# A CONVOLUTIVE CLASS OF MONOTONE LIKELIHOOD RATIO FAMILIES

BY S. G. GHURYE AND DAVID L. WALLACE

*University of Chicago*

**1. Introduction.** This note stems from the following problem posed us by J. Loevinger.[1] Let $X_1, \cdots X_n$ and $Y$ be real-valued random variables such that, conditionally on $Y$, the $\{X_i\}$ are mutually independent with

$$p_i(Y) = \Pr\{X_i = 1 \mid Y\} = 1 - \Pr\{X_i = 0 \mid Y\}$$

and $p_i(y)$ is nondecreasing in $y$. Let $S = X_1 + \cdots + X_n$. Is $E\{Y \mid S = r\}$ a nondecreasing function of $r$? The answer, yes, will follow from showing that $\Pr\{S = r + 1 \mid Y\}/\Pr\{S = r \mid Y\}$ is a nondecreasing function of $Y$ for each $r$. Here we have a simple case of the convolution of families of distributions with monotone likelihood ratios (hereafter MLR) being an MLR family. It is easy to see that the convolution of two MLR families is not necessarily MLR. In Section 2, a sufficient condition on MLR families is given that their convolution be MLR. In Section 3, some special results are given for multidimensional distributions. The problem leading to this work is discussed in Section 4.

The MLR property is identical with the Pólya type 2 property (cf. [2]). The definitions used here extend to Pólya type $m$ but the extended results, except for Lemma 4, are not generally true for $m > 2$.

**2. Convolutions of MLR families.** Let $G$ be an ordered additive group, let $\Theta$ be an ordered set, and let $\mu$ be an invariant, $\sigma$-finite measure on $G$. Throughout this section, a family $f$ will mean a real-valued, nonnegative function on $G \times \Theta$, such that $f(x, \theta)$ is measurable in $x$ for each $\theta$ and

$$0 < \textstyle\int_G f(x, \theta)\, d\mu(x) < \infty.^2$$

Ordinarily, $f$ is a family of probability densities relative to $\mu$ for a random variable with range contained in $G$ and with parameter space $\Theta$. The convolution family $f * g$ of two families $f$ and $g$ is defined by

$$f * g(x, \theta) = \int_G f(x - u, \theta) g(u, \theta)\, d\mu(u).$$

The spaces $G$ and $\Theta$ must be ordered for the definitions which follow and $G$ must be at least a semigroup for convolutions to be defined in the same space.

Received January 14, 1959; revised April 15, 1959.

[1] The consultation with Dr. Loevinger (Jewish Hospital of St. Louis) and the research connected with this paper were performed under a grant from the Rockefeller Foundation.

[2] For some purposes, such as in Section 3, it would be convenient to permit the integral to be zero for some $\theta$. Lemma 3 and the theorems of this section clearly hold under this extension.

The group requirement is only a slight restriction. $G$ will ordinarily be the real line or the integers, and is taken as an ordered group primarily because it permits a simple unified treatment at no extra cost.

*Definition.* A nonnegative function $h$ defined on the product of two ordered sets $X$, $Y$, is Pólya type 2 if, for all $x, x' \varepsilon X$ and $y, y' \varepsilon Y$ such that $x \leqq x'$, $y \leqq y'$,

$$h(x, y)h(x', y') - h(x, y')h(x', y) \geqq 0.$$

*Definition A.* A family $f$ has property $A$ if, as a function on $G \times \Theta$, it is Pólya type 2.

*Definition B.* A family $f$ has property $B$ if, for each $\theta \varepsilon \Theta$, the function $h_\theta$ on $G \times G$ defined by

$$h_\theta(x, \xi) = f(x - \xi, \theta)$$

is Pólya type 2.

*Definition C.* A family $f$ has property $C$ if, for all $x, x', y, y' \varepsilon G$ such that $y \leqq x \leqq y'$ and $x + x' = y + y'$ and all $\theta, \theta' \varepsilon \Theta$ such that $\theta \leqq \theta'$,

$$f(x, \theta)f(x', \theta') \geqq f(y', \theta)f(y, \theta').$$

Property $A$ is the monotone likelihood ratio or Pólya type 2 property for the family $f$. Property $B$ is the monotone likelihood ratio property for the location parameter family generated by $f(\cdot, \theta)$ for each fixed $\theta$. Provided all quantities used as divisors are positive, the definitions of properties $A$ and $B$ can be expressed in the more intuitive form:

$A$: $f(x, \theta')/f(x, \theta)$ nondecreasing in $x$ for all $\theta < \theta'$, or

$A$: $f(x + h, \theta)/f(x, \theta)$ nondecreasing in $\theta$ for all $x$ and all $h > 0$, and

$B$: $f(x + h, \theta)/f(x, \theta)$ nonincreasing in $x$ for all $\theta$ and all $h > 0$.

Note that on taking $x = y$ and $x' = y'$ in $C$, one obtains $A$; that on taking $\theta = \theta'$ in $C$, one obtains $B$. We shall now show that property $C$ is, in fact, equivalent to $A$ and $B$ together, and that it is invariant under convolution.

It may be helpful to note that all results and methods of this paper are unaffected if any $f(x, \theta)$ is multiplied by any positive function of $\theta$. Multiplication by a positive function of $x$ does not destroy MLR, but does affect the convolution and its MLR properties.

LEMMA 1: *If $f$ has property $B$, then the set $I_f(\theta) = \{x: f(x, \theta) > 0\}$ is, for every $\theta \varepsilon \Theta$, an interval of $G$; i.e., $y \varepsilon I$, $y' \varepsilon I$ imply $x \varepsilon I$ for all $x \varepsilon G$ such that $y \leqq x \leqq y'$.*

PROOF: Suppose $f(y, \theta)f(y', \theta) > 0$ and $y < y'$. To any $x \varepsilon G$ such that $y \leqq x \leqq y'$, there corresponds an $x' \varepsilon G$ such that $x + x' = y + y'$ and by property $B$, $f(x, \theta)f(x', \theta) \geqq f(y, \theta)f(y', \theta) > 0$.

Thus, for each $\theta$, there is a decomposition of $G$ into three intervals $M(\theta)$, $I(\theta)$, $M'(\theta)$ such that $x \varepsilon M$, $y \varepsilon I$, $z \varepsilon M'$ imply $x < y < z$ and $f(x, \theta) = f(z, \theta) = 0$, $f(y, \theta) > 0$. For all $\theta$, $I_f(\theta)$ is nonempty, though it may contain only one point. $M(\theta)$ and $M'(\theta)$ may be empty.

LEMMA 2: *If $f$ has properties $A$ and $B$, then for any $\theta$, $\theta' \varepsilon \Theta$ such that $\theta < \theta'$, $M(\theta) \subset M(\theta')$ and $M'(\theta) \supset M'(\theta')$.*

PROOF: For any $\theta$, choose $x' \varepsilon I(\theta)$. Then for any $x \varepsilon M(\theta)$, $x < x'$ and using property $A, f(x, \theta') = 0$. Since this holds for all $y \leqq x$, then $x \varepsilon M(\theta')$. A similar proof holds for $M'(\theta)$.

LEMMA 3: *$f$ has property $C$ if and only if it has properties $A$ and $B$.*

PROOF. That $C$ implies $A$ and $B$ is immediate. Property $C$ is nontrivial only when $y' \varepsilon I(\theta)$ and $y \varepsilon I(\theta')$. Since $y < y'$, we know from Lemma 2 that $f(y, \theta) > 0$. Using successively $A$ and $B$,

$$f(y, \theta)f(x, \theta)f(x', \theta') \geqq f(y, \theta')f(x, \theta)f(x', \theta) \geqq f(y, \theta')f(y, \theta)f(y', \theta)$$

and $C$ follows by division.

LEMMA 4. (Schoenberg [5]). *If $f$ and $g$ have property $B$, then $f * g$ has property $B$.* Schoenberg's proof for the real line extends immediately to the group $G$. He proves this result in its Pólya type $m$ form.

THEOREM 1: *If $f$ and $g$ have property $C$, then $f * g$ has property $C$.*

PROOF: Using Lemmas 3 and 4, it remains only to show that $f * g$ has property $A$, i.e., for $x < x' = x + h$, $\theta < \theta'$,

$$\Delta_2 \equiv [f * g(x, \theta)][f * g(x', \theta')] - [f * g(x', \theta)][f * g(x, \theta')] \geqq 0.$$

Throughout the proof, write $f(x) = f(x, \theta)$ and $f'(x) = f(x, \theta')$.

$$\Delta_2 = \int [f(u)g(x - u)f'(v)g'(x' - v) - f'(u)g'(x - u)f(v)g(x' - v)] \cdot d[\mu(u) \times \mu(v)] = I_1 + I_2 + I_3$$

in which $I_1, I_2, I_3$ are respectively the integrals over the sets, $u > v, u = v, u < v$.

Interchange $u$ and $v$ in $I_3$ and incorporate with $I_1$.

$$I_1 + I_3 = \int_{u>v} \{f(u)f'(v)[g(x - u)g'(x' - v) - g(x' - u)g'(x - v)] + f'(u)f(v)[g(x - v)g'(x' - u) - g'(x - u)g(x' - v)]\} \cdot d[\mu(u) \times \mu(v)].$$

For $u > v$, the quantity in the second brackets is nonnegative by $C$ and its coefficient $f'(u)f(v) \geqq f(u)f'(v)$. Then,

$$I_1 + I_3 \geqq \int_{u>v} f(u)f'(v)[g(x - u)g'(x' - v) - g'(x - u)g(x' - v) + g(x - v)g'(x' - u) - g(x' - u)g'(x - v)]\, d[\mu(u) \times \mu(v)]$$

and

$$\Delta_2 \geqq \int_{u>v} f(u)f'(v)[g(x - u)g'(x' - v) - g'(x - u)g(x' - v)]\, d[\mu(u) \times \mu(v)] + \int_{u>v+2h} f(u)f'(v)[g(x - v)g'(x' - u) - g(x' - u)g'(x - v)]\, d[\mu(u) \times \mu(v)]$$

$$+ \int_{0 \leqq u-v \leqq 2h} f(u)f'(v)[g(x-v)g'(x'-u) - g(x'-u)g'(x-v)] \cdot d[\mu(u) \times \mu(v)]$$

$= J_1 + J_2 + J_3$ respectively.

In $J_2$ make the transformation $u = u' + h$, $v = v' - h$, suppress the primes on $u'$, $v'$, and recall that $h = x' - x > 0$. Then

$$J_1 + J_2 = \int_{u>v} [f(u)f'(v) - f(u+h)f'(v-h)] \cdot [g(x-u)g'(x'-v) - g'(x-u)g(x'-v)]\, d[\mu(u) \times \mu(v)] \geqq 0.$$

Break $J_3$ into three integrals respectively over the sets $0 \leqq u - v < h$, $h < u - v \leqq 2h$, $u - v = h$. The third integral vanishes, and on making the transformation $u = v' + h$, $v = u' - h$ in the second and suppressing the primes on $u'$, $v'$, we get

$$J_3 = \int_{0 \leqq u-v<h} [f(u)f'(v) - f(v+h)f'(u-h)] \cdot [g(x-v)g'(x'-u) - g(x'-u)g'(x-v)]\, d[\mu(u) \times \mu(v)] \geqq 0.$$

Hence $\Delta_2 \geqq 0$, and the theorem is proved.

*Remark.* This result would appear to be subsumed under Theorem 3 of Lehmann [3], taking $g_\theta(x, \xi) = f(x - \xi, \theta)$ and $d\lambda_\theta(\xi) = g(\xi, \theta)\, d\mu(\xi)$. However, in lines 6 to 8 of page 410 of his proof, an additional assumption is needed, which is not met in our case.[3]

COROLLARY 1: *If* $\Theta = G$, $f(x, \theta) = f(x - \theta)$, $g(x, \theta) = g(x - \theta)$ *and* $f$ *and* $g$ *have property* $A$, *then* $f * g$ *has property* $A$. (This result for location parameter families is known and is just the Schoenberg result of Lemma 4.)

COROLLARY 2: *If* $G$ *is the set of integers, if for each* $i = 1, \cdots, n$,

$$f_i(x, \theta) = \begin{cases} p_i(\theta) & x = 1 \\ 1 - p_i(\theta) & x = 0 \\ 0 & \text{otherwise} \end{cases}$$

*and* $p_i(\theta)$ *is nondecreasing in* $\theta$, *then each* $f_i$ *and the convolution* $f_1 * \cdots * f_n$ *have property* $A$.

That $B$ is not a necessary property for the convolution of two MLR families to be MLR is shown by the construction below based on the following theorem, whose proof is a simple computation.

THEOREM 2: *If* $f$ *has property* $A$ *and if, for each* $\theta$, *the range of* $x$ *for which* $f(x, \theta) > 0$ *is contained in* $(0, 1, 2)$, *then* $f * f$ *has property* $A$.

This result does not extend in general to nonidentical convolutions, to threefold identical convolutions, or to fourpoint ranges.

A family $f$ which satisfies Theorem 2 but does not have property $B$ is easily

---

[3] We wish to thank Professor S. Karlin for calling this fact to our attention.

constructed by taking $\Theta$ as the real line, $a(\theta)$, $b(\theta)$ as increasing functions on $\Theta$ such that $0 < a(\theta) < b(\theta) < \infty$, and letting $f(\cdot, \theta)$ be the distribution with probabilities at 0, 1 and 2 respectively given by

$$c(\theta), \qquad a(\theta)c(\theta), \qquad a(\theta)b(\theta)c(\theta)$$

with $c(\theta) = [1 + a(\theta) + a(\theta)b(\theta)]^{-1}$.

**3. Some results for multivariate distributions.** A family of generalized densities $f(x, \theta)$, where $x$ is a vector is said to be MLR (or Pólya type 2) if it is MLR along each increasing curve, i.e., if for every vector function $x(t)$ of the real-parameter $t$ for which the components are nondecreasing functions of $t$, $g(t, \theta) = f\{x(t), \theta\}$ is MLR in $t$ and $\theta$. (Cf. Lehmann [3], Pratt [4].) The definition can also be stated in the form:

$f(x_1, \cdots, x_K, \theta)$ is MLR if, for all $x_i \leqq x_i'$, $i = 1, \cdots, K$, $\theta \leqq \theta'$,

$$f(x_1, \cdots, x_K, \theta)f(x_1', \cdots, x_K', \theta') \geqq f(x_1', \cdots, x_K', \theta)f(x_1, \cdots, x_K, \theta').$$

We consider only the simplest problem of extending Corollary 2 to families of distributions on the vertices of the cube or the simplex in $K$ dimensions. In two dimensions already, two MLR families on the points (0, 0), (0, 1), (1, 0) need not have an MLR convolution (Counterexample 1). Restricting consideration to $n$-fold convolutions of a single family, the $n$-fold convolution of an MLR family on the vertices of the square is MLR (Theorem 3), but even the two-fold convolution of an MLR family on the vertices of the three-dimensional cube need not be MLR (Counterexample 2). However, the $n$-fold convolution of an MLR family on the vertices of the $K$-dimensional simplex is MLR for all $n$ and $K$ (Theorem 2).

*Counterexample* 1. The convolution of two MLR families $f_1$ and $f_2$ on the points (0, 0), (1, 0), (0, 1) need not be MLR: Let $a(\theta)$ be a positive, increasing function of $\theta$ and let $f_1$ place nonzero probabilities only on the three points (0, 0), (1, 0), (0, 1) proportional, respectively, to 1, 2, $a(\theta)$. Let $f_2(x, \theta) = \frac{1}{3}$ at each point. Both are MLR. Then at the two points (0, 1) and (1, 1), $f_1 * f_2$ has probabilities proportional, respectively, to $[1 + a(\theta)]$ and $[2 + a(\theta)]$, and hence $f_1 * f_2$ is not MLR.

*Counterexample* 2. The two-fold convolution of an MLR family on the vertices of the three-dimensional cube need not be MLR: Let $f$ place nonzero probabilities only on the five points (0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0) proportional, respectively, to 1, $a(\theta)$, $a(\theta)$, 2, $a(\theta)$, with $a(\theta)$ as above. Then $f$ is MLR but $f * f$ is not, since at the points (1, 1, 0) and (1, 1, 1) the probabilities are proportional, respectively, to $[1 + a(\theta)]$ and 2.

THEOREM 3: *If $f$ is an MLR family on the four points* (0, 0), (1, 0), (0, 1), (1, 1), *then, for every $n$, the $n$-fold convolution of $f$ with itself is MLR.*

PROOF. Let $f(x, \theta) = p_{ij}(\theta)$, for $x = (i, j)$, and let $q_{ij}(\theta)$ be the value of the $n$-fold convolution of $f$ at $(i, j)$, which is given by

$$\sum_{i,j} q_{ij}(\theta)t^iu^j = [\sum_{i,j} p_{ij}(\theta)t^iu^j]^n. \tag{1}$$

We are given that for all $i \leqq i', j \leqq j', \theta \leqq \theta'$,

$$p_{ij}(\theta)p_{i'j'}(\theta') \geqq p_{ij}(\theta')p_{i'j'}(\theta), \tag{2}$$

and must show that for all $r \leqq r'$, $s \leqq s'$, $\theta \leqq \theta'$,

$$q_{rs}(\theta)q_{r's'}(\theta') \geqq q_{rs}(\theta')q_{r's'}(\theta). \tag{3}$$

From (1), it follows that, for given $s$, the sequence $\{q_{rs}(\theta), r = 0, 1, \cdots\}$ has the generating function

$$P_s(t) = \binom{n}{s}\{p_{00}(\theta) + p_{10}(\theta)t\}^{n-s}\{p_{01}(\theta) + p_{11}(\theta)t\}^s$$

and, for given $r$, the sequence $\{q_{rs}(\theta), s = 0, 1, \cdots\}$ has the generating function

$$Q_r(u) = \binom{n}{r}\{p_{00}(\theta) + p_{01}(\theta)u\}^{n-r}\{p_{10}(\theta) + p_{11}(\theta)u\}^r.$$

Both represent convolutions of two-point, one-dimensional families, MLR by (2), and hence, by Corollary 2 to Theorem 1,

$$q_{rs}(\theta)q_{r's}(\theta') \geqq q_{r's}(\theta)q_{rs}(\theta')$$

and

$$q_{rs}(\theta)q_{rs'}(\theta') \geqq q_{rs'}(\theta)q_{rs}(\theta').$$

The desired conclusion (3) follows easily if at least one of $q_{r's}(\theta)$, $q_{r's}(\theta')$, $q_{rs'}(\theta)$, $q_{rs'}(\theta')$ is positive. (3) is trivially true unless $q_{rs}(\theta')q_{r's'}(\theta) > 0$ and is one of the above special cases unless $r' > r$ and $s' > s$. If $r' < s'$, then $q_{r's'}(\theta) > 0$ implies either $p_{11}(\theta)p_{01}(\theta) > 0$ or $p_{10}(\theta)p_{01}(\theta) > 0$. In either case, $q_{rs'}(\theta) > 0$ and (3) follows. A similar argument holds if $r' > s'$ and also if $r' = s'$ except when $p_{11}(\theta) > 0$, $p_{10}(\theta) = p_{01}(\theta) = 0$. But then, by the MLR property, also $p_{10}(\theta') = p_{01}(\theta') = 0$ and at $\theta$, $\theta'$ the distributions are one-dimensional along the diagonal. Hence Corollary 2 (for the group of diagonal integers) applies directly.

THEOREM 4: *If $f$ is an MLR family on the $K + 1$ vertices of the unit simplex in $K$ dimensions, then, for every $n$, the $n$-fold convolution of $f$ with itself is MLR.*

PROOF: Let $\{p_j(\theta), j = 0, 1, \cdots K\}$ be the values of the family $f$ at the origin and unit points of the $K$ axes respectively. Let $q_r(\theta)$ denote the value of the $n$-fold convolution at the point $r = (r_1, \cdots, r_K)$. We must show that for $\theta \leqq \theta'$, and $r$, $r'$ such that $r_i \leqq r_i'$, $i = 1, \cdots, K$,

$$q_r(\theta)q_{r'}(\theta') \geqq q_r(\theta')q_{r'}(\theta). \tag{4}$$

A generating function argument similar to that used above easily proves the result when $r$ and $r'$ differ in only one coordinate. But if $q_{r'}(\theta) > 0$, then $p_j(\theta) > 0$ for all coordinates such that $r_j' > 0$, and hence $q_s(\theta) > 0$ for all $s$ such that $s_j \leqq r_j'$, $j = 1, \cdots, K$. Division by $q_s(\theta)$ is permissible and (4) follows easily by repeated application of the result for changes in a single coordinate.

**4. An application.** The problem mentioned in the introduction arose in the following way. Let $X_1, \cdots, X_n$ denote the scores made by an individual on

$n$ test items with the value 1 for correct, 0 for incorrect. Let $S = X_1 + \cdots + X_n$. Let $Y$ be a real random variable representing the individual's (unobservable) position on a single scale (latent continuum) assumed to determine his performance on the test according to the model: For each $i$,

$$f_i(1, Y) = p_i(Y) = \Pr\{X_i = 1 \mid Y\} = 1 - \Pr\{X_i = 0 \mid Y\}$$

and conditionally on $Y$, the $\{X_i\}$ are mutually independent. Let $\phi$ be the probability density function for $Y$, representing, perhaps, the distribution of the ability $Y$ over some population. If it is assumed only that each $f_i$ is a nondecreasing function, what can be said about the individual value of $Y$ conditionally on the sum $S$ of the scores of the $n$ items? The answer is that the conditional distribution function of $Y$ given $S = a$ lies to the left of that for $S = b > a$. Hence, the conditional mean (or median or quantile) of $Y$ given $S$ is a nondecreasing function of $S$.

(The result would not be true without restrictions on the functions $p_i$ if, for example, the difference between a correct and an incorrect score differed from item to item.)

The result is an application of Corollary 2 to Theorem 1 and of the following lemma.

LEMMA. *If $X$, $Y$ are real random variables, if $Y$ has density $\phi(y)$ relative to the measure $\nu$, if $X$ given $Y = y$ has the conditional density $f(x, y)$ relative to the measure $\mu$, and if the family $f$ is MLR, then*

$$\Pr\{Y \leqq a \mid X = x\} \geqq \Pr\{Y \leqq a \mid X = x'\}$$

*for all $a$ and all $x \leqq x'$.*

The lemma can be proved by relatively simple calculations and is equivalent to a result of Good [1].

## REFERENCES

[1] I. J. GOOD, "A note on positive determinants," *J. London Math Soc.*, Vol. 22 (1947), pp. 92–95.

[2] SAMUEL KARLIN, "Decision theory for Pólya type distributions. Case of two actions, I," *Proceedings of the Third Berkeley Symposium on Probability and Statistics*, Vol. 1, University of California Press, 1955, pp. 115–128.

[3] E. L. LEHMANN, "Ordered families of distributions," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 399–419.

[4] J. W. PRATT, "Some results in the decision theory of one-parameter multivariate Pólya type distributions," *Stanford Technical Report No. 37*, 21 October 1955.

[5] I. J. SCHOENBERG, "On Pólya frequency functions, I. The totally positive functions and their Laplace transforms," *J. d'Analyse Math.*, Vol. 1 (1951), pp. 331–374.