ON THE UNIMODALITY OF HIGH CONVOLUTIONS OF DISCRETE DISTRIBUTIONS

BY A. M. ODLYZKO AND L. B. RICHMOND

AT&T Bell Laboratories and University of Waterloo

It is shown that if $\{p_j\}$ is a discrete density function on the integers with support contained in $\{0, 1, \cdots, d\}$, and $p_0 > 0$, $p_1 > 0$, $p_{d-1} > 0$, $p_d > 0$, then there is an n_0 such that the n-fold convolution $\{p_j\}^{*n}$ is unimodal for all $n \geq n_0$. Examples show that this result is nearly best possible, but weaker results are proved under less restrictive assumptions.

1. Introduction. The unimodality of distribution functions has been of substantial interest, especially in connection with the question of whether all class L distributions are unimodal (which was finally answered in the affirmative by Yamazato, 1978). In view of the (at that time unproved, but widely conjectured) unimodality of the limiting distributions of class L, A. Rényi (see Medgyessy, 1977) conjectured that something stronger ought to hold for a discrete distribution $\{p_j\}$ on the integers, namely that for each such distribution there ought to be an integer n_0 such that the n-fold convolution $\{p_j\}^{*n}$ is unimodal for all $n \geq n_0$. Medgyessy (1977) extended this conjecture to continuous distribution functions. However, the Rényi and Medgyessy conjectures are both false, as was recently shown by Brockett and Kemperman (1982) and by Ushakov (1982). Their counterexamples show that it is hard to guarantee unimodality even for high convolutions of a discrete distribution if the distribution has infinite support. However, Brockett and Kemperman conjectured that if $p_0, p_1, \cdots, p_d > 0$ and $p_k = 0$ for $k < 0$ and $k > d$, then for $n \geq n_0$ the n-fold convolution $\{p_j\}^{*n}$ is unimodal, and they proved this conjecture for $d = 2$. A similar question was raised by B. McKay (unpublished). This paper proves a result stronger than that conjectured by Brockett and Kemperman, namely that $\{p_j\}^{*n}$ is even strongly unimodal.

A probability distribution $\{p_j\}$ on the integers is called unimodal if the sequence $\{p_{j+1} - p_j\}_{-\infty}^{\infty}$ has exactly one change of sign. Various results about unimodal distributions are contained in Karlin (1968) and Medgyessy (1977). A more restrictive concept than that of unimodality is that of strong unimodality; a discrete distribution $\{p_j\}$ is strongly unimodal if $\{p_j\} * \{q_j\}$ is unimodal for any unimodal discrete distribution $\{q_j\}$. A strongly unimodal distribution is unimodal, but not conversely. A discrete distribution $\{p_j\}$ is strongly unimodal if and only if it is log concave; i.e., $p_j^2 \geq p_{j-1}p_{j+1}$ for all $j \in Z$ (see Keilson and Gerber, 1971). We prove:

THEOREM 1. *If $\{p_j\}$ is a discrete distribution with $p_j = 0$ for $j < 0$ and $j > d$,*

while $p_0 > 0$, $p_1 > 0$, $p_{d-1} > 0$, $p_d > 0$, then there exists an integer n_0 such that for $n > n_0$ the n-fold convolution $\{p_j\}^{*n}$ is strongly unimodal.

THEOREM 2. *If $\{p_j\}$ is a discrete distribution with $p_j = 0$ for $j < 0$ and $j > d$, while $p_0 > 0$, $p_d > 0$, and*

$$(1.1) \qquad\qquad \gcd\{j : p_j \neq 0\} = 1, \qquad\qquad .$$

then for any $\delta > 0$ there is an $n_0 = n_0(\delta)$ such that if $a_{k,n}$ denotes the value of the n-fold convolution $\{p_j\}^{*n}$ at k, then for $n \geq n_0$,

$$a_{k,n}^2 \geq a_{k-1,n} a_{k+1,n}$$

for $\delta n \leq k \leq (d - \delta)n$.

The greatest common divisor condition (1.1) of Theorem 2 is obviously necessary for the conclusions of that theorem to hold (as otherwise the distribution and all multiple convolutions of it with itself are concentrated on multiples of that greatest common divisor), but it is not sufficient to obtain the conclusions of Theorem 1. In Section 2 we show that for any $\varepsilon > 0$, there is a distribution satisfying the hypotheses of Theorem 2, and for which the inequalities

$$(1.2) \qquad\qquad a_{k,n} > a_{k+1,n}, \quad a_{k+1,n} < a_{k+2,n}$$

hold for k as large as $n^{1-\varepsilon}$ and $n \geq n_0(\varepsilon)$.

It is possible to obtain results stronger than those of our Theorems 1 and 2 by more careful analysis. For example, it can be shown that high convolutions of distributions satisfying the hypotheses of Theorem 1 have stronger variation-diminishing properties than that guaranteed by strong unimodality. (See Karlin, 1968, for a discussion of such properties.)

Our proofs also provide quantitative information about the distribution $\{p_j\}^{*n}$. For example, it can be deduced easily from our proofs that if the $\{p_j\}$ satisfy the conditions of Theorem 1, if $k \to \infty$ in such a way that $nd - k \to \infty$, and α is defined as the unique positive solution to

$$(1.3) \qquad\qquad e^\alpha p'(e^\alpha)/p(e^\alpha) = k/n,$$

where

$$(1.4) \qquad\qquad p(z) = \sum_{j=0}^{d} p_j z^j,$$

then $a_{k,n}$, the value of $\{p_j\}^{*n}$ at k, satisfies

$$(1.5) \qquad\qquad a_{k,n} \sim \frac{e^{-\alpha k} p(e^\alpha)^n}{2\sqrt{\pi n \beta_k(n)}} \quad \text{as} \quad n \to \infty,$$

where

$$(1.6) \qquad\qquad \beta_k(n) = \tfrac{1}{2}(\partial^2/\partial x^2)\log p(e^x)\,|_{x=\alpha}.$$

Finally we mention that related results and references to many unimodality results from combinatorial theory are contained in Odlyzko and Richmond (1982).

The authors thank the referee for several useful comments and corrections.

2. Examples and elementary proofs.

In this section we show that under the hypotheses of Theorem 2, its conclusions cannot be strengthened significantly. We also prove Theorem 1 for $0 \le k \le n^{1/4}$ and $dn - n^{1/4} \le k \le dn$, provided n is large enough.

To show that Theorem 2 is nearly best possible, consider the distribution

$$p_0 = p_2 = p_m = \frac{1}{3}, \quad p_j = 0 \quad \text{for} \quad j \ne 0, 2, m,$$

where m is an odd integer ≥ 3. Then condition (1.1) is satisfied. We will show that for this distribution, (1.2) holds for $k \le n^{1-2/m}$ if n is large enough. This result can also be proved by a more elementary argument that uses estimates of multinomial coefficients, but we prefer to use the analytic proof given below, since it introduces the techniques which we find necessary to use in later sections.

The value of the n-fold convolution $\{p_j\}^{*n}$ at the integer k is $3^{-n}a_{k,n}$ where $a_{k,n}$ is the coefficient of z^k in $p(z)^n$, $p(z) = 1 + z^2 + z^m$. Now $a_{k,n}$ is given by

$$(2.1) \qquad a_{k,n} = \frac{1}{2\pi i} \int_{|z|=r} p(z)^n z^{-k-1} \, dz,$$

where $r > 0$ is any constant. Choose $r = k^{1/2}(2n - k)^{-1/2}$, $0 \le k \le n$. Then, on $|z| = r$, for $k \le n$, we have

$$p(z) = 1 + z^2 + O(r^m) = (1 + z^2)(1 + O(r^m)),$$

and so for $nr^m = O(1)$, say, which we assume from now on,

$$p(z)^n = (1 + z^2)^n(1 + O(r^m))^n = (1 + z^2)^n(1 + O(nr^m)).$$

Therefore for $nr^m = O(1)$,

$$(2.2) \quad a_{k,n} = \frac{1}{2\pi i} \int_{|z|=r} (1 + z^2)^n z^{-k-1} \, dz + O\left(nr^{m-k-1} \int_{|z|=r} |1 + z^2|^n \, dz\right).$$

Now the first integral above is just the coefficient of z^k in $(1 + z^2)^n$, which equals $\binom{n}{k/2}$ if k is even and 0 otherwise. On the other hand,

$$(2.3) \qquad \int_{|z|=r} |1 + z^2|^n \, dz = O(r(1 + r^2)^n) = O\left(r \exp\left(\frac{k}{2}\right)\right).$$

If $h = [k/2]$ (the greatest integer $\le k/2$), then the last term in (2.2) is

$$O(n^{(3-m)/2}h^{(m-1)/2}(ne/h)^h \exp(-9h^2/(10n))).$$

But $h! \sim (2\pi h)^{1/2}(h/e)^h$ as $h \to \infty$, so for large h,

$$\binom{n}{h} = \frac{n(n-1) \cdots (n-h+1)}{h!}$$

$$\ge (3\pi h)^{-1/2}\left(\frac{ne}{h}\right)^h \prod_{j=0}^{h-1}\left(1 - \frac{j}{n}\right) \ge (100h)^{-1/2}\left(\frac{ne}{h}\right)^h \exp\left(\frac{-2h^2}{3n}\right).$$

Therefore the quantity in (2.3) is

$$O\left(n^3 \left(\frac{h}{n}\right)^{m/2} \binom{n}{h} \exp\left(\frac{-h^2}{5n}\right)\right).$$

This is $o(\binom{n}{h})$ as $n \to \infty$ if $h = O(n^{1-2/m})$, (which guarantees $nr^m = O(1)$), and so in that range

$$a_{2h-2,n} > a_{2h-1,n}, \quad a_{2h-1,n} < a_{2h,n},$$

which shows that the sequence oscillates in that range.

We next prove Theorem 1 for k very small. Suppose that $p_0, p_1 > 0$,

$$p(z) = \sum_{j=0}^{d} p_j z^j,$$

and we are interested in the value $a_{k,n}$ of the n-fold convolution $\{p_j\}^{*n}$ at k. Then $a_{k,n}$ is again given by (2.1). This time we choose $r = p_0 p_1^{-1} k(n - k)^{-1}$. On $|z| = r$, as $n \to \infty$, $k = o(\sqrt{n})$,

$$p(z) = p_0 + p_1 z + O(k^2 n^{-2}),$$

$$p(z)^n = (p_0 + p_1 z)^n(1 + O(k^2 n^{-1})),$$

so

$$a_{k,n} = \frac{1}{2\pi i} \int_{|z|=r} (p_0 + p_1 z)^n z^{-k-1} \, dz + O\left(k^2 n^{-1} \int_{|z|=r} |p_0 + p_1 z|^n |z|^{-k-1} \, dz\right)$$

$$= \binom{n}{k} p_0^{n-k} p_1^k + O(k^{2-k} n^{k-1} p_0^{n-k} p_1^k e^k).$$

Now for $n \to \infty$, $k = o(\sqrt{n})$,

$$\binom{n}{k} = (1 + o(1)) \frac{n^k}{k!} \geq (ck)^{-1/2}\left(\frac{ne}{k}\right)^k,$$

for some constant $c > 0$, so

$$a_{k,n} = \binom{n}{k} p_0^{n-k} p_1^k (1 + O(k^{5/2} n^{-1})).$$

Hence if $k = o(n^{2/7})$, then $a_{k,n}^2 > a_{k-1,n} a_{k+1,n}$ for large n, which is the desired result. (By a more careful analysis, the range of values of k for which this inequality holds can be extended.) Note that in this part of the proof we did not use the fact that $p_{d-1} > 0$.

To conclude this section, we only have to consider the range $dn - n^{1/4} \leq k \leq dn$. However, this range corresponds to the range $0 \leq k \leq n^{1/4}$ for the n-fold convolution of the distribution $p_j^* = p_{d-j}$, $0 \leq j \leq d$, and so is covered by the preceding discussion. (Note that this part of the proof uses $p_{d-1} > 0$ but not $p_1 > 0$.)

3. Main part of the proofs of Theorems 1 and 2. In view of the preceding results, it will suffice to prove that $a_{k,n}^2 \geq a_{k-1,n} a_{k+1,n}$ holds for $n^{1/4} \leq k \leq dn - n^{1/4}$ under the conditions of Theorem 1, and for $\delta n \leq k \leq (d - \delta)n$ under the conditions of Theorem 2. From Cauchy's theorem we have

$$(3.1) \qquad a_{k,n} = \frac{1}{2\pi i} \int_{|z|=e^\alpha} p(z)^n z^{-k-1} \, dz,$$

where α is any constant. We can write this as

$$(3.2) \qquad a_{k,n} = \frac{1}{2\pi} e^{-\alpha k} \int_{-\pi}^{\pi} \exp(n \log p(e^{\alpha+i\theta}) - ik\theta) \, d\theta.$$

Eq. (3.2) now defines $a_{k,n}$ as a real function of the real variable k for any fixed value of α. (There is a mistake in Odlyzko and Richmond (1982) on this point in the proof of Theorem 2 of that paper, but it is easily corrected along the lines used in this paper.) It is immediate from the definition that as a function of k, $a_k = a_{k,n}(\alpha) \in C^\infty(-\infty, \infty)$. To prove our results it suffices to show that

$$(3.3) \qquad a_k^2 \geq a_{k-1} a_{k+1}$$

for k in the appropriate ranges. To prove (3.3) for $k = k_0$, we choose $\alpha = \alpha(k_0)$ by

$$(3.4) \qquad e^\alpha p'(e^\alpha)/p(e^\alpha) = k_0/n,$$

and, defining $a_k = a_{k,n}$ by (3.2) with α defined by (3.4), show that

$$(3.5) \qquad (\partial^2/\partial k^2)\log a_k < 0$$

for $k \in [k_0 - 1, k_0 + 1]$.

To prove (3.5) with the α given by (3.4), we define for $m = 0, 1$, and 2,

$$J_m = \int_{-\pi}^{\pi} \theta^m \exp(n \log p(e^{\alpha+i\theta}) - ik\theta) \, d\theta.$$

Note that J_0 and J_2 are real, whereas J_1 is purely imaginary. Inequality (3.5) is equivalent to

$$(3.6) \qquad J_0 J_2 > J_1^2.$$

Since J_1 is purely imaginary, (3.6) will follow if we show $J_0 > 0$, $J_2 > 0$. To prove (3.6), we estimate the J_m. We first consider $p(z)$ that satisfies the conditions of Theorem 1, i.e., $\deg p(z) = d$, $p_0, p_1, p_{d-1}, p_d > 0$. It is also sufficient to consider $n^{1/4} \leq k_0 \leq 3dn/4$, since the range $k_0 > 3dn/4$ can be treated by considering the polynomial $z^d p(1/z)$.

Define, for any $k \in [k_0 - 1, k_0 + 1]$,

$$(3.7) \qquad \theta_0 = n^{1/30} k^{-1/2}.$$

Since $k_0 \leq 3dn/4$, $\alpha \leq c$ for some constant c, and so for $\theta \in [\theta_0, 2\pi - \theta_0]$,

$$| p(e^{\alpha+i\theta}) | \leq | p_0 + p_1 e^{\alpha+i\theta} | + \sum_{j=2}^{d} p_j e^{j\alpha}.$$

But

$$| p_0 + p_1 e^{\alpha + i\theta} |^2 = p_0^2 + 2p_0 p_1 e^{\alpha} \cos \theta + p_1^2 e^{2\alpha}$$

$$= (p_0 + p_1 e^{\alpha})^2 + 2p_0 p_1 e^{\alpha}(\cos \theta - 1)$$

$$\leq (p_0 + p_1 e^{\alpha})^2 (1 - c' e^{\alpha} \theta_0^2)^2$$

for some constant $c' > 0$, and so

(3.8) $$| p(e^{\alpha + i\theta}) | \leq p(e^{\alpha}) \exp(-c'' e^{\alpha} \theta_0^2)$$

for some $c'' > 0$. Therefore if

(3.9) $$J_m^* = \int_{-\theta_0}^{\theta_0} \theta^m \exp(n \log p(e^{\alpha + i\theta}) - ik\theta) \, d\theta,$$

then

(3.10) $$J_m = J_m^* + O(p(e^{\alpha})^n \exp(-c'' n e^{\alpha} \theta_0^2)).$$

Next, we consider J_m^*. Since $\theta_0 \to 0$ as $n \to \infty$, for $| \theta | \leq \theta_0$ we have

(3.11) $$\log p(e^{\alpha + i\theta}) = \log p(e^{\alpha}) + i\theta e^{\alpha} \frac{p'(e^{\alpha})}{p(e^{\alpha})} - \theta^2 \beta + O(| \theta |^3 \gamma),$$

where

(3.12) $$\beta = \beta(k_0) = \frac{1}{2} \frac{\partial^2}{\partial x^2} \log p(e^x) \bigg|_{x=\alpha},$$

(3.13) $$\gamma = \gamma(k_0) = \max_{|\theta| \leq \theta_0} \left| \frac{\partial^3}{\partial y^3} \log p(e^{\alpha + iy}) \right| \bigg|_{y=0}.$$

Since $p_0 > 0$, $p_1 > 0$, and α is bounded above, $\beta e^{-\alpha}$, $\gamma e^{-\alpha} \in (a_1, a_2)$ for some constants a_1, a_2, $0 < a_1 < a_2 < \infty$. Therefore, by (3.4) and (3.11),

$$J_m^* p(e^{\alpha})^{-n} = \int_{-\theta_0}^{\theta_0} \theta^m \exp(-n\theta^2 \beta + O(n | \theta |^3 \gamma) + i(k_0 - k)\theta) \, d\theta.$$

By (3.4), $\gamma n/k \in (a_1', a_2')$ for some constants a_1', a_2', $0 < a_1' < a_2' < \infty$, so

$$n | \theta |^3 \gamma = O(| \theta | n^{1/15})$$

in $| \theta | \leq \theta_0$. Hence

$$J_m^* p(e^{\alpha})^{-n} = \int_{-\theta_0}^{\theta_0} \theta^m \exp(-n\beta \theta^2 + O(| \theta | n^{1/15})) \, d\theta$$

$$= \int_{-\theta_0}^{\theta_0} \theta^m \exp(-n\beta \theta^2) \, d\theta + O\left(\int_{-\theta_0}^{\theta_0} n^{1/15} | \theta |^{m+1} \exp(-n\beta \theta^2) \, d\theta \right)$$

$$= c_m (n\beta)^{-(m+1)/2} + O(n^{1/15} (n\beta)^{-(m+2)/2}),$$

where $c_1 = 0$ (since the integrand is odd), $c_0 = \pi^{1/2}$, $c_2 = \pi^{1/2}/2$. Therefore $J_0 > 0$,

$J_2 > 0$, and (3.6) holds for n sufficiently large, $n^{1/4} \leq k \leq 3dn/4$, and this completes the proof of Theorem 1.

The estimates of the J_m obtained above yield easily the estimates for the coefficients $a_{k,n}$ that were mentioned in the Introduction.

The proof of Theorem 2 is very similar, and will not be presented in detail. The major difference is that α is bounded below as well as above, and so an estimate like (3.8) can be obtained (under the assumptions of Theorem 2) even when $p_1 = 0$. To see this we first show that if (1.1) holds then for large n, for each θ satisfying $\theta_0 \leq |\theta| \leq \pi$ (θ_0 defined by (3.7) as before) there is at least one j with $p_j > 0$ such that the inequality

$$(3.14) \qquad \cos(\theta j) \leq 1 - \theta_0^2/4$$

holds. If there were no such j, then for each $j > 0$ with $p_j \neq 0$ there would be an integer $m_j \neq 0$ such that $|m_j| < j$ and

$$|\theta j - 2\pi m_j| < \theta_0.$$

But then

$$|\theta - 2\pi m_j/j| \leq \theta_0/j \leq \theta_0$$

for all $j > 0$ with $p_j \neq 0$, and so if n is sufficiently large (and $\delta n \leq k \leq (d - \delta)n$) then

$$(3.15) \qquad |m_i j - m_j i| < \tfrac{1}{2}$$

for all $i > 0$, $j > 0$ with $p_i \neq 0$, $p_j \neq 0$. But (3.15) means that

$$m_i j = m_j i$$

for all $i > 0$, $j > 0$ with $p_i \neq 0$, $p_j \neq 0$. If j_0 denotes the smallest $j > 0$ for which $p_j \neq 0$, then for each $i > 0$ with $p_i \neq 0$ we have

$$(3.16) \qquad i = m_i j_0/m_{j_0}.$$

Now $|m_j| < j$, so if D is the greatest common divisor of j_0 and m_{j_0}, then $j_0/D \neq 1$, and by (3.16), j_0/D divides all $i > 0$ with $p_i \neq 0$, which contradicts (1.1). Hence we have shown that (3.14) holds for every θ, $\theta_0 \leq |\theta| \leq \pi$, and some j with $p_j \neq 0$.

Once (3.14) is established, an estimate of the form (3.8) is easily obtained for $\theta_0 \leq |\theta| \leq \pi$. Theorem 2 then follows easily from the estimates for J_0 obtained in the proof of Theorem 1.

REFERENCES

BROCKETT, P. L. and KEMPERMAN, J. H. B. (1982). On the unimodality of high convolutions. *Ann. Probab.* **10** 270–277.

KARLIN, S.(1968). *Total Positivity*, Vol. 1. Stanford Univ Press.

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

MEDGYESSY, P. (1977). *Decomposition of Superpositions of Density Functions and Discrete Distributions.* Wiley, New York.

ODLYZKO, A. M. and RICHMOND, L. B. (1982). On the unimodality of some partition polynomials. *European J. Combin.* **3** 69–84.

USHAKOV, N. G. (1982). On a problem of Rényi. *Theory Probab. Appl.* **27** 361–362.
YAMAZATO, M. (1978). Unimodality of infinitely divisible distribution functions of class L. *Ann. Probab.* **6** 523–531.

AT&T BELL LABORATORIES   UNIVERSITY OF WATERLOO
600 MOUNTAIN AVENUE   WATERLOO, ONTARIO
MURRAY HILL, NEW JERSEY 07974   CANADA