cdots \Gamma_{n_s}^{\alpha_s}$, $\alpha_i > 0$. The numbers $n_i$ are such that the intersection of the kernels of the representations $Q_{n_i}$ is equal to $e_1$.

**Proof.** Let $\mathfrak{M}_0^{(n)}$ be the semigroup of measures in $\mathfrak{M}_0$ for which $Q_i(\mu) = 0$ when $i \geq n+1$. $\mathfrak{M}_0 = \bigcup_{n=1}^{\infty} \mathfrak{M}_0^{(n)}$. Consider the mapping $\varphi : f \mapsto f + 1$ for $f \in \bigcup_{i=1}^{n} R_i = R^{(n)}$. The mapping $\varphi$ is one-to-one, continuous, and $\varphi(f_1 f_2) = \varphi(f_1)\varphi(f_2)$ (the product is understood as convolution). Every measure from $\mathfrak{M}_0^{(n)}$ is represented in a unique way in the form $f+1$, or symbolically $\varphi^{-1}(\mu) = f$. If $f \in R^{(n)}$ is sufficiently small in modulus, then $\lambda f + 1$, $|\lambda| \leq 1$, will always be the density of some measure from $\mathfrak{M}_0^{(n)}$. Thus $\varphi^{-1}(\mathfrak{M}_0^{(n)})$ contains a neighborhood of zero in $R^{(n)}$. Let $D$ be given on $\mathfrak{M}$. By virtue of the density of $\mathfrak{M}_0 = \bigcup_{n=1}^{\infty} \mathfrak{M}_0^{(n)}$ in $\mathfrak{M}$, $D$ is determined by its specification on $\mathfrak{M}_0$. Consequently, for some $n$, $D$, considered on $\mathfrak{M}_0^{(n)}$, is not identically equal to zero. Define $D$ on $\varphi^{-1}(\mathfrak{M}_0^{(n)})$ by putting $D(\varphi^{-1}(\mu)) = D(\mu)$. Then $D$ is a continuous homomorphism of the neighborhood $\varphi^{-1}(\mathfrak{M}_0^{(n)})$ into the nonnegative numbers. By standard arguments it is shown that $D$ extends uniquely from $\varphi^{-1}(\mathfrak{M}_0^{(n)})$ to all of $R^{(n)}$. But then, by virtue of Proposition 6, $D$ has the form $\Gamma_{n_1}^{\alpha_1} \cdots \Gamma_{n_s}^{\alpha_s}$, $\alpha_i > 0$. Consequently, this is also an expression for $D$ on $\mathfrak{M}$, since $n$ is arbitrary. Since a dispersion is a weak dispersion with kernel $e_1$, it follows, by Lemma 2, that the intersection of the kernels $\Gamma_{n_i}$, $i = 1, s$, equal to the kernels $Q_{n_i}$, is $e_1$. The theorem is proved.

In conclusion I express my gratitude to V. Ya. Kozlov for his attention.

Institute of Chemical Physics
Academy of Sciences of the USSR
Moscow

Received
13 XI 1969

## References


1. V. M. Maksimov, *Theory of Probability and Its Applications*, **12**, 1 (1967).

2. B. M. Kloss, ibid., **4**, 3 (1959).

3. L. S. Pontryagin, *Continuous Groups*, 1954.

4. V. V. Sazonov, V. N. Tutubalin, *Theory of Probability and Its Applications*, **11**, 1 (1966).

5. M. A. Naimark, *Normed Rings*, "Nauka," 1968.


*Note: Figure translations are in progress. See original paper for figures.*