ON MULTIVARIATE UNIMODAL DISTRIBUTIONS

By

Tao Dai

B. Sc. (Mathematics), Wuhan University, 1982

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

DEPARTMENT OF STATISTICS

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

June 1989

67

Department of Statistics

The University of British Columbia
Vancouver, Canada

Date June 28, 1989

ABSTRACT

In this thesis, Kanter's representation of multivariate unimodal distributions is shown equivalent to the usual mixture of uniform distributions on symmetric, compact and convex sets. Kanter's idea is utilized in several contexts by viewing multivariate distributions as mixtures of uniform distributions on sets of various shapes. This provides a unifying viewpoint of what is in the literature and gives some important new classes of multivariate unimodal distributions. The closure properties of these new classes under convolution, marginality and weak convergence, etc. and their relationships with other notions of multivariate unimodality are discussed. Some interesting examples and their 2- or 3-dimensional pictures are presented.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

ACKNOWLEDGEMENT

I would like to thank Professor Albert W. Marshall for his guidance and assistance in the producing of this thesis. I would also like to express my gratitude to Professor Harry Joe, who kindly gave me helpful advice and carefully read this thesis.

On Multivariate Unimodal Distributions

1. Introduction

For univariate distributions there is a generally accepted definition of unimodality, which does not require the existence of densities. The term of unimodality is considered here as an indication of concentration about the centre rather than as the shape of a (density) function. Khintchine's representation (Khintchine, 1938) gives several equivalent ways to define unimodality on the line; these definitions when extended to higher dimensions are not equivalent. Several attempts were made to translate the geometric notion of unimodality in R^n into analytic forms. The earliest attempt was made by Anderson (1955). During the last thirty years, and especially since 1970, multivariate unimodality has received a lot of attention in the literature. More definitions of multivariate unimodality were given and several desirable properties (such as closure under convolution, mixture, marginality, product measures and weak convergence) that a class of multivariate unimodal distributions may or may not satisfy were extensively examined. See Anderson (1955), Sherman (1955), Olshen and Savage (1970), Ghosh (1975), Das Gupta (1976a, 1976b), Dharmadhikari and Jogdeo (1976), Kanter (1977), Wells (1978), Uhrin (1984).

Application of multivariate unimodality include concepts of dependence and probability inequalities. Some concepts of dependence can be combined with suitable concepts of dependence to yield useful probability inequalities. Anderson type of inequalities are another example. Arguments involving unimodality have been used quite often in statistical inference for proving the unbiasedness of some multivariate tests and constructing minimum volume confidence regions. For more recent comprehensive discussions on multivariate unimodality and its applications, see Dharmadhikari and Joag-dev (1988).

Kanter (1977) generalized Khintchine's representation of univariate unimodality to higher dimensions in the symmetric case; he defines symmetric multivariate unimodal distributions on R^n as generalized mixtures (in the sense of integrating on a probability measure space) of uniform distributions on symmetric, compact and convex sets (symmetric intervals when $n = 1$) in R^n. In this thesis we show that Kanter's representation of multivariate unimodal distributions as a generalized mixture is equivalent to the usual mixture of uniform distributions on symmetric, compact and convex sets. We utilize Kanter's idea in several contexts by viewing multivariate distributions as mixtures of uniform distributions on sets of various shapes such as star-shaped, connected, or the unions of convex sets, etc. This provides a unifying viewpoint of what is in the literature and gives some important new classes of multivariate unimodal distributions. The closure properties of these new classes under convolution, marginality and weak convergence, etc. and their relationships with other notions of multivariate unimodality are also discussed in this thesis. Some interesting examples and their 2- or 3-dimensional pictures are given. All pictures are drawn using graphical and perspective commands in S (Becker and Chambers, 1984)..

Section 2 contains a brief discussion of some basic results involving unimodal distributions on the real line. In Section 3 we review the various definitions of multivariate unimodal distributions and their inter-relationships in the literature. In Section 4 we obtain an equivalent form of Kanter's representation using the usual mixture approach, which we use later to unify several notions of multivariate unimodality. In this section we also show Kanter's unimodality is equivalent to central convex unimodality (Dharmadhikari and Joag-dev (1988) obtain independently the same result). In Section 5 we use the approach obtained in Section 4 to generalize Anderson's unimodality, and a decomposition theorem of generalized Anderson unimodality is obtained. Furthermore we show that the class of generalized Anderson unimodal distributions is closed under weak convergence when the dimension $n = 2$ but not when $n > 2$. In Section 5, we also show that, by an example, the notion of star or n-unimodality (due to Olshen and Savage, 1970; see also Dharmadhikari and Joag-dev, 1988) is somewhat unnatural in the geometrical sense. To avoid this drawback, we present in Section 6 the notion called strong star unimodality and show the class of strongly star unimodal distributions is closed under projection to a subspace under some mild symmetry conditions. The various properties about this unimodality are discussed. In the last Section the preservation properties for the various definitions are examined.

2. Univariate Case

On the real line the term "unimodal density" sometimes refers to a density function f that has a maximum at unique point $x = a$ and strictly decreases as x goes away from a in either direction. The normal and the Cauchy distributions are unimodal in this sense. If we want to include distributions whose support is only a part of of the real line, like the gamma and the uniform distributions, it becomes clear that one will have to allow the density f to be just nonincreasing as x goes away from the mode a, which needs not to be unique. Because densities are not unique, it is desirable to have a definition in terms of distribution functions.

2. 1. Definition (Khintchine, 1938). A probability distribution F on R is said to be *unimodal* if there exists a number a, called mode, such that F is convex on (∞, a) and concave on (a, ∞).

It is known that a unimodal distribution function has right and left derivatives everywhere except possibly at the mode of the distribution (the point a) and that these derivatives increase monotonically for $x < a$ and decrease monotonically for $x > a$.

If F is unimodal and has a density function, then some version f of the density is non-decreasing for $x < a$ and non-increasing for $x > a$. In the other words, the sets $\{x: f(x) \geq u\}$ are intervals for all u, $0 < u < \infty$.

Furthermore Khintchine (1938) shows the following well known representation theorem :

2. 2. Theorem (Khintchine, 1938). *A real random variable Z has a unimodal distribution with mode at* 0 *if and only if* $Z \sim UX$ *(that is, Z is distributed like UX), where U is uniform on* $[0, 1]$ *and U and X are independent.*

Let Φ denote the set of all distributions on R which are unimodal about 0. It is well known that Φ is closed under weak convergence. In other words, the limit of a sequence of unimodal distributions is itself unimodal (Gnedenko and Kolmogorov, 1954). Clearly Φ is also convex (under mixture). Let Γ denote the set of all uniform distributions on intervals with 0 as one end point. Then an equivalent statement of Khintchine's theorem is that Φ is the closed (in the sense of weak topology) convex hull of Γ.

If Z is also symmetric, i.e. $Z \sim -Z$, we can modify Khintchine's representation of a unimodal distribution about 0 in a slightly different way: $Z \sim UX$, U is uniform on $[-1, 1]$, and $X \geq 0$. In this case, S, the set of all symmetric unimodal distributions on R is the closed convex hull of W, where W is the set of all uniform distributions on symmetric intervals about 0.

Wintner (1938) showed that the convolution of two symmetric and unimodal distributions is symmetric and unimodal. The conclusion is not true in general if the assumption of symmetry is dropped; see Chung's Appendix II of Gnedenko and Kolmogorov (1954), also Feller (1966, page 164) and Ibragimov (1956, page 255). Ibragimov (1956) proved a theorem that offers some solace for the disappointment of Chung's discovery. The theorem says that the convolution of a unimodal distribution function F with any unimodal distribution function is unimodal if and only if F is continuous and $\Phi(x) = \log F'(x)$ is a concave function; then F is said to be strongly unimodal.

Finally, it is clear that the mixture $F = \int F_\theta \, dG(\theta)$ of unimodal distributions with a common mode $\boldsymbol{a}$ is also unimodal with the same mode. If each F_θ is symmetric about $\boldsymbol{a}$, then so is F.

3. Multivariate Definitions of Unimodality, A Review

As we have seen in the previous section, it is rather obvious how a unimodal distribution should be defined on the real line. However, the choice of a definition of unimodality based on the distribution function in higher dimensions is not so clear, even if attention is restricted to the symmetric case. Several attempts were made to translate the geometric notion of unimodality in R^n into analytic form. In this section we will review some of these multivariate extensions and see how they relate to each other. The earliest attempt was made by Anderson (1955):

3. 1. Definition (Anderson, 1955). A probability distribution in R^n is *Anderson unimodal* if it possesses a density f with respect to Lebesgue measure μ_n such that for all $u \geq 0$ the sets $\{x: f(x) \geq u, x \in R^n\}$ are convex and symmetric about the origin whenever they are non-empty.

It should be noted that Anderson unimodality is equivalent to symmetric unimodality when $n = 1$.

By introducing the above definition, Anderson shows the following theorem:

3.2. Theorem (Anderson, 1955). *Let* E *be a symmetric (i.e.,* $E = -E$ *) convex set in* R^n *and let* f *be a density function on* R^n *with respect to Lebesgue measure* μ_n. *Then for any fixed* $y \in R^n$ *and* $0 \leq \lambda \leq 1$,

$$\int_E f(x + \lambda y)\, d\mu_n(x) \geq \int_E f(x + y)\, d\mu_n(x). \tag{1}$$

or equivalently,

$$\int_{E+\lambda y} f(x)\, d\mu_n(x) \geq \int_{E+y} f(x)\, d\mu_n(x) \tag{2}$$

where $E + \lambda y = \{ x + \lambda y : x \in E \}$.

It should be pointed out that (2) is equivalent to

$$P(E+\lambda y) \geq P(E+y). \tag{3}$$

There a simple geometric explanation for (3): the probability over a symmetric convex set E is nondecreasing if the centre of E is moving away from the origin along a given direction.

The class of Anderson unimodal distributions in R^n ($n > 1$) is not closed under convolution. Sherman (1955) gives the following example due to Anderson (1955):

3.3. Example (Sherman,1955). *Let* $f = 2\chi_A + \chi_B$, *where* $A:=\{x \in R^2: |x_1| \leq 1, |x_2| \leq 1\}$, $B:=\{x \in R^2: |x_1| \leq 1, |x_2| \leq 5\}$, *and where* χ_A *and* χ_B *are indicator functions. Then,* f *and* χ_A *are Anderson unimodal, but* $f * \chi_A$ *is not.*

Sherman (1955) modified Anderson's definition by considering f as a member of the closure (with respect to the maximum of the sup-norm and the L_1-norm) of the convex cone generated by the indicator functions of compact, symmetric convex sets in R^n containing 0 in their interiors. He showed that Anderson's theorem still holds for f in his generalized class, and moreover his class of unimodal distributions is closed under convolution.

There are some drawbacks to Sherman's definition as well as to Anderson's. First, they both have to assume that probability measures have density functions, which is not required in the univariate case. Secondly, it seems that it is not easy to deal with the norms Sherman uses, and they are not natural in this context.

Kanter (1977) defined a probability measure on R^n to be symmetric unimodal if it is a generalized mixture (in the sense of integrating with respect to a probability measure) of all uniform probability measures on symmetric, compact, convex sets in R^n. This definition is based on Sherman's idea. But, it avoids the drawback of Sherman's definition.

We need some notation before presenting details of Kanter's definition. We will follow Kanter's notations.

Notation 1. If L is a linear subspace of R^n of dimension $k > 0$, we let Vol_L stand for k-dim Lebesgue measure on L, with the normalizing condition that $\mathrm{Vol}_L(B(1)) = \pi^{k/2} / \Gamma(1 + k/2)$, where $B(1) = \{y: y \in R^n, \|y\| \leq 1\}$. Let $\mathrm{Vol}_n = \mathrm{Vol}_{R^n}$. And if $L = \{\mathbf{0}\}$, we let $\mathrm{Vol}_L(A) = 1$, if $\mathbf{0} \in A$; $\mathrm{Vol}_L(A) = 0$, otherwise, for all $A \in B^n$, where B^n is the class of Borel subsets of R^n.

Notation 2. If K is a nonempty compact, convex subset of R^n we define the probability measure λ_K on the Borel subsets of R^n by

$$\lambda_K(A) = \mathrm{Vol}_L(A \cap K) / \mathrm{Vol}_L(K)$$

where L is the affine subspace spanned by K.

We often deal with sets S which consist of a subset of the collection of all probability measures on R^n and we will need to have a uniform convention for the concept of Borel subsets of S. Our convention will be to endow S with the Borel structure consisting of the smallest σ-field containing all those subsets U of S, which are open in the topology of weak convergence of probability measures.

Notation 3. Let L be a linear subspace of R^n, and let W_L stand for the set of all probability measures on L of the form λ_K (defined as before) where K is a compact symmetric convex subset of L. We give W_L the Borel structure mentioned above. We will let W_n stand for W_{R^n}.

Notation 4. Let U_L stand for the set of all the probability measures F on R^n which are of the form

$$F = \int_{W_L} \lambda_K \, d\nu(\lambda_K),$$

where ν is a probability measure on W_L. For notational convenience, let $U_n = U_{R^n}$.

3.4. Definition (Kanter, 1977). A probability measure F on R^n is called *Kanter unimodal* if F is in U_n.

3.5. Theorem (Kanter, 1977). *The class of Kanter unimodal probability measures on* R^n *is closed under convolution , marginality , product measure and weak convergence.*

We will discussed Kanter's definition more fully in Section 4 and we will use Kanter's approach to somehow unify several kinds of definitions of multivariate unimodality.

Dharmadhikari and Jogdeo (1976) introduced the notion of a *central convex unimodal* distribution, the definition of which is also based on Sherman's idea.

3.6. Definition (Dharmadhikari & Jogdeo, 1976). A distribution in R^n is called *central convex unimodal* if it is in the closed (in the sense of the weak convergence) convex hull of the set of all uniform distributions on symmetric compact convex bodies in R^n.

We will show in the next section that the class of all *central convex unimodal* distributions on R^n is equivalent to the class of all unimodal distributions in the sense of Kanter's definition on R^n.

It should be noted that the main difference between Definition 3.6. and Sherman's definition is that instead of dealing with density functions, it deals with the class of probability measures, which is a metric space under the weak topology. The following extension of Anderson's theorem was first proved in the absolutely continuous case by Sherman (1955).

3.7. Theorem (Sherman, 1955). *If a distribution* P *on* R^n *is central convex unimodal and* E *is a symmetric (i.e.,* E = − E *) convex set in* R^n*, then for any fixed* y $\in R^n$*) and* $0 \le \lambda \le 1$ *, we have*

$$P(E + \lambda y) \geq P(E + y). \qquad (4)$$

Based on the above result, another definition of multivariate unimodality was also formulated by Dharmadhikari and Jogdeo (1976) as follows:

3. 8. Definition (Dharmadhikari and Jogdeo, 1976). A probability measure on R^n is called (symmetric) *monotone unimodal (SMUM)* if for every convex set E in R^n symmetric about 0, the quantity $P(E + \lambda y)$ is nonincreasing in $\lambda \geq 0$ for every fixed non-zero vector y in R^n.

It is easy to see by Theorem 3.7 that every *central convex unimodal* distribution is *monotone unimodal.* So, Sherman (1955) conjectured the converse might be true. But Dharmadhikari and Jogdeo (1976) suggested a possible counterexample, which was indeed proved to be a counterexample by Wells (1978). The following is the counterexample:

3.9. Example. Let $A_1, \dots, A_6$ be the vertices of a regular hexagon centered at the origin. Let T_1, T_2 respectively, be the triangles $A_1A_3A_5$ and $A_2A_4A_6$. The set $T_1 \cup T_2$ consists of six outer triangles and an inner hexagon. Let P be a distribution with support $T_1 \cup T_2$ having density α on the outer triangles and β on the inner hexagon, where $2\alpha \leq \beta < 3\alpha$.

Dharmadhikari and Jogdeo (1976) as well as Ghosh (1974) introduced the so-called *linear unimodal* distributions.

3.10. Definition (Ghosh, 1974; Dharmadhikari and Jogdeo, 1976). A random variable X in R^n is *linear unimodal (LUM)* if for every vector a in R^n, the distribution of a' X is unimodal in the univariate sense.

This definition, however is somewhat unnatural, as Dharmadhikari and Jogdeo themselves and other authors have pointed out, because the density of such a "unimodal" distribution may not become maximum at the mode of univariate unimodality. Dharmadhikari and Jogdeo (1976) gave the following example:

3.11. Example. *Consider the bivariate density*

$$f(x,y) = ke^{-(x^2+y^2)/2}\left[e^{\alpha(x^2+y^2)/2} - \beta\right] \qquad x, y \in R$$

where $0 < \alpha < 1$, $\beta < 1$ *and k is so chosen that f is a density. It can be verified that if* $\beta > (1 - \alpha)$ *then there a* $\delta > 0$ *such that* $f(x, x)$ *is strictly increasing in* $x \in [0, \delta]$ *. On the other hand if* $\beta < (1 - \alpha)^{1/2}$*, then the x-marginal of f is unimodal. The circular symmetry of the distribution shows that if* $(1 - \alpha) < \beta < (1 - \alpha)^{1/2}$ *then f is linear* *unimodal about* 0, *but the density is not maximized at* (0, 0). *See Figure 3.12 for the cut of the bivariate density.*

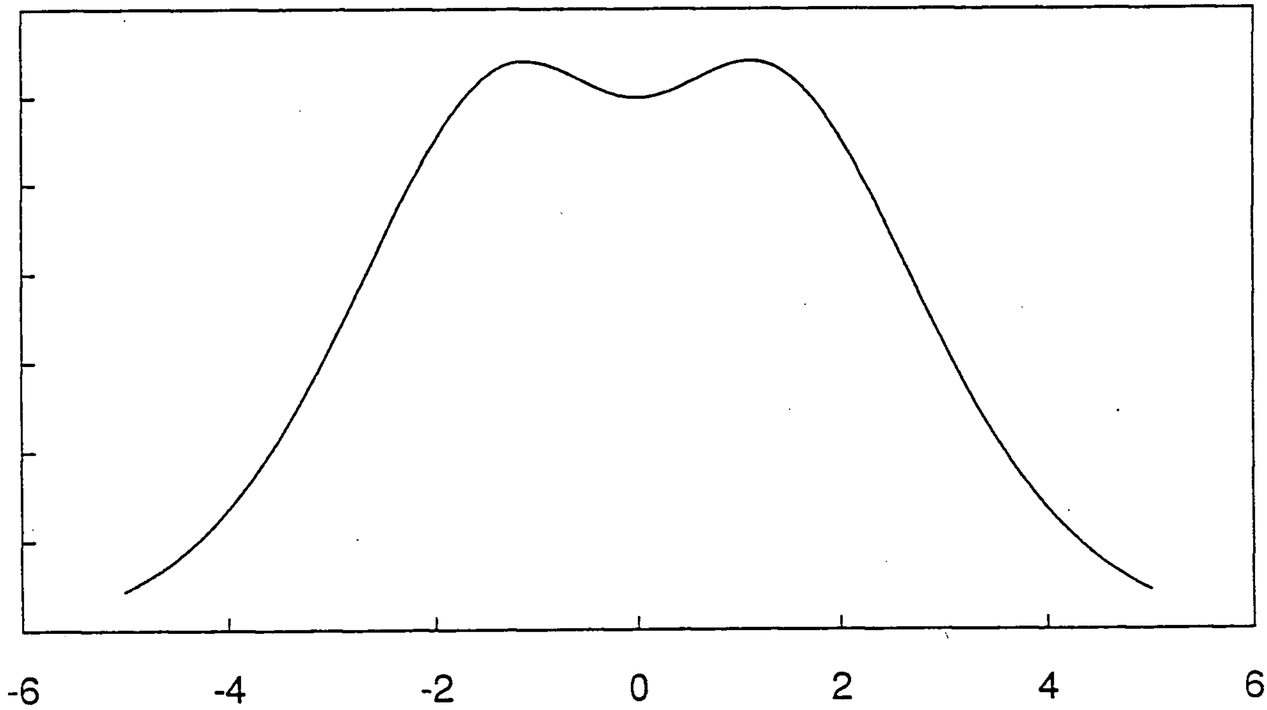

Figure 3.11

The next definition is due to Olshen and Savage (1970):

3.12. Definition (Olshen & Savage, 1970). A random variable X in R^n is called α-unimodal about 0 if for all real, bounded, nonnegative Borel function g on R^n the function $t^{\alpha}E(g(tX))$ is non-decreasing for all fixed X as t increases in $[0, \infty)$.

When an α-unimodal random variable X in R^n has a density f, Olshen and Savage (1970) show that it has the following property:

3.13. Theorem (Olshen & Savage, 1970). *If a random variable X in* R^n *possesses a density function* f *with respect to Lebesgue measure* μ_n *on* R^n*, then X is* α*-unimodal if and only if* $t^{n-\alpha} f(tx)$ *is decreasing in* $t \geq 0$.

When $\alpha = n$ in the above theorem, $f(tx)$ is decreasing in t for all fixed x . So , the sets $\{f \geq u\}$, $u > 0$, are starshaped (centered at 0). Some authors refer to n-unimodality (n is dimension of the random variable) as star unimodality.

Dharmadhikari and Jogdeo (1976) discussed the relationships of the classes of monotone unimodal (MUM), linear unimodal (LUM) and star unimodal (n-UM) distributions on R^n. They gave the following theorem:

3.14. Theorem (Dharmadhikari and Jogdeo, 1976). *Every monotone unimodal distribution is linear unimodal and n-unimodal.*

But, there is no implication relating n-unimodal and linear unimodal. Example 3.11 shows that a linear unimodal distribution is not necessary an n-unimodal distribution . Since univariate marginals of an n-unimodal distribution on R^n need not be unimodal in the univariate sense (i.e. 1-unimodal), it is easy to see that an n-unimodal distribution need not be linear unimodal. We will see in Section 7 that the kdimensional marginals of an n-unimodal distributions are n-unimodal for $1 \leq k < n$, but no lower index of unimodality can be asserted.

Some other definitions of generalized multivariate unimodality were also introduced by Das Gupta (1976) and Uhrin (1984).

Das Gupta (1980) reviewed some of the above definitions of multivariate unimodality and gave the following description of the inter-relationships of those definitions:

$$
\begin{array}{ccccccc}
\text{A-UM(Anderson)} & \Rightarrow & \text{SUM(Sherman)} & \Rightarrow & & \text{SUM(Kanter)} & \\
 & & & & & \Downarrow & \\
 & & & & & \text{SMUM(D\&J)} & \\
 & & & & \Downarrow & & \Downarrow \\
 & & & & \text{LUM} & \not\Rightarrow & \text{n-UM(O\&S)}
\end{array}
$$

All the implications are strict.

Now we give some examples of the well-known densities and see which notions of multivariate unimodality can be fitted in.

3.15. Example (a) Mutivariate normal distribution.

Let

$$f(x) = C \exp(-x' \Sigma^{-1} x/2), \qquad x \in R^n .$$

where C is a suitable constant and Σ is a fixed positive definite, symmetric matrix of order $n \times n$. Since Σ is positive definite, the function $x' \Sigma^{-1} x$ is convex in x. Thus the sets $\{x: f(x) \geq u, x \in R^n\}$ are convex for all $u > 0$.

(b) Wishart distribution.

Let

$$g(A) = C\,(\det A)^{(n-p-1)/2} \exp[-(\mathrm{tr} A\Sigma^{-1})/2],$$

where A is positive definite and $g(A) = 0$, otherwise. Here A is a symmetric matrix variable of order $p \times p$ and Σ is a fixed positive definite, symmetric matrix of order $p \times p$. Assume $n \geq (p + 1)$. If we write $A = (a_{ij})$, it can be shown that det A is logconcave in the a_{ij}'s. This is true by the following well known inequality

$$\det[\theta A_1 + (1 - \theta)A_2] \geq (\det A_1)^{\theta} (\det A_2)^{1-\theta},$$

where A_1, A_2 are positive definite and $0 < \theta < 1$. See, for example, Marshall and Olkin (1979, page 476). Since $\mathrm{tr} A\Sigma^{-1}$ is linear in the a_{ij}'s, g is logconcave. Thus the sets $\{x: g(x) \geq u, x \in R^{p(p+1)/2}\}$ are convex for all $u > 0$ for $n \geq (p + 1)$.

(c) Mutivariate t distribution with m degrees of freedom.

On R^n, let

$$h(x) = K \frac{1}{\left(1 + \frac{x'x}{m}\right)^{(n+m)/2}} .$$

By an argument similar to (a), it follows that *h* is Anderson unimodal.

All three examples of above are Anderson unimodal, hence as well as unimodal in all other senses introduced above.

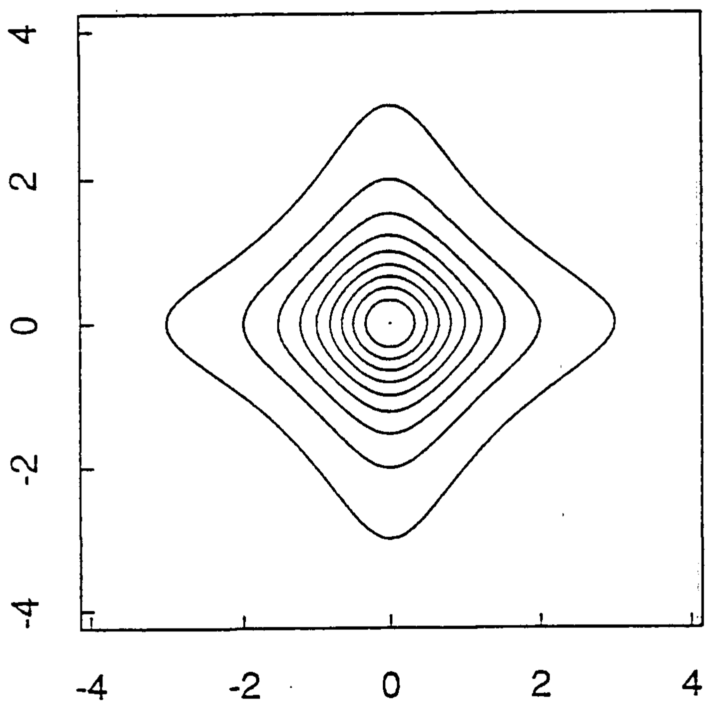

Figure 3.16

3.16. Example. Let

$f(x, y) = 1 / [\pi^2(1 + x^2)(1 + y^2)]$, $\quad x, y \in R$. For $(x, y) \in R^2$, The marginals of f are Anderson unimodal, hence f is Kanter unimodal (Theorem 3.4), but it is not Anderson unimodal , since contour sets $\{f \geq u\}$ are starshaped but not convex for small $u > 0$. See Figure 3.15.

4. More About Kanter's Unimodality

In this section, we will establish an equivalent form of Kanter's unimodality by using the usual mixture over an index set in the real line instead of an abstract metric space of all probability measures on R^n. We will use the same notations as in Section 3.

4.1. Theorem. *A probability measure* F *on* R^n *is in the class of Kanter's unimodal distributions if and only if* F *can be represented as*

$$F(A) = \int_0^1 \lambda_{K_\theta}(A)\ dP(\theta) \qquad \textit{for all } A \in B^n \tag{5}$$

where for all $\theta \in [0, 1]$, K_θ *is a compact symmetric convex subset of* R^n, B^n *is the class of Borel subsets of* R^n, P *is a probability measure on* R *with* $P(0) = 0$, $P(1) = 1$, *and* $\lambda_{K_\theta}(A)$ *is measurable in* θ *for all* $A \in B^n$

Proof. Let S be the set that consists of all probability measures on B^n. It is well known(see Parthasarathy, 1968) that S can be made into a complete separable metric space. By Theorem 6.6.6 (Ash, 1972; page 265), S is Borel equivalent to a Borel subset E of [0, 1]. (The metric space Ω is said to be Borel equivalent to a subset of the metric space Ω' if and only if there is a one-to-one map f: $\Omega \rightarrow \Omega'$ such that $E = f(\Omega) \in B(\Omega')$ and f and f^{-1} are Borel measurable.) So, F in Definition 3.4. can be represented as (5). Q. E. D.

By introducing the above equivalent representation of Kanter's unimodal distributions, we can now more easily deal with this kind of unimodality. In the next section, we will use this idea to formulate other notions of unimodality in R^n. Now we use our new approach to discuss an interesting example.

4.2. Example. Consider the density f on R^2 defined by

$$f(x, y) = \begin{cases} \dfrac{2}{\pi\sqrt{x^2 + y^2}} & x^2+y^2 \leq 1 \text{ and } xy \geq 0, \\ \infty & x = y = 0, \\ 0 & \text{otherwise.} \end{cases}$$

See Figure 4.2, where the picture is drawn for the function $\min\{f(x,y), c\}$, here c is a suitable constant. Since the sets $\{ f \geq u \}$, $u > 0$, are not convex, f is not Anderson unimodal. Now we show that it is unimodal in Kanter's sense.

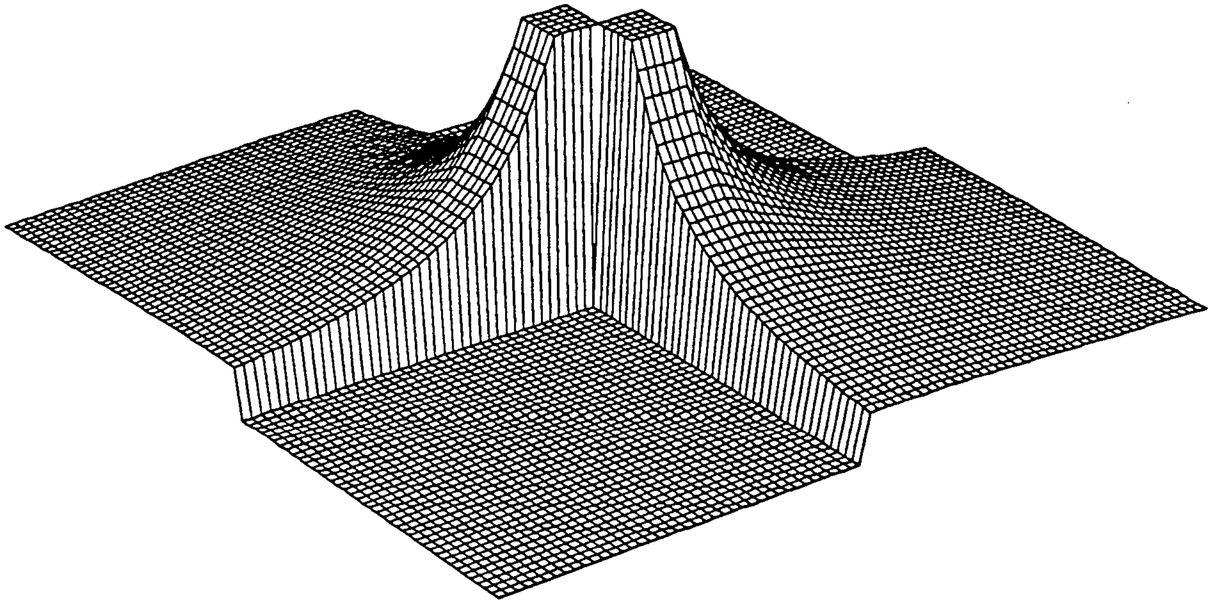

Figure 4.2

Let $K_\theta = \{(x, y): y = x \tan\theta , x^2+y^2 \le 1\}, 0 \le \theta < \frac{\pi}{2}$

$K_{\pi/2} = \{(0, y): |y| \le 1\}$

Define λ_{K_θ} as usual , i.e. λ_{K_θ} is the uniform distribution on K_θ . Let L_θ be the affine subspace spanned by K_θ (L_θ is the real line: $y = x \tan\theta$ on R^2). Then λ_{K_θ} has a density function (with respect to Vol_{L_θ}):

$$f_{K_\theta}(x, y) = \begin{cases} \frac{1}{2} & (x, y) \in K_\theta , \\ 0 & \text{otherwise.} \end{cases}$$

Let $A = \{ (x, y): x^2 + y^2 \le r^2 , x/y \le \tan\varphi \}$, where, $0 \le r \le 1$, $0 \le \varphi \le \pi/2$. Then we have

$$\lambda_{K_\theta}(A) = \begin{cases} \frac{r}{2} & 0 \le \theta < \varphi, \\ 0 & \theta > \varphi. \end{cases}$$

Taking $dP(\theta) = 2d\theta/\pi$, we get

$$\lambda(A) = \frac{2}{\pi}\int_0^{\pi/2} \lambda_{K_\theta}(A)\, d\theta = \frac{2}{\pi}\int_0^{\varphi} \frac{r}{2} d\theta = \frac{r\varphi}{\pi}$$

This means that the probability measure λ has a constant density under polar coordinates. Let $x = r\cos\varphi$, $y = r\sin\varphi$, then $dxdy = r\, drd\varphi$, so under the usual coordinates, λ has the density function

$$f_\lambda(x, y) = \frac{c}{r} = \frac{c}{\sqrt{x^2 + y^2}} \qquad x^2 + y^2 \le 1, xy \ge 0$$

where $c = 2/\pi$ is a constant.

4.3. Remark. An interesting thing about Example 4.4. is that for every θ, λ_{K_θ} is not absolutely continuous with respect to μ_2 (Lebesgue measure on R^2), but their mixture measure λ has a density function with respect to μ_2. If we modify Example 4.4 by allowing $-\pi/2 < \theta < \pi/2$, the density of λ becomes

$$f(x,y)=\begin{cases}\dfrac{2}{\pi\sqrt{x^2+y^2}} & x^2+y^2\le 1\\ \infty & x=y=0,\\ 0 & \text{otherwise.}\end{cases}$$

Clearly this is Anderson unimodal. Let $K_\alpha = \{ f \geq \alpha \}$, $\alpha > 0$, then K_α is convex and symmetric, and

$$f(x,y) = \int_0^\infty I_{K_\alpha}(x,y)\, d\alpha \;.$$

By Fubini theorem, it follows that

$$\lambda(A) = \int_A f(x,y)\, dxdy = \int_0^\infty \mathrm{Vol}(A \cap K_\alpha)\, d\alpha = \int_0^\infty \lambda_{K_\alpha}(A)\mathrm{Vol}(K_\alpha)\, d\alpha \;.$$

Let $P(\alpha) = \int_0^\alpha \mathrm{Vol}(K_u)\, du$, so λ has another representation $\lambda = \int \lambda_{K_\alpha}\, dP(\alpha)$, In other words, the representation in Definition 4.1. or Definition 3.4. is not unique.

4.4. Theorem. F *is a symmetric unimodal distribution on* R^n *in Kanter's sense if and only if* F *is central convex unimodal in* R^n.

Proof. Let W be the set of all uniform distributions on symmetric compact convex bodies with non-zero Lebesgue measures in R^n. And let $W_n = W_{R^n}$ (see Notation 3). Clearly, $W_n \subset \mathrm{cl}(W)$, defined as the closure of W_n under the weak topology. Let U_n be the class of all symmetric unimodal distributions in Kanter's sense as before, and C be the class of CCUM. By Theorem 3.5. , U_n is weakly compact and convex. It is easily to see that W_n = ext (U_n) consists of the extreme points of U_n. By the Krein-Milman theorem (see Holmes,1975), U_n is the closed convex hull of W, i.e. U_n = C. Q. E. D.

5. Generalized Anderson unimodality, Star-Unimodality, and other Definitions

At first we discuss a notion called generalized Anderson unimodality, which is based on Anderson's idea. However, we do not restrict ourselves to multivariate distributions which are absolutely continuous or symmetric.

In this section, we again let λ_K stand for the uniform distribution on $K \subset R^n$, where K is a compact convex subset of R^n, not necessarily symmetric, but only satisfies $0 \in K$. Let W_n stand for the class of all such distributions λ_K, $K \subset R^n$.

5.1. Definition. A probability distribution F on R^n is called *generalized Anderson unimodal* (GA-UM) if there exists a subset M of W_n such that

(a) for all λ_{K_1}, $\lambda_{K_2} \in$ M, either $K_1 \subset K_2$ or $K_2 \subset K_1$,

(b) $$F = \int_M \lambda_K d\nu(\lambda_K) \qquad (6)$$

where ν is a probability measure on W_n with $\nu(M) = 1$.

5.2. Theorem. *Suppose that the probability distribution* F *on* R^n *is absolutely continuous. Then* (a) F *is generalized Anderson unimodal if and only if* (b) *it possesses a density f with respect to Lebesgue measure* μ_n *such that the sets* $\{f \geq \theta\}$, $\theta > 0$, *are convex whenever they are non-empty.*

Proof. Suppose (b). For some version of density of F, consider $D_\theta = \{x: f(x) \geq \theta; x \in R^n\}, \theta > 0$, and $K_\theta = cl(D_\theta)$. Then the K_θ's are convex, compact, and nested. By following exactly the proof of Lemma 3.1.of Kanter (Kanter, 1977; page 70-71), it follows that F can be represented in the form (5).

Next suppose (a). We need to show that there exists a density f of F such that sets $D_\theta = \{x: f(x) \geq \theta; x \in R^n\}, \theta > 0$, are convex . By the definition of GA-UM, F has the representation (6). Since F is also absolutely continuous, it is clear that one version of f is:

$$f(x) = \int_M \frac{1}{Vol(K)} I_K(x)\, d\nu(\lambda_K)$$

where I_K are the indicator functions of convex sets K . For all y, z $\in D_\theta$, $0 < \alpha < 1$, let $M_y = \{ \lambda_K : I_K(y) = 1\}$, and $M_z = \{\lambda_K : I_K(z) = 1\}$. Since all of the convex sets K are nested, we have either $M_y \subset M_z$ or $M_y \supset M_z$. We assume that $M_y \subset M_z$

without loss of generality. Let $f_K(x) = \frac{1}{\mathrm{Vol}(K)} I_K(x)$. It is clear that $M_y \subset M_{\alpha y+(1-\alpha)z}$,$\forall\ 0 < \alpha < 1$, hence,

$$f(\alpha y + (1-\alpha)z) = \int_{M_{\alpha y+(1-\alpha)z}} f_K(x)\, dv(\lambda_K) \geq \int_{M_y} f_K(x)\, dv(\lambda_K) = f(y) \geq \theta$$

i.e. $\alpha y + (1-\alpha)z \in D_\theta$. So, D_θ is convex. Q. E. D.

5.3. Theorem. *If* F *has a density function with respect to* μ_n *and* F *is symmetric, then* F *is generalized Anderson unimodal if and only if* F *is Anderson unimodal.*

Proof. This follows from the symmetry condition on F and Theorem 5.2. Q. E. D.

The following theorem gives us the structure of generalized Anderson unimodal distributions:

5.4. Theorem. *Let* F *be a generalized Anderson unimodal probability distribution on* R^n. *Then there exists* (a) *linear subspaces* L_i *of* R^n *of dimension* i, i = 0, 1, . . ., n *such that* $L_i \subset L_{i+1}$, i = 0, 1, . . ., n-1, *and* (b) *for such* i *that* $\alpha_i > 0$, *probability distributions* F_i *having respective densities* f_i *with respect to* Vol_{L_i} *such that sets of the form* $\{ f_i \geq u \}$ *are convex and compact in* L_i, i = 0, 1. . . , n, *and*

$$F = \sum_{i=0}^{n} \alpha_i F_i \tag{7}$$

where $\alpha_i \geq 0, \sum_{i=0}^{n} \alpha_i = 1$.

Proof. Suppose F is a generalized Anderson unimodal distribution on R^n as in Definition 5.1. Let $M_i = \{\lambda_K: \lambda_K \in M; \dim(K) = i\}$, and $K_i = \{K: \lambda_K \in M_i\}$, $i = 0, 1, \ldots, n$. By (a) in Definition 5.1, it follows that $M = M_0 \cup M_1 \cup \ldots \cup M_n$ and for each K_i, all sets in K_i span the same affine subspace L_i. Then F can be written as

$$F = \sum_{i=0}^{n} \int_{M_i} \lambda_K dv(\lambda_K) .$$

For i such that $\alpha_i = v(M_i) > 0$, let

$$F_i = \frac{1}{v(M_i)} \int_{M_i} \lambda_K dv(\lambda_K).$$

Then

$$F = \sum_{i=0}^{n} \alpha_i F_i .$$

Clearly, for $\alpha_i > 0$, F_i is absolutely continuous with respect to Lebesgue measure on L_i. By Theorem 5.2., F_i must have density f_i such that the set $\{f_i \geq u\}$ is convex and compact in L_i. Since $K_i \subset K_{i+1}$ it follows that $L_i \subset L_{i+1}$. Q. E. D.

5.5. Remark. When $n = 1$, generalized Anderson unimodality is equivalent to univariate unimodality in Definition 2.1, and Theorem 5.4 reduces to the statement that a univariate distribution F is unimodal if and only if $F = \alpha F_0 + (1 - \alpha)F_1$, where F_0 is a point mass at a, $\alpha \in [0, 1]$ and F_1 is absolutely continuous unimodal distribution on the line.

As in Theorem 4.1 we can rewrite (6) in the usual index form and the uniform measures are, in this particular case, continuous (with respect to weak convergence) and monotone.

5.6. Theorem. *A probability distribution* F *on* R^n *is generalized Anderson unimodal if and only if it can be represented as* (5), *where the* K_θ *are compact convex subsets of* R^n *such that* $0 \in K_\theta$ *and* $0 \leq \theta < \theta' \leq 1$ *implies* $K_\theta \subset K_{\theta'}$. *Moreover* λ_{K_θ} *is continuous in* θ *(under weak topology).*

Proof. Let F be a generalized Anderson unimodal distribution on R^n. By Theorem 5.4, we have representation (7), where for such i that $\alpha_i > 0$, F_i has density f_i with respect to the Lebesgue measure on L_i such that the sets $\{f_i \geq u\}$ are convex and compact in L_i, $i = 0, \ldots, n$. Let $D_\theta^i = \{x: f_i(x) \geq 1/\theta; x \in R^n\}$ and $K_\theta^i = \text{cl}\,(D_\theta^i)$, $0 < \theta < 1$. Clearly, $0 < \theta < \theta' < 1$ implies $K_\theta^i \subset K_{\theta'}^i$, for each $i = 0, \ldots, 1$. And $0 \leq i < i' \leq n$ implies $K_\theta^i \subset K_{\theta'}^{i'}$ for all $0 < \theta, \theta' < 1$. By following the similar proof of Theorem 5.2, we have, for each $i = 0, \ldots, 1$,

$$F_i = \int_0^1 \lambda_{K_\theta^i} dP_i(\theta) . \tag{8}$$

Moreover we can make $\lambda_{K_\theta^i}$ continuous by inserting some classes of sets K_θ^i at the discontinuous points and let P_i be zero on those classes. Then the theorem follows by combining F_i by some linear transformations of the index and inserting some continulization classes of sets K_θ^i if necessary. The converse of the theorem is obvious. Q. E. D.

We now prove that the class of symmetric generalized Anderson unimodal distributions on R^n is closed under weak convergence when $n = 2$. Since the class of Anderson unimodal distributions is not closed under weak convergence, this is one reason for introducing the notion of generalized Anderson unimodality.

5.7. Theorem. *The class of all symmetric generalized Anderson unimodal distributions on* R^2 *is closed under weak convergence.*

Proof. Let $\{F_n\}$ be a sequence of symmetric generalized Anderson unimodal distributions on R^2 and F_n converge weakly to F. By Theorem 5.6 it follows that

$$F_n = \int_0^1 \lambda_{K_\theta^n} dP_n(\theta) .$$

Clearly $\{P_n\}$ is tight since $P_n\{[0, 1]\} = 1$ for all $n \geq 1$. It follows by Prokhorov's theorem (Ash, 1972) that $\{P_n\}$ is relatively compact, namely there exists a subsequence $\{P_{n_m}\}$ such that P_{n_m} converges weakly to a probability measure P on [0, 1]. Then for fixed n, it follows (see Chandra, 1977) that

$$\int_0^1 \lambda_{K_\theta^n} dP_{n_m}(\theta) \to \int_0^1 \lambda_{K_\theta^n} dP(\theta)$$

as m goes to infinite. Again by Prokhorov's theorem, by going to a subsequence if necessary, we may assume that , for all fixed θ, $\lambda_{K_\theta^n}$ converges weakly to some probability μ_θ on R^2. So it easily follows that

$$F_{n_m} \to \int_0^1 \mu_\theta \, dP(\theta).$$

Since F_n also converges to F, we see that

$$F = \int_0^1 \mu_\theta \, dP(\theta).$$

So to show the theorem, it suffices to show that for any two sequences of symmetric, compact and convex sets $K^n \subset K'^n \subset R^2$, $n = 1, \ldots, m, \ldots$, if $\{\lambda_{K^n}\}$, $\{\lambda_{K'^n}\}$ converge weakly to μ and μ' respectively, where μ and μ' can be written as

$$\mu = \int_0^1 \lambda_{K_\theta} \, dP(\theta), \quad \mu' = \int_0^1 \lambda_{K'_\theta} \, dP'(\theta),$$

then the sets in the class $\{K_\theta\} \cup \{K'_\theta\}$ are nested. (note that K_θ and K'_θ are increasing in θ respectively.) Since the nested nature of K^n and K'^n, we only need to prove this following three cases:

(a) μ and μ' are absolutely continuous with respect to Lebesgue measure on R^2.

In this case μ and μ' must be uniform distributions on K and K', respectively, where $K = \lim K^n$ and $K' = \lim K'^n$. Clearly $K \subset K'$.

(b) μ and μ' are absolutely continuous with respect to Lebesgue measure on a subspace L with dimension < 2. In this case the sets in $\{K_\theta\} \cup \{K'_\theta\}$ must be some symmetric intervals or $\{0\}$ on L. Thus they must be nested.

(c) μ' is absolutely continuous with respect to Lebesgue measure on R^2 and μ is absolutely continuous with respect to Lebesgue measure on a subspace L with dimension < 2. In this case we have $K_\theta \subset K = \lim K_\theta^n \subset K' = \lim K'^n$, for all $0 < \theta < 1$. Since $\{K'_\theta\} = \{K'\}$, the sets in the class $\{K_\theta\} \cup \{K'_\theta\}$ are nested. Q. E. D.

5.8. Remark. Theorem 5.7 is not true in general without the assumption of symmetry or without the assumption n < 3 as the following two examples show.

5.9. Example. In R^2 let

$$K_{1n} = \text{conv}(\{(0,0), (1,0), (1,1/n)\}),$$
$$K_{2n} = \text{conv}(\{(0,0), (1/2,0), (1/2,1/2n)\}).$$

(conv(A) is defined as the smallest convex set containing A.) A simple computation shows that $\lambda_{K_{1n}} \to \mu_1$, $\lambda_{K_{2n}} \to \mu_2$, where μ_1 is the probability measure on $K_1 = \{(x, y): x \in [0, 1]; y = 0\}$ with density $f_1(x) = 2x$, and μ_2 is the probability measure on $K_1 = \{(x, y): x \in [0, 1/2]; y = 0\}$ with density f_2 is 8x. Let $F_n = (\lambda_{K_{1n}} + \lambda_{K_{2n}})/2$, then $F_n \to F = (\mu_1 + \mu_2)/2$ where F is the probability measure on L with density $f = (f_1 + f_2)/2$. Clearly, the sets $\{f_i \geq u\}$ are not necessary convex. Therefore F is not generalized Anderson unimodal distribution.

5.10. Example. In R^3 let

$$K_n = \{(x, y, z): x^2 + y^2 + n^2 z^2 \le 2\},$$

$$K' = \text{conv}\{(1, 1, 0), (1, -1, 0), (-1, 1, 0), (-1,-1,0)\}.$$

A simple computation shows that $\lambda_{K_n} \to \mu$, where μ is the probability measure on $L = \{(x, y, z): z = 0\}$ whose density with respect to Lebesgue measure on L is $c(2 - x^2 - y^2)$ on $K = \{(x, y, z): x^2 + y^2 \le 2, z = 0\}$; otherwise it is 0. Let

$$F_n = (\lambda_{K_n} + \lambda_{K'})/2$$

then $F_n \to F = (\lambda + \lambda_{K'})/2$. Figure 5.10 (a) and (b) provide the pictures for the density function and it contours. Since the contours are not all convex, F is not generalized Anderson unimodal.

The following notion of multivariate unimodality is due to Dharmadhikari and Joag-dev (1988):

5.11. Definition (Dharmadhikari and Joag-dev, 1988). A probability distribution F on R^n is called *star unimodal* about 0 if it belongs to the closed convex hull of the set of all uniform distributions on sets in R^n which are star-shaped about 0.

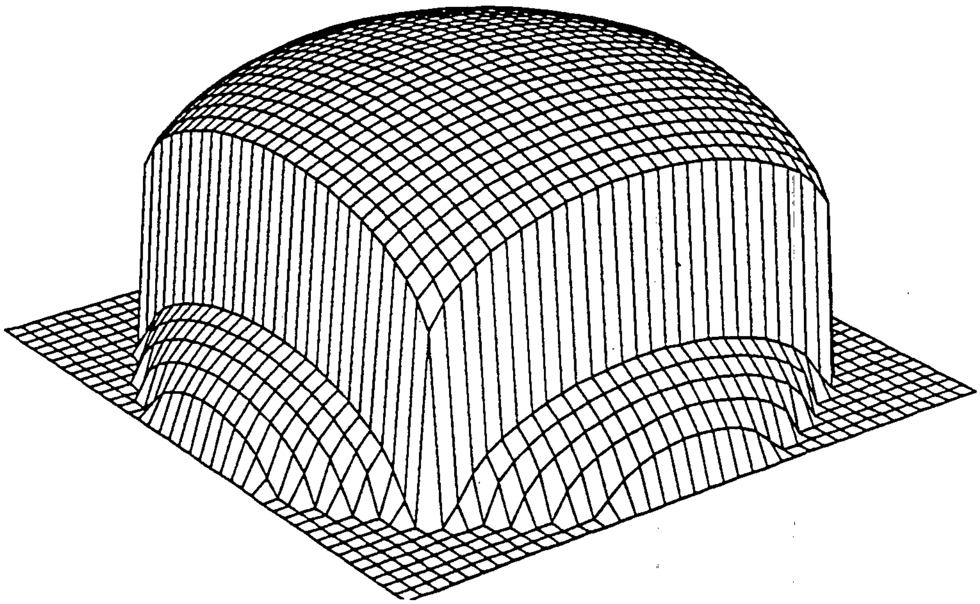

Figure 5.10 (a)

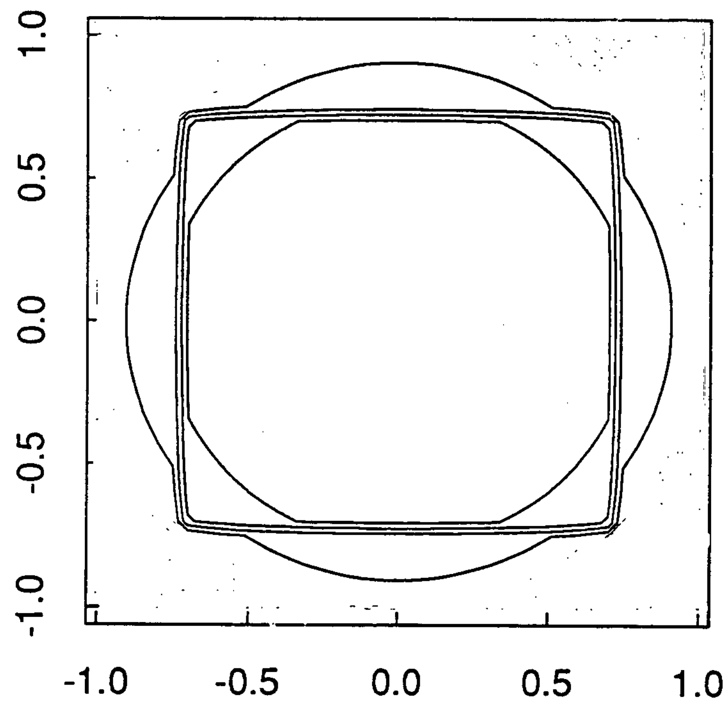

Figure 5.10 (b)

By introducing the above definition, they prove the following theorem which says that star unimodality is a special case of α-unimodality introduced by Olshen and Savage (1970).

5.12. Theorem (Dharmadhikari and Joag-dev, 1988). *A probability distribution* F *on* R^n *is star unimodal about* 0 *if and only if* F *is* n-*unimodal (Definition 3.13)*.

5.13. Remark. Dharmadhikari and Joag-dev (1988, page 75) obtain a representation for star unimodal distributions, but it is not mixture of uniform distributions (on star-shaped sets), i.e. it is not of form (5) or (6). The representation (6) is generally not true in this case. The following example shows that there are densities on R which are not unimodal, but when regarded as singular distributions on R^2, they are star unimodal.

5.14. Example. In R^2 let $K = \{(x, y): x \in [-1, 1], y = 0\}$ so that K spans the subspace $L = \{(x, y): y = 0\}$. Consider a probability distribution F on R^2 which has a density $f(x) = |x|$ on K with respect to Vol_L. Although it has a "V-shaped" density, F is 2-unimodal, i.e. star unimodal, because F is the weak limit of λ_{K_n}, where $K_n = K' \cup -K'$, and $K' = conv(\{(0, 0), (1, 1/n), (1, -1/n)\})$. But it is clear that it is impossible to represent F as the form of (5), i.e. a mixture of uniform distributions of star-shaped sets in R^2.

Although the representation of form (5) for star unimodal distributions is not true in general as Example 5.14 shows, the following theorem shows that under the assumption of absolute continunity, as in the convex situation, there is still a representation of form (5) for star unimodal distributions:

5.15. Theorem. *An absolutely continuous probability distribution* F *on* R^n *is star unimodal if and only if it can be represented as (5), where the* K_θ *'s are compact star-shaped (about* 0*) subsets of* R^n *with* $Vol_n(K_\theta) > 0$, *where* $0 \leq \theta < \theta' \leq 1$ *implies* $K_\theta \subset K_{\theta'}$, *and moreover* λ_{K_θ} *is continuous in* θ *(under weak topology).*

Proof. If a star unimodal distribution F on R^n possesses a density function f, by Theorem 3.13 and Theorem 5.12, it follows that the sets $\{f \geq u\}$, $u > 0$, are star-shaped about 0. The theorem follows by going through an argument similar to the proof of Theorem 5.6. Q. E. D.

Some authors have suggested that when a probability distribution on R^n has a density with respect to Vol_n the condition that F be star unimodal is the most general general notion of unimodality. This is not true since this notion does not include the situations that contour sets of densities are connected (e.g. see Figure 5.16) instead of star-shaped. That motivates us to give the following definition.

5.16. Definition. Let F be a probability distribution on R^n which is absolutely continuous with respect to Vol_n. Then F is called *C-unimodal* if there exists a version

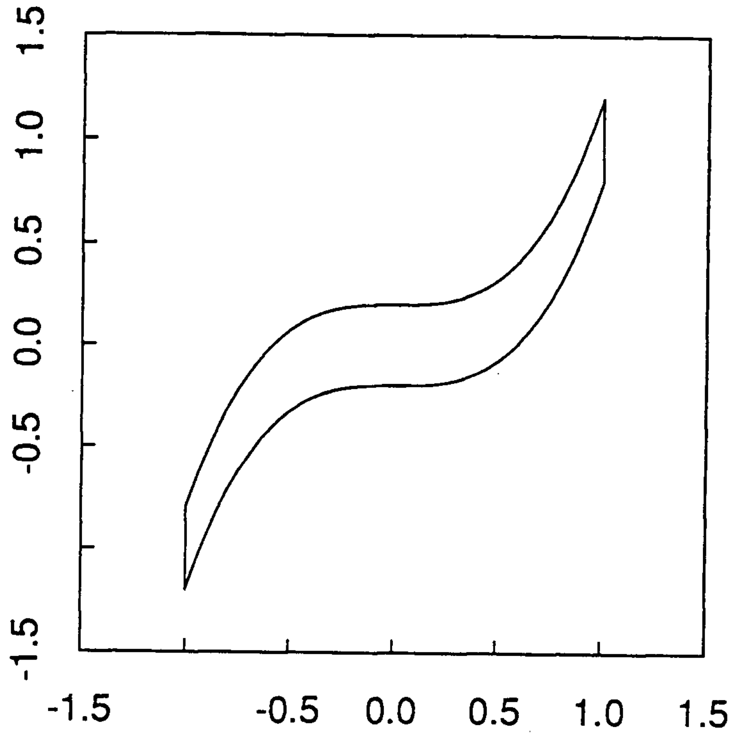

Figure 5.16

f of its density such that the sets $\{f \geq u\}$, $u > 0$, have connected interiors and are bounded whenever they are non-empty.

The following example shows that the limit distribution(under weak convergence) of a sequence of (symmetric) C-unimodal distributions on R^n need not be C-unimodal even when the limit distribution is absolutely continuous with respect to Vol_n.

5.16. Example. In R^2, let $K_1 = conv\{(1, 1), (2, 1), (1, -1), (2, -1)\}$, $K_2 = -K_1$, $K'_n = conv\{(-1, 1/n), (1, 1/n), (-1,-1/n), (1, -1/n)\}$,and let $K_n = K_1 \cup K_2 \cup K'_n$. Then $\lambda_{K_n} \rightarrow \lambda_{K_1 \cup K_2}$. Clearly, $K_1 \cup K_2$ is not connected.

6. Strong Star Unimodality

As we pointed out in Remark 5.13, a star unimodal distribution on R^n ($n > 1$) may have a V-shaped density with respect to Lebesgue measure on a linear subspace L (with $\dim(L) < n$) of R^n. In this case, a star unimodal distribution is not unimodal in any reasonable sense, i.e. it is not concentrated at a centre. It is a main concern for a "unimodal" distribution to be concentrated at a single centre. Besides, the marginals of a star unimodal distribution need not to be star unimodal. In this section we will show that a subclass (under additional symmetric restrictions) of star unimodal distributions on R^n, which we called *strong star unimodality*, is closed under marginality, convolution as well as weak convergence.

6.1. Definition. A subset K of R^n is called *totally symmetric* (*sign invariant*) if K is symmetric about all coordinates, i.e., if $(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_n) \in K \Rightarrow (x_1,\ldots,x_{i-1},-x_i,x_{i+1},\ldots,x_n) \in K$, for all $1 \leq i \leq n$. A random variable $X = (X_1,\ldots,X_n)$ taking values in R^n is called *totally symmetric* if $(X_1,\ldots,X_{i-1},X_i,X_{i+1},\ldots,X_n)$ has the same distribution as $(X_1,\ldots,X_{i-1},-X_i,X_{i+1},\ldots,X_n)$ for all $1 \leq i \leq n$. Such a distribution is also called *totally symmetric* .

6. 2. Definition. A Borel subset K of R^n is called *strictly star-shaped* about 0 if there exits a class of symmetric convex subsets $\{K_t, t \in T\}$ of R^n, where T is a

Borel subset of R, such that $K = \cup K_t$ and $K_t, t \in T$, all span the same affine subspace.

6. 3. Definition. A Borel subset K of R^n is called *strongly star-shaped* about 0 if there exits a class of totally symmetric convex subsets $\{K_t, t \in T\}$ of R^n, where T is a Borel subset of R, such that $K = \cup K_t$ and $K_t, t \in T$, all span the same affine subspace.

It is clear that if K is a Borel subset of R^n then K is strongly star-shaped $\Rightarrow$ K is strictly star-shaped . But the converse is not true. The following example shows that a totally symmetric star-shaped set need not be strongly star-shaped.

6.4. Example. In R^2, let

$$K_1 = \{(x, y): (x+y)^2/8 + (x-y)^2/2 \le 1\},$$

$$K_2 = \{(x, y): (x+y)^2/2 + (x-y)^2/8 \le 1\}$$

and $K = K_1 \cup K_2$. Clearly K is strictly star-shaped, but not strongly star-shaped.

Notation 5. Let Ω_n be the set of all uniform distributions on strictly star-shaped subsets of R^n and let Θ_n be the set of all uniform distributions on strongly star-shaped subsets of R^n.

6. 5. Definition. A probability distribution F on R^n is called a *strictly star unimodal* distribution if F can be represented as

$$F = \int_{\Omega_n} \lambda_K dv(\lambda_K) \qquad (7)$$

where ν is a probability measure on Ω_n.

6. 6. Definition. A probability distribution F on R^n is called a *strongly star unimodal* distribution if F can be represented as

$$F = \int_{\Theta_n} \lambda_K d\nu(\lambda_K) \qquad (8)$$

where ν is a probability measure on Θ_n.

6. 7. Remark. The class of all strongly (strictly) star unimodal distributions is closed under mixture.

6. 8. Remark. A probability distribution F on R^n is totally symmetric and central convex unimodal $\Rightarrow$ F is strongly star unimodal $\Rightarrow$ F is totally symmetric and star unimodal.

The following example shows that strong star unimodality does not imply monotone unimodality.

6. 9. Example. In R^2, let

$K_1 = \text{conv}(\{(3, 1), (3, -1), (-3, 1), (-3, -1)\},$

$K_2 = \text{conv}(\{(1, 3), (1, -3), (-1, -3), (-1, 3)\}).$

Suppose that λ is the uniform distribution on $K_1 \cup K_2$. Let

$$C = \mathrm{conv}(\{(-1,\ 1-\varepsilon),\ (-1+\varepsilon,\ 1),\ (1,\ -1+\varepsilon),\ (1-\varepsilon,\ -1)\}).$$

When $\varepsilon > 0$ is very small, it is clear that $\lambda\ (C + k\ (1,\ 1))$ is increasing in $|k|$ for small $|k|$. See Figure 6.9.

6. 10. Theorem. *Marginals of strongly star unimodal distributions are strongly star unimodal.*

Proof. It suffices to show that for the uniform distribution λ_K on any K, where K is defined as in Definition 6.3 satisfying that $\mathrm{Vol}(K_t) > 0$, for $t \in T$, the (n-1)-dimensional marginals, are strongly star unimodal on R^{n-1}. For $K \subset R^n$, define

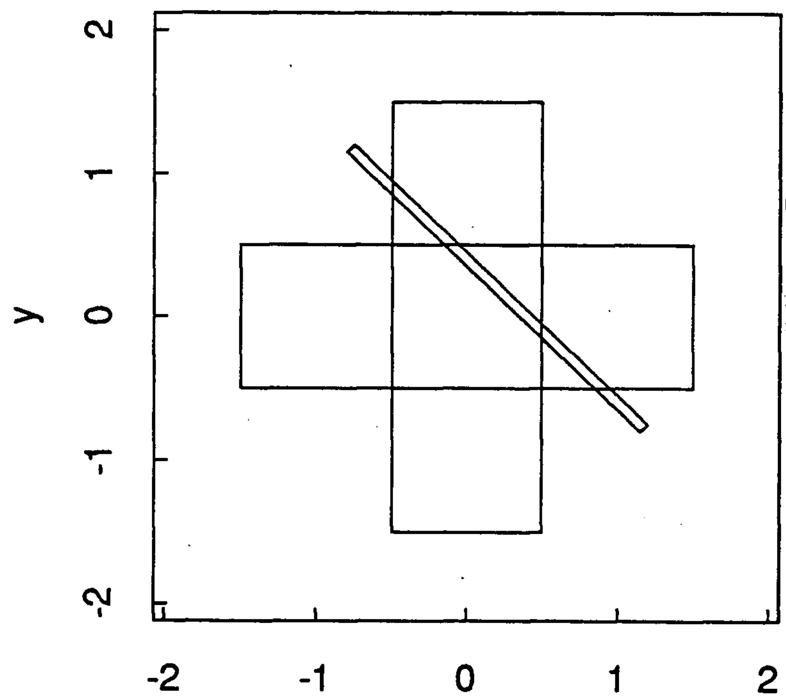

Figure 6.9

$$K(x_n^0) = \{(x_1,\dots,x_{n-1}) \mid (x_1,\dots,x_{n-1},x_n^0) \in K\}.$$

It is clear that if K is totally symmetric, compact and convex, then $K(x_n^0)$ is also totally symmetric, compact and convex for all x_n^0. So if $K = \cup K_t$, where K_t, $t \in T$, are compact, totally symmetric and convex bodies of R^n (namely K is strongly star-shaped), $K(x_n^0) = \cup K_t(x_n^0)$ is, by definition, strongly star-shaped bodies of R^{n-1}. The density of λ_K is given by

$$f(x) = I_K(x)/\mathrm{Vol}(K), \qquad x = (x_1, \dots, x_n) \in R^n.$$

The (n-1)-dimensional marginal density is

$$\begin{aligned} f_{n-1}(x_1,\dots,x_{n-1}) &= \int f(x)\, dx_n \\ &= \frac{1}{\mathrm{Vol}(K)} \int I_K(x)\, dx_n . \end{aligned}$$

For all $A \in R^{n-1}$, the marginal distribution F_{n-1} is

$$\begin{aligned} F_{n-1}(A) &= \int_A f_{n-1}(x_1,\dots,x_{n-1})\, dx_1 \dots dx_{n-1} \\ &= \frac{1}{\mathrm{Vol}(K)} \int_A dx_1 \dots dx_{n-1} \int I_K(x)\, dx_n . \end{aligned}$$

By Fubini's Theorem, it follows that

$$
\begin{aligned}
F_{n-1}(A) &= \frac{1}{\mathrm{Vol}(K)} \int dx_n \int_A I_K(x)\, dx_1 \ldots dx_{n-1} \\
&= \frac{1}{\mathrm{Vol}(K)} \int \mathrm{Vol}_{n-1}(A \cap K(x_n))\, dx_1 \ldots dx_{n-1} \\
&= \int \lambda_{K(x_n)}(A)\, dv(x_n) ,
\end{aligned}
$$

where $dv(x_n) = \mathrm{Vol}_{n-1}(K(x_n))\, dx_n / \mathrm{Vol}(K)$ is a probability density. By Definition 6.6, it follows that F_{n-1} is a strongly star unimodal on R^{n-1}. Q. E. D.

6.11. Definition. Let L be a linear subspace of R^n, and let $L^\perp$ be its orthogonal complement. Then each $x \in R^n$ can be written in the form $x = v + w$, where $v \in L$ and $w \in L^\perp$. A Borel subset K of R^n is said to be L-*symmetric* if $x = v + w \in K$ implies $x' = v - w$ and $x'' = w - v$ are also in K. A star-shaped subset K of R^n is said to be L-*symmetric* if there exits a class of convex subsets $\{K_t, t \in T\}$ of R^n, where T is a Borel subset of R, such that $K = \cup K_t$ and $K_t, t \in T$, are L-symmetric and all $K_t, t \in T$, span the same affine subspace.

6.12. Definition. Let L be a linear subspace of R^n. A probability distribution F on R^n is called a L-*symmetric star unimodal* distribution if F can be represented as

$$F = \int_{\Lambda_n} \lambda_K dv(\lambda_K) \qquad (9)$$

where v is a probability measure on Λ_n, and Λ_n is the set of all uniform distributions on L-symmetric, star-shaped subsets of R^n.

The following more generalized theorem can be similarly proven by using the similar argument in the proof of Theorem 10.

6.13. Theorem. Let F be a L - symmetric star unimodal distribution on R^n. The projection distribution of F into L is strictly star unimodal.

We have a strong feeling that the following conjecture is true.

6.14. Conjecture. The class of all strongly star unimodal distributions on R^n is closed under convolution and weak convergence.

The following example shows that the class of all strictly star unimodal distributions on R^2 is not closed under weak convergence. It is easy to generalize this to n > 2.

6.15. Example. In R^2, let

K_{1n} = conv({(1, 3/2n), (1, 1/2n), (-1, 1/2n), (-1, -3/2n)},

K_{2n} = conv({(-1, 3/2n), (-1, 1/2n), (1, 1/2n), (1, -3/2n)}).

and $K_n = K_{1n} \cup K_{2n}$. Clearly is K_n totally symmetric and star-shaped. Let $\lambda_{K_n} \to F$, then F has a "V-shaped" density on the subspace L = {(x, y): y = 0}.

7. Summaries of Definitions and Closure Properties

We have used Kanter's representation form to unify all kinds of definition of multivariate unimodality in the previous sections. Now we give the following table to summarize those representations.

Table 1. Kanter's Representation for Various Definitions

Definition	sets K
Anderson(3.1)	symmetric, convex, nested and spanning R^n
G-Anderson(5.1)	convex and nested
Symmetric G-Anderson(n>2)	symmetric, convex and nested
Kanter(3.5)	symmetric and convex
Symmetric Monotone(3.9)	no presentation
Symmetric Linear UM(3.11)	no presentation
Absolutely Cont. & Star UM	star-shaped and spanning R^n
Strict star UM	unions of symmetric, convex sets
Strong star UM	unions of totally symmetric, convex sets
L-symmetric Star UM	unions of L-symmetric, convex sets
C-UM	connected and bounded

The inter-relationships of those definitions for symmetric case are given as follows:

Anderson UM $\Rightarrow$ G-Anderson UM $\Rightarrow$ Kanter UM

$\Downarrow$ $\Downarrow$

?

Stict Star UM $\not\Rightarrow$ Monotone UM

$\Downarrow$ $\Downarrow$

LUM $\not\Rightarrow$ Star UM(O&S)

$\Downarrow$

C-UM

For the class of all distributions that are unimodal according to a given definition, it is of interest to determine closure properties under such operations as convolution, mixture, marginality, product measures and weak convergence. We have discussed some closure properties in the previous sections. In this section we will examine more of these properties for the various notions of unimodality give a table to summarize those closure properties.

It is known that Anderson's definition (Definition 3. 1.) for unimodality does not meet any of these requirements whereas Kanter's definition (namely, CCUM) meets all of them.

We have seen in Section 3 the counterexample for convolution (Example 3.3) and it is easy to see that the class of Anderson's unimodal distributions is not closed under mixture and weak convergence. Now we give the following counterexamples for product measures and marginality.

7. 1. Example (Das Gupta, 1976a; Kanter,1977). Let

$$f(x, y) = 1 / [\pi^2(1 + x^2)(1 + y^2)], \qquad x, y \in R.$$

The marginals of f are Anderson unimodal, but f is Kanter unimodal (Theorem 3.5), not Anderson unimodal , since contour sets $\{f \geq u\}$ are starshaped for small $u > 0$. See Figure 3.15.

We know that the marginal distribution of a bivariate Anderson unimodal distribution is symmetric, absolutely continuous and unimodal in the univariate sense. So it is also Anderson unimodal by the remark following Definition 3.1. The following example shows for dimension $n > 2$, the maginal distributions with dimension > 1 of an Anderson unimodal distribution are not necessary Anderson unimodal. We give an example for $n = 3$, and it is easy to generalize the example to any higher dimension.

7. 2. Example. In R^3, let

$$A = \{(x, y, z): x^2 + y^2 + z^2 \leq 8\},\ B = \{(x, y, z): |x| \leq 1, |y| \leq 1, |z| \leq 1\}.$$

Let g be the uniform density on A and h be the uniform density on B. Define $f = (g + h)/2$. Then f is Anderson unimodal on R^3. Define

$$f_{12}(x, y) = \int_R f(x, y, z)\, dz .$$

It follows that f_{12} is not Anderson unimodal by using an argument similar to that of Example 5.9.

7.3. Remark. Another counterexample about the marginal closure property of Anderson unimodality is given by Das Gupta (1976b). But our example is much simpler and easy to generalize to higher dimensions. Moreover our example is totally symmetric, which means the marginal distributions of a totally symmetric Anderson unimodal distribution need not to be Anderson unimodal.

The following theorem is due to Dharmadhikari and Jogdeo .

7. 4. Theorem (Dharmadhikari and Jogdeo, 1976). *If* F_1 *and* F_2 *are monotone unimodal and Kanter (CCUM) unimodal distributions respectively in* R^n, *then the convolution* $F_1 * F_2$ *is monotone unimodal.*

7. 5. Remark. It is still not known whether the convolution of two monotone unimodal distributions is monotone unimodal.

7. 6. Theorem (Olshen and Savage, 1970). *The convolution of an a-unimodal distribution with an a'-unimodal distribution in* R^n *is* $(a + a')$*-unimodal. No lower index of unimodality can be asserted, even for a new origin.*

7. 7. Theorem. *The set of all a-unimodal distributions in* R^n *is closed under mixture.*

Proof. Let F_θ be a-unimodal distributions in R^n for all $\theta \in R$ and $G(\theta)$ be an any probability measure on R. For all real, bounded, nonnegative Borel functions g in R^n, by the definition of a-unimodality ,

$$t^{n-\alpha}\int g(tx)\, dF_\theta(x)$$

is decreasing in t for all θ. Let

$$F(x) = \int F_\theta(x)\, dG(\theta).$$

By Neveu(1965, page 77), it follows that

$$t^{n-\alpha}\int g(tx)\, dF(x) = t^{n-\alpha}\int g(tx) d[\int F_\theta(x)\, dG(\theta)] = t^{n-\alpha}\int [\int g(tx)\, d F_\theta(x)]\, dG(\theta)$$

is decreasing in t. Therefore the mixture F is also a-unimodal. Q. E. D.

We conclude with Table 2, which summarizes the closure properties for various definitions of unimodality.

Table 2 Closure Properties

Definition	convolution	mixture	marginality	product measure	weak convergence	monotonity (Anderson Thm)
Anderson(3.1)	no{3.3}	no	no{7.2}	no{7.1}	no	yes
G-Anderson(5.1)	no	no	no	no	no{5.9}	no
Symmetric G-Anderson(n=2)	no	no	yes	no	yes{5.7}	yes
Symmetric G-Anderson(n>2)	no	no	no	no	no{5.10}	yes
Kanter(3.5)	yes[14]	yes[14]	yes[14]	yes[14]	yes[14]	yes[14]
Symmetric Monotone(3.9)	?	yes[9]	yes[9]	yes[9]	yes[9]	yes
Symmetric Linear UM(3.11)	yes	yes	yes	yes	yes	no
Star UM(5.11)	no	yes	no	yes	yes	no
Strict star UM	?	no	yes	no	no{6.15}	?
Strong star UM	?	no	yes	no	?	?
C-UM	no	yes	no	?	no{5.16}	no

Numbers in braces { } refer to examples in this paper; numbers in brackets [ ] refer to references in the bibliography.

BIBLIOGRAPHY

[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] Ash, R. B. (1972). *Real Analysis and Probability*. Wiley, New York.

[3] Becker, R. A. and Chambers, J. M. (1984). *S: An Interactive Environment for Data Analysis and Graphics*. Bell Tel. Lab., Inc., New York.

[4] Chandra, S. (1977). On the mixtures of probability distributions. *Scand. J. Statist.,* **4**, 105-112.

[5] Das Gupta, S. (1976a). A generalization of Anderson's theorem on unimodal functions. *Proc. Amer. Math. Soc.*, **60**, 85-91.

[6] Das Gupta, S. (1976b). S-unimodal functions: related inequalities and statistical applications. *Sankhyā*, B, **38**, 301-314.

[7] Das Gupta, S. (1980). Brunn-Minkowski inequality and its aftermath. *J. Multiv. Anal.*, **10**, 296-318.

[8] Dharmadhikari, S and Joag-dev, K. (1988). *Unimodality, Convexity, and Applications*. Academic Press, New York.

[9] Dharmadhikari, S and Jogdeo, K. (1976). Multivariate Unimodality. *Ann. Statist.*, **4**, 607-613.

[10] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. **2**, Second Edition. Wiley, New York.

[11] Ghosh, P. (1974). On generalized unimodality. *Comm. Statist.*, **3**, 567-580.

[12] Gnedenko, BV. and Kolmogorov, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Addision-Wesley, Cambrige, Massachusetts.

[13] Ibragimov, I. A. (1956). On the composition of unimodal distributions. *Theor. Probability Appl.*, **1**, 255-266.

[14] Kanter, M. (1977). Unimodality and dominance for symmetric random vectors. *Proc. Amer. Math. Soc.*, **229**, 65-85.

[15] Khintchine, A. Y. (1938). On unimodal distributions. *Izv. Nauchno-Isled. Inst. Mat. Mech. Tomsk. Gos. Univ.*, **2**, 1-7.

[16] Holmes, R. B. (1975). *Geometric Functional Analysis and its Applications*. New York: Springer-Verlag.

[17] Marshall, A. W. and Olkin, I. (1979). *Inequalities: The Theory of Majorization with Applications to Combinatorics, Probability, Statistics, and Matrix*. Academic Press, New York.

[18] Neveu, J. (1965). *Mathematical Foundations of the Calculus of Probability*. Holden-Day, San Francisco.

[19] Olshen, R. A. and Savage, L. J. (1970). A generalized unimodality. *J. Appl. Probability*, **7**, 21-34.

[20] Parthasarathy, K. R. (1967). *Probability Measures on Metric Spaces.* Academic Press, New York.

[21] Sherman, S. (1955). A theorem on convex sets with applications. *Ann. Math. Statist.*, **26**, 763-766.

[22] Uhrin, B. (1984). Some remarks about the convolutions of unimodal functions. *Ann. Probability,* **12**, 640-645.

[23] Wells, D. R. (1978). A result in multivariate unimodality. *Ann. Statist.*, **6**, 926-931.

[24] Wintner, A. (1938). *Asymptotic Distributions and Infinite Convolutions.* Edwards Brothers, Ann Arbor, Michigan.