

Logarithmic Concave Measures and Related Topics

András Prékopa

Technological University and Computer Automation Institute of the
Hungarian Academy of Sciences, Budapest, Hungary

Abstract

Results obtained in the last few years are summarized in connection with logarithmic concave measures and related topics. A new and simple proof is presented concerning the basic theorem of logarithmic concave measures stating that if the probability measure P in \mathbb{R}^m is generated by a logarithmic concave density, A, B are convex subsets of \mathbb{R}^m and $0 < \lambda < 1$, then $P(\lambda A + (1-\lambda)B) \geq [P(A)]^\lambda [P(B)]^{1-\lambda}$. The most important convolution theorems belonging to the subject are collected. The notion of convex measures and an important theorem concerning these are formulated. Special distributions are analysed. Properties of special constraint and objective functions which are derived from the theorems presented in the paper are described.

1 Introduction

Optimizing a function subject to constraints where at least one of the constraints is a probabilistic inequality, or maximizing a probability subject to inequality constraints, are important stochastic programming problems with many practical applications. When formulating a mathematical programming model belonging to one of the above-mentioned categories, it is very important to clarify, from the point of view of the numerical solution of the model, whether it is a convex problem or one of a related type. The research work of the past few years aimed at producing efficient analytic tools for handling sophisticated functions in these models has led to the introduction of new mathematical concepts and the proof of some basic theorems which will probably be useful not only in stochastic programming but in other fields of mathematics. The purpose of the present paper is to summarize these results together with earlier results with which they are mathematically interconnected.

We restrict ourselves to the case of absolutely continuous probability distributions in the space \mathbb{R}^m . Proofs will be omitted except for special cases. One exception is the proof of Theorem 2, whose original proof is sophisticated. Here we give a simple proof which should prove useful for classroom presentation.

In Section 2 we deal with logarithmic concave measures and prove the basic theorem. Section 3 is devoted to investigations concerning convolutions of functions defined in \mathbb{R}^m . In Section 4 convex measures are considered. Special absolute and conditional distributions will be analysed in Section 6. Finally, in Section 7 some special functions appearing in stochastic programming models will be investigated.

2 Logarithmic Concave Measures

The notion of a logarithmic concave measure was introduced in [19], [20]. A measure P defined on the Borel-measurable subsets of \mathbb{R}^m is said to be *logarithmic concave* if for every pair A, B of convex subsets of \mathbb{R}^m and every $0 < \lambda < 1$ we have

$$P(\lambda A + (1 - \lambda)B) \geq [P(A)]^\lambda [P(B)]^{1-\lambda}.$$

Here the symbol $+$ means Minkowski addition of sets, i.e. $C + D = \{c + d \mid c \in C, d \in D\}$, and the constant multiple of a set is defined as $kC = \{kc \mid c \in C\}$ for every fixed real number k .

A non-negative point function given by $h(x)$, $x \in \mathbb{R}^m$, is said to be *logarithmic concave* if for every pair $x_1, x_2 \in \mathbb{R}^m$ and every $0 < \lambda < 1$ we have the inequality

$$h(\lambda x_1 + (1 - \lambda)x_2) \geq [h(x_1)]^\lambda [h(x_2)]^{1-\lambda}.$$

If P is a logarithmic concave measure in \mathbb{R}^m , then the function of the variable $x \in \mathbb{R}^m$ given by

$$P(A + x)$$

is logarithmic concave. In particular, if A is the set $A = \{t \mid t \leq 0\}$, then we have that the function defined by

$$F(x) = P(\{t \mid t \leq x\}),$$

i.e. the probability distribution function, is logarithmic concave in \mathbb{R}^m .

In his classical work [4] Brunn proved the following theorem.

THEOREM 1 (Brunn–Minkowski Inequality). *If A and B are non-empty convex subsets of \mathbb{R}^m and $0 < \lambda < 1$, then the following inequality holds:*

$$(2.1) \quad \mu^{1/m}(\lambda A + (1 - \lambda)B) \geq \lambda \mu^{1/m}(A) + (1 - \lambda) \mu^{1/m}(B),$$

where μ denotes the Lebesgue measure.

By the arithmetic-geometric means inequality it readily follows from (2.1) that

$$\mu(\lambda A + (1 - \lambda)B) \geq [\mu(A)]^\lambda [\mu(B)]^{1-\lambda},$$

so that Lebesgue measure is logarithmic concave.

If P is a logarithmic concave measure in \mathbb{R}^m and C is a fixed convex subset of \mathbb{R}^m , then the measure $P(A \cap C)$ defined on the measurable subsets A of \mathbb{R}^m , is also logarithmic concave. In fact if A, B are convex subsets of \mathbb{R}^m and $0 < \lambda < 1$, then we have that

$$[\lambda A + (1 - \lambda)B] \cap C \supset \lambda[A \cap C] + (1 - \lambda)[B \cap C]$$

whence the required inequality follows. Thus we see that a uniform probability distribution over a convex set is a logarithmic concave probability measure.

Theorem 2 expresses the most important fact in connection with logarithmic concave measures.

THEOREM 2. *Let P be a probability measure in \mathbb{R}^m generated by a probability density of the form*

$$f(x) = e^{-Q(x)}, \quad x \in \mathbb{R}^m,$$

where Q is a convex function (i.e., f is a logarithmic concave point function). Then P is a logarithmic concave probability measure.

Remark. The original proof of this theorem in [19] and [20] is based on the inequality of Brunn–Minkowski and the integral inequality

$$(2.2) \quad \int_{-\infty}^{\infty} r(t) dt \geq \left[\int_{-\infty}^{\infty} f^2(x) dx \right]^{1/2} \left[\int_{-\infty}^{\infty} g^2(y) dy \right]^{1/2},$$

where

$$(2.3) \quad r(t) = \sup_{(x+y)/2=t} f(x)g(y), \quad -\infty < t < \infty$$

and f, g are non-negative Borel-measurable functions. Leindler [15] generalized the inequality (2.2) to the case using $\lambda, 1 - \lambda$ in place of $1/2, 1/2$. This was again generalized in [23] for the case of functions of m variables. Thus we have the inequality

$$(2.4) \quad \int_{\mathbb{R}^m} r(t) dt \geq \left[\int_{\mathbb{R}^m} f^{1/\lambda}(x) dx \right]^\lambda \left[\int_{\mathbb{R}^m} g^{1/(1-\lambda)}(x) dx \right]^{1-\lambda},$$

where

$$(2.5) \quad r(t) = \sup_{\lambda x + (1-\lambda)y=t} f(x)g(y)$$

and f, g are non-negative Borel-measurable functions which implies the Lebesgue-measurability of r . (This is proved in [20] for the case of $m = 1$ and the same proof can be used for $m \geq 2$. The formulation of Theorem 1 in [20] has to be corrected in such a way that we assume f and g to be Borel-measurable and infer that r is Lebesgue-measurable. For this corrected theorem, the proof given in [20], can be used. Just note that on p. 306 of [20] E_1, \dots, E_N are Borel-measurable, hence their projections, H_1, \dots, H_N , are Lebesgue-measurable.)

Inequality (2.4) gives an immediate proof for Theorem 2. The proof of the inequality is, however, very sophisticated. It can be established relatively easily for the case of logarithmic concave functions f, g which will suffice for the proof of Theorem 2 and leads to a simplification in the proof of the theorem.

We remark that if a function defined on \mathbb{R}^m is logarithmic concave, then the set of vectors on which the function is positive is convex. This implies that the function is continuous in the interior of this convex set and it follows that the function is Borel-measurable on \mathbb{R}^m . It is easy to see that when the functions f and g are logarithmic concave, the function r defined by (2.5) is also logarithmic concave and hence is Borel-measurable.

Proof of Theorem 2. First we prove inequality (2.4) for the case of $m = 1$. We assume for the moment that both f and g are bounded. Let us introduce the notation

$$\sup_{x \in \mathbb{R}} f(x) = U, \quad \sup_{y \in \mathbb{R}} g(y) = V.$$

It follows from this that

$$\sup_{t \in \mathbb{R}} r(t) = UV.$$

If at least one of the numbers U, V is equal to 0, then (2.4) holds trivially. Thus we can assume that $U > 0, V > 0$.

We remark that if h is a measurable function satisfying the inequality $0 \leq h(x) \leq 1$ for every $x \in \mathbb{R}$, then we have

$$(2.6) \quad \int_{-\infty}^{\infty} h(x) \, dx = \int_0^1 H(z) \, dz,$$

where

$$H(z) = \mu[\{x \mid h(x) \geq z\}], \quad 0 \leq z \leq 1.$$

Let $0 \leq z < 1, 0 < \lambda < 1$, and define

$$F(z) = \mu \left[\left\{ x \mid \frac{1}{U} f(x) \geq z^\lambda \right\} \right],$$

$$G(z) = \mu \left[\left\{ y \mid \frac{1}{V}g(y) \geq z^{1-\lambda} \right\} \right],$$

$$R(z) = \mu \left[\left\{ t \mid \frac{1}{UV}r(t) \geq z \right\} \right].$$

We have the following relation;

$$(2.7) \quad \{t \mid r(t) \geq z\} \supset \lambda\{x \mid f(x) \geq z^\lambda\} + (1 - \lambda)\{y \mid g(y) \geq z^{1-\lambda}\}.$$

All sets participating in this relation are non-empty and they are intervals since f, g, r are logarithmic concave. Relation (2.7) implies that

$$(2.8) \quad R(z) \geq \lambda F(z) + (1 - \lambda)G(z).$$

Integrating (2.8) on both sides between 0 and 1 and using Equation (2.6) we conclude that

$$\int_{-\infty}^{\infty} \frac{1}{UV}r(t) dt \geq \lambda \int_{-\infty}^{\infty} \left[\frac{1}{U}f(x) \right]^{1/\lambda} dx + (1 - \lambda) \int_{-\infty}^{\infty} \left[\frac{1}{V}g(y) \right]^{1/(1-\lambda)} dy.$$

This implies (2.4) by the arithmetic-geometric means inequality.

If at least one of the functions f, g is unbounded, then we define the functions

$$f_U(x) = \begin{cases} f(x) & \text{if } f(x) < U, \\ U & \text{if } f(x) \geq U, \end{cases}$$

$$g_V(y) = \begin{cases} g(y) & \text{if } g(y) < V, \\ V & \text{if } g(y) \geq V, \end{cases}$$

and take limits as $U \rightarrow \infty, v \rightarrow \infty$ in the inequality

$$\int_{-\infty}^{\infty} r(t) dt \geq \int_{-\infty}^{\infty} \sup_{\lambda x + (1-\lambda)y=t} f_U(x)g_V(y) dt$$

$$\geq \left\{ \int_{-\infty}^{\infty} [f_U(x)]^{1/\lambda} dx \right\}^\lambda \left\{ \int_{-\infty}^{\infty} [g_V(y)]^{1/(1-\lambda)} dy \right\}^{1-\lambda}.$$

This has already been established since f_U, g_V are bounded logarithmic concave functions for all $U > 0, V > 0$. Thus we obtain Inequality (2.4) for all logarithmic concave functions f and g .

We suppose that we have already proved Inequality (2.4) for functions f, g of at most $m - 1$ variables and prove it for functions of m variables. Let $h(u, v)$ be a logarithmic concave function of the vector variables $u \in \mathbb{R}^{m_1}, v \in \mathbb{R}^{m_2}$. The logarithmic concavity of the function implies that

$$h(\lambda u_1 + (1 - \lambda)u_2, v) \geq [h(u_1, v)]^\lambda [h(u_2, v)]^{1-\lambda}$$

for every u_1, u_2, v_1, v_2 , $0 < \lambda < 1$, where $v = \lambda v_1 + (1 - \lambda)v_2$. If $m_2 \leq m - 1$, then we can apply Inequality (2.4) and obtain

$$\begin{aligned} \int_{\mathbb{R}^{m_2}} h(\lambda u_1 + (1 - \lambda)u_2, v) dv &\geq \int_{\mathbb{R}^{m_2}} \sup_{\lambda v_1 + (1-\lambda)v_2 = v} [h(u_1, v_1)]^\lambda [h(u_2, v_2)]^{1-\lambda} dv \\ &\geq \left[\int_{\mathbb{R}^{m_2}} h(u_1, v) dv \right]^\lambda \left[\int_{\mathbb{R}^{m_2}} h(u_2, v) dv \right]^{1-\lambda}. \end{aligned}$$

This means that the integral of $h(u, v)$ with respect to v is a logarithmic concave function of u .

Now let f and g be logarithmic concave functions of m variables and partition the variables as $x = (x_1, x_2)$, $y = (y_1, y_2)$ so that $x_1, y_1 \in \mathbb{R}^{m_1}$, $x_2, y_2 \in \mathbb{R}^{m_2}$, where $1 \leq m_1 \leq m - 1$, $1 \leq m_2 \leq m - 1$, $m_1 + m_2 = m$. Then we can write

$$\begin{aligned} &\int_{\mathbb{R}^m} \sup_{\lambda x + (1-\lambda)y = t} f(x)g(y) dt \\ &= \int_{\mathbb{R}^{m_1+m_2}} \sup_{\substack{\lambda x_1 + (1-\lambda)y_1 = t_1 \\ \lambda x_2 + (1-\lambda)y_2 = t_2}} f(x_1, x_2)g(y_1, y_2) dt_1 dt_2 \\ &\geq \int_{\mathbb{R}^{m_2}} \sup_{\lambda x_2 + (1-\lambda)y_2 = t_2} \left[\int_{\mathbb{R}^{m_1}} \sup_{\lambda x_1 + (1-\lambda)y_1 = t_1} f(x_1, x_2)g(y_1, y_2) dt_1 \right] dt_2 \\ &\geq \int_{\mathbb{R}^{m_2}} \sup_{\lambda x_2 + (1-\lambda)y_2 = t_2} \left[\int_{\mathbb{R}^{m_1}} f^{1/\lambda}(x_1, x_2) dx_1 \right] \\ &\quad \times \left[\int_{\mathbb{R}^{m_1}} g^{1/(1-\lambda)}(y_1, y_2) dy_1 \right]^{1-\lambda} dt_2 \\ &\geq \left[\int_{\mathbb{R}^{m_1+m_2}} f^{1/\lambda}(x_1, x_2) dx_1 dx_2 \right]^\lambda \left[\int_{\mathbb{R}^{m_1+m_2}} g^{1/(1-\lambda)}(y_1, y_2) dy_1 dy_2 \right]^{1-\lambda} \\ &= \left[\int_{\mathbb{R}^{m_1}} f^{1/\lambda}(x) dx \right]^\lambda \left[\int_{\mathbb{R}^{m_2}} g^{1/(1-\lambda)}(y) dy \right]^{1-\lambda}. \end{aligned}$$

Thus we have proved Inequality (2.4) for logarithmic concave functions.

Let A, B be convex sets in \mathbb{R}^m and $0 < \lambda < 1$ and define the functions f_1, f_2, f_3 in terms of the probability density f as follows:

$$\begin{aligned} f_1(x) &= f(x) \quad \text{if } x \in A \text{ and } f_1(x) = 0 \text{ otherwise,} \\ f_2(x) &= f(x) \quad \text{if } x \in B \text{ and } f_2(x) = 0 \text{ otherwise,} \\ f_3(x) &= f(x) \quad \text{if } x \in \lambda A + (1 - \lambda)B \text{ and } f_3(x) = 0 \text{ otherwise.} \end{aligned}$$

Since f is a logarithmic concave function, we can write

$$f_3(t) \geq \sup_{\lambda x + (1-\lambda)y = t} [f_1(x)]^\lambda [f_2(y)]^{1-\lambda}.$$

By Inequality (2.4) this implies that

$$\begin{aligned} \int_{\lambda A + (1-\lambda)B} f(t) dt &= \int_{\mathbb{R}^m} f_3(t) dt \geq \left[\int_{\mathbb{R}^m} f_1(x) dx \right]^\lambda \left[\int_{\mathbb{R}^m} f_2(y) dy \right]^{1-\lambda} \\ &= \left[\int_A f(x) dx \right]^\lambda \left[\int_B f(y) dy \right]^{1-\lambda} \end{aligned}$$

Thus Theorem 2 is proved. \square

In the course of the proof we also proved the following theorem, first published in [23].

THEOREM 3. *If $f(x, y)$ is a logarithmic concave function of all variables in the vectors $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, then*

$$\int_{\mathbb{R}^m} f(x, y) dy$$

is a logarithmic concave function of the variable x .

Finally, let us make one remark in connection with logarithmic convex functions. (A function f is *logarithmic convex* if $1/f$ is logarithmic concave.) Let A be a convex subset of \mathbb{R}^m and consider the sets $A + x_1$, $A + x_2$, $A + [\lambda x_1 + (1 - \lambda)x_2]$. If f is a logarithmic convex function defined on a convex subset of \mathbb{R}^m which contains the above three sets, then a simple application of the Hölder inequality shows that

$$P(A + [\lambda x_1 + (1 - \lambda)x_2]) \leq [P(A + x_1)]^\lambda [P(A + x_2)]^{1-\lambda}.$$

3 Convolutions

Logarithmic concave functions were first investigated by Fekete [8] in 1912. He introduced the notion of a *multiple positive sequence*; we may call a logarithmic concave sequence a twice positive sequence, or, in other terms, a discrete logarithmic concave function. The sequence $\{p_i\}$ is said to be *logarithmic concave* if

$$p_i \geq (p_{i-1}, p_{i+1})^{1/2}, \quad i = 0, \pm 1, \pm 2, \dots$$

from which one can derive the inequality

$$p_i \geq p_{i-j}^{k/(j+k)} p_{i+k}^{j/(j+k)}$$

for all positive integers j, k . Fekete proved that the convolution of two logarithmic concave sequences is also logarithmic concave. From this one can derive the theorem of Schoenberg [27] stating that if f and g are two logarithmic concave functions

defined on \mathbb{R} , then their convolution is also logarithmic concave on \mathbb{R} . The generalization of this theorem for logarithmic concave functions defined on \mathbb{R}^m was proved first by Davidovich, Korenblum and Hacet [7]. See also [23] where this statement is derived from Theorem 3 of Section 2. In fact, if $f(x)$ and $g(y)$ are logarithmic concave functions of $x, y \in \mathbb{R}^m$, then the function

$$f(x - y)g(y)$$

is logarithmic concave in \mathbb{R}^{2m} . Hence by Theorem 3 its integral with respect to y is a logarithmic concave function of the remaining variable x .

Results in connection with convolutions are important in this framework because using convolution properties we can prove theorems for functions of the type

$$P(A + x) = \int_{A+x} f(t) dt, \quad x \in \mathbb{R}^m,$$

which is the convolution of the function f and the indicator function of the set $-A$. If A is a convex set, then the indicator function of $-A$ is logarithmic concave. Hence the theorem that the convolution of two logarithmic concave functions in \mathbb{R}^m is also logarithmic concave implies that if f is a logarithmic concave function in \mathbb{R}^m , then $P(A + x)$ is also logarithmic concave in \mathbb{R}^m — which is an important special case of Theorem 2. For the case of the normal distribution the logarithmic concavity of $P(A + x)$, where A is a convex set, was proved first by Zalgaller [33].

In 1956 Ibragimov [11] published a theorem that contains the following

THEOREM 4. *Let $f \geq 0$ be a quasi-concave, and g a logarithmic concave function defined on \mathbb{R} . Then the convolution of these functions is quasi-concave on \mathbb{R} .*

One can give a simple counter-example showing that the corresponding statement in m -dimensional space does not hold. Though we are not dealing with discrete probability distributions, it is interesting to mention that Keilson and Gerber [14] proved the discrete version of Ibragimov's theorem in the one-dimensional case. Ibragimov's Theorem 4 implies

THEOREM 5. *Let $f \geq 0$ be a quasi-concave function defined on \mathbb{R} and $I \subset \mathbb{R}$ be an interval. Then the function*

$$\int_{I+x} f(t) dt$$

is a quasi-concave function of the variable x .

As noted in [9], the convolution of two univariate, unimodal distributions is not necessarily unimodal. However, Wintner [32] proved that the convolution of two univariate symmetric unimodal distributions is unimodal. A generalization of this result is the following theorem, from Sherman [28].

THEOREM 6. Let us introduce the norm $\|f\|_3$ for real-valued functions defined on \mathbb{R}^m by the equation

$$\|f\|_3 = \max\{\|f\|, \|f\|_1\},$$

where

$$\|f\| = \sup_{x \in \mathbb{R}^m} \{|f(x)|\}, \quad \|f\|_1 = \int_{\mathbb{R}^m} |f(x)| dx.$$

Let C be the closed (with respect to $\|\cdot\|_3$) convex cone generated by indicator functions of convex symmetric sets in \mathbb{R}^m . Then for $f, g \in C$, we have $f * g \in C$, where

$$(f * g)(x) = \int_{\mathbb{R}^m} f(x-t)g(t) dt.$$

The above theorem implies Anderson's inequality [1] which we formulate in the following theorem.

THEOREM 7. Let f be a quasi-concave probability density defined on \mathbb{R}^m satisfying $f(x) = f(-x)$ and let D be a convex symmetric subset of \mathbb{R}^m . Then for every $y \in \mathbb{R}^m$ and $0 \leq \lambda \leq 1$ we have

$$(3.1) \quad P(D + \lambda y) = \int_{D+\lambda y} f(x) dx \geq \int_D f(x) dx.$$

Proof. We show that Theorem 6 implies Theorem 7. In fact for the indicator function g of a symmetric convex set the inequality $g(\lambda y) \geq g(y)$ holds for every $y \in \mathbb{R}^m$ and $0 \leq \lambda \leq 1$. It follows that the same inequality holds for convex combinations of such functions and also for functions which are uniform limits of the latter, i.e. for all functions belonging to C . Since

$$\int_{D+\lambda y} f(x) dx$$

is the convolution of f and the indicator function of the set $-D$, Theorem 6 implies that it belongs to C hence (3.1) follows. \square

4 Convex Measures in \mathbb{R}^m

Following the work in [20]; Borell developed the notion of "convex measure" in \mathbb{R}^m and proved important theorems [3]. We mention here only the definition and one theorem which seems to be the most important for stochastic programming. The probability measure P defined on the (Borel) measurable subsets of \mathbb{R}^m will be said

to be *convex (of order s)* if for every nonempty pair A, B of convex subsets of \mathbb{R}^m and every $0 < \lambda < 1$ we have

$$P(\lambda A + (1 - \lambda)B) \geq \{\lambda[P(A)]^s + (1 - \lambda)[P(B)]^s\}^{1/s},$$

where $-\infty < s < \infty$, $s \neq 0$. The cases $s = -\infty$, $s = +\infty$ and $s = 0$ are interpreted by continuity. Thus if $s = -\infty$, the right hand side equals $\min(P(A), P(B))$ and for $s = 0$ we obtain as a special case the notion of a logarithmic concave measure. Convex measures of order $s = -\infty$ will be called *quasi-concave*. In his paper [3] Borell proved the following result.

THEOREM 8. *If f is the probability density of a continuous probability distribution in \mathbb{R}^m and $f^{-1/m}$ is convex in the entire space, then the probability measure*

$$P(C) = \int_C f(x) dx$$

defined on the (Borel) measurable subsets C of \mathbb{R}^m is quasi-concave.

In the next section we show that among the well known probability distributions there are some which are quasi-concave and not logarithmically concave. This emphasizes the importance of Theorem 8.

5 Special Joint and Conditional Probability Distributions

In this section we make some general remarks concerning special probability distributions. A few of them will be analysed in detail. We include a consideration of conditional distributions, since they frequently appear in stochastic programming models.

Let $\zeta \in \mathbb{R}^N$ and $\eta \in \mathbb{R}^n$ be two vector-valued random variables and suppose that their joint distribution is absolutely continuous with probability density $f(x, y)$ where x refers to ζ and y refers to η . It follows that the joint distribution of the components of η is absolutely continuous with probability density

$$g(y) = \int_{\mathbb{R}^N} f(x, y) dy.$$

If f is a logarithmic concave function in \mathbb{R}^{N+n} then, by Theorem 3, the function g is a logarithmic concave function in \mathbb{R}^n . Thus the marginal densities belonging to a logarithmic concave density are logarithmic concave. This statement was first formulated in [23].

The conditional probability density of ζ given $\eta = y$ is given by the following

$$f(x | y) = \frac{f(x, y)}{g(y)}.$$

Obviously $f(x | y)$ is logarithmic concave in x for any fixed y provided $f(x, y)$ is logarithmic concave. Thus for conditional distributions we can derive formulae similar to those mentioned in the previous sections. Also if f satisfies the condition of Theorem 8, i.e. $[f(x, y)]^{1/n}$ is convex in \mathbb{R}^{N+n} , then $[f(x | y)]^{-1/n}$ is convex in \mathbb{R}^N for any fixed y and thus the conditional probability measure is quasi-concave.

We give an example in which $f(x, y)$ is logarithmic concave in \mathbb{R}^{N+n} and the same holds for $f(x | y)$ as a function of both arguments. A further example will show that this is not always the case.

Let us consider the non-degenerate normal distribution in \mathbb{R}^m . Its probability density function is given by

$$f(z) = \left(\frac{\det C^{-1}}{(2\pi)^m} \right)^{1/2} e^{-(1/2)(z-\mu)'C^{-1}(z-\mu)}, \quad x \in \mathbb{R}^m,$$

where μ is the vector of expectations and C is the covariance matrix. C is non-singular since the distribution is non-degenerate. Hence f is a logarithmic concave function so that the corresponding probability distribution is logarithmic concave.

Let us now consider a non-degenerate normal distribution in m dimensional space ($m = N+n$) and suppose that it is the joint distribution of the N -component random vector ζ and the n -component random vector η . Partition C and μ accordingly in the following manner

$$C = \begin{pmatrix} S & U \\ U' & T \end{pmatrix},$$

$$\mu = \begin{pmatrix} \nu \\ \tau \end{pmatrix}.$$

The conditional distribution of ζ is given $\eta = y$ is again a normal distribution (see e.g. [2]) with expectation vector

$$\nu + UT^{-1}(y - \tau)$$

and covariance matrix

$$S - UT^{-1}U'.$$

Hence the conditional probability density has the following form

$$(5.1) \quad f(x | y) = K \exp \left\{ -\frac{1}{2} [x - \nu - UT^{-1}(y - \tau)]' \right. \\ \left. \times [S - UT^{-1}U']^{-1} [x - \nu - UT^{-1}(y - \tau)] \right\},$$

where

$$K = \left[\frac{\det(S - UT^{-1}U)}{(2\pi)^N} \right]^{1/2},$$

Function (5.1) is logarithmic concave in both variables x and y . Using this and the theorems of the previous sections, various statements can be proved. For example, it is a consequence of Theorem 3 that if G is a convex subset of \mathbb{R}^N , then

$$\int_G f(x | y) dx$$

is a logarithmic concave function of y .

Consider now the Dirichlet distribution – for the sake of simplicity in \mathbb{R}^2 . The probability density of this distribution is given by

$$f(x, y) = \frac{\Gamma(p_1 + p_2 + p_3)}{\Gamma(p_1)\Gamma(p_2)\Gamma(p_3)} x^{p_1-1} y^{p_2-1} (1 - x - y)^{p_3-1},$$

if $x > 0$, $y > 0$, $1 - x - y > 0$ and $f(x, y) = 0$ otherwise. The numbers p_1 , p_2 , p_3 are supposed to be positive. Suppose that $p_1 \geq 1$, $p_2 \geq 1$, $p_3 \geq 1$. Then $f(x, y)$ is logarithmic concave in \mathbb{R}^2 . If this is the joint distribution of the random variables ζ and η then the probability density of η equals

$$g(y) = \int_{-\infty}^{\infty} f(x, y) dx = \frac{\Gamma(p_1 + p_2 + p_3)}{\Gamma(p_2)\Gamma(p_1 + p_3)} y^{p_2-1} (1 - y)^{p_1+p_3-1}$$

if $0 < y < 1$ and $g(y) = 0$ otherwise. Thus the conditional density of ζ given $\eta = y$ has the form

$$f(x | y) = \frac{\Gamma(p_1 + p_3)}{\Gamma(p_1)\Gamma(p_3)} \left(\frac{x}{1 - y} \right)^{p_1-1} \left(1 - \frac{x}{1 - y} \right)^{p_1+p_3-1},$$

for $0 < x < 1 - y$ and $f(x, y) = 0$ otherwise. It is easy to see that there are parameters p_1 , p_2 , p_3 for which this function is not a logarithmic concave function in \mathbb{R}^2 .

As mentioned in [20], the multivariate beta, Dirichlet, and Wishart distributions are logarithmic concave (for suitable parameter values). Many other distributions belong to this category. The reader may consult [13] and check that logarithmic concavity is quite often a property of probability densities.

Examples of probability densities satisfying the condition of Theorem 8 are given in [3]; they include the multivariate t and the F densities and one of the multivariate Pareto densities introduced in [16]. The Pareto density in question is given by

$$f(z) = a(a + 1) \dots (a + m - 1) \prod_{j=1}^m d_j \left(\sum_{j=1}^m \frac{z_j}{d_j} - m + 1 \right)^{-(a+m)}$$

for $z_j > d_j$, $j = 1, \dots, m$ and $f(z) = 0$ otherwise, where a, d_1, \dots, d_m are positive constants. It is interesting to note that this function is also logarithmic convex in the domain $z_j > d_j$, $j = 1, \dots, m$, and hence the remark made in Section 1 in connection with logarithmic convex functions applies for this density.

6 Special Functions Appearing in Stochastic Programming Models

In this section we analyse certain functions which appear as constraint or objective functions in stochastic programming models. The theorems formulated below contain statements expressing analytic properties of certain functions which are advantageous when numerically solving the problems in which these functions appear. The first theorem was published in [21].

THEOREM 9. *If $g_1(x, y), \dots, g_r(x, y)$ are concave functions in \mathbb{R}^{m+q} , where x is an m -component and y is a q -component vector, and ζ is a q -component random vector whose probability distribution is logarithmic concave in \mathbb{R}^q , then the function*

$$(6.1) \quad h(x) = P(g_1(x, \zeta) \geq 0, \dots, g_r(x, \zeta) \geq 0), \quad x \in \mathbb{R}^m$$

is logarithmic concave on \mathbb{R}^m .

Proof. Let us consider the following family of sets

$$H(x) = \{y \mid g_i(x, y) \geq 0, \quad i = 1, \dots, r\},$$

where $x \in \mathbb{R}^m$ is a parameter. If for some $x_1, x_2 \in \mathbb{R}^m$ we have $H(x_1) \neq \emptyset$ and $H(x_2) \neq \emptyset$, then for every $0 < \lambda < 1$ we also have $H(\lambda x_1 + (1 - \lambda)x_2) \neq \emptyset$. In fact if $y_1 \in H(x_1)$ and $y_2 \in H(x_2)$, then, since g_1, \dots, g_r are concave functions, it follows that

$$\begin{aligned} g_i(\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2) &\geq \\ \lambda g_i(x_1, y_1) + (1 - \lambda)g_i(x_2, y_2) &\geq 0, \quad i = 1, \dots, r, \end{aligned}$$

hence $\lambda y_1 + (1 - \lambda)y_2 \in H(\lambda x_1 + (1 - \lambda)x_2)$. From this we also see that the set

$$H = \{x \mid H(x) \neq \emptyset\}$$

is convex. If H is empty then (6.1) is identically zero and the assertion holds trivially. If $H \neq \emptyset$ then we observe that as a consequence of the above reasoning, the family $H(x)$ is concave on H , i.e. for $x_1, x_2 \in H$ and $0 < \lambda < 1$,

$$H(\lambda x_1 + (1 - \lambda)x_2) \supset \lambda H(x_1) + (1 - \lambda)H(x_2).$$

Then using the equation

$$h(x) = P(\zeta \in H(x)),$$

valid for every $x \in \mathbb{R}^m$, and Theorem 2, we derive

$$\begin{aligned} h(\lambda x_1 + (1 - \lambda)x_2) &= P(\zeta \in H(\lambda x_1 + (1 - \lambda)x_2)) \\ &\geq P(\zeta \in \lambda H(x_1) + (1 - \lambda)H(x_2)) \\ &\geq [P(\zeta \in H(x_1))]^\lambda [P(\zeta \in H(x_2))]^{1-\lambda} \\ &= [h(x_1)]^\lambda [h(x_2)]^{1-\lambda} \end{aligned}$$

This means that $h(x)$ is logarithmic concave on the convex set H . Since $h(x) = 0$ if $x \notin H$ it follows that $h(x)$ is logarithmic concave on the entire space \mathbb{R}^m . \square

The next theorem can be proved by the same method except that instead of Theorem 2 we must use Theorem 8.

THEOREM 10. *Suppose that the functions $g_1(x, y), \dots, g_r(x, y)$ are as in Theorem 9. If ζ is a q -component random vector having a quasi-concave probability distribution on \mathbb{R}^q , then the function $h(x)$ defined by (6.1) is quasi-concave on \mathbb{R}^m .*

These theorems can be applied to many important stochastic programming models. We refer to the reservoir system design models described in [22] and [21]. In these cases we were able to show that some fairly complex optimization problems were convex programming problems.

The above theorems have important theoretical consequences as well. For example, let us consider the function of the variables $x_1, x_2 \in \mathbb{R}$ given by

$$(6.2) \quad P(x_1 \leq \zeta \leq x_2),$$

where ζ is a random variable. Jagannathan [12] shows that if ζ has a uniform or exponential distribution then this function is logarithmic concave in \mathbb{R}^2 . Now from Theorem 9 it follows that the function given by (6.2) is logarithmic concave if ζ has an absolutely continuous distribution with logarithmic concave density. In fact if we introduce the functions of the variables x_1, x_2, y defined by

$$\begin{aligned} g_1(x, y) &= y - x_1 \\ g_2(x, y) &= x_2 - y, \end{aligned}$$

where $x \in \mathbb{R}^2$ is the vector having components x_1, x_2 , then clearly

$$P(x_1 \leq \zeta \leq x_2) = P(g_1(x, \zeta)) \geq 0, \quad g_2(x, \zeta) \geq 0$$

and our assertion follows.

THEOREM 11. *Let ζ be an absolutely continuous random variable having a quasi-concave density and let $c \geq 0$, $d \geq 0$ be constants and $a \in \mathbb{R}^m$ be a constant vector. Then the function of the variable vector x given by*

$$(6.3) \quad P(\zeta - c \leq a'x \leq \zeta + d)$$

is quasi-concave on \mathbb{R}^m .

Proof. The proof is based on Theorem 5. Since f is a quasi-concave function, Theorem 5 implies that

$$\int_{z-d}^{z+c} f(x) dx$$

is quasi-concave in $-\infty < z < \infty$. Thus

$$(6.4) \quad P(\zeta - c \leq z \leq \zeta + d) = P(z - d \leq \zeta \leq z + c) = \int_{z-d}^{z+c} f(x) dx$$

is quasi-concave in $-\infty < z < \infty$. The function given by (6.3) arises from (6.4) if we replace z by $a'x$. Since a quasi-concave function of a linear function is quasi-concave, our theorem is proved. \square

Van de Panne and Popp [30] proved that $\zeta_1, \dots, \zeta_n, \zeta$ are random variables having a joint normal distribution, then the set

$$\{x \mid P(\zeta_1 x_1 + \dots + \zeta_n x_n \leq \zeta) \geq p\}$$

is convex on \mathbb{R}^m provided $p \geq \frac{1}{2}$. In the original proof ζ was supposed to be a constant, but the generalization to the above situation is trivial. A similar result has been obtained for the case of random variables having stable distributions [29], but independent and identically distributed random variables ζ_i are required. We shall consider the more general function given by

$$(6.5) \quad P(\mathbf{A}x \leq \zeta)$$

in which some or all entries of the matrix \mathbf{A} are allowed to be random as well. \mathbf{A} will be supposed to be an $m \times n$ matrix, its columns will be denoted by ζ_1, \dots, ζ_n and for ζ we shall use the alternative notation $\zeta = -\zeta_{n+1}$. The following three theorems are proved in [19], [24].

THEOREM 12. *Suppose that the $m(n+1)$ components of the random vectors $\zeta_1, \dots, \zeta_{n+1}$ have a joint normal distribution where the cross-covariance matrices of ζ_i and ζ_j are constant multiples of a fixed covariance matrix C , i.e.*

$$E[(\zeta_i - \mu_i)(\zeta_j - \mu_j)'] = s_{ij}C, \quad i, j = 1, \dots, n+1,$$

where

$$\mu_i = E(\zeta_i), \quad i = 1, \dots, n+1.$$

Then the set

$$(6.6) \quad \{x \mid P(\mathbf{A}x \leq \zeta) \geq p\}$$

is convex for every fixed $p \geq \frac{1}{2}$.

The assumption concerning the cross-covariances is a very special one. There are, however, important cases where such a condition is satisfied. If for example only one of the vectors $\zeta_1, \dots, \zeta_{n+1}$ is random, the others are constants and the components of the random vector have a joint normal distribution, then the assumptions of Theorem 12 are valid.

THEOREM 13. *Let \mathbf{A}_i denote the i th row of the random $m \times n$ matrix \mathbf{A} and let ζ_i denote the i th component of ζ , $i = 1, \dots, m$. Suppose that the random $(n+1)$ -component row vectors*

$$(\mathbf{A}_i, -\zeta_i), \quad i = 1, \dots, m$$

are independent, normally distributed and have covariance matrices which are constant multiples of a fixed covariance matrix C . Then the set (6.6) is convex for every fixed $p \geq \frac{1}{2}$.

Before formulating the third theorem, consider again the probability

$$G(x_1, \dots, x_n) = P(\zeta_1 x_1 + \dots + \zeta_n x_n \geq \zeta).$$

If ζ_1, \dots, ζ_n are positive-valued random variables and the joint distribution of $\eta_1 = \log \zeta_1, \dots, \eta_n = \log \zeta_n$ and ζ is logarithmic concave, then a simple application of Theorem 9 shows that

$$G(e^{z_1}, \dots, e^{z_n}) = P(-e^{\eta_1+z_1} - \dots - e^{\eta_n+z_n} + \zeta \geq 0)$$

is logarithmic concave in z comprising the components z_1, \dots, z_n . Since every set of positive numbers x_1, \dots, x_n can be represented in the form e^{z_1}, \dots, e^{z_n} , this result is important from the point of view of mathematical programming. The portfolio selection problem with log normally distributed returns seems to be one possible field of application. For the formulation of the general theorem we need further notation.

Let \mathbf{a}_{ij} , $i = 1, \dots, m$; $j = 1, \dots, n$ be the elements of the matrix \mathbf{A} and for the sake of simplicity suppose that the columns of \mathbf{A} are numbered so that those which contain random variables come first. Let r be the number of these columns and introduce the following notation:

J denotes the set of those ordered pairs (i, j) for which \mathbf{a}_{ij} is a random variable, $1 \leq i \leq m, 1 \leq j \leq r$,

L denotes the set of those subscripts i for which ζ_i is random variable, where $1 \leq i \leq m$.

THEOREM 14. *Suppose that the random variables \mathbf{a}_{ij} $(i, j) \in J$ are positive with probability 1 and the constant a_{ij} , $1 \leq i \leq m, 1 \leq j \leq r, (i, j) \notin J$ are non-negative. Suppose further that the joint distribution of the random variables*

$$\alpha_{ij}, (i, j) \in J, \beta_i, i \in L$$

is a logarithmic concave probability distribution, where $\alpha_{ij} = \log \mathbf{a}_{ij}, (i, j) \in J, \beta_i = \log \zeta_i, i \in L$. Under these conditions the function

$$G(e^{x_1}, \dots, e^{x_r}, x_{r+1}, \dots, x_n)$$

is logarithmic concave on \mathbb{R}^n , where G is the function defined by (6.5).

The proofs of the last three theorems are based on special cases of Theorem 9.

The following theorem helps in checking quasi-concavity of probability distribution functions. A function g defined on an m -dimensional interval is said to be *concave in the positive direction* if for every x, y belonging to this interval satisfying $x \leq y$ we have

$$g(\lambda x + (1 - \lambda)y) \geq \lambda g(x) + (1 - \lambda)g(y),$$

where $0 \leq \lambda \leq 1$.

THEOREM 15. *Let $F(x, y)$ be a probability distribution function on \mathbb{R}^m , where x is an m_1 -component vector, y is an m_2 -component vector and $1 \leq m_1, m_2 \leq m - 1, m_1 + m_2 = m$. If $F(x, y)$ is concave in the positive direction with respect to x (respectively y) in an m_1 -dimensional (respectively m_2 -dimensional) interval, then $F(x, y)$ is quasi-concave on the Cartesian product of these intervals.*

This theorem is proved in [18]. In order to show how one can apply it, let us consider the two-dimensional extreme-value distribution [13] given by

$$F(x, y) = \exp[-e^{-x} - e^{-y} + d(e^x + e^y)^{-1}],$$

where $-\infty < x, y < \infty$ and d is a constant, $0 \leq d \leq 1$. It is easy to see that for fixed values of the remaining argument, $F(x, y)$ is concave in the finite intervals $\{x \mid x \geq 0\}$ and $\{y \mid y \geq 0\}$, respectively. Hence by Theorem 15, the function $F(x, y)$ is quasi-concave in the two-dimensional interval $\{(x, y) \mid x \geq 0, y \geq 0\}$.

Finally, we formulate a consequence of Anderson's Theorem (Theorem 7) which states that a certain set is star shaped. A set K in \mathbb{R}^m is said to be *star shaped* with respect to the point $b \in \mathbb{R}^m$ if the intersection of K and the ray $\{b + \lambda z \mid \lambda \geq 0\}$ is an interval for every fixed $z \in \mathbb{R}^m$.

THEOREM 16. Let $\zeta \in \mathbb{R}^m$ be a random vector having an absolutely continuous probability distribution. Suppose that the probability density f of ζ is quasi-concave and satisfies the equality $f(-z) = f(z)$ for every $z \in \mathbb{R}^m$. Then the set

$$K = \{x \mid P(\zeta - c \leq Ax \leq \zeta + d) \geq p\}$$

is of a star shape with respect to any x_0 which satisfies $Ax_0 = \frac{1}{2}(d - c)$, provided K is not empty and such an x_0 exists. Here $c \geq 0$, $d \geq 0$ are constant vectors, A is an $m \times n$ constant matrix and p is a fixed probability level, $0 < p < 1$.

Proof. Theorem 7 implies that the set

$$L = \left\{ z \mid P\left(z - \frac{1}{2}(c + d) \leq \zeta \leq z + \frac{1}{2}(c + d)\right) \leq p \right\}$$

is star shaped with respect to the origin. Hence the set

$$L + \frac{1}{2}(d - c) = \{z \mid P(z - d \leq \zeta \leq z + c) \geq p\} = \{z \mid P(\zeta - c \leq z \leq \zeta + d) \geq p\}$$

is star shaped with respect to the point $\frac{1}{2}(d - c)$. Since

$$K = \left\{ x \mid Ax \in L + \frac{1}{2}(d - c) \right\} = \{x_0 + x \mid Ax \in L\}$$

and $\{x \mid Ax \in L\}$ is obviously star shaped with respect to the origin, our theorem is proved. \square

The star shaped property of the set of feasible vectors may be helpful when developing an algorithm for solving the relevant stochastic programming problem.

References

- [1] Anderson, T. W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.*, **6**, 170–176.
- [2] Anderson, T. W. (1958). “An Introduction to Multivariate Statistical Analysis”. Wiley, New York.
- [3] Borell, Chr. (1973). “Convex set functions in d -space”. Uppsala Univ. Dept. of Math., Report No. 8.
- [4] Brunn, H. (1887), “Über Ovale und Eiflächen”. Inaugural Dissertation, München.

- [5] Charnes, A., W. W. Cooper and G. H. Symonds (1958). Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil production. *Management Sci.* **4**, 235–263.
- [6] Charnes, A. and M. J. L. Kirby (1966). Optimal decision rules for the E-model of chance-constrained programming. *Cahiers Centre Études Recherche Opér.* **8**, 5–44.
- [7] Davidovich, Yu. S., B. J. Korenblum and B. J. Hacet (1969). On a property of logarithmic concave functions. *Dokl. Math. Acad. Nauk.* **185**, 1215–1218.
- [8] Fekete, M. (1912). Über ein problem von Laguerre. *Rend. Circ., Mat. Palermo* **34**, 89–100, 110–120.
- [9] Gnedenko, B. V. and A. N. Kolmogorov (1954), “Limit Distributions for Sums of Independent Random Variables”. Addison-Wesley, Cambridge, Mass. (Translated from Russian.)
- [10] Hadwiger, H. (1957). “Vorlesungen über Inhalt, Oberfläche und Isoperimetrie”. Springer Verlag, Berlin, Göttingen, Heidelberg.
- [11] Ibragimov, I. A. (1956). On the composition of unimodal distributions. *Theor. Probability App.* **1** 255–260 (in Russian).
- [12] Jagannathan, R. and M. R. Rao (1973). A class of nonlinear chance-constrained programming models with joint constraints. *Operations Res.* **21**, 360–364.
- [13] Johnson, N. L. and S. Kotz (1972). “Distributions in Statistics: Continuous Multivariate Distributions”. Wiley, New York.
- [14] Keilson, J. and H. Gerber (1971). Some results for discrete unimodality. *J. Amer. Statist. Assoc.* **66**, 386–389.
- [15] Leindler, L. (1972). On a certain converse of Hölder’s inequality II. *Acta. Sci. Math.* **33** 217–223.
- [16] Mardia, K. V. (1962). Multivariate Pareto distributions. *Ann. Math. Statist.* **33**, 1008–1015.
- [17] Markovitz, H. M. (1959). “Portfolio Selection: Efficient Diversification of Investments”. Wiley, New York.
- [18] Prékopa, A. (1970). On probabilistic constrained programming. In “Princeton Symp. on Math. Prog.” Princeton University Press, Princeton, N. J., 113–138.
- [19] Prékopa, A. (1970). On the optimization problems of stochastic systems. Hungarian Academy of Sciences, Budapest (in Hungarian).
- [20] Prékopa, A. (1971). Logarithmic concave measures with application to stochastic programming. *Acta Sci. Math.* **32**, 301–316.
- [21] Prékopa, A. (1972). A class of stochastic programming decision problems. *Math. Operationsforsch. Statist.* **3**, 349–354.
- [22] Prékopa, A. (1973a). Stochastic programming models for inventory control and water storage problems, Inventory Control and Water Storage. *Coll. Math. Soc. János Bolyai* **7**, 229–245.

- [23] Prékopa, A. (1973b). On logarithmic concave measures and functions. *Acta. Sci. Math.* **34**, 335–343.
- [24] Prékopa, A. (1974). Programming under probabilistic constraints with a random technology matrix. *Math. Operationsforsch. Statist.* **5**, 109–116.
- [25] Prékopa, A. (1975). Optimal control of a storage level using stochastic programming. *Problems of Control and Information Theory* **4**, 193–204.
- [26] Raike, W. M. (1974). Solutions to triangular conditional E-models arising from inventory problems. Presented at the International Conference on Stochastic Programming, Oxford, July 15–17.
- [27] Schoenberg, I. J. (1951). On Pólya frequency functions I. The totally positive functions and their Laplace transforms. *J. Analyse Math.* **1**, 331–374.
- [28] Sherman, S. (1955). A theorem on convex sets with applications. *Ann. Math. Statist.* **26**, 763–767.
- [29] Vajda, S. (1972). “Probabilistic Programming”. Academic Press, New York, London.
- [30] Van de Panne, C. and W. Popp (1963). Minimum-cost cattle feed under probabilistic protein constraints. *Management Sci.* **9**, 405–430.
- [31] Wilks, S. S. (1962). “Mathematical Statistics”. Wiley, New York, London.
- [32] Wintner, A. (1938). “Asymptotic Distributions and Infinite Convolutions”. Edwards Brothers, Ann Arbor, Mich.
- [33] Zalgaller, V. A. (1967). Mixed volumes and the probability of falling into convex sets in case of multivariate normal distributions. *Math. Zametki* **2**, 97–104 (in Russian).