The Annals of Probability
1976, Vol. 4, No. 6, 1020-1026

ON CONVEXITY OF MEASURES[1]

BY YOSEF RINOTT

The Weizmann Institute of Science, Israel and Cornell University

A simple geometric proof and some applications are given to results of C. Borell providing necessary and sufficient conditions that a density in R^n generates a measure satisfying a convexity property of the type

$$P(\theta A_0 + (1-\theta)A_1) \geqq \{\theta[P(A_0)]^s + (1-\theta)[P(A_1)]^s\}^{1/s}.$$

1. Introduction. Let P be a probability measure defined on the Borel sets in R^n. We say that P belongs to the class $\mathscr{M}_s$ if for all $\theta \in [0, 1]$, $A_0, A_1 \subset R^n$

$$P(\theta A_0 + (1-\theta)A_1) \geqq \{\theta[P(A_0)]^s + (1-\theta)[P(A_1)]^s\}^{1/s} \tag{1}$$

whenever the sets involved are measurable. Here

$$\theta A_0 + (1-\theta)A_1 = \{\theta a_0 + (1-\theta)a_1 : a_i \in A_i,\ i = 0, 1\}.$$

C. Borell [3], proved that if a measure satisfies the convexity property (1) for some $s \in [-\infty, 1/n)$ then it is absolutely continuous with respect to Lebesgue measure in R^n. Borell gave necessary and sufficient conditions for (1) in terms of the density (Theorem 1, below). The case $s = 0$ in which the right-hand side of (1) is interpreted by continuity as $[P(A_0)]^\theta[P(A_1)]^{1-\theta}$ was considered by Prékopa [8]. In the case $s = 1/n$, (1) holds for Lebesgue measure and is known as the Brunn–Minkowski inequality. (For $s = -\infty$ the right-hand side of (1) is $\min\{P(A_0), P(A_1)\}$.) A host of applications and examples for these results have already appeared in [2], [3], [4], [5], [6] and [9].

In this paper we give a simple new proof of Theorem 1 for the cases $s \in [-\infty, 1/(n+1))$ using a convexity argument that reduces the problem to special cases (Lemma 1 and Lemma 2) which can be computed in an easy way. We believe our approach throws some light on Prékopa's and Borell's somewhat obscure and lengthy proofs. Some new applications are given.

2. Characterization of convex measures.

THEOREM 1. *Let P be a probability measure on R^n generated by a density f,* i.e., *$P(A) = \int_A f(\mathbf{x})\,d\mathbf{x}$ for any Borel set $A \subset R^n$. $P \in \mathscr{M}_s$ if and only if there exists a version g of the density f such that*

$$g^{s/(1-sn)} \text{ is convex } \quad (-\infty \leqq s < 0),$$
$$\log g \text{ is concave } \quad (s = 0),$$
$$g^{s/(1-sn)} \text{ is concave } \quad (0 < s < 1/n).$$

Received August 11, 1975; revised February 2, 1976.

[1] Research supported in part by NSF Grant GP 24438.

AMS 1970 *subject classifications.* Primary 62E10, 62H10; Secondary 26A51, 26A86.

Key words and phrases. Convexity, Brunn–Minkowski inequality.

REMARK. The idea of the proof is to replace integrals of f over sets in R^n by a certain measure μ (to be defined below) of epigraphs in R^{n+1}. We prove directly that the measure μ is in $\mathscr{M}_s$ and Theorem 1 follows by a simple geometric property of the epigraphs. As a result of the embedding in R^{n+1} our proof is valid only for the cases $s \in [-\infty, 1/(n+1))$

To prove Theorem 1 we need the following lemmas.

LEMMA 1. *Let μ be the measure defined on R^{n+1} by $d\mu(x_1, \cdots, x_{n+1}) = e^{-x_{n+1}} dx_1 \cdots dx_{n+1}$. Then for $A_0, A_1 \subset R^{n+1}$, $\theta \in [0, 1]$ we have*

$$\mu(\theta A_0 + (1-\theta)A_1) \geqq [\mu(A_0)]^\theta [\mu(A_1)]^{1-\theta} . \tag{2}$$

LEMMA 2. *Let μ be the measure on R^{n+1} defined by*

$$d\mu(x_1, \cdots, x_{n+1}) = \left|\frac{1-sn}{s}\right| x_{n+1}^{((1-sn)/s)-1} dx_1 \cdots dx_{n+1} , \qquad x_{n+1} > 0 .$$

Then for $s \in [-\infty, 1/(n+1))$, $s \neq 0$ and $\theta \in [0, 1]$

$$\mu(\theta A_0 + (1-\theta)A_1) \geqq \{\theta[\mu(A_0)]^s + (1-\theta)[\mu(A_1)]^s\}^{1/s} . \tag{3}$$

REMARK. The lemmas can be obtained by first proving them for rectangular sets, i.e., products of intervals, for which the μ-measure can be computed directly. The result will follow for any sets using the Hadwiger–Ohman argument referred to by Borell [3]. We have included a different proof at the end of this section.

PROOF OF THEOREM 1. We start with the "if" statement, and the case $s = 0$. For a real valued function g defined on R^n and a set $A \subset R^n$ we define the epigraph in R^{n+1}

$$E_A(g) = \{(\mathbf{x}, \alpha) : \mathbf{x} \in A, \alpha \in R \text{ and } g(\mathbf{x}) \leqq \alpha\} .$$

Denoting

$$A^* = E_A(-\log f) , \qquad A^* \subset R^{n+1}$$

we have for the measure μ defined in Lemma 1,

$$P(A) = \mu(A^*) \tag{4}$$

and the convexity of $-\log f$ implies

$$[\theta A_0 + (1-\theta)A_1]^* \supseteqq \theta A_0 + (1-\theta)A_1^* . \tag{5}$$

Taking μ measure of the sets on both sides of (5) and applying Lemma 1 and (4) we obtain

$$P(\theta A_0 + (1-\theta)A_1) \geqq [P(A_0)]^\theta [P(A_1)]^{1-\theta}$$

completing the proof for $s = 0$.

For $s < 0$ repeat the above argument with

$$A^* = E_A(f^{s/(1-sn)}) .$$

Relations (4) and (5) and the rest of the proof continue to be true in this notation where μ is now defined as in Lemma 2. For $s > 0$ define

$$A^* = H_A(f^{s/(1-sn)})$$

where the hypograph $H_A(g)$ is defined here to be

$$H_A(g) = \{(\mathbf{x}, \alpha): \mathbf{x} \in A, \alpha \in R \text{ and } 0 \leqq \alpha \leqq g(\mathbf{x})\}$$

and the proof follows as above.

In order to prove the "only if" statement we start with the case $s = 0$, so that P, generated by f satisfies

$$P(\theta A_0 + (1 - \theta)A_1) \geqq [P(A_0)]^\theta [P(A_1)]^{1-\theta} . \tag{6}$$

Let $B_k(\mathbf{x})$ be the sphere in R^n centered at $\mathbf{x}$ with radius $1/k$. Define

$$f_k(\mathbf{x}) = \frac{1}{|B_k(\mathbf{x})|} \int f(\mathbf{s})\, d\mathbf{s}$$

where the integral is taken over the sphere $B_k(\mathbf{x})$, and $|B_k(\mathbf{x})|$ denotes its Lebesgue measure.

Now (6) implies

$$f_k(\theta \mathbf{x} + (1 - \theta)\mathbf{y}) \geqq f_k^{\theta}(\mathbf{x}) f_k^{\theta}(\mathbf{y})$$

and therefore the function g defined by $g(\mathbf{x}) = \liminf_{k\to\infty} f_k(\mathbf{x})$ satisfies $g(\theta\mathbf{x} + (1 - \theta)\mathbf{y}) \geqq g^\theta(\mathbf{x}) g^{1-\theta}(\mathbf{y})$, i.e., g is log-concave.

By differentiation of the integral argument we have $f = g$ almost everywhere implying f is log-concave a.e., completing the case $s = 0$.

For $s \neq 0$ we conclude by a similar argument that $P \in \mathscr{M}_s$ implies f^s is convex for $s < 0$ and concave for $0 < s < 1/n$. In particular f is continuous for $-\infty < s < 1/n$ (and it is easy to see that f can be uniformly approximated by an a.e. continuous density g satisfying $g(\theta\mathbf{x} + (1 - \theta)\mathbf{y}) \geqq \min\{g(\mathbf{x}), g(\mathbf{y})\}$ in the case $s = -\infty$).

Consider the sets $A_i \subset R^{n+1}$, $i = 0, 1$, defined by $A_i = B_i \times (a_i, \infty)$ where B_i are spheres in R^n, $a_i > 0$. Denote by v_{B_i} the Lebesgue measure of B_i. Then a direct computation shows that if

$$\frac{v_{B_0}^{1/n}}{v_{B_1}^{1/n}} = \frac{a_0}{a_1}$$

then the sets A_i will satisfy (3) of Lemma 1 with equality for $s < 0$ (define $A_i = B_i \times (0, a_i)$ for $s > 0$). If $f^{s/(1-sn)}$ is *not* convex (in the case $s < 0$) we can construct A_i as above with $A_i \subset E_{B_i}(f^{s/(1-sn)})$ but $\theta A_0 + (1 - \theta)A_1 \supset E_{\theta B_0 + (1-\theta)B_1}(f^{s/(1-sn)})$ and as in the proof of Theorem 1 we obtain from the above relations

$$\begin{aligned} P(\theta B_0 + (1 - \theta)B_1) &< \mu(\theta A_0 + (1 - \theta)A_1) = \{\theta[\mu(A_0)]^s + (1 - \theta)[\mu(A_1)]^s\}^{1/s} \\ &< \{\theta[P(B_0)]^s + (1 - \theta)[P(B_1)]^s\}^{1/s} \end{aligned}$$

so that $P \notin \mathscr{M}_s$ thus proving the "only if" for $s < 0$, and similarly for $s > 0$. This completes the proof of Theorem 1.

Prékopa's [7] original proof of the case $s = 0$ of Theorem 1 was based on the following integral inequality, a "reversal" of Hölder's inequality, which we derive easily from Lemma 1.

COROLLARY 1. *Let f_0, f_1 be positive functions on R^n and for t in R^n let*

$$r(\mathbf{t}) = \sup\{f_0(\mathbf{x}_0)f_1(\mathbf{x}_1): \theta\mathbf{x}_0 + (1-\theta)\mathbf{x}_1 = \mathbf{t}\}.$$

Then

$$\int_{R^n} r(\mathbf{t})\,d\mathbf{t} \geqq [\int_{R^n} f_0^{1/\theta}(\mathbf{x})\,d\mathbf{x}]^\theta[\int_{R^n} f_1^{1/(1-\theta)}(\mathbf{x})\,d\mathbf{x}]^{1-\theta}.$$

PROOF. For any two functions $g_0(\mathbf{x})$, $g_1(\mathbf{x})$, $\mathbf{x} \in R^n$ define

$$(g_0 \square g_1)(\mathbf{t}) = \inf\{\theta g_0(\mathbf{x}_0) + (1-\theta)g_1(\mathbf{x}_1): \theta\mathbf{x}_0 + (1-\theta)\mathbf{x}_1 = \mathbf{t}\}.$$

Setting $g_0(\mathbf{x}) = -\log f_0^{1/\theta}(\mathbf{x})$, $g_1(\mathbf{x}) = -\log f_1^{1/(1-\theta)}(\mathbf{x})$ we have

$$r(\mathbf{t}) = e^{-(g_0 \square g_1)(\mathbf{t})}.$$

Since ([10], Section 5)

$$E_{R^n}(g_0 \square g_1) = \theta E_{R^n}(g_0) + (1-\theta)E_{R^n}(g_1) \tag{7}$$

we obtain invoking the measure μ in Lemma 1, (2), (4) and (7)

$$\begin{aligned}\int_{R^n} r(\mathbf{t})\,d\mathbf{t} &= \mu\{E_{R^n}(g_0 \square g_1)\}\\ &= \mu\{\theta E_{R^n}(g_0) + (1-\theta)E_{R^n}(g_1)\}\\ &\geqq \{\mu[E_{R^n}(g_0)]\}^\theta\{\mu[E_{R^n}(g_1)]\}^{1-\theta}\\ &= [\int_{R^n} f_0^{1/\theta}(\mathbf{x})\,d\mathbf{x}]^\theta[\int_{R^n} f_1^{1/(1-\theta)}(\mathbf{x})\,d\mathbf{x}]^{1-\theta}.\end{aligned}$$

COROLLARY 2. *Let f_i $i = 0, 1$ be nonnegative functions defined on R^n. Let $g_i = f_i^{s/(1-sn)}$ and define*

$$h(\mathbf{t}) = (g_0 \square g_1)(\mathbf{t}) = \inf\{\theta g_0(\mathbf{x}_0) + (1-\theta)g_1(\mathbf{x}_1): \theta\mathbf{x}_0 + (1-\theta)\mathbf{x}_1 = \mathbf{t}\}$$

then for $s < 0$

$$\int_{R^n} h^{(1-sn)/s}(\mathbf{t})\,d\mathbf{t} \geqq \{\theta(\int_{R^n} f_0(\mathbf{x})\,d\mathbf{x})^s + (1-\theta)(\int_{R^n} f_1(\mathbf{x})\,d\mathbf{x})^s\}^{1/s}.$$

PROOF. Invoking the measure μ defined in Lemma 2 we have by (3) and (7)

$$\begin{aligned}\int_{R^n} h^{(1-sn)/s}(\mathbf{t})\,d\mathbf{t} &= \mu\{E_{R^n}(h(\mathbf{t}))\}\\ &= \mu\{\theta E_{R^n}(g_0) + (1-\theta)E_{R^n}(g_1)\}\\ &\geqq \{\theta\mu^s(E_{R^n}(g_0)) + (1-\theta)\mu^s(E_{R^n}(g_1))\}^{1/s}\\ &= \{\theta(\int_{R^n} f_0(\mathbf{x})\,d\mathbf{x})^s + (1-\theta)(\int_{R^n} f_1(\mathbf{x})\,d\mathbf{x})^s\}^{1/s}.\end{aligned}$$

PROOF OF LEMMA 1. Part of the method of proof is close to that in the proof of the Brunn-Minkowski inequality in [1]. For a set $A \subset R^{n+1}$ let $T_A(z)$ denote the n dimensional section of A on the hyperplane $\{x_{n+1} = z\}$, i.e.,

$$T_A(z) = \{(x_1, \cdots, x_n, z) \in A: (x_1, \cdots, x_n) \in R^n\}.$$

Let $v_A(z)$ be the n-dimensional volume of $T_A(z)$. Denote $m_i = \mu(A_i)$ $i = 0, 1$, and set $B = \theta A_0 + (1-\theta)A_1$. By the definition of μ we have

$$m_i = \int_{-\infty}^{\infty} v_{A_i}(x)e^{-x}\,dx \qquad i = 0, 1.$$

For $i = 0, 1$ and $0 \leqq \tau \leqq 1$ we define $z_i(\tau)$ by the relations

$$m_i\tau = \int_{-\infty}^{z_i(\tau)} v_{A_i}(x)e^{-x}\,dx$$

and differentiate to obtain

$$m_i = v_{A_i}(z_i)e^{-z_i}\frac{dz_i}{d\tau} \qquad i = 0, 1 . \tag{8}$$

Now

$$\mu(B) = \int_{-\infty}^{\infty} v_B(z)e^{-z}\,dz$$

and we substitute $z = \theta z_0(\tau) + (1-\theta)z_1(\tau)$ in the last integral, using (8) to obtain

$$\begin{aligned}\mu(B) &= \int_0^1 v_B(\theta z_0 + (1-\theta)z_1)e^{-[\theta z_0 + (1-\theta)z_1]}\left[\frac{\theta m_0}{v_{A_0}(z_0)e^{-z_0}} + \frac{(1-\theta)m_1}{v_{A_1}(z_1)e^{-z_1}}\right]d\tau \\ &\geqq m_0^{\theta}m_1^{1-\theta}\int_0^1 v_B(\theta z_0 + (1-\theta)z_1)\frac{1}{v_{A_0}^{\theta}(z_0)v_{A_1}^{1-\theta}(z_1)}\,d\tau\end{aligned} \tag{9}$$

where the last inequality follows by arithmetic-geometric means inequality.

Since the sections satisfy

$$T_B(\theta z_0 + (1-\theta)z_1) \supseteqq \theta T_{A_0}(z_0) + (1-\theta)T_{A_1}(z_1)$$

we have by the Brunn–Minkowski inequality and the arithmetic-geometric means inequality respectively:

$$\begin{aligned}v_B(\theta z_0 + (1-\theta)z_1) &\geqq \{\theta[v_{A_0}(z_0)]^{1/n} + (1-\theta)[v_{A_1}(z_1)]^{1/n}\}^n \\ &\geqq v_{A_0}^{\theta}(z_0)v_{A_1}^{1-\theta}(z_1) .\end{aligned} \tag{10}$$

Thus the last integrand in (9) is $\geqq 1$ implying

$$\mu(B) \geqq m_0^{\theta}m_1^{1-\theta}$$

thus completing the proof of Lemma 1.

Proof of Lemma 2. We use the same notation as above and with the same substitution we obtain after invoking the first inequality in (10) (i.e., the Brunn–Minkowski inequality) and setting $\beta = (1 - s(n+1))/s$,

$$\mu(B) \geqq \int_0^1 [\theta v_{A_0}^{1/n} + (1-\theta)v_{A_1}^{1/n}]^n(\theta z_0 + (1-\theta)z_1)^{\beta}\left[\frac{\theta m_0}{v_{A_0}(z_0)z_0^{\beta}} + \frac{(1-\theta)m_1}{v_{A_1}(z_1)z_1^{\beta}}\right]d\tau .$$

The last integrand is easily shown to be $\geqq [\theta m_0^s + (1-\theta)m_1^s]^{1/s}$ by Hölder's inequality, proving Lemma 2.

3. Applications.

(1) Consider the problem of testing hypotheses concerning a location parameter η of the density $f(\mathbf{x} - \eta)$, $\mathbf{x}, \eta \in R^n$. Let H_0 be a convex subset of R^n, and let A be a convex acceptance region, Let the measure P generated by f be in $\mathscr{M}_s$ for some $s \in [-\infty, 1/n)$. If the test determined by A: $\phi(\mathbf{x}) = 1 - \chi_A(\mathbf{x})$ is α-similar on the boundary of H_0 then it is an unbiased level α test.

(2) In the above setting suppose P is in $\mathscr{M}_s$ for some $s \in [-\infty, 1/n)$ and satisfies (1) with a strict inequality (this would follow if the density satisfies one of the convexity conditions of Theorem 1 strictly). If H_0 is not strictly convex then tests which are α-similar on the boundary of H_0, with convex acceptance region A, do not exist.

(1) and (2) follow from the fact that the power function $\beta(\boldsymbol{\eta})$ of a test based on a random vector $\mathbf{X}$ with joint probability measure $P \in \mathscr{M}_s$, and having a convex acceptance region A, i.e., $\beta(\boldsymbol{\eta}) = P(\mathbf{X} - \boldsymbol{\eta} \in A)$ satisfies:

$$\begin{aligned}\beta(\theta\boldsymbol{\eta}_0 + (1-\theta)\boldsymbol{\eta}_1) &\geqq \{\theta\beta^s(\boldsymbol{\eta}_0) + (1-\theta)\beta^s(\boldsymbol{\eta}_1)\}^{1/s} \\ &\geqq \min\{\beta(\boldsymbol{\eta}_0), \beta(\boldsymbol{\eta}_1)\},\end{aligned}$$

which holds with strict inequality in (2). If the power function is continuous, (2) implies that level α unbiased tests do not exist. As an example consider $n = 2$, $\boldsymbol{\eta} = (\eta_1, \eta_2)$. If f is a bivariate normal density then a convex set in R^2 can never be an acceptance region of an unbiased test of level α for H_0 like $|\eta_1| + |\eta_2| < C$.

(3) Let $X_1, \cdots, X_n$ be random variables having a joint probability measure in $\mathscr{M}_s$ for some $s \in [-\infty, 1/n)$. Define $\eta(t_1, \cdots, t_{n-1})$ by the implicit relation

$$P\{X_1 \geqq t_1, \cdots, X_{n-1} \geqq t_{n-1}, X_n \geqq \eta(t_1, \cdots, t_{n-1}\} = \alpha$$

($0 \leqq \alpha \leqq 1$ fixed). Then $\mu(t_1, \cdots, t_{n-1})$ is concave on the convex region R^{n-1} where it is defined. (The special case of bivariate normal distribution was proved in [11].)

(4) Let $\mathbf{X} = X_1, \cdots, X_n$ be as in (3) and let ϕ be a concave function on R^n. Set $\bar{G}(t) = P(\phi(\mathbf{X}) \geqq t)$. Then

$$\bar{G}(\theta t_1 + (1-\theta)t_2) \geqq \{\theta[\bar{G}(t_1)]^s + (1-\theta)[\bar{G}(t_2)]^s\}^{1/s}.$$

In the case of $s = 0$ it would follow that $\log \bar{G}(t)$ is concave which is equivalent to the distribution having an increasing failure rate (IFR) in the context of reliability theory. Choosing $\phi(\mathbf{x}) = \min x_i$ we see that a system which consists of a series of components with joint log concave density, is IFR.

If ψ is convex in $\mathbf{x} \in R^n$ then the function $F(t) = P(\psi(\mathbf{X}) \leqq t)$ will satisfy

$$F(\theta t_1 + (1-\theta)t_2) \geqq \{\theta[F(t_1)]^s + (1-\theta)[F(t_2)]^s\}^{1/s}. \tag{11}$$

Thus the distribution function of the range statistic $\max X_i - \min X_i$ and the sample standard deviation $\{\sum_{i=1}^n (X_i - \bar{X})^2\}^{\frac{1}{2}}$, for example, satisfy (11). ψ can also be taken to be vector valued, i.e., $\psi(\mathbf{x}) = (\psi_1(\mathbf{x}), \cdots, \psi_m(\mathbf{x}))$ with each ψ_i convex.

(5) If $X(t)$ is a Gaussian process then the joint density of $X(t_0), \cdots, X(t_n)$ is log concave. It follows as in (4) that the distribution function of $\sum_{j=1}^n X^2(t_j)(t_j - t_{j-1})$ is log concave. When $\{t_j\}_{j=1}^n$ is a sequence of partitions becoming dense in (0, 1), the above sum approaches $W = \int_0^1 X^2(t)\,dt$ in distribution implying that the distribution function of W is log concave. Similarly the variable $U = \sup_{0 \leq t \leq 1} |X(t)|$ has a log concave distribution function since (assuming the process is defined to be separable) the variable U can be approximated by $\max_{j=0,\cdots,n} |X(t_j)|$. U and W are the limiting distributions of certain statistics for testing goodness of fit.

Acknowledgments. This work is based on part of a thesis submitted at the Weizmann Institute of Science. I wish to thank my thesis advisor, Professor S. Karlin, for his help and guidance. I am also grateful to Professor M. D. Perlman and Professor R. J. B. Wets for helpful comments and references.

REFERENCES

[1] Bonnesen, T. and Fenchel, W. (1948). *Theorie der konvexen Korper*. Chelsea, New York.

[2] Borell, C. (1974). Convex measures on locally convex spaces. *Ark. Mat.* **12** 239–252.

[3] Borell, C. (1973). Convex functions in d-space. Uppsala Univ. Dept. of Math. Report No. 8. To appear in *Period. Math. Hungar.*

[4] Brascamp, H. J. and Lieb, E. H. On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. To appear.

[5] Eaton, M. L. and Perlman, M. D. (1974). A monotonically property of the power functions of some invariant tests for manova. *Ann. Statist.* **2** 1022–1028.

[6] Gupta, S. D. (1974). S-Unimodal functions: related inequalities and statistical applications. Tech. Report No. 149, Stanford Univ.

[7] Prékopa, A. (1971). Logarithmic concave measures with application to stochastic programming. *Acta Sci. Math.* (*Szeged*) **32** 301–315.

[8] Prékopa, A. (1973). On logarithmic concave measures and functions. *Acta Sci. Math.* (*Szeged*) **34** 335–343.

[9] Rinott, Y. (1974). Some multivariate concepts of monotonicity, convexity and total positivity with applications to probability inequalities. Ph. D. thesis, Weizmann Inst. of Science, Israel.

[10] Rockafellar, R. T. (1970). *Convex Analysis*. Princeton Univ. Press.

[11] Tihanski, D. P. (1972). Properties of the bivariate normal cumulative distribution. *J. Amer. Statist. Assoc.* **67** 903–905.

Department of Statistics
The University of Chicago
Chicago, Illinois 60637