The Annals of Statistics
1991, Vol. 19, No. 4, 2120–2144

MIXTURES WITH A LIMITED NUMBER OF MODAL INTERVALS[1]

By J. H. B. Kemperman

Rutgers University

We derive necessary and sufficient conditions in order that each mixture of a given family of probability densities have no more than s modal intervals, with special attention to ordinary unimodality and strong unimodality of such mixtures.

1. Introduction. In the sequel, unless otherwise stated, each density f is a nonzero integrable function $f(x) \geq 0$, $x \in \mathbf{R}$, all whose discontinuities (if any) are of the first kind. We further let $c(x) = \max(f(x-0), f(x+0))$. What really matters is the associated finite and nonzero measure $d\mu = f\,dx$, thus, one may assume that for each x, either

$$f(x-0) \leq f(x) \leq f(x+0) \quad \text{or} \quad f(x+0) \leq f(x) \leq f(x-0).$$

Let further $F(x) = \mu((-\infty, x])$ be the associated distribution function (d.f.).

What exactly is meant by a mode x_0 of a density f? One cannot define a mode as any local maximum x_0 of f. For instance, when $f(x) = 1$ for $0 \leq x \leq 1$; $f(x) = 2$ for $1 < x \leq 2$; $f(x) = 0$, otherwise, one would not consider *each* point $0 < x_0 < 1$ to be a mode of f.

We will certainly regard $x_0 \in \mathbf{R}$ to be a mode of f if $f(x) \leq c = c(x_0)$ holds throughout some neighborhood of x_0 and, moreover, $f(x) < c$ for at least one point in every left and every right neighborhood of x_0. In this situation, $J(x_0) = [x_0, x_0] = \{x_0\}$ will be considered to be a degenerate modal interval of f.

A nondegenerate compact interval $[a, b]$, $a < b$ is defined to be a *modal interval* of f, if, first, $f(x) = c > 0$ on (a, b); second, $f(x) \leq c$ in an entire open neighborhood of $[a, b]$; and third, $f(x) < c$ for at least one point in every left neighborhood of a and at least one point in every right neighborhood of b. There always is at least one modal interval (Lemma 1) and distinct modal intervals must be disjoint. We define a *mode* of f to be any point x_0 which belongs to some (necessarily unique) modal interval $[a, b]$; in that case we also write $J(x_0) = [a, b]$. See Definition 2.3 for further details.

Let s be a fixed positive integer and let F_0 be a given family of densities f on $\mathbf{R}$. In this paper, we will be especially concerned with conditions that are necessary and sufficient in order that an arbitrary superposition of densities $f \in F_0$ have at most s modal intervals. Besides its greater generality, a further compelling reason to also admit densities f having discontinuities, is the fact

Received September 1985; revised November 1990.

[1]Research was supported in part by NSF Grant DMS-84-01608.

AMS 1980 *subject classifications*. Primary 62E10; secondary 26A48.

Key words and phrases. Superposition, mixture, mode, modal interval, unimodal, strongly unimodal, piecewise monotone functions.

that a mixture of continuous densities can easily be discontinuous; see the example following (4.2).

By osc(s), we will denote the class of all densities f such that its (secant) slope has at most $s - 1$ changes of sign from strictly negative to strictly positive (when moving to the right); see Definition 3.1 for a more precise statement. If $f \in C^1$, then $f \in \operatorname{osc}(s)$ is equivalent to its slope having no more than s changes of sign from strictly positive to strictly negative. By Osc(s) we denote the class of all continuous d.f.'s F such that the second difference $\Delta_h^2 F(x)$ has at most $s - 1$ changes of sign from strictly negative to strictly positive (when moving to the right); see Definition 3.6 for a precise statement.

The condition $f \in \operatorname{osc}(s)$ is obviously sufficient for the density f to have at most s modal intervals. It turns out (Theorems 1 and 3) that the latter property is in fact equivalent to $f \in \operatorname{osc}(s)$ and also to $F \in \operatorname{Osc}(s)$. Moreover (Theorem 2), each $F \in \operatorname{Osc}(s)$ is the integral of an essentially unique $f \in \operatorname{osc}(s)$. We also describe the precise structure of a density f which has exactly s modal intervals.

Let F_0 be any given class of densities. Then (Theorem 6) in order that *each* mixture g of densities $f \in F_0$ have at most s modal intervals, it is (necessary and) sufficient that this be true for each special mixture $g = \Sigma p_i f_i$, $p_i > 0$, which involves at most $2s$ densities $f_i \in F_0$. Here, the integer $2s$ cannot be replaced by a smaller one. A dual type of condition, which is necessary and sufficient for each F_0-mixture to be in osc(s), is given in Theorem 7.

The remaining results all concern ordinary unimodality or strong unimodality. A density f is said to be unimodal if it has precisely one modal interval, that is, $f \in \operatorname{osc}(1)$. In order that each mixture of F_0 be unimodal, it is necessary and sufficient (Theorem 4) that this be true for each special mixture $p_1 f_1 + p_2 f_2$. Applicable necessary and sufficient conditions for the latter property are given by Theorem 8 and Proposition 1. Under strong side assumptions, related necessary and sufficient conditions are due to Kakiuchi (1981). Analogous results hold for discrete density functions.

A density f is said to be strongly unimodal if $J = \{x\colon f(x) > 0\}$ is an interval and, moreover, $\log f(x)$ is concave on J. In order that each mixture of a family F_0 be *strongly unimodal*, it is (necessary and) sufficient that each special mixture $p_1 f_1 + p_2 f_2$ be strongly unimodal (where $f_i \in F_0$ and $p_i > 0$); see Theorem 5. A useful necessary and sufficient condition for the latter is given in Theorem 9.

Sections 2–6 largely consist of definitions and results, together with a running discussion and many applications. Most of the proofs have been collected in Section 7.

2. Modal intervals. In the sequel, F denotes the d.f. $F(x) = \mu((-\infty, x])$ of a finite nonzero measure μ on $\mathbf{R}$. Each point x_0 with $\mu(\{x_0\}) > 0$ should be regarded as a mode. Replacing μ by its continuous component, we may and will assume from now on that F is continuous. Thus, by a d.f. F, we will mean any nondecreasing continuous function on $\mathbf{R}$ such that $F(-\infty) = 0$ and $0 < F(\infty) < \infty$. Usually, we take F of the form $dF = f\,dx$ with f as a density. In that case

$$F(x) = \int_{-\infty}^{x} f(u)\,du. \tag{2.1}$$

Here, by a density we mean any nonnegative measurable function on $\mathbf{R}$ such that $0 < \int f(x)\,dx < \infty$. In addition, unless otherwise stated, we also assume that the discontinuities of f (if any) are of the first kind. That is, for all $x \in \mathbf{R}$, both left-hand limit $f(x-0)$ and right-hand limit $f(x+0)$ exist; these are equal for all but countably many x, the so-called jump points of f. The actual values $f(x)$ at these jump points are irrelevant as to F. Since they also should be irrelevant as to the modal properties of f, we may and will assume that

$$\min(f(x-0), f(x+0)) \le f(x) \le \max(f(x-0), f(x+0)) \tag{2.2}$$

for all $x \in \mathbf{R}$. Below, a density f *as usual* will be one having the above properties. It suffices that f be a continuous density.

The above assumptions are satisfied when the density f happens to be *piecewise monotone*. The latter means that $\mathbf{R}$ is the union of finitely many intervals J_r such that, for each r, the restriction of f to $\mathrm{int}(J_r)$ is either nondecreasing or nonincreasing. Besides upward or downward jumps of f inside these intervals, we also allow f to have an upward or downward jump at the junction x_0 of a pair of adjoining intervals J_r [the value $f(x_0)$ itself must satisfy (2.2) but is otherwise irrelevant]. For instance, f might be increasing on (a, x_0) and also increasing on (x_0, b), but not on the full interval (a, b), due to a downward jump of f at x_0.

If f is piecewise monotone, then the associated d.f. F is *piecewise convex-concave*. Here, a continuous d.f. F will be said to be piecewise convex-concave if $\mathbf{R}$ can be partitioned into finitely many intervals J_r such that F is either convex on J_r or concave on J_r, this for each r. Consequently, the (nonnegative) left- and right-hand slopes $F'(x-0)$ and $F'(x+0)$ exist everywhere, and they are equal for all but at most countably many x. These slope functions are monotone on $\mathrm{int}(J_r)$, and may have jumps there. In addition, they may have an upward or downward jump at the junction of any pair of adjoining intervals J_r.

Such a continuous piecewise convex-concave d.f. F is absolutely continuous implying that, conversely, F is simply of the form (2.1) with f as a piecewise monotone function; thus, $F'(x-0) = f(x-0)$ and $F'(x+0) = f(x+0)$. For, suppose a d.f. F is convex on the interval $[a, b]$, thus, its slope is nondecreasing there. Letting $0 < \varepsilon < b - a$ and $K(\varepsilon) = (F(b) - F(b-\varepsilon))/\varepsilon \ge 0$, one has

$$|F(x) - F(y)| \le K(\varepsilon)|x - y| \quad \text{for all } x, y \in [a, b-\varepsilon],$$

showing that F is absolutely continuous on $[a, b]$; similarly, when F is concave.

(2.3) Definition. The precise definition of a mode and a modal interval is not entirely trivial. Let f be a density as above, having only discontinuities of the first kind. Let $x_0 \in \mathbf{R}$ and define

$$c = c(x_0) = \max(f(x_0 - 0), f(x_0 + 0)), \tag{2.4}$$

thus, $c = f(x_0)$ if f is continuous at x_0. We say that x_0 is a *mode* of f if $c = c(x_0) > 0$ and, moreover, the associated quantities

$$\begin{aligned} a(x_0) &= \sup\{x < x_0 \colon f(x) < c\}, & A(x_0) &= \sup\{x < x_0 \colon f(x) > c\}, \\ b(x_0) &= \inf\{x > x_0 \colon f(x) < c\}, & B(x_0) &= \inf\{x > x_0 \colon f(x) > c\} \end{aligned} \tag{2.5}$$

are such that

$$\begin{gathered} A(x_0) < a(x_0) \quad \text{and} \quad b(x_0) < B(x_0), \quad \text{hence,} \\ A(x_0) < a(x_0) \le x_0 \le b(x_0) < B(x_0). \end{gathered} \tag{2.6}$$

In this case, the corresponding finite closed interval

$$J(x_0) = [a(x_0), b(x_0)] \tag{2.7}$$

will be regarded as the *modal interval* associated to the mode x_0. Distinct modal intervals are easily seen to be disjoint. In many applications, $a(x_0) = b(x_0) = x_0$ in which case $J(x_0) = \{x_0\}$ is degenerate.

(2.8) REMARK 1. Let us describe the situation in some more detail. Consider a mode x_0 of f as above, thus, $c = c(x_0) > 0$. Put $J = J(x_0) = [a, b] = [a(x_0), b(x_0)]$. Then $c(x) = c$ for each $x \in J$, thus, $f(x) = c > 0$ for each $x \in \operatorname{int} J = (a, b)$. Moreover, $f(x) \le c$ throughout an entire (open) neighborhood of J [with $(A(x_0), B(x_0))$ as the largest such neighborhood]. Finally, inside *each* (open) left neighborhood of a and *each* (open) right neighborhood of b, there are points with $f(x) < c$. These properties characterize the notion of $J = [a, b]$ being a modal interval. Note that each $x \in J = [a, b]$ is also a mode of f having $c(x) = c$ and $J(x) = J$.

(2.9) REMARK 2. Suppose $J = [a, b]$ is a modal interval of f with f as a *piecewise monotone* density, in particular, f is monotone (either nonincreasing or nondecreasing) in a sufficiently small left neighborhood of a. Therefore, either $f(a - 0) < f(a + 0) = c$ or else f is nondecreasing and nonconstant (that is, F is convex but not linear) on some left neighborhood U of a, in such a way that $f(x) < f(a - 0) = c$, for all $x \in U$. Similarly, either $f(b + 0) < f(b - 0)$ or else the density f will be nonincreasing and nonconstant (thus F is concave but not linear) on some right neighborhood V of b. In both cases, $f(x) < c$ on a sufficiently small right neighborhood V of b. Possibly $c = \infty$ but then $a = b$.

However, if $f(a - 0) < f(a + 0) = c$ (the slope of F has an upward jump at a), then it is possible for F to be concave (and not convex) in a left neighborhood of a. Similarly, if $f(b + 0) < f(b - 0)$, then it is possible for F to be convex (and not concave) in a right neighborhood of b.

Definition 2.3 still makes sense when $f \geq 0$ is not integrable and has an open interval K of $\mathbf{R}$ as its domain, still assuming that f has only discontinuities of the first kind. In this context, the notion of a modal interval is invariant under monotone transformations. More precisely, if ϕ: $K \to \mathbf{R}$ is strictly increasing and continuous, then the subinterval $[a, b]$ of K is a modal interval for a function $f \geq 0$ on K if and only if $[\phi(a), \phi(b)]$ is a modal interval for the associated function $g(x) = f(\phi^{-1}(x))$ on $\phi(K)$.

LEMMA 1. *Let f be a density as usual (having only discontinuities of the first kind). Let further $I = (\alpha, \beta)$ be a finite or infinite open interval such that*

$$f(x) > f(\alpha + 0) \quad \text{and} \quad f(x) > f(\beta - 0) \quad \text{for some } x \in I. \tag{2.10}$$

Then I contains at least one modal interval.

PROOF. Let $c = \sup_{x \in I} f(x)$ thus $f(\alpha + 0) < c$ and $f(\beta - 0) < c$. Let $x_n \in I$ be such that $f(x_n) \to c$ and $\lim_n x_n = x_0$. Clearly, $\alpha < x_0 < \beta$.

Necessarily, $c = c(x_0)$, with $c(x_0)$ as in (2.4). The quantities (2.5) are easily seen to satisfy $A(x_0) \leq \alpha < a(x_0) \leq x_0 \leq b(x_0) < \beta \leq B(x_0)$. Hence, $J = [a(x_0), b(x_0)]$ is a modal interval contained in $I = (\alpha, \beta)$. □

REMARK. Lemma 1 also holds for a closed interval $I = [\alpha, \beta]$, provided in (2.10) we replace $f(\alpha + 0)$ by $f(\alpha - 0)$ and $f(\beta - 0)$ by $f(\beta + 0)$. Similarly for half open intervals $(\alpha, \beta]$ and $[\alpha, \beta)$.

COROLLARY 1. *Each density f (as usual) has at least one modal interval.*

PROOF. Let x' be fixed such that $f(x') > 0$. Since $f \geq 0$ is integrable, there exists $\alpha < x'$ with $f(\alpha + 0) < f(x')$ and also $\beta > x'$ with $f(\beta - 0) < f(x')$. Now apply Lemma 1. □

COROLLARY 2. *Let f be as usual and suppose that*

$$\Delta_{h_1} f(x_1) > 0 \quad \text{and} \quad \Delta_{h_2} f(x_2) < 0, \quad \text{where } x_1 < x_1 + h_1 < x_2 < x_2 + h_2. \tag{2.11}$$

Then the interval $l = [x_1, x_2 + h_2]$ contains at least one modal interval of f. If f is right-continuous at x_1 and left-continuous at $x_2 + h_2$, then even the open interval $(x_1, x_2 + h_2)$ contains a modal interval of f.

REMARK. Most results of the present paper can also be applied to the restrictions of d.f.'s F and densities f to a fixed open interval $K = (p, q)$ (possibly $p = -\infty$ or $q = +\infty$). In this connection, it is useful to consider the density $\tilde{f}$ defined by $\tilde{f}(x) = f(x)$ for $x \in K$ and $\tilde{f}(x) = 0$, otherwise, and its associated d.f. $\tilde{F}$. Clearly, f and $\tilde{f}$ have the same modal intervals J insofar as J is entirely contained in $K = (p, q)$. If $J = [a, b]$ is a modal interval for f with $a \leq p < b \leq q$, then $[p, b]$ is a modal interval for $\tilde{f}$ (it is essential that

$p < b$). The converse is false. For instance, a modal interval of $\tilde{f}$ of the form $[p, b]$ need not derive from a modal interval of f.

3. The classes osc(c) and Osc(s). Let s be a fixed positive integer. We will be interested in densities f having at most s modal intervals (necessarily disjoint). Such a function f does not admit more than s disjoint intervals as described in Lemma 1 or Corollary 2. The class of densities having the latter property is denoted by osc(s). A more precise description is as follows.

(3.1) DEFINITION. We will say that $f \in \text{osc}(s)$ if f is a density as usual (having only discontinuities of the first kind), such that it is *impossible* to find numbers x_r and h_r, $r = 1, \dots, 2s$, satisfying

$$(3.2)\quad x_1 < x_2 < \cdots < x_{2s}, \qquad 0 < h_r < x_{r+1} - x_r, \qquad r = 1, \dots, 2s;$$

$x_{2s+1} = +\infty$, and

$$(3.3)\qquad (-1)^r \Delta_{h_r} f(x_r) > 0 \quad \text{for } r = 1, \dots, 2s.$$

An equivalent definition obtains when one insists that, in addition, all the $4s$ points x_r and $x_r + h_r$ be continuity points of f and/or that $h_r < \varepsilon$ for all r, with $\varepsilon > 0$ preassigned. Essentially, we are requiring that the slope of f changes its sign at most $s - 1$ times from strictly negative to strictly positive (when moving to the right).

LEMMA 2. *Suppose that f is a continuous density which is differentiable at each point x outside a discrete subset D of* $\mathbf{R}$. *Then $f \in \text{osc}(f)$ if and only if it is* impossible *to find a set of $2s$ points $x_r \notin D$ such that*

$$(3.4)\quad x_1 < x_2 < \cdots < x_{2s} \quad \textit{and} \quad (-1)^r f'(x_r) > 0, \qquad r = 1, \dots, 2s.$$

PROOF. The necessity follows from the definition of a derivative and the sufficiency from the mean value theorem. □

LEMMA 3. *A density f belongs to* osc(s) *if and only if it is impossible to find numbers z_r, $r = 0, \dots, 2s$, such that*

$$(3.5)\quad z_{r-1} < z_r, \qquad (-1)^r \big(f(z_r) - f(z_{r-1})\big) > 0 \quad \textit{for } r = 1, \dots, 2s.$$

PROOF. See Section 7.

REMARK. In many applications, there exists a finite interval (c, d) such that f is nondecreasing on $(-\infty, c]$ and nonincreasing on $[d, \infty)$. Since (3.4) requires that $x_1 < x_{2s}$ and $f'(x_1) < 0$, $f'(x_{2s}) > 0$, one may assume that $c < x_1 < \cdots < x_{2s} < d$. An analogous remark applies to (3.2) and (3.5).

THEOREM 1. *In order that a density f have at most s different modal intervals, it is necessary and sufficient that $f \in \text{osc}(s)$.*

PROOF. See Section 7. The necessity of $f \in \operatorname{osc}(s)$ follows from Lemma 1.

(3.6) DEFINITION. Let s be a positive integer. We say that $F \in \operatorname{Osc}(s)$ if F is a continuous d.f. such that it is impossible to find numbers x_r and h_r, $r = 1, \dots, 2s$, satisfying

$$(3.7)\quad x_1 < x_2 < \cdots < x_{2s}, \qquad 0 < h_r < (x_{r+1} - x_r)/2, \qquad r = 1, \dots, 2s,$$

$x_{2s+1} = \infty$, and

$$(3.8)\qquad (-1)^r \Delta^2_{h_r} F(x_r) > 0 \quad \text{for } r = 1, \dots, 2s.$$

Here, $\Delta^2_h F(x) = F(x + 2h) - 2F(x + h) + F(x)$.

THEOREM 2. *Each $f \in \operatorname{osc}(s)$ is piecewise monotone and each $F \in \operatorname{Osc}(s)$ is piecewise convex-concave.*

THEOREM 3. *If $f \in \operatorname{osc}(s)$, then the associated d.f. as in* (2.1) *belongs to* $\operatorname{Osc}(s)$. *Conversely, each $F \in \operatorname{Osc}(s)$ is the integral of a density $f \in \operatorname{osc}(s)$.*

PROOF. See Section 7.

What exactly is the *structure* of a density f having precisely s modal intervals? First consider the case $s = 1$. From Theorem 1, since every density has at least one modal interval (Corollary 1 of Lemma 1), this case is equivalent to $f \in \operatorname{osc}(1)$. Which requires precisely that a strict decrease of f is never followed by a strict increase of f. In other words, there exists $x_0 \in \mathbf{R}$ such that f is nondecreasing for $x < x_0$ and nonincreasing for $x > x_0$. This coincides with the usual definition of a unimodal density.

When $s \geq 2$, we have from Theorem 1 that a density f has precisely s modal intervals if and only if both $f \in \operatorname{osc}(s)$ and $f \notin \operatorname{osc}(s - 1)$. In particular (Theorem 2) such a density is piecewise monotone.

The case where f has exactly s modal intervals can be described as follows. There exists a partition of $\mathbf{R}$ into $2s$ disjoint nonempty (open, closed or half open; finite or infinite) intervals $J_1, \dots, J_{2s}$ with J_{k+1} immediately to the right of J_k. Moreover, on each *odd* interval J_{2r-1} (*even* interval J_{2r}) the density f is *nondecreasing* (*nonincreasing*) and nonconstant.

For $1 \leq k < 2s$, let x_k be the right end point of J_k (it belongs to either J_k or J_{k+1}). The precise values $f(x_k)$ are irrelevant, except that we insist on condition (2.2). It will be allowed that J_k is degenerate, but only when f has a jump at x_k. In that case, $J_k = \{x_k\}$ while $x_{k-1} = x_k$ does not belong to either J_{k-1} or J_{k+1}. More precisely, if $J_k = \{x_k\}$ with k odd, we require that f has a nonzero upward jump at x_k (note that f is nonincreasing on each of the adjoining even intervals J_{k-1} and J_{k+1}). In this special case, by convention, we will regard f as being nonconstant on J_k; similarly, when k is even.

If J_k is nondegenerate and $x_k \in J_k$ and further f has a nonzero jump at x_k, then we insist that this be an upward (downward) jump when k is odd (even). In such a case, by convention, we always regard f to be nonconstant on J_k [it would be for most choices of the value $f(x_k)$]. Similarly when $x_{k-1} \in J_k$.

The rigorous proof, that the above structure is necessary and sufficient, for f to have exactly s modal intervals, is rather long and will be omitted. The partition on hand is not always unique, due to the presence of continuity endpoints x_k and/or intervals of constancy, which can be included in either an even or an odd interval.

EXAMPLE. Let f be defined by

$$f(x) = \begin{cases} (x - [x])/s, & \text{if } 0 \le x \le s, \\ 0, & \text{otherwise}. \end{cases} \tag{3.9}$$

Then f has precisely s modal intervals, namely, the one-point sets $\{1\}, \{2\}, \dots, \{s\}$, thus, there are also precisely s modes. Note that f has a downward jump at each of the s modes but, otherwise, is nondecreasing everywhere. The latter structure intervals $J_1, \dots, J_{2s}$ can presently be chosen as $J_1 = (-\infty, 1)$; $J_2 = \{1\}$; $J_3 = (1, 2)$; $J_4 = \{2\}$; $\dots$; $J_{2s-2} = \{s - 1\}$; $J_{2s-1} = (s - 1, s)$; $J_{2s} = [s, +\infty)$.

4. Mixtures. In the sequel, F denotes a nonempty family of continuous nonzero d.f.'s, that is, each $G \in F$ is a nondecreasing continuous function on $\mathbf{R}$ such that $G(-\infty) = 0$ and $0 < G(+\infty) < \infty$.

A d.f. G is said to be a mixture of F (or also an F-mixture) if it is of the form

$$G(x) = \int F(t, x)\rho(dt). \tag{4.1}$$

Here, t runs through a measurable space T while ρ is a σ-finite nonzero measure on T. Further, $F(t, x)$ is any function which is measurable in t and which, as a function of x, belongs to F. We also require that $G(\infty) < \infty$. From the dominated convergence theorem, G is always continuous.

By a density f, we mean any measurable function $f \ge 0$ on $\mathbf{R}$ with $0 < \int f\,dx < \infty$. In most applications below, each $G \in F$ is absolutely continuous and then the corresponding (nonempty) family of densities will be denoted as F_0. In this case, if G is a mixture of F as in (4.1), then (Fubini) G is absolutely continuous with a density of the form

$$g(x) = \int f(t, x)\rho(dt). \tag{4.2}$$

Here, $f(t, x) \ge 0$ is jointly measurable while $f(t, \cdot) \in F_0$ is a density for $F(t, \cdot)$. Thus, the mixture G of F has a mixture g of F_0 as its density.

Below, we always assume that F is a subset of $\mathrm{Osc}(s)$ for some $s \ge 1$. But then we know (Theorem 2) that each $G \in F$ is piecewise convex-concave and further that $dG = g\,dx$ with $g \in \mathrm{osc}(s)$. In this situation, we may and will assume that F_0 is a subset of osc(s). In particular, all $f \in F_0$ are then piecewise monotone and thus densities as usual.

EXAMPLE. Even if each $f \in F_0$ were continuous, it does not follow that the mixture g in (4.2) will be continuous. For instance,

$$f(t, x) = \frac{1}{t}\left[1 - \left|1 - \frac{x}{t}\right|\right]_+, \qquad x \in \mathbf{R}; t > 0,$$

is a continuous probability density with support $[0, 2t]$. Nevertheless, the mixture

$$g(x) = \sum_{t=1}^{\infty} 2^{-t} f(2^{-t}, x) = 2(1 - x), \quad \text{if } 0 < x \leq 1,$$

$[g(x) = 0$, otherwise$]$, is discontinuous at 0.

THEOREM 4. *In order that each F-mixture G be unimodal, it is (necessary and) sufficient that each special F-mixture*

$$G(x) = p_1 F_1(x) + p_2 F_2(x) \tag{4.3}$$

be unimodal. Here, $F_j \in F$, $p_j \geq 0$ $(j = 1, 2)$, $p_1 + p_2 > 0$.

APPLICATION. Behboodian (1970) showed that each mixture of two normal (m_1, σ_1) and (m_2, σ_2) densities is unimodal when $|m_1 - m_2| \leq 2\min(\sigma_1, \sigma_2)$. Let $\Omega = [-1, +1] \times [1, \infty)$ thus $|u_1 - u_2| \leq 2 \leq 2\min(v_1, v_2)$ if $(u_r, v_r) \in \Omega$ $(r = 1, 2)$. Combining Behboodian's result with Theorem 4, we conclude that

$$g(x) = \int \exp\left(-(x - u)^2/(2v^2)\right)\rho(du, dv)$$

is a unimodal density for each choice of the σ-finite measure ρ on Ω, at least when g is finite on $\mathbf{R}$. More generally, the same proof shows that each such mixture g is unimodal on any open interval on which it is finite.

Theorem 4 follows from the special case $s = 1$ of Theorem 6 below. A d.f. G is said to be unimodal if it is continuous and has only one modal interval. Equivalently, $G \in$ Osc(1). That is, G is continuous and $x_0 \in \mathbf{R}$ exists such that G is convex for $x \leq x_0$ and concave for $x \geq x_0$. Equivalently, G has a density g which is unimodal in the usual sense.

A d.f. F will be said to be *strongly unimodal* if the convolution $F * H$ is unimodal for each choice of the continuous unimodal d.f. H. A result due to Ibragimov (1956) states that F has this property if and only if it is absolutely continuous with a log-concave density f. That is, $f(x) = \exp(-u(x))$ with $u\colon \mathbf{R}(-\infty, +\infty]$ as a convex function. The support $\{x\colon f(x) > 0\} = \{x\colon u(x) < \infty\}$ is always an interval.

THEOREM 5. *In order that each F-mixture G be strongly unimodal it is (necessary and) sufficient that each special F-mixture as in* (4.3) *be strongly unimodal.*

PROOF. It is given that each special F-mixture (4.3) is strongly unimodal. We must prove that any given mixture G of F is strongly unimodal. Let H be any fixed continuous unimodal d.f. It suffices to show that $G * H$ is unimodal.

Since $G * H$ is clearly (Fubini) a mixture of the family $F_H = \{F * H: F \in F\}$, it suffices (Theorem 4) to verify that any special F_H-mixture is unimodal. Indeed, if $p_i \geq 0$ and $F_i \in F$, $i = 1, 2$; $p_1 + p_2 > 0$, then

$$p_1(F_1 * H) + p_2(F_1 * H) = (p_1 F_1 + p_2 F_2) * H$$

is unimodal, simply because $p_1 F_1 + p_2 F_2$ is strongly unimodal. □

REMARK 1. It is useful to reformulate Theorem 5 in terms of log-concave densities. Namely, if a family F_0 of densities is such that each special F_0-mixture $p_1 f_1 + p_2 f_2$ ($p_j > 0$; $f_j \in F_0$) is log-concave, then each F_0-mixture is log-concave.

Formulated this way, the result even carries over to log-concave densities on $\mathbf{R}^n$ ($n \geq 1$). For, let g be a mixture of a given family F_0 of densities on $\mathbf{R}^n$. In showing that g is log-concave, it suffices to show that, for each choice of $x_0, y_0 \in \mathbf{R}^n$, the associated function $h(s) = g(x_0 + sy_0) \geq 0$ is log-concave in $s \in \mathbf{R}$. From (4.2), h itself is a mixture of the family $F_0(x_0, y_0) = \{f(x_0 + sy_0): f \in F_0\}$ of densities on $\mathbf{R}$. In view of Theorem 5, reformulated as above, it suffices that each special mixture of $F_0(x_0, y_0)$ be log-concave on $\mathbf{R}$. And the latter property, for all $x_0, y_0 \in \mathbf{R}^n$, is equivalent to the condition that each special mixture of F_0 be log-concave on $\mathbf{R}^n$.

REMARK 2. Theorems 4 and 5 have obvious analogues for discrete densities $\{p_n\}_{n \in \mathbf{Z}}$, ($\mathbf{Z}$ = integers). Here, $\{p_n\}$ is said to be unimodal if, for some n_0, $p_{n+1} \geq p_n$ when $n < n_0$ and $p_{n+1} \leq p_n$ when $n > n_0$. Equivalently, there exist no $m, n \in \mathbf{Z}$ with $m < n$, $p_{m+1} < p_m$ and $p_{n+1} > p_n$, and strong unimodality property is equivalent to having $p_{n-1} p_{n+1} \leq (p_n)^2$ for all $n \in \mathbf{Z}$, a result due to Keilson and Gerber (1971).

THEOREM 6. *Let $s \geq 1$. In order that each mixture G of a family F of continuous d.f.'s have at most s modal intervals, that is,* $G \in \mathrm{Osc}(s)$, *it is (necessary and) sufficient that this be true for each special mixture*

$$G(x) = p_1 F_1(x) + \cdots + p_{2s} F_{2s}(x) \tag{4.4}$$

involving at most $2s$ members $F_i \in F$, ($p_j \geq 0$; $\sum_i p_j > 0$). The analogous result is true for a mixture g of a family F_0 of densities relative to the property $g \in \mathrm{osc}(s)$.

PROOF. See Section 7.

Theorem 4 follows from the case $s = 1$ of Theorem 6. For a good understanding of Theorem 6, the reader should also consult Theorems 1, 2 and 3. For instance, the last assertion of Theorem 6 immediately follows from the first because of Theorem 3. Theorem 6 is only interesting when each $G \in F$

has at most s modal intervals, that is, $G \in \mathrm{Osc}(s)$. Equivalently, each $G \in F$ is absolutely continuous, which admits a corresponding density $g \in \mathrm{osc}(s)$, in particular and is piecewise monotone.

SHARPNESS. The integer $2s$ in (4.4) cannot be replaced by any smaller one. To show this, it suffices to exhibit a family F_0 consisting of $2s$ continuous densities f_j, $j = 1, \dots, 2s$, such that

$$(4.5)\qquad g(x) = p_1 f_1(x) + \cdots + p_{2s} f_{2s}(x), \qquad p_j \geq 0, \textstyle\sum_j p_j > 0,$$

has at most s modal intervals as soon as $p_j = 0$ for some $1 \leq j \leq 2s$ while, on the other hand, the density (4.5) has more than s modal intervals when $p_j > 0$ for all j.

For $k \geq 1$, let $\phi_k(x)$ denote the continuous trapezoidal density defined by $\phi_k(x) = 1$ for $1 \leq x \leq 2k - 1$; furthermore $\phi_k(x) = x_+$ for $x \leq 1$ and $\phi_k(x) = (2k - x)_+$ if $x \geq 2k - 1$ [where $z_+ = \max(0, z)$]. In particular, $\phi_k(x) = 0$ when $x \notin (0, 2k)$. For $r = 1, \dots, s$, let $f_{2r-1}(x) = \phi_r(x)$ and let $f_{2r}(x) = \phi_{s+1-r}(2s + 2 - x)$ [having support $(0, 2r)$ and $(2r, 2s + 2)$, respectively].

The resulting density g as in (4.5) is continuous and piecewise linear. As is easily seen, g has slope

$$(4.6)\qquad g'(x) = (-1)^j p_j \quad \text{on the interval } [j, j + 1], \qquad j = 1, \dots, 2s.$$

In addition, g is linear with slope $p_1 + p_3 + \cdots + p_{2s-1}$ on $[0, 1]$ and linear with slope $-(p_2 + p_4 + \cdots + p_{2s})$ on $[2s + 1, 2s + 2]$. Finally, $g(x) = 0$ if $x \notin [0, 2s + 2]$.

If $p_j = 0$ for some j, then $g \in \mathrm{osc}(s)$ in view of Lemma 2. On the other hand, suppose that $p_j > 0$ for all j. Then $g \notin \mathrm{osc}(s)$, again from Lemma 2. In fact, (4.6) clearly shows that the slope $g'(x)$ does have s changes of sign from strictly negative to strictly positive. And no more such changes, thus, $g \in \mathrm{osc}(s + 1)$. Presently, g has the $s + 1$ (degenerate) modal intervals $\{2j - 1\}$, $j = 1, \dots, s + 1$. They are separated by the 'valleys' $\{2j\}$, $j = 1, 2, \dots, s$.

(4.7) CONDITION. Let F_0 be a given subset of $\mathrm{osc}(s)$, $(s \geq 1)$. The present condition requires that, for each choice of the numbers x_r and h_r, $r = 1, \dots, 2s$, such that $x_r < x_r + h_r < x_{r+1}$ $(x_{2s+1} = \infty)$, there exist *nonnegative* numbers a_r, $r = 1, \dots, 2s$, not all zero, such that

$$(4.8)\qquad \sum_{r=1}^{2s} (-1)^r a_r \Delta_{h_r} f(x_r) \leq 0 \quad \text{for all } f \in F_0.$$

THEOREM 7. *Let F_0 be a (nonempty) subset of* $\mathrm{osc}(s)$. *Then Condition* 4.7 *is necessary and sufficient in order that each mixture of F_0 [as in* (4.2)] *belong to* $\mathrm{osc}(s)$.

THEOREM 7′. *Let F_0 be a subset of* $\mathrm{osc}(s)$. *Then in order that each mixture of F_0 belong to* $\mathrm{osc}(s)$, *it is necessary and sufficient that, for each choice of the* *numbers* $z_0 < z_1 < \cdots < z_{2s}$, *there exist nonnegative numbers* a_r, $r = 1, \dots, 2s$, *not all zero such that*

$$\sum_{r=1}^{2s} (-1)^r a_r [f(z_r) - f(z_{r-1})] \leq 0 \quad \text{for all } f \in F_0. \tag{4.9}$$

THEOREM 7″. *Let* F_0 *be a given subset of* osc(s) *and suppose that each* $f \in F_0$ *is everywhere continuous and, moreover, has a derivative* $f'(x)$ *at each* $x \notin D$. *Here,* D *is a discrete subset of* $\mathbf{R}$ *which is independent of* f.

Then in order that each mixture of F_0 *belong to* osc(s), *it is necessary and sufficient that, for each choice of the* $x_r \notin D$, $r = 1, \dots, 2s$, *with* $x_r < x_{r+1}$, *there exist nonnegative numbers* a_r, $r = 1, \dots, 2s$, *not all zero such that*

$$\sum_{r=1}^{2s} (-1)^r a_r f'(x_r) \leq 0 \quad \text{for all } f \in F_0. \tag{4.10}$$

PROOF. See Section 7.

Note that (4.8) implies that the analogous inequality holds for each mixture g of F_0. Hence, in view of Definition 3.1 of osc(s), it is obvious that Condition 4.7 is sufficient in order that each F_0-mixture belongs to osc(s). The necessity of Condition 4.7 is less trivial. Similar remarks hold for Theorems 7′ and 7″, using Lemma 3 or Lemma 2, respectively.

5. Unimodality. Let f be any unimodal density, that is, $f \in \text{osc}(1)$. Its unique modal interval will be denoted as $[m(f), M(f)]$, $-\infty < m(f) \leq M(f) < \infty$. Thus, f is nondecreasing for $x < M = M(f)$ and nonincreasing for $x > m = m(f)$, therefore, $f(x)$ is constant for $m < x < M$. Further $f(m+0) = f(M-0) = c > 0$ and $f(x) = c$ for $m < x < M$. Finally, $f(x) < c$ for all $x < m$ and all $x > M$.

Let F_0 be a family of unimodal densities. We will be interested in sufficient conditions in order that each mixture g of F_0 be unimodal. From Theorem 4, this is true if and only if each special mixture

$$g(x) = p_1 f_1(x) + p_2 f_2(x) \tag{5.1}$$

is unimodal. Here, $f_i \in F_0$ and $p_i \geq 0$; $p_1 + p_2 > 0$.

Let f_1, f_2 be fixed unimodal densities. We want necessary and/or sufficient conditions in order that each mixture g as in (5.1) be unimodal. One may as well assume that $p_i > 0$, $i = 1, 2$. Let $[m(f_i), M(f_i)]$ be the unique modal interval of f_i. Let further

$$\alpha = \min(M(f_1), M(f_2)), \qquad \beta = \max(m(f_1), m(f_2)). \tag{5.2}$$

Then f_1, f_2 and thus g are nondecreasing for $x < \alpha$ and nonincreasing for $x > \beta$. Thus g is always unimodal when $\beta \leq \alpha$ and it only remains to consider the case that $\alpha < \beta$. Our major concern is then with the behavior of g on the interval $[\alpha, \beta]$. Interchanging f_1, f_2 if necessary, one may suppose that

$$M(f_1) \le M(f_2), \tag{5.3}$$

thus $\alpha = M(f_1)$. Further $m(f_1) \le M(f_1) = \alpha < \beta$ thus $\beta = m(f_2)$, yielding that

$$M(f_1) < m(f_2), \tag{5.4}$$

which in turn implies (5.3).

We may assume that $f_i(\alpha) = \max(f_i(\alpha - 0), f_i(\alpha + 0))$ and similarly for $f_i(\beta)$. One has $f_1(x) < f_1(\alpha)$ for all $x > \alpha$ and $f_2(x) < f_2(\beta)$ for all $x < \beta$. In addition, f_1 is *nonincreasing* for $x \ge \alpha$ (hence on $[\alpha, \beta]$) and f_2 is *nondecreasing* for $x \le \beta$ (hence, on $[\alpha, \beta]$).

For $p_i > 0$ fixed, the mixture g is not unimodal [that is, $g \notin \operatorname{osc}(1)$], if and only if some strict decrease of g is followed on the right by some strict increase of g, in the sense of either (3.3) or (3.5). More precisely, $g \notin \operatorname{osc}(1)$ is equivalent to the existence of numbers x_r such that

$$x_1 < x_2 \le x_3 < x_4 \quad \text{and} \quad g(x_2) < g(x_1);\ g(x_4) > g(x_3). \tag{5.5}$$

From the above properties of g, if the situation (5.5) occurs at all, then (5.5) can also be realized with the x_r such that $\alpha \le x_1 < x_4 \le \beta$. If desired one may insist on having $x_2 = x_3$, see (3.5).

The situation (5.5) definitely occurs when $f_1(\alpha + 0) < f_1(\alpha - 0) = f_1(\alpha)$ while f_2 is continuous at α. For, then $g(x)$ *always* has a downward jump at α while, for suitable points $x_3 < x_4 \le \beta$ close to β, one will have $g(x_3) < g(x_4)$, as soon as p_1/p_2 is sufficiently small. Similarly, when $f_2(\beta - 0) < f_2(\beta + 0) = f_2(\beta)$ and f_1 is continuous at β. This happens, for instance, when $f_1(x) = e^{-x}$ for $x > \mathbf{0}$, $f_1(x) = 0$ for $x < 0$ and $f_2(x) = f_1(x - \beta)$, $\beta > 0$.

THEOREM 8. *Let f_1, f_2 be unimodal densities such that $M(f_1) \le M(f_2)$. Then in order that each mixture g as in* (5.1) *be unimodal, it is necessary and sufficient that*

$$\begin{vmatrix} f_1(z_0) & f_1(z_1) & f_1(z_2) \\ f_2(z_0) & f_2(z_1) & f_2(z_2) \\ 1 & 1 & 1 \end{vmatrix} \ge 0, \tag{5.6}$$

whenever $M(f_1) \le z_0 < z_1 < z_2 \le m(f_2)$.

LEMMA 4. *Let u_i and v_i, $i = 1, 2$, be given nonnegative numbers. Then in order that there exist numbers $p_i \ge 0$, $i = 1, 2$, such that*

$$p_1 u_1 - p_2 v_1 > 0 > p_1 u_2 - p_2 v_2, \tag{5.7}$$

it is necessary and sufficient that $u_1 v_2 - u_2 v_1 > 0$.

PROOF (Necessity). Clearly, (5.7) implies that $p_1 > 0$, $p_2 > 0$ and $u_1 > 0$, $v_2 > 0$. Thus the assertion is obvious unless $u_2 > 0$, $v_1 > 0$. But then

$$v_1/u_1 < p_1/p_2 < v_2/u_2. \tag{5.8}$$

(Sufficiency.) Suppose $u_1 v_2 - u_2 v_1 > 0$ thus $u_1 > 0$, $v_2 > 0$. If $u_2 = 0$, then choose $p_2 > 0$ and $p_1 > p_2 v_1 / u_1$. Otherwise, choose $p_1, p_2 > 0$ so as to satisfy (5.8). □

PROOF OF THEOREM 8. Use criterion (5.5) with $x_2 = x_3$ and $\alpha \le x_1 < x_4 \le \beta$, where $\alpha = M(f_1)$ and $\beta = m(f_2)$. Thus, in order that some mixture g as in (5.1) *not* be unimodal it is necessary and sufficient that numbers $\alpha \le z_0 < z_1 < z_2 \le \beta$ exist such that for some choice of $p_1 \ge 0$, $p_2 \ge 0$ one has

$$g(z_0) > g(z_1) \quad \text{and} \quad g(z_2) > g(z_1), \quad \text{where } g(x) = p_1 f_1(x) + p_2 f_2(x).$$

For fixed values z_j, this is equivalent to condition (5.7) with $u_j = f_1(z_{j-1}) - f_1(z_j)$ and $v_j = f_2(z_j) - f_2(z_{j-1})$, $j = 1, 2$. One has $u_j \ge 0$ and $v_j \ge 0$, since f_1 is nonincreasing on $[\alpha, \beta]$ and f_2 is nondecreasing on $[\alpha, \beta]$. By Lemma 4, the system (5.7) admits a solution $p_1 \ge 0$, $p_2 \ge 0$ if and only if $u_1 v_2 - u_2 v_1 > 0$. The latter inequality is easily seen to be equivalent to the negation of (6). □

REMARK. It also follows from the proof that, in the case $M(f_1) \le M(f_2)$, the set of numbers $\rho > 0$ such that $g = \rho f_1 + f_2$ is *not* unimodal is an open set equal to the union of all open intervals of the form

$$\frac{f_2(z_1) - f_2(z_0)}{f_1(z_0) - f_1(z_1)} < \rho < \frac{f_2(z_2) - f_2(z_1)}{f_1(z_1) - f_1(z_2)}, \tag{5.9}$$

one for each choice of the numbers $M(f_1) \le z_0 < z_1 < z_2 \le m(f_2)$ such that $f_1(z_0) > f_1(z_1)$ and $f_2(z_1) < f_2(z_2)$ [none if $m(f_2) \le M(f_1) \le M(f_2)$]. If $f_1(z_1) = f_1(z_2)$, then the right-hand side of (5.9) is to be interpreted as $+\infty$. Condition (5.6) requires precisely that the interval (5.9) be always empty.

THEOREM 8′. *Let f_1, f_2 be unimodal densities such that $M(f_1) \le M(f_2)$. Suppose further that f_1 and f_2 are continuous for $M(f_1) \le x \le m(f_2)$, and differentiable for $M(f_1) < x < m(f_2)$. Then in order that each mixture g as in* (5.1) *be unimodal, it is necessary and sufficient that*

$$\begin{vmatrix} f_1'(x) & f_1'(y) \\ f_2'(x) & f_2'(y) \end{vmatrix} \ge 0, \quad \textit{whenever } M(f_1) < x < y < m(f_2). \tag{5.10}$$

PROOF. Exactly as the proof of Theorem 8, but this time using the criterion (3.4) instead of (3.5) (with $s = 1$). Or else use the generalized mean value theorem in showing that (5.6) and (5.10) are equivalent. □

REMARK 1. Provided f_1'' and f_2'' exist, condition (5.10) is equivalent to

$$\begin{vmatrix} f_1'(x) & f_1''(x) \\ f_2'(x) & f_2''(x) \end{vmatrix} \ge 0, \quad \text{whenever } M(f_1) < x < m(f_2).$$

REMARK 2. The following is a related result for arbitrary $s \geq 1$. Namely, let $f_j \in \operatorname{osc}(s)$ be differentiable and put $u_j(x) = f_j'(x)$, $j = 1, \dots, 2s$. Then in order that a density $g = p_1 f_1 + \cdots + p_{2s} f_{2s}$, $p_i \geq 0$, $\Sigma_i p_i > 0$, have no more than s modal intervals, it is clearly sufficient that the derivative $g' = p_1 u_1 + \cdots + p_{2s} u_{2s}$ have no more than $2s - 1$ distinct zeros (if there were $s + 1$ modal intervals there would be also s valleys; we also used that a derivative g' always has the mean value property). Equivalently, if for each choice of $x_1 < \cdots < x_{2s}$, the system $\Sigma_j p_j u_j(x_r) = 0$, $r = 1, 2s$, has no non-trivial nonnegative solution. It would be sufficient that

$$\det(u_i(x_r); i, r = 1, \dots, 2s) \neq 0 \quad \text{when } x_1 < \cdots < x_{2s}.$$

In the latter case, $\{u_1, \dots, u_{2s}\}$ would be a type of Tchebycheff system, see Karlin and Studden (1966). If moreover $u_j = f_j'$ is continuous, then the latter determinant would be of one sign, such as always strictly positive. The special case $s = 1$ of this sufficient condition coincides with the *strict* version of (5.10).

PROPOSITION 1. *Here, we make the same assumptions as in Theorem 8′. Let E be the set of points $M_1(f) < x < m(f_2)$, such that either $f_1'(x) \neq 0$ or $f_2'(x) \neq 0$. Define*

$$\phi(x) = \left| f_1'(x)/f_2'(x) \right| \quad \text{for each } x \in E. \tag{5.11}$$

Here, $\phi(x) = \infty$ if $f_1'(x) \neq 0$ and $f_2'(x) = 0$.

Then in order that each mixture g as in (5.1) *be unimodal, it is necessary and sufficient that ϕ: $E \to [0, \infty]$ be nondecreasing.*

In particular, the following conditions are necessary.

(i) *If $x \in E$; $f_2'(x) = 0$, then $f_2'(y) = 0$ for all $y > x$; $y \in E$.*
(ii) *If $y \in E$; $f_1'(y) = 0$, then $f_1'(x) = 0$ for all $x < y$; $x \in E$.*

PROOF. Note that $f_1' \leq 0$ and $f_2' \geq 0$ when $M_1(f) < x < m(f_2)$. Since (5.10) is trivially true if either $x \notin E$ or $y \notin E$, one only needs to verify that, for each fixed pair $x, y \in E$, condition (5.10) is equivalent to $\phi(x) \leq \phi(y)$. Both are trivially true when $f_1'(x) = 0$ [thus $\phi(x) = 0$] and also when $f_2'(y) = 0$ [thus $\phi(y) = +\infty$]. Thus one may assume that $f_1'(x) < 0$ and $f_2'(y) > 0$, thus $\phi(x) > 0$ and $\phi(y) < \infty$. If $f_2'(x) = 0$, thus $\phi(x) = \infty$, then both (5.10) and $\phi(x) \leq \phi(y)$ are false. If also $f_2'(x) > 0$, the equivalence follows by dividing (5.10) by $f_2'(x) f_2'(y)$. The last assertion amounts to saying that if $x, y \in E$; $x < y$, then $\phi(x) = \infty$ must imply $\phi(y) = \infty$ and $\phi(y) = 0$ must imply $\phi(x) = 0$. □

REMARK. The discrete analogue of Proposition 1 is as follows. Let $\{p_n\}$, $\{q_n\}$, $n \in \mathbf{Z}$ be discrete unimodal densities. Thus $\max_j p_j$ is assumed for $m(p) \leq n \leq M(p)$, while p_n is nondecreasing for $n \leq m(p)$ and nonincreasing for $n \geq M(p)$. Similarly for $\{q_n\}$. One may assume that $M(p) \leq M(q)$. Since

$\{p_n + \theta q_n\}$, $\theta > 0$ is obviously unimodal if $M(p) \geq m(q) - 1$, let us assume that $M(p) \leq m(q) - 2$.

In that case, we have that $\{p_n + \theta q_n\}$ is unimodal, for *all* $\theta > 0$, if and only if $|(p_{n+1} - p_n)/(q_{n+1} - q_n)|$ is nondecreasing for $M(p) \leq n \leq m(q) - 1$. Here, we ignore the integers n with both $p_{n+1} = p_n$ and $q_{n+1} = q_n$. The above condition requires, in particular, that for $M(p) + 1 \leq n \leq m(q) - 1$, one cannot have both $p_{n+1} = p_n$ and $q_{n+1} > q_n$. And neither both $p_{n-1} > p_n$ and $q_{n-1} = q_n$.

PROPOSITION 2. *Again we make the same assumptions as in Theorem* 8′. *Let* $\theta(f_1, f_2)$ *denote the set of all numbers* $\theta > 0$ *such that* $g = f_1 + \theta f_2$ *is not unimodal. Let* E *and* ϕ *be as in Proposition* 1. *We claim that* $\theta(f_1, f_2)$ *is precisely the union of all open intervals* $(\phi(y), \phi(x))$, *over all pairs* $x, y \in E$ *such that* $x < y$ *and* $\phi(y) < \phi(x)$.

PROOF. Clearly, $\theta(f_1, f_2)$ consists of all $\theta > 0$ such that $g'(x) < 0 < g'(y)$ for some choice of x, y with $M(f_1) < x < y < m(f_2)$. In particular, $x, y \in E$. Equivalently, $\theta(f_1, f_2)$ is the union of all intervals

$$\{\theta\colon f_1'(x) + \theta f_2'(x) < 0 < f_1'(y) + \theta f_2'(y)\} = \{\theta\colon \phi(y) < \theta < \phi(x)\},$$

one for each choice of $x, y \in E$ with $x < y$. □

EXAMPLE. Let $f_1(x) = \exp(-|x|)$, $f_2(x) = \exp(-|x - a|)$, $a > 0$. Then $M(f_1) = 0$ and $m(f_2) = a$. Further $\phi(x) = \exp(a - 2x)$ for $0 < x < a$, yielding that $\theta(f_1, f_2) = (e^{-a}, e^{a})$.

APPLICATIONS. Let f and g be fixed densities, each unimodal about 0. Thus $m(f) \leq 0 \leq M(f)$ and $m(g) \leq 0 \leq M(g)$. Put $M = M(f)$ and $m = -m(g)$, thus $m, M \geq 0$. We shall be interested in the *largest* value $c^* \geq 0$ such that

$$h(x) = pf(x) + qg(x - c) \tag{5.12}$$

is unimodal for *all* $p, q > 0$ and *all* $0 \leq c \leq c^*$. Since h is trivially unimodal when $0 \leq c \leq M + m$ (whatever $p, q > 0$), one only needs to study the case $c > M + m$. Thus, letting $f_1(x) = f(x)$ and $f_2(x) = g(x - c)$, one has $M(f_1) = M < c - m = m(f_2)$.

We will assume that, for $x > M$, the function f is continuous with two derivatives and such that $f'(x) < 0$. Similarly, for $x < -m$, we assume g to be continuous on $x \leq -m$, with two derivatives and such that $g'(x) > 0$. Let $u(x)$ and $v(x)$ be defined by

$$-f'(M + x) = \exp(-u(x)), \quad g'(-m - x) = \exp(-v(x)), \qquad x > 0. \tag{5.13}$$

We claim that

$$c^* = m + M + \inf\{s + t\colon s > 0; t > 0; u'(s) + v'(t) > 0\}. \tag{5.14}$$

PROOF. With f_1, f_2 as above and ϕ as in (5.11), one easily sees that

$$\phi(x) = \exp[-u(x-M) + v(c-m-x)] \quad \text{if } M < x < c-m. \tag{5.15}$$

From Proposition 1, h is unimodal, for all $p,q > 0$, if and only if, on the interval $M < x < c - m$, the function $\phi(x)$ is nondecreasing, that is, $u(x - M) - v(c - m - x)$ is nonincreasing. Letting $s = x - M$ and $t = c - m - x$, this leads to the necessary and sufficient condition that $u'(s) + v'(t) \leq 0$ when $s > 0$; $t > 0$; $u + v = c - M - m$. Therefore, the largest possible such value c is given by (5.14).

Often $m = M = 0$ and then $u'(s) = -f''(s)/f'(s)$ and $v'(t) = g''(-t)/g'(-t)$, $s,t > 0$. In the further special case that $g(-x) = f(x)$ for $x \geq 0$, one has $v'(t) = u'(t)$, $t > 0$. In that case, the infimum (5.14) is often assumed with $s = t = s^*$, hence, $c^* = 2s^*$, with s^* satisfying $u'(s^*) = 0$. That is, $f''(s^*) = 0$ thus s^* is an inflection point of the density f.

Assuming $f = g$, this happens for the Cauchy density $f(x) = 1/(1 + x^2)$, with $s^* = 3^{-1/2}$ and $c^* = 2s^* = 1.1547$. Also for the normal density $f(x) = \exp(-x^2/2)$, with $s^* = 1$ and $c^* = 2$. Similarly for $f(x) = \exp(-|x|^a)$ when $1 < a < 2$.

The situation is more complicated when $f(x) = g(x) = \exp(-x^4)$. Then the infimum (5.14) is assumed for $s = (-1 + 3^{1/2})/2$ and $t = (1 + 3^{1/2})/2$, thus, $c^* = 3^{1/2} = 1.732$. Choosing $c = 1.74$ (just slightly larger than c^*), it turns out (somewhat surprisingly) that the mixture

$$h(x) = p\exp(-x^4) + q\exp\left(-(x - 1.74)^4\right)$$

is unimodal when $q/p = 1$ but not when $q/p = 0.655$. In the latter case, one has the two modes $x_1 = 0.262$ and $x_2 = 0.500$. □

NORMAL CASE. Suppose

$$f_1(x) = c_1\exp(-x^2/2), \qquad f_2(x) = c_2\exp\left(-(x-\mu)^2/(2\sigma^2)\right)$$

are normal probability densities with $\mu > 0; \sigma > 0$. It is easily seen then that the mixture g as in (5.1) cannot have more than two modes. Robertson and Fryer (1969) already determined precisely what mixtures g are unimodal; their results are also described in Titterington, Smith and Makov (1985). They would also follow quite easily from Propositions 1 and 2, using that

$$\log\phi(x) = d + \log\frac{x}{\mu - x} - \frac{x^2}{2} + \frac{(\mu - x)^2}{2\sigma^2}, \quad \text{when } 0 < x < \mu,$$

with d as a known constant.

GAMMA CASE. Suppose $f_1 = x^a e^{-x}$ and $f_2 = x^b e^{-x}$, where $0 \leq a < b$ and put $c = b - a$ [$x > 0$; $f_i(x) = 0$, otherwise]. Then $\alpha = M(f_1) = a$ and $\beta = m(f_2) = b$. In order that each mixture (5.1) be unimodal, $\phi(x)$ must be nondecreasing for $a \leq x \leq b$. From (5.11), we find that $\phi(x) = x^{-c}(x - a)/$

$(b - x)$. Calculating $\phi'(x)$, one is led to the condition that $(a - x)(x - b) \leq x$ for $a \leq x \leq b$. It is easily checked that this is true if and only if $c \leq 1 + \sqrt{4a}$. Equivalently, if $\sqrt{b} \leq 1 + \sqrt{a}$. Thus *each* mixture of $f = x^a e^{-x}$ relative to the parameter a is unimodal (even strongly unimodal, see Section 6), provided we restrict this parameter to a fixed interval $[A, B]$, such that $0 \leq A \leq B$ and $B \leq (1 + \sqrt{A})^2$. The latter bound is best possible.

6. Strong unimodality. In this section, F_0 is a given family of strongly unimodal densities, all having the same open interval K as their common support [such as (0, 1) or $(0, \infty)$ or $(-\infty, \infty)$]. Equivalently, each $f \in F_0$ is of the form $f = e^{-u}$ with $u = u(x)$ as a convex function on $\mathbf{R}$ which is finite on K and infinite off K. Let U denote the corresponding class of convex functions u. For convenience, we will assume that each $u \in U$ has two derivatives on K, thus, $u''(x) \geq 0$.

LEMMA 5. *Let* $f_i = \exp(-u_i) \in F_0$ $(i = 1, 2)$. *Then the mixture*

$$g = p_1 f_1 + p_2 f_2 = p_1 \exp(-u_1) + p_2 \exp(-u_2) \tag{6.1}$$

is strongly unimodal for all $p_1, p_2 > 0$, *if and only if*

$$|u_1'(x) - u_2'(x)| \leq u_1''(x)^{1/2} + u_2''(x)^{1/2} \quad \text{for all } x \in K. \tag{6.2}$$

PROOF. See Section 7.

THEOREM 9. *In order that each* F_0*-mixture be strongly unimodal, it is necessary and sufficient that for each* $x \in K$,

$$\sup_{u \in U} \left[u'(x) - u''(x)^{1/2} \right] \leq \inf_{u \in U} \left[u'(x) + u''(x)^{1/2} \right]. \tag{6.3}$$

PROOF. Condition (6.3) says that (6.2) holds for all pairs $f_1, f_2 \in F_0$. Hence from Lemma 5, condition (6.3) is necessary and sufficient in order that each *special* mixture (6.1) be strongly unimodal. Hence, it is certainly necessary in order that each F_0-mixture be strongly unimodal. That (6.3) is also sufficient follows from Theorem 5, which states that each F_0-mixture is strongly unimodal as soon as each special mixture (6.1) is strongly unimodal. Since the proof of Theorem 5 is not entirely simple, also because of the use of Ibragimov's theorem, we include in Section 7 a more elementary sufficiency proof which is valid for a wide class of mixtures, including all finite mixtures. □

COROLLARY. *Let* $f = e^{-u}$ *be a strongly unimodal density of class* C^2 *on* $\mathbf{R}$. *Let* $F_0 = \{f_c(x) = f(x - c)\colon 0 \leq c \leq c^*\}$. *Then the largest value* c^* *such that each mixture of* F_0 *be strongly unimodal, is given by*

$$c^* = \inf\left\{ |x - y| : u'(x) - u'(y) > u''(x)^{1/2} + u''(y)^{1/2};\ x, y \in \mathbf{R} \right\}. \tag{6.4}$$

PROOF. Immediate from Theorem 9. Note that (6.4) is the analogue of formula (5.14) for the unimodal case. □

APPLICATIONS. Choosing $f(x) = \exp(-x^2/2)$, one has $u'(x) = x$, $u''(x) = 1$ and (6.4) yields that $c^* = 2$. Thus *each* mixture of $\exp(-(x-c)^2/2)$, with c restricted to $[0, 2]$, is strongly unimodal and this interval cannot be replaced by a larger one (even when one only requires that each mixture be simply unimodal, see Section 5).

More generally, suppose there exist positive constants A and B such that $A^2 \le u''(x) \le B^2$ for all $x \in \mathbf{R}$. Then $|u'(x) - u'(y)| \le B^2|x - y|$ and (6.4) yield that $c^* \ge c_0$, where $c_0 = 2A/B^2$. This estimate is sharp in the normal case $u(x) = x^2/2$.

As an application of (6.2) or (6.3), suppose F_0 is some class of gamma densities $f(x) = x^a e^{-rx}$, $a > -1$; $r > 0$. Provided $a \ge 0$, this density is always strongly unimodal with $K = (0, \infty)$ and $u(x) = -a \log x - rx$ [if $x > 0$; $u(x) = +\infty$, otherwise]. Choosing

$$f_1(x) = x^a e^{-rx}, \qquad f_2(x) = x^b e^{-sx}, \qquad x > 0; a \ge 0; b \ge 0,$$

one finds that condition (6.2) is *never* satisfied when $r \ne s$; that is, if $r \ne s$, then no nontrivial mixture of f_1, f_2 is strongly unimodal. On the other hand, when $r = s$, then (6.2) holds if and only if

$$\left|\sqrt{a} - \sqrt{b}\right| \le 1. \tag{6.5}$$

It follows that, for any choice of $q \ge 0$, $r > 0$, the family

$$F_0 = \left\{f = f(x) = x^a e^{-rx}, \text{ where } q^2 \le a \le (q+1)^2\right\} \tag{6.6}$$

has the property that each F_0-mixture is strongly unimodal. The a-interval on hand cannot be enlarged. In fact, it happens to coincide with the maximal interval we already found (Section 5) relative to the property of simple unimodality.

REMARK 1. Lemma 5 and Theorem 9 *carry over* to log-concave densities $f = e^{-u}$ of class C^2 on an open convex subset K of $\mathbf{R}^n$. Namely, a mixture

$$g = p_1 f_1 + p_2 f_2 = p_1 e^{-u} + p_2 e^{-v}, \qquad p_1 > 0; p_2 > 0,$$

is log-concave on K if and only if it is log-concave (that is, strongly unimodal), on each line segment in K. It follows easily from Lemma 5 that each such mixture is log-concave on K if and only if the inequality

$$\left|\sum_i (u_i - v_i)\xi_i\right| \le \left[\sum_i \sum_j u_{ij}\xi_i\xi_j\right]^{1/2} + \left[\sum_i \sum_j v_{ij}\xi_i\xi_j\right]^{1/2} \tag{6.7}$$

holds for all $x \in K$ and $(\xi_1, \dots, \xi_n) \in \mathbf{R}^n$. Here, $i, j \in 1, 2, \dots, n$ while u_i, u_{ij} denote the first and second order partial derivatives of $u(x)$; similarly v_i and v_{ij}.

As one application, suppose f_1 is the density of $X = (X_1, \dots, X_n)$ with the X_i as independent standard normal random variables. Thus $f(x) =$

$a\exp(-\|x\|^2/2)$ with $\|x\|$ as the Euclidean norm of $x \in \mathbf{R}^n$. Let $f_2(x) = f_1(x-c)$, $c \in \mathbf{R}^n$. Then (6.7) holds if and only if $\|c\| \le 2$. Let E be any subset of $\mathbf{R}^n$. It follows that *each* mixture of the associated family of normal densities $\{f_1(x-c)\colon c \in E\}$ is log-concave *if and only if* E *has diameter* ≤ 2. For example, E could be the closed ball of radius 1 about the origin. However, adding even a single point to that ball, one would obtain a set E with diameter larger than 2 and thereby destroy the latter mixing property.

REMARK 2. There also exists an analogue of Lemma 5 and Theorem 9 for integer valued random variables X. The discrete density $p_n = \Pr(X = n)$, $n \in \mathbf{Z}$ is strongly unimodal if and only if, for each $n \in \mathbf{Z}$, one has $(p_n)^2 \ge p_{n-1}p_{n+1}$. Equivalently,

$$p_{n-1}w_n^2 - 2p_n w_n + p_{n+1} \le 0 \quad \text{for some } w_n \in \mathbf{R}.$$

Moreover, if F_0 is a family of strongly unimodal densities $\{p_n\}_{n\in\mathbf{Z}}$, then each F_0-mixture is strongly unimodal if and only if w_n can be chosen so as to be independent of the particular member $\{p_n\}$ of F_0 (the sufficiency is obvious). Equivalently, if and only if

$$\left|\frac{p_n}{p_{n-1}} - \frac{q_n}{q_{n-1}}\right| \le \frac{\left[(p_n)^2 - p_{n-1}p_{n+1}\right]^{1/2}}{p_{n-1}} + \frac{\left[(q_n)^2 - q_{n-1}q_{n+1}\right]^{1/2}}{q_{n-1}}$$

for all $n \in \mathbf{Z}$ and for all pairs $\{p_n\}$, $\{q_n\}$ in F_0 as long as both $p_{n-1} > 0$ and $q_{n-1} > 0$.

7. Proofs.

PROOF OF LEMMA 3 (Necessity). Suppose first that the z_r are as in (3.5). One may assume that the z_r are continuity points of f. Introduce $x_r = z_{r-1}$ and $h_r = z_r - z_{r-1} - \varepsilon$, $r = 1, 2, \dots, 2s$. Choosing $\varepsilon > 0$ sufficiently small, one has properties (3.2) and (3.3), thus, $f \notin \operatorname{osc}(s)$.

(Sufficiency.) Suppose the x_r and h_r are as in (3.2) and (3.3). Let S denote the set of $4s$ distinct points x_r and $x_r + h_r$, $r = 1, \dots, 2s$. Consider any choice of the $w_j \in S$, $j = 1, \dots, 4s$ such that $w_1 \le w_2 \le \cdots \le w_{4s}$ and

$$(7.1) \qquad (-1)^r\big(f(w_{2r}) - f(w_{2r-1})\big) > 0 \quad \text{for } r = 1, \dots, 2s,$$

thus, $w_{2r-1} < w_{2r}$. One example would be $w_{2r-1} = x_r$ and $w_{2r} = x_r + h_r$, $r = 1, \dots, 2s$.

Each pair w_{2r}, w_{2r+1} with $w_{2r} < w_{2r+1}$, $1 \le r \le 2s-1$ will be regarded as a *gap*. Now suppose the w_j have been chosen such that, moreover, the number of gaps is as small as possible. We assert then that there are no gaps at all, that is, $w_{2r} = w_{2r+1}$ for all $r = 1, \dots, 2s-1$. If it is true, then (7.1) implies that (3.5) holds with $z_r = w_{2r}$, $r = 1, \dots, 2s-1$ and $z_0 = w_1$; $z_{2s} = w_{4s}$.

If there were such a gap, then there exists $1 \le t \le 2s-1$ such that $w_{2t} < w_{2t+1}$. If $(-1)^t(f(w_{2t+1}) - f(w_{2t})) > 0$, then property (7.1) remains valid on replacing w_{2t} by w_{2t+1} (leaving the w_j with $j \neq 2t$ unchanged). In the remaining case, property (7.1) remains valid on replacing w_{2t+1} by w_{2t} (leaving the w_j with $j \neq 2t+1$ unchanged). In each case, the number of gaps would be reduced, yielding a contradiction. □

PROOF OF THEOREM 1 (Necessity). It suffices to show that f has at least $s+1$ modal intervals when there exist numbers x_r, h_r, $r = 1, \dots, 2s$ satisfying (3.2) and (3.3). One may assume that the x_r and $x_r + h_r$ are continuity points of f.

There exist numbers x_0 and h_0 such that $x_0 < x_0 + h_0 < x_1$ and $\Delta_{h_0} f(x_0) > 0$. For, if not, then $f(x)$ would be nonincreasing for $x < x_1$. Since $f \geq 0$ is integrable, one would have $f(x) = 0$ for all $x < x_1$, hence, $f(x_1) = f(x_1 - 0) = 0$. But then we have $f(x_1 + h_1) = \Delta_{h_1} f(x_1) < 0$, a contradiction. One may assume that x_0 and $x_0 + h_0$ are continuity points of f.

Similarly, there exist numbers x_{2s+1} and $h_{2s+1} > 0$ with $x_{2s+1} > x_{2s} + h_{2s}$ and $\Delta_{h_{2s+1}} f(x_{2s+1}) < 0$. One may assume that x_{2s+1} and $x_{2s+1} + h_{2s+1}$ are continuity points of f. Using Corollary 2 of Lemma 1, we conclude that each of the $s+1$ disjoint open intervals $(x_{2r}, x_{2r+1} + h_{2r+1})$, $r = 0, 1, \dots, s$ contains at least one modal interval of f. This proves the stated assertion.

(Sufficiency.) Suppose there exists $s+1$ distinct (and thus disjoint) modal intervals $[a_i, b_i]$, $i = 0, 1, \dots, s$. Here, $a_i \leq b_i$ and $b_i < a_{i+1}$, $i = 0, 1, \dots, s$; $a_{s+1} = \infty$. It suffices to show that $f \notin \operatorname{osc}(f)$. Equivalently, that there exist numbers x_r and h_r, $r = 1, \dots, 2s$ satisfying (3.2) and (3.3). Which amounts to (at least) s changes of sign from strictly negative to strictly positive for the slope of f.

Consider the modal interval $J(x_0)] = [a, b] = [a_i, b_i]$, where $0 \leq i \leq s$ is *fixed*. Here, $a = a(x_0)$ and $b = b(x_0)$ are as in Definition 2.3 while $x_0 \in [a, b]$ is arbitrary. Let $\delta > 0$ be arbitrary but fixed. It suffices to show that the slope of f has at least one change of sign from strictly positive to strictly negative in $(a - \delta, b + \delta)$. More precisely, it suffices to show that there exist numbers ξ, u and η, v satisfying

$$a - \delta < \xi < \xi + u < \eta < \eta + v < b - \delta, \qquad \Delta_u f(\xi) > 0 \quad \text{and } \Delta_v f(\eta) < 0.$$

We will use the notations of Definition 2.3. Thus $f(x) = c$ for $a < x < b$ and

$$\max(f(a-0), f(a+0)) = \max(f(b-0), f(b+0)) = c.$$

Here, $c > 0$. Choosing $\delta > 0$ sufficiently small, one has $A(x_0) < a - \delta$ and $b + \delta < B(x_0)$, therefore, $f(x) \leq c$ throughout $(a - \delta, b + \delta)$.

Suppose no numbers ξ and u exist such that $a - \delta < \xi < \xi + u < a$ and $\Delta_u f(\xi) > 0$. Then f would be nonincreasing on $(a - \delta, a)$. Because $a = \sup(x < a\colon f(x) < c)$ [see (2.5) with $x_0 = a$], it follows that $f(a-0) < c$, thus, $f(a+0) = c$. Therefore, $\Delta_u f(\xi) > 0$ will hold by choosing ξ slightly to the left of a and $\xi + u$ slightly to the right of a.

Similarly, if no numbers η and v exist such that $b < \eta < \eta + v < b + \delta$ and $\Delta_v f(\eta) < 0$, then f is nondecreasing on $(b, b + \delta)$ thus $f(b+0) < c$, hence, $f(b-0) = c$. And in the latter case, one can attain $\Delta_v f(\eta) < 0$ by choosing η slightly to the left of b and $\eta + v$ slightly to the right of b. There is no difficulty to achieve that, moreover, $\xi + u < \eta$. For, the two above exceptional situations cannot occur simultaneously unless $a < b$. □

Proof of Theorem 2. Suppose $F \in \operatorname{Osc}(s)$. We assert that F is piecewise convex-concave. Let $-\infty < x_0 \leq \infty$ and denote by $Q = Q(x_0)$ the set of values with $-\infty \leq q < x_0$ and such that either F is convex on $[q, x_0]$ or F is concave on $[q, x_0]$. Suppose Q were empty. Then F is not concave on $(-\infty, x_0]$, hence, there exist x_1 and h_1 such that $x_1 < x_1 + 2h_1 < x_0$ and $\Delta^2_{h_1} F(x_1) < 0$. Similarly, since F is not concave on $[x_1 + 2h_1, x_0]$, there exist x_2, h_2 with $x_1 + 2h_1 < x_2 < x_2 + 2h_2 < x_0$ and $\Delta^2_{h_2} F(x_2) < 0$ and so on. Since the process could be continued indefinitely, it would follow that $F \notin \operatorname{Osc}(s)$ (for all s), a contradiction.

Thus Q is nonempty. It is easily seen that also $\alpha(x_0) = \inf Q$ belongs to Q. In a similar way, one has, for each $-\infty \leq x_0 < \infty$, that there exists a largest value $\beta = \beta(x_0)$ with $x_0 < \beta \leq \infty$ and such that either F is convex on $[x_0, \beta]$ or F is concave on $[x_0, \beta]$.

Choosing $x_0 = -\infty$ and letting $b_1 = \beta(-\infty)$, $b_1 > -\infty$, we find that F is convex on $(-\infty, b_1]$ (it is impossible for F to be concave and not $\equiv 0$ on $(-\infty, b_1]$; after all, F is nondecreasing and bounded below). Moreover, $b_1 < \infty$, since F is bounded and not constant.

Next, let $b_2 = \beta(b_1)$ and, in general, define $b_{n+1} = \beta(b_n)$ as long as b_n is finite. Note that $b_n < b_{n+1}$. If $b_n < b_{n+1} = +\infty$, for some $n \geq 1$, then F is clearly piecewise convex-concave. Suppose on the contrary that $b_n < \infty$ for all $n \geq 1$ and consider $b_\infty = \lim b_n$ and $a = \alpha(b_\infty) < b_\infty$. Thus either F is either convex on $[a, b_\infty]$ or F is concave on $[a, b_\infty]$. For n sufficiently large, $b_n \geq a$, thus $b_{n+1} = \beta(b_n) \geq b_\infty$ and we have a contradiction.

Next, suppose $f \in \operatorname{osc}(s)$ and let F be the associated integral as in (2.1). We assert that $F \in \operatorname{Osc}(s)$. If not, then there exist numbers x_r and h_r, $r = 1, \dots, 2s$, satisfying (3.7) and (3.8). In particular, the $2s$ intervals $I_r = [x_r, x_r + 2h_r]$ are disjoint. From (3.8), $(-1)^r F(x)$ is not concave on I_r, thus $(-1)^r f(x)$ is not non-increasing on I_r, that is, there exist $\xi_r \in I_r$ and $u_r > 0$ such that $\xi_r + u_r \in I_r$ and $(-1)^r \Delta_{u_r} f(\xi_r) > 0$, $r = 1, \dots, 2s$. But this contradicts the assumption that $f \in \operatorname{osc}(s)$. Knowing that $F \in \operatorname{Osc}(s)$, it follows that F is piecewise convex-concave, hence, its integrand f is piecewise monotone. □

Proof of Theorem 3. We just proved the first assertion. Conversely, let $F \in \operatorname{Osc}(s)$. We know from Theorem 2 that F is piecewise convex-concave. From the remark preceding (2.3), we have that F is absolutely continuous, that is, F is of the form (2.1) with f as a piecewise monotone function. We claim that $f \in \operatorname{osc}(s)$. If not, then there exist numbers x_r and h_r, $r = 1, \dots, 2s$ satisfying (3.2) and (3.3). One may even assume that x_r and $x_r + h_r$ are continuity points of f. From (2.1) and (3.3), the absolutely continuous function $(-1)^r F(x)$ has at $x_r + h_r$ a strictly larger slope than at x_r and thus is not concave on $[x_r, x_r + h_r]$. Therefore, there exist numbers ξ_r and u_r such that

$$x_r < \xi_r < \xi_r + 2u_r < x_r + h_r \quad \text{and} \quad (-1)^r \Delta_{u_r}^2 F(\xi_r) > 0, \qquad r = 1, \dots, 2s.$$

And it would follow (Definition 3.6) that $F \notin \operatorname{Osc}(s)$. □

PROOF OF THEOREM 6 (Sufficiency). Consider a general mixture G of F as in (4.1), where $F(t, \cdot) \in F$ and suppose that $G \notin \operatorname{Osc}(s)$. From Definition 3.6, since $G \notin \operatorname{Osc}(s)$, there exist numbers x_r and h_r, $r = 1, \dots, 2s$, satisfying

$$x_r < x_r + 2h_r < x_{r+1} \quad \text{and} \quad c_r := (-1)^r \Delta_{h_r}^2 G(x_r) > 0, \qquad r = 1, \dots, 2s;$$

$x_{2s+1} = \infty$. Further define measurable functions ϕ_r: $T \to \mathbf{R}$ by

$$\begin{aligned} \phi_r(t) &= (-1)^r \Delta_{h_r}^2 F(t, x_r) \\ &= (-1)^r \left[F(t, x_r + 2h_r) - 2F(t, x_r + h_r) + F(t, x_r) \right], \end{aligned}$$

$r = 1, \dots, 2s$, and put $\phi(t) = (\phi_1(t), \dots, \phi_{2s}(t))$ and $c = (c_1, \dots, c_{2s})$. Thus ϕ: $T \to \mathbf{R}^{2s}$ and $c \in \mathbf{R}^{2s}$. It follows from (4.1) that $c = \int \phi(t) \rho(dt)$, implying that c belongs to the convex cone K in $\mathbf{R}^{2s}$ which is spanned by all the vectors $\phi(t)$, $t \in T$. [The assertion $c \in K$ follows easily from a proof by induction with respect to the dimension $m = 2s$, see Kemperman (1968), page 95 for a closely related result]. As is well known, since $c \in K$, there must exist $p_j \geq 0$ and $t_j \in T$, $j = 1, \dots, 2s$, such that $c = p_1 \phi(t_1) + \cdots + p_{2s} \phi(t_{2s})$. Equivalently, $c_r = \sum_j p_j \phi_r(t_j)$ for $r = 1, \dots, 2s$. Since $c_r > 0$, one has $\sum_j p_j > 0$.

Finally, let $\overline{G}(x) = p_1 F(t_1, x) + \cdots + p_{2s} F(t_{2s}, x)$, thus, $\overline{G}$ is a special F-mixture as in (4.4). It satisfies

$$(-1)^r \Delta_{h_r}^2 \overline{G}(x_r) = \sum_j p_j (-1)^r \Delta_{h_r}^2 F(t_j, x_r) = \sum_j p_j \phi_r(t_j) = c_r > 0$$

(for all $r = 1, \dots, 2s$), showing that $\overline{G} \notin \operatorname{Osc}(s)$. This completes the proof. □

PROOF OF THEOREM 7. We will arrange the proof so as to also yield a new proof of Theorem 6. We want to show that the following assertions are equivalent. Here, $f_j \in F_0$ and $p_j \geq 0$, $\sum_j p_j > 0$.

(i) Each special F_0-mixture $g = p_1 f_1 + \cdots + p_{2s} f_{2s}$ belongs to $\operatorname{osc}(s)$.
(ii) Condition (4.7) holds for each subfamily of F_0 of size $\leq 2s$.
(iii) Condition (4.7) holds for F_0 itself.
(iv) Each mixture g of F_0 belongs to $\operatorname{osc}(s)$.

We already showed that (iii) ⇒ (iv), see the comments following (4.10), while it is trivial that (iv) ⇒ (i). Hence, it suffices to show that (i) ⇒ (ii) ⇒ (iii).

Proof that (ii) ⇒ (iii): Let $x_1 < x_2 < \cdots < x_{2s}$ be fixed as well as the numbers $h_r > 0$, such that $x_r + h_r < x_{r+1}$, $r = 1, \dots, 2s$; $x_{2s+1} = \infty$. Let further S denote the simplex which consists of all vectors $a = (a_1, \dots, a_{2s})$ such that $a_r \geq 0$; $\Sigma_r a_r = 1$. For each $f: \mathbf{R} \to \mathbf{R}$, put

$$S(f) = \left\{ a \in S: \sum_{r=1}^{2s} (-1)^r a_r \Delta_{h_r} f(x_r) \leq 0 \right\}.$$

Thus, assertion (iii) says that the collection $\{S(f): f \in F_0\}$ of compact and convex subsets of S has a nonempty intersection. Since $\dim(S) = 2s - 1$, it follows from Helly's theorem [see Valentine (1964) page 70] that the latter is true as soon as each subcollection, which consists of at most $2s$ sets $S(f_j)$, $f_j \in F_0$, has a nonempty intersection. Clearly, the latter property is nothing but a restatement of (iii).

Proof that (i) ⇒ (ii): Let $f_j \in F_0$, $j = 1, \dots, 2s$, and let the x_r, h_r be as in Condition 4.7. Further put $b_{rj} = (-1)^r \Delta_{h_r} f_j(x_r)$, $r, j = 1, \dots, 2s$. It follows from (i) and Definition 3.1 of osc(s) that there do *not* exist numbers $p_j \geq 0$, $j = 1, \dots, 2s$, such that $\Sigma_j b_{rj} p_j > 0$, $r = 1, \dots, 2s$. While (ii) says that there do exist numbers $a_r \geq 0$ not all zero such that $\Sigma_r a_r b_{rj} \leq 0$, $j = 1, \dots, 2s$. And it is well known that these two properties are equivalent, see, for instance, Fishburn [(1985) page 24]. Geometrically, this equivalence means: In order that the convex cone in $\mathbf{R}^{2s}$ which is spanned by the columns of $B = (b_{rj})$ be disjoint from the open first quadrant, it is necessary and sufficient that this cone be contained in some half space $\{z \in \mathbf{R}^{2s}: \Sigma_r a_r z_r \leq 0\}$ with $a_r \geq 0$ and not all zero. □

The proofs of Theorems 7′ and 7″ are completely analogous, except that now one uses the criterion furnished by Lemma 3 or Lemma 2, respectively.

PROOF OF LEMMA 5. We want that the mixture g be always log-concave on K, that is,

$$(d/dx)^2 \log[p_1 \exp(-u_1(x)) + p_2 \exp(-u_2(x))] \leq 0$$

for all $x \in K$; $p_1, p_2 > 0$. Let $x \in K$ be fixed. Putting $q = (p_1/p_2)\exp(-u_1 + u_2)$, this is equivalent to

$$u_1'' q^2 - 2Bq + u_2'' \geq 0 \quad \text{for all } q > 0. \tag{7.2}$$

Here, $2B = (u_1' - u_2')^2 - u_1'' - u_2''$. Note that $u_i'' \geq 0$, $i = 1, 2$. Condition (7.2) holds if and only if either $B \leq 0$ or else $B^2 - u_1'' u_2'' \leq 0$. Equivalently, if and only if $B \leq (u_1'' u_2'')^{1/2}$. And the latter inequality is equivalent to (6.2). □

PROOF OF THEOREM 9. We are still assuming that $u \in C^2$ and $\{x: u(x) < \infty\} = K$ for all $u \in U$. As we observed already, Theorem 9 immediately follows from Lemma 5 and Theorem 5. For instance, the necessity of (6.3) follows from Lemma 5. The following is a more elementary proof that, in the case on hand, condition (6.3) is sufficient for a large class of F_0-mixtures g to be log-concave on K. This will include all mixtures

$$g = \int f_t \rho(dt) \quad \text{such that} \quad g' = \int f_t' \rho(dt) \quad \text{and} \quad g'' = \int f_t'' \rho(dt). \tag{7.3}$$

Here, $f_t \in F_0$, $t \in T$ while ρ is a finite nonzero measure on some measurable space T. Property (7.3) is certainly true if ρ has a finite support, that is, for all mixtures involving only finitely many $f_t \in F_0$.

Let $x \in K$ be fixed. We want to show that g as in (7.3) is log-concave at x, that is, $(g'(x))^2 - g(x)g''(x) \geq 0$. Equivalently,

$$w^2 g(x) + 2wg'(x) + g''(x) \leq 0$$

for at least one number $w = w(x)$. From (7.3), it suffices that there exists a constant $w = w(x)$ such that

$$w^2 f(x) + 2wf'(x) + f''(x) \leq 0 \quad \text{for all } f \in F_0. \tag{7.4}$$

If $f = e^{-u}$, then $u' = -f'/f$ and $u'' = (f'^2 - ff'')/f^2$. One easily verifies that (7.4) demands precisely that $w = w(x)$ belongs to all intervals $J(f)$, $f \in F_0$, where

$$J(f) = \left[u'(x) - u''(x)^{1/2}, u'(x) + u''(x)^{1/2}\right].$$

Clearly, such a point w exists if and only if (6.3) holds. □

REFERENCES

BEHBOODIAN, J. (1970). On the modes of a mixture of two normal distributions. *Technometrics* **12** 131–139.

FISHBURN, P. C. (1985). *Interval Orders and Interval Graphs: A Study of Partially Ordered Sets.* Wiley, New York.

IBRAGIMOV, I. A. (1956). On the composition of unimodal distributions. *Theory Probab. Appl.* **1** 255–260.

KAKIUCHI, I. (1981). Unimodality conditions of the distribution of a mixture of two distributions. *Math. Sem. Notes Kobe Univ.* **9** 315–325.

KARLIN, S. and STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics.* Wiley, New York.

KEILSON, J. and GERBER, H. (1971). Some results for discrete unimodality. *J. Amer. Statist. Assoc.* **66** 386–389.

KEMPERMAN, J. H. B. (1968). The general moment problem, a geometric approach. *Ann. Math. Statist.* **39** 93–122.

ROBERTSON, C. A. and FRYER, J. G. (1969). Some descriptive properties of normal mixtures. *Scand. Actuar. Tidskr.* **52** 137–146.

TITTERINGTON, D. M., SMITH, A. F. M. and MAKOV, U. E. (1985). *Statistical Analysis of Finite Mixture Distributions.* Wiley, New York.

VALENTINE, F. A. (1964). *Convex Sets.* McGraw-Hill, New York.

DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903