Journal of Mathematical Sciences, Vol. 93, No. 3, 1999

CHARACTERIZATION OF GAUSSIAN MEASURES BY THE ISOPERIMETRIC PROPERTY OF HALF-SPACES

S. G. Bobkov and C. Houdré

UDC 519.2

If the half-spaces of the form $\{x \in \mathbf{R}^n : x_1 \le c\}$ are extremal in the isoperimetric problem for a product-measure μ^n, $n \ge 2$, then the marginal distribution of μ is Gaussian. Bibliography: 8 *titles.*

§1. Introduction

Let μ be a probability measure on the real line $\mathbf{R}$. We denote by μ^n its nth power on $\mathbf{R}^n$ and by A^h the open h-neighborhood (in the sense of the Euclidean distance) of the subset $A \subset \mathbf{R}^n$, $h > 0$:

$$A^h = \{x \in \mathbf{R}^n : \|x - a\|_2 < h \text{ for some } a \in A\}.$$

The isoperimetric problem for $(\mathbf{R}^n, \mu^n)$ consists of minimizing the value

$$\mu^n(A^h), \tag{1.1}$$

over all Borel subsets $A \subset \mathbf{R}^n$ with $\mu^n(A) \ge p$, where $p \in (0,1)$ and $h > 0$ are fixed.

If μ is Gaussian, then the minimum of (1.1) is attained at any subspace of measure p. This fact may be written as the isoperimetric inequality

$$\mu^n(A^h) \ge \mu^n(B^h), \tag{1.2}$$

where B is the standard half-space of the form

$$\{x \in \mathbf{R}^n : x_1 \le c\}$$

and c is completely determined by p. This deep property of Gaussian measures was discovered by V. N. Sudakov and B. S. Tsirel'son [7] and independently by C. Borell [3]. Their proofs were based on the isoperimetric property of balls on the sphere (the Levi–Schmidt theorem). Isoperimetric methods were used for the first time in the theory of Gaussian processes by H. J. Landau and L. A. Shepp [5] for establishing an extremal property of half-spaces in another problem; they also used the Levi–Schmidt theorem. Another proof of isoperimetric inequality (1.2) based on techniques of symmetrization of subsets in the Gaussian probability space was later found by A. Ehrhard [4]. In this paper, we prove that only Gaussian measures satisfy (1.2) in the class of all product-measures.

Theorem 1.1a. *Let $n \ge 2$. If for all $p \in (0,1)$ and all positive h the minimal value of* (1.1) *is attained at a standard half-space, then μ is Gaussian (possibly degenerate, i.e., concentrated at the origin).*

The case $n = 1$ differs substantially from the case $n \ge 2$. Indeed, many interesting probability distributions on the real line satisfy (1.2).

In the case where the measure μ has continuous positive density, necessary and sufficient conditions for (1.2) to be satisfied are known (see [2, Sec. 13]). In particular, the measure μ must be symmetric with respect to its median and have a finite exponential moment. Actually, these two conditions are necessary without any additional regularity requirements (see Proposition 2.6 below). Moreover, the hypotheses of Theorem 1.1a can be weakened if we assume that μ is symmetric and has a finite second moment.

Translated from *Zapiski Nauchnykh Seminarov POMI*, Vol. 228, 1996, pp. 31–38. Original article submitted December 12, 1995.

 1072-3374/99/9303-0270$22.00

Theorem 1.1b. *Let $n \geq 2$ and $p = 1/2$. Assume that a probability measure μ on the real line $\mathbf{R}$ is symmetric with respect to some point and has a finite second moment. If for all positive h the minimal value of* (1.1) *is attained at a standard half-space, then the measure μ is Gaussian* (*possibly degenerate*).

We stress the important role of the Euclidean distance in this characterization. For example, if the distance $\|x-a\|_2$ in the definition of the set A^h is replaced by the maximum-distance $\|x-a\|_\infty$, then (1.2) is satisfied for a wide class of logarithmically concave distributions [1] (cf. [2, Sec. 15]). Inequalities similar to (1.2) with various definitions of extension of a set were studied by many authors (see, e.g., M. Talagrand [8] and M. Ledoux [6]) in connection with the phenomenon of concentration of measure.

It is obvious that inequality (1.2) gets stronger as the dimension n grows. This means that in fact we may prove the assertions of Theorems 1.1a and 1.1b for the case of the plane ($d = 2$). Moreover, under the hypotheses of Theorem 1.1b, the Gaussian property follows from inequality (1.2) if we apply it to the half-plane

$$A(t) = \left\{ (x_1, x_2) : \frac{x_1 + x_2}{\sqrt{2}} \leq t \right\}, \quad t = 0.$$

Proof of Theorem 1.1b. Let ξ and η be independent identically distributed random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with a common distribution μ symmetric with respect to zero. Then under the condition $\mu(B) \geq 1/2$, the minimal value of the right-hand side of (1.2) is attained at $B = \{ x \in \mathbf{R}^2 : x_1 \leq 0 \}$ and is equal to $\mathbf{P}\{\xi < h\}$. Likewise, under the condition $\mu(A(t)) \geq 1/2$, the minimal value of the left-hand side of (1.2) is attained at $t = 0$. Since $(A(t))^h = A(t+h)$, we get

$$\mathbf{P}\{(\xi+\eta)/\sqrt{2} < h\} \geq \mathbf{P}\{\xi < h\}$$

for all positive h. Thus,

$$\mathbf{D}\left(\frac{\xi+\eta}{\sqrt{2}}\right) = 4\int_0^{+\infty} \mathbf{P}\left\{\frac{\xi+\eta}{\sqrt{2}} > h\right\} h\, dh \leq 4\int_0^{+\infty} \mathbf{P}\{\xi > h\}\, h\, dh = \mathbf{D}(\xi).$$

Since

$$\mathbf{D}\left((\xi+\eta)/\sqrt{2}\right) = \mathbf{D}(\xi),$$

it follows that

$$\mathbf{P}\left\{(\xi+\eta)/\sqrt{2} > h\right\} = \mathbf{P}\{\xi > h\}$$

for almost all (with respect to the Lebesgue measure) positive h. Clearly, this relation holds for all positive h and so the random variables $(\xi+\eta)/\sqrt{2}$ and ξ are identically distributed. Hence, the characteristic function f of the random variable ξ satisfies $f^2(t/\sqrt{2}) = f(t)$ for all real t. It is readily seen that this equation holds only for Gaussian random variables. This completes the proof.

In order to prove Theorem 1.1a, let us consider the one-dimensional case and, in particular, show that condition (1.2) implies the hypotheses of Theorem 1.1b. We do not know whether or not the hypothesis of finiteness of the second moment can be omitted.

§2. Necessary conditions for the case $n = 1$

We introduce the following notation. Let μ be a probability measure on the real line. Let
$F(x) = \mu((-\infty, x])$, $x \in (-\infty, +\infty]$;
$\mathrm{Im}(F) = \{ F(x) > 0 : x \in (-\infty, +\infty] \}$;
$S(F) = \{ x \in (-\infty, +\infty] : F(y) < F(x) \text{ for all } y < x \}$;
$F^{-1}(p) = \inf\{ x \in (-\infty, +\infty] : F(x) \geq p \}$, $p \in (0, 1]$.

Note that $F^{-1}(p)$ is the minimal quantile of order p. Since the distribution function F is continuous on the right, the infimum in the definition of $F^{-1}(p)$ can be replaced by the minimum. In particular, $F^{-1}(p)$ is the least solution of the equation $F(x) = p$ for $p \in \mathrm{Im}(F)$. Hence, $F(F^{-1}(p)) \geq p$ for all $p \in (0, 1]$ and $F(F^{-1}(p)) = p$ for all $p \in \mathrm{Im}(F)$. The set $S(F)$ (without the point $x = +\infty$) is a subset of the (closed) support of μ. It follows that $\mu(S(F)) = 1$.

Lemma 2.1. *The map F is an increasing bijection of $S(F)$ to $\mathrm{Im}(F)$, and the restriction of the function F^{-1} to $\mathrm{Im}(F)$ is the inverse map for it. Moreover, the function F^{-1} is continuous on the left on $(0,1)$.*

Lemma 2.2. *For all $p \in (0,1]$, we have the following equivalences:*
(a) $F(F^{-1}(p)) = p \iff p \in \mathrm{Im}(F)$;
(b) $x \geq F^{-1}(p) \iff F(x) \geq p$ *for all* $x \in (-\infty, +\infty]$;
(c) $x \leq F^{-1}(p) \iff F(x) \leq p$ *for all* $x \in S(F)$.

Both lemmas are elementary, and their proofs are omitted.

Now, let F and G be the distribution functions of the probability measures μ and ν.

Lemma 2.3. *The map $U = G^{-1}(F)$ takes μ to ν if and only if $\mathrm{Im}(G) \subset \mathrm{Im}(F)$.*

Proof. Consider the restriction of U to $S(F)$. Let $p = F(x)$, $x \in S(F)$, and $q = G(t)$, $t \in (-\infty, +\infty]$, so that $p \in \mathrm{Im}(F)$ and $q \in \mathrm{Im}(G)$ and, hence, $q \in \mathrm{Im}(F)$. By Lemma 2.2 (b), (c), we have

$$U(x) = G^{-1}(F(x)) = G^{-1}(p) \leq t \iff G(t) \geq p \iff F(x) \leq q \iff x \leq F^{-1}(q).$$

It follows that $\mu\{U \leq t\} = F(F^{-1}(q)) = q$ since $q \in \mathrm{Im}(F)$. The converse assertion is obvious.

Lemma 2.4. *Assume that for all $p \in (0,1)$ and all positive h we have*

$$F(F^{-1}(p) + h) \leq G(G^{-1}(p) + h). \tag{2.1}$$

Then the map $U = G^{-1}(F)$ takes μ to ν and for all $x \in S(F)$ and positive h we have

$$U(x+h) \leq U(x) + h. \tag{2.2}$$

Proof. Let $h \to 0$ in (2.1). Then for all $p \in (0,1)$ we have

$$F(F^{-1}(p)) \leq G(G^{-1}(p)). \tag{2.3}$$

Since $F(F^{-1}(1)) = G(G^{-1}(1)) = 1$, relation (2.3) also holds for $p = 1$. Let $p \in \mathrm{Im}(G)$. By (2.3), we have $F(F^{-1}(p)) \leq p$. As noted above, the inequality $F(F^{-1}(p)) \geq p$ always holds, and hence $F(F^{-1}(p)) = p$. Once again using Lemma 2.2a, we see that $p \in \mathrm{Im}(F)$. Thus, $\mathrm{Im}(G) \subset \mathrm{Im}(F)$, and in view of Lemma 2.3 the map U takes μ to ν. Now, take an arbitrary $x \in S(F)$. Then $F^{-1}(F(x)) = x$ by Lemma 2.1. Applying (2.1) to $p = F(x)$, we obtain

$$F(x+h) \leq G(U(x) + h).$$

Further,

$$U(x+h) \leq G^{-1}(G(U(x) + h))$$

because $F(x+h) \geq F(x) > 0$ and the function G^{-1} is nondecreasing. It remains to show that $G^{-1}(G(U(x)+h)) \leq U(x) + h$. To this end, note that $G^{-1}(G(y)) \leq y$ for all y such that $G(y) > 0$. Moreover, for $y = U(x) + h$, we have

$$G(y) \geq G(U(x)) = G(G^{-1}(F(x))) \geq F(x) > 0$$

because $G(G^{-1}(p)) \geq p$ for all $p \in (0,1]$ and F is positive on $S(F)$. This completes the proof of Lemma 2.4.

Let $m_p(\cdot)$ denote the minimal quantile of order p.

Proposition 2.5. *The inequality*

$$\mathbf{P}\{\lambda \leq m_p(\lambda) + h\} \geq \mathbf{P}\{\xi \leq m_p(\xi) + h\} \tag{2.4}$$

holds for random variables ξ and λ for all $p \in (0,1)$ and all positive h if and only if there exists a nondecreasing Lipschitz function $U\colon \mathbf{R} \to \mathbf{R}$ (a nonstrict contraction) such that the random variables λ and $U(\xi)$ are identically distributed.

Proof. Let (2.4) hold, i.e., let (2.1) hold for the distribution functions F and G of the random variables ξ and λ. By Lemma 2.4, the map $U = G^{-1}(F)$ restricted to $S = S(F) \setminus \{+\infty\}$ takes the distribution of ξ

to that of λ. (Recall that $\mathbf{P}\{\xi \in S\} = 1$.) The function U is nondecreasing and, according to (2.3), it is finite and Lipschitz on S with Lipschitz (semi-)norm $K \leq 1$. By the Kirsbraun–McShein (Hahn–Banach) theorem, U can be extended to the whole of $\mathbf{R}$ without increasing the Lipschitz norm K of U. Moreover, such an extension can be chosen to be a nondecreasing function. Indeed, U can be uniquely extended to the closure clos(S) by continuity, and hence we can assume that U is Lipschitz and nondecreasing on clos(S). The complement $T = \mathbf{R} \setminus \mathrm{clos}(S)$ is open and hence can be represented as the union of not more than a countable family of pairwise disjoint open intervals. For every finite interval (a, b) of this family, we define the extension of U as the linear function such that its limits at a and b coincide with the corresponding values of U. We also define the extension of U to the intervals $(a, +\infty)$ and $(-\infty, b)$ as the linear functions $U(a) + K(x - a)$ and $U(b) + K(x - b)$, respectively. Obviously, this extension is a nondecreasing Lipschitz function on $\mathbf{R}$.

The proof of the converse assertion is trivial.

Proposition 2.6. *Let μ be a probability measure on $\mathbf{R}$ such that* (1.1) *attains its minimal value at intervals of the form $A = (-\infty, x]$ for all $p \in (0,1)$ and all positive h. Then the measure μ is symmetric with respect to its median and has a finite exponential moment.*

Proof. Applying (1.2) to minimizing intervals $A = [a, +\infty)$ and $B = (-\infty, x]$ of measure greater than or equal to p, we obtain (2.4) for random variables ξ and $\lambda = -\xi$ under the assumption that μ is the distribution of ξ. By Proposition 2.5, there exists a nondecreasing Lipschitz function U such that the random variables λ and $U(\xi)$ are identically distributed. For $V(x) = U(-x)$, the random variables λ and $V(\lambda)$ are also identically distributed, where V is now a nonincreasing Lipschitz function. Let λ' be an independent copy of λ. Since

$$|V(\lambda') - V(\lambda)| \leq |\lambda' - \lambda|$$

and both sides of this inequality are identically distributed random variables, it follows that $|V(x) - V(y)| = |x - y|$ for almost all (x, y) with respect to the product-measure $\nu \otimes \nu$, where ν is the distribution of the random variable λ. By Fubini's theorem, there exists a point y_0 such that

$$|V(x) - V(y_0)| = |x - y_0|$$

for ν-almost all x. Since V is nonincreasing, it follows that $V(x) = -x - 2a$ for some a for ν-almost all x. In other words, the distribution of the random variable $\lambda + a$ and, hence, that of the random variable $\xi - a$ is symmetric with respect to zero.

In order to prove the exponential integrability, assume that the measure μ is nondegenerate and symmetric with respect to zero, and that the infimum in the expression

$$R_h(p) = \inf_{\mu(A) \geq p} \mu(A^h), \quad 0 < p < 1, \quad h > 0, \tag{2.5}$$

is attained at the intervals $A = (-\infty, x]$, where $x = F^{-1}(p)$ and F is the distribution function of the measure μ. Since in this case $A^h = (-\infty, x + h)$, we have

$$R_h(p) = F(F^{-1}(p) + h - 0).$$

Let us check the inequality

$$R_h(p + q) \leq R_h(p) + R_h(q) \tag{2.6}$$

for all positive h and for all p and q such that $0 < p, q < 1$, and $p + q < 1$. Indeed, let $A = (-\infty, x]$ be an extremal subset in (2.5) for p. Since μ is symmetric, let an extremal subset for q be taken in the form $B = [y, \infty)$ with the maximal possible value of y. The assumption $p + q < 1$ implies that $x \leq y$. If $x = y$, then $A \cup B = \mathbf{R}$ and hence

$$R_h(p) + R_h(q) = \mu(A^h) + \mu(B^h) \geq \mu(A) + \mu(B) \geq 1.$$

Therefore, (2.6) holds. If $x < y$, the measure of $A \cup B$ is equal to $p+q$, and hence, in view of the identity $(A \cup B)^h = A^h \cup B^h$, we get

$$R_h(p+q) \le \mu((A \cup B)^h) \le \mu(A^h) + \mu(B^h) = R_h(p) + R_h(q).$$

Now, using (2.6), let us show that

$$\liminf_{p \to 0+} R_h(p)/p > 1$$

for all sufficiently large positive h. Indeed, suppose that this lower limit is equal to one. In this case, the set E_ε of all $p \in (0,1)$ satisfying the inequality $R_h(p) \le (1+\varepsilon)p$ is infinite for any positive ε, and, moreover, zero is a limit point of E_ε. Hence, for every $p \in (0,1)$, there exists a sequence $p_n \in E_\varepsilon$ (some of its elements may coincide) such that

$$r_n = p_1 + \cdots + p_n \to p \quad \text{as} \quad n \to \infty.$$

Applying (2.6) to r_n, we obtain

$$R_h(r_n) \le R_h(p_1) + \cdots + R_h(p_n) \le (1+\varepsilon)(p_1 + \cdots + p_n) \le (1+\varepsilon)p.$$

Passing to the limit as $n \to \infty$, taking into account the continuity of the function R_h on the left, and using the last assertion in Lemma 2.1, we obtain the inequality $R_h(p) \le (1+\varepsilon)p$, which now holds for any p. Since ε is an arbitrary positive number, we have $R_h(p) \le p$, and so $R_h(p) = p$ for all $p \in (0,1)$. But this is impossible for sufficiently large h. Indeed, since the measure μ is nondegenerate, the numbers $x, y \in \mathbf{R}$ can be chosen so that $0 < F(x) < F(y)$, and then $R_h(p) > p$ for $p = F(x)$ and $h > y - F^{-1}(p)$.

Thus, we can find $h > 0$, $p_0 \in (0,1)$, and $c > 0$ such that $R_h(p) = F(F^{-1}(p) + h - 0) \ge cp$ for all $p \in (0, p_0]$ ($cp_0 \le 1$). Hence, $F(F^{-1}(p) + 2h) \ge cp$. Substituting both sides of the last inequality for the argument in F^{-1} and taking into account that $F^{-1}(F(x)) \le x$ if $F(x) > 0$, we obtain

$$F^{-1}(cp) - F^{-1}(p) \le 2h.$$

In particular, if $c^{n-1}p \le p_0$, then

$$F^{-1}(c^k p) - F^{-1}(c^{k-1}p) \le 2h \quad \text{for all} \quad k = 1, \ldots, n.$$

Summing over all k, we get

$$F^{-1}(c^n p) - F^{-1}(p) \le 2nh.$$

Substituting $p_0 c^{-n}$ for p in this inequality, we obtain the inequality

$$F^{-1}(p_0 c^{-n}) \ge -2nh + F^{-1}(p_0),$$

which holds for all n. This easily implies an estimate of the form

$$F(x) \ge \exp(ax), \quad a > 0, \quad x \to -\infty.$$

This research was partially supported by the International Science Foundation, grant NXZ000–NXZ300, and by the Russian State Committee for Higher Education, grant 94-1.4-57. The research was also partially supported by an NSF Postdoctoral Fellowship. The second author enjoyed the hospitality of the Steklov Mathematical Institute (St. Petersburg Branch) and of the Department of Mathematics, University of Syktyvkar, Russia, where part of this research was done.

Translated by V. N. Sudakov.

REFERENCES

1. S. G. Bobkov, "Extremal properties of half-spaces for log-concave distributions," *Ann. Probab.* (to appear).
2. S. G. Bobkov and C. Houdré, "Some connections between Sobolev-type inequalities and isoperimetry," Preprint (1995).
3. C. Borell, "The Brunn–Minkowski inequality in Gauss space," *Invent. Math.*, **30**, 207–211 (1975).
4. A. Ehrhard, "Symétrisation dans l'espace de Gauss," *Math. Scand.*, **53**, 281–301 (1983).
5. H. J. Landau and L. A. Shepp, "On the supremum of a Gaussian process," *Sankhya, Ser. A*, **32**, 369–378 (1970).
6. M. Ledoux, "Isoperimetry and Gaussian analysis," Preprint, École d'été Probab. Saint-Flour (1994).
7. V. N. Sudakov and B. S. Tsirel'son, "Extremal properties of half-spaces for spherically symmetric measures," *Zap. Nauchn. Semin. LOMI*, **41**, 14–24 (1974).
8. M. Talagrand, "Concentration of measure and isoperimetric inequalities in product spaces," Preprint (1994).