Canad. Math. Bull. Vol. 24 (2), 1981

A PROBABILISTIC APPROACH TO THE CONVOLUTION TRANSFORM

BY
LOUIS-PAUL RIVEST*

ABSTRACT. The inversion and the characterization of the convolution transform is derived via the concept of unimodality introduced by Khintchine (1938). This method yields simple and intuitively appealing proofs.

Introduction. According to Olshen and Savage (1970), a distribution F is unimodal (with vertex at 0) if and only if F can be regarded as the distribution of UY where U is a $U[0, 1]$ random variable distributed independently of Y and Y has distribution $F(x)-xf(x)$.

In the first section this result is generalized to characterize random variables of the type $U^{1/\alpha}Y$. Taking logarithm convolutions of the form $Y+X/\alpha$ are characterized where X is an exponential random variable distributed independently of Y. In section 2 successive applications of this result yield the inversion and the characterization of the convolution transform.

1. **Preliminary results.** In this section the concept of α-unimodality introduced by Olshen and Savage (1970) is characterized:

DEFINITION. A distribution F with a well defined (except possibly at 0) density $f(x)$ is said to be α-unimodal ($\alpha \neq 0$) if $f(x)/\alpha\, |x|^{\alpha-1}$ is increasing for $x<0$ and decreasing for $x>0$.

REMARK. For $\alpha=1$, this reduces to the unimodality of Khintchine (1938).

NOTATION. Let $U_1, U_2, \ldots, U_n$ be independent $U[0, 1]$ random variables, $\{Y_i\}_{i=1}^n$ denote a random sample independent of the U_i's $\psi_\alpha(x)=\operatorname{sgn} x\, |x|^\alpha$.

THEOREM 1. *The following statements are equivalent:*

(i) *F is 1-unimodal.*

(ii) *$G(x)=F(x)-xf(x)$ is a distribution function (without loss of generality assume that f is right continuous so that G is also right continuous as should be a distribution function).*

(iii) *F can be regarded as the distribution of UY. Furthermore if Y has distribution $G(x)$, UY has distribution F.*

Received by the editors September 27, 1979 and, in revised form, February 13, 1980.
* Research supported by the NSERC grant, Number 3-656-147-30.

https://doi.org/10.4153/CMB-1981-032-5 Published online by Cambridge University Press

Proof. See section 2 of Olshen and Savage (1970) and Gnedenko and Kolomogorov (1954) p. 154.

For α-unimodality, Theorem 1 becomes:

THEOREM 2. *The following statements are equivalent:*

(i) *F is α-unimodal.*

(ii) *$G(x)=F(x)-xf(x)/\alpha$ is a distribution function.*

(iii) *F can be regarded as the distribution of $U^{1/\alpha}Y$. Furthermore if Y has distribution $G(x)$, $U^{1/\alpha}Y$ has distribution $F(x)$.*

Proof. Assume that F is α-unimodal for $\alpha>0$. By the definition of α-unimodality it is easily seen that this assumption is equivalent to $F(\psi_{1/\alpha}(x))$ is 1-unimodal. Applying Theorem 1, $F(x)$ is α-unimodal if and only if:

$$F(\psi_{1/\alpha}(x))-x\,|x|^{1/\alpha-1}f(\psi_{1/\alpha}(x))/\alpha$$

is a distribution, if $\psi_{1/\alpha}(x)=y$, the last expression is equal to $G(y)$ and (i) $\Leftrightarrow$ (ii).

Theorem 1 implies that $F(\psi_{1/\alpha}(x))$ can be viewed as the distribution of UY hence $F(x)$ can be viewed as the distribution of $U^{1/\alpha}Y^{1/\alpha}$ where $Y^{1/\alpha}$ has distribution G and the theorem is proved for $\alpha>0$.

For $\alpha<0$, assume without loss of generality that F does not give probability mass to 0 (if $F(0)-\lim_{x\uparrow 0}F(x)=\delta$ write F as $\delta I(x)+(1-\delta)F_0(x)$ and work with F_0). The proof is similar to the case α positive, it relies on the fact that F is α-unimodal ($\alpha<0$) if and only if:

$$H(x)=\begin{cases}F(0)-F(x^{-1}) & x\le 0\\ F(0)+1-F(x^{-1}) & x>0\end{cases}$$

is $(-\alpha)$-unimodal. Q.E.D.

α-unimodality can also be characterized via characteristic functions:

THEOREM 3. *A distribution function F with characteristic function ϕ is α-unimodal if and only if: $\gamma(t)=\phi(t)+r\phi'(t)/\alpha$ is a characteristic function; furthermore if $\gamma(t)$ is the characteristic function of Y, $U^{1/\alpha}\,Y$ has distribution F.*

Proof. Take α positive, according to Theorem 2 F is α-unimodal if and only if F can be seen as the distribution of $U^{1/\alpha}Y$. Suppose F is the distribution of $U^{1/\alpha}Y$ then

$$\phi(t)=\int_0^1 \nu(tx^{1/\alpha})\,dx$$

where ν is the characteristic function of Y. For $t>0$ and $tx^{1/\alpha}=w$,

$$\phi(t)=\alpha t^{-\alpha}\int_0^t \nu(w)w^{\alpha-1}\,dw$$

Differentiating both sides:

$$\phi(t)+t\phi'(t)/\alpha=\nu(t).$$

A similar result holds for $t<0$, hence $\gamma(t)$ is a characteristic function. Conversely if $\gamma(t)$ is the characteristic function

$$\phi(t)=\int_0^1 \gamma(tx^{1/\alpha})\,dx \tag{1.1}$$

and ϕ is the characteristic function of an α-unimodal random variable. (1.1) also proves the last statement of the theorem. The proof for $\alpha<0$ is derived with a similar argument. Q.E.D.

As a first application of these results, consider the problem investigated by Williamson (1956) and Levy (1962) about the characterization of:

DEFINITION. n-Motone function.

A function $f(x)$ defined on $(0,\infty)$ is said to be n-monotone if $(-1)^k f^{(k)}(x)\geq 0$ for $k=0,1,\ldots,n$ where $f^{(k)}(x)=d^k/(dx^k)f(x)$.

In a probabilistic context, Williamson and Levy's result can be stated as:

COROLLARY 1. *Let $F(x)$ be a distribution with a density $f(x)$ satisfying $u(x)=f(x)/x^{\alpha-1}$ is n-monotone* ($\alpha>0$); *then F can be seen as the distribution of*

$$\prod_{i=0}^{n-1} U_i^{1/(\alpha+i)}Y$$

where Y has density

$$h(x)=(-1)^n x^{\alpha+n-1}u^{(k)}(x)/\alpha(\alpha+1)\cdots(\alpha+k-1).$$

Proof. Since $f(x)/x^{\alpha+1}$ decreases F is α-unimodal and can be seen as the distribution of $U_0^{1/\alpha}Y_0$ where Y_0 has distribution $H_0(x)=F(x)-xf(x)/\alpha$ and density

$$h_0(x)=f(x)(1-1/\alpha)-xf'(x)/\alpha=-x^\alpha u^{(1)}(x)/\alpha.$$

Now H_0 is $(\alpha+1)$-unimodal hence F can be seen as the distribution of

$$U_0^{1/\alpha}U^{1/(\alpha+1)}Y_1$$

where Y_1 has density $h_1(x)=x^{\alpha+1}u^{(2)}(x)/\alpha(\alpha+1)$. The result is proved by induction. Q.E.D.

Note that $\prod_{i=0}^{n-1} U_i^{1/(\alpha+i)}$ is distributed Beta with parameters α and n (Rao (1973), p. 168) hence if $u(x)$ is n monotone,

$$f(x)=[B(\alpha,n)]^{-1}\int_x^\infty ((x/y)^{\alpha-1}(1-x/y)^{n-1}/yh(y))\,dy.$$

https://doi.org/10.4153/CMB-1981-032-5 Published online by Cambridge University Press

2. The convolution transform

NOTATION. Let $\{X_j\}_{j=1}^{\infty}$ be a sequence of independent exponential random variables distributed independently of $\{Y_j\}_{j=1}^{\infty}$.

Convolutions with an exponential random variable are now characterized.

THEOREM 4. *F can be viewed as the distribution of $Y+X_1/a$, $(a\neq 0)$ if and only if $F(X)+f(x)/a$ is a distribution function. Furthermore if Y has distribution $F(x)+f(x)/a$, $Y+X_1/a$ has distribution F.*

Proof. F can be viewed as the distribution of $Y+X_1/a$ $(a\neq 0)$ if and only if

$$F(x)=\int_{-\infty}^{x}(1-e^{-a(x-t)})\,dG(t)$$
$$=\int_{-\infty}^{x}\int_{-\infty}^{y} ae^{-(y-t)}\,dG(t)\,dy.$$

by Fubini's theorem. This implies $F(x)+f(x)/a=G(x)$. To prove the converse the differential equation $F(x)+f(x)/a=G(x)$ has to be solved subject to the condition $F(-\infty)=0$. Using the general solution for first order equation (see Simmons (1972) p. 48).

$$F(x)=\int_{-\infty}^{x}(1-e^{-a(x-t)})\,dG(t)$$

and Theorem 4 is true for $a>0$.

The argument is similar for $a<0$. Q.E.D.

REMARK. This result could have been proved using Theorem 2 since F is the distribution of $Y+X_1/a$ if and only if $F(\log(x))$ is $(-a)$-unimodal.

Using D as the differential operator, the distribution of Y is $(1+D/a)F$. Iterating Theorem 4 yields:

COROLLARY 2. *F is the distribution of $\sum_{j=1}^{n} X_j/a_j+Y$ if and only if $\prod_{j=1}^{n}(1+D/a_j)F(x)$ is a distribution function. If Y has distribution $\prod_{j=1}^{n}(1+D/a_j)F(x)$, $\sum_{j=1}^{n} X_j/a_j+Y$ has distribution F.*

Corollary 2 implies that there is a unique distribution which is a solution of the following differential equation:

$$\prod_{j=1}^{n}(1+D/a_j)F(x)=H(x)$$

where $H(x)$ is the distribution of Y, namely the distribution of $\sum_{j=1}^{n} X_j/a_j+Y$.

https://doi.org/10.4153/CMB-1981-032-5 Published online by Cambridge University Press

As n tends to infinity Corollary 2 becomes the inversion theorem for the convolution transform:

THEOREM 5. *Let $\{a_j\}_{j=1}^{\infty}$ be a sequence of real numbers such that $\sum a_j^{-2}<\infty$. Then F can be seen as the distribution of $Y+\sum_{j=1}^{\infty}(X_j-1)/a_j$ if and only if: $\prod_{j=1}^{n}(1+D/a_j)F(x)$ is a distribution for each n. Furthermore, in this case, the distribution of Y is equal to:*

$$\prod_{j=1}^{\infty}(1+D/a_j)e^{-D/a_j}F(x)$$

where

$$e^{-D/a}F(x)=F(x-a^{-1}).$$

Proof. According to Feller (1971) theorem (p. 266) $\sum_{j=1}^{\infty}(X_j-1)/a_j$ is a well defined random variable. Since $Y+\sum_{j=1}^{\infty}(X_j-1)/a_j$ is distributed according to $F(x)\prod_{j=1}^{n}(1+D/a_j)F(x)$ is a distribution using Corollary 2.

Conversely, suppose that $\prod_{j=1}^{n}(1+D/a_j)e^{-D/a_j}F(x)$ is the distribution of a random variable, say Y_n. In order to prove the theorem, it suffices to prove that the sequence $\{Y_n\}$ converges in distribution.

For $m>n$, $Y_m-Y_n=\sum_{j=n+1}^{m}(X_j-1)/a_j$ has 0 expectation and variance

$$\sum_{j=n+1}^{m}a_j^{-2}<\sum_{j=n+1}^{\infty}a_j^{-2}.$$

Therefore given ε there exists a constant n_0 such that $m,n>n_0$ implies $E((Y_m-Y_n)^2)<\varepsilon$. Hence $\{Y_n\}$ converges in distribution to a random variable Y with distribution

$$\prod_{j=1}^{\infty}(1+D/a_j)e^{-D/a_j}F \qquad \text{Q.E.D.}$$

REMARK. If $f(x)$ is completely monotone (i.e. n-monotone for each n), Corollary 1 and Theorem 5 imply that $F(x)=\int_0^x f(x)\,dx$ is the distribution of $\prod_{i=1}^{\infty}(eU_i)^{1/i}Y$ where $-\log Y$ has distribution $1-\prod_{i=1}^{\infty}(1+D/i)e^{-D/i}F(-\log x)$. Since $\prod_{i=1}^{\infty}(eU_i)^{1/i}e^{-\gamma}$ (γ is the Euler constant) is exponentially distributed, for $a_j=j$ Theorem 5 yields the real inversion formula for the Laplace transform: $f(x)$ is completely monotone implies that $F(x)$ is the distribution of X_1Y_1 where $\log Y_1$ is distributed according to $1-\prod_{i=1}^{\infty}(1+D/i)e^{-D/i}e^{-D\gamma}F(-\log x)$.

Dealing with characteristic functions, note that

$$\phi(t)+t\phi'(t)/a=(1+D/a)\phi(e^x)\,\big|_{x=\log t}$$

Corollary 2 becomes:

COROLLARY 3. *ϕ is the characteristic function of $Y\prod_{j=1}^{n}U_j^{1/a_j}$ if and only if*

$$\nu(t)=\prod_{j=1}^{n}(1+D/a_j)\phi(e^x)\,\big|_{x=\log t}$$

is a characteristic function. If Y has characteristic function $\nu(t)$ then $Y\prod_{j=1}^{n} U_j^{1/a_j}$has characteristic function $\phi(t)$.

Along the lines of Theorem 5, one proves:

THEOREM 6. *Let $\{a_j\}_{j=1}^{\infty}$ be an R-sequence satisfying $\sum a_j^{-2}<\infty \cdot \phi(t)$ can be viewed as the characteristic function of $Y\prod_{j=1}^{\infty}(eU_j)^{1/a_j}$ if and only if:*

$$\prod_{j=1}^{n}(1+D/a_j)\phi(e^x)\,\big|_{x=\log t}$$

is a characteristic function for each n. Furthermore if the characteristic function Y is

$$\nu(t)=\prod_{j=1}^{\infty}(1+D/a_j)e^{-D/a_j}\phi(e^x)\,\big|_{x=\log t}$$

then $\phi(t)$ is the characteristic function of $Y\prod_{j=1}^{\infty}(eU_j)^{1/a_j}$.

COMMENTS. Analytical proofs of the inversion theorem for the convolution transform are given in Chapter 3 of Hirschman and Widder (1955) and in Karlin (1968) p. 355, see also Widder (1971).

The characterization of the convolution transform of Theorem 5 is similar to the results of Chapter 7 of Hirschman and Widder (1955) while the characterization part of Theorem 6 is new.

EXAMPLE. The general birth process.

In a general birth process, let Y_n be the random variable representing the time spent at state n and let $F_n(t)$ be the distribution of $\sum_{j=0}^{n} Y_j$.

Now

$$\begin{aligned} P_n(t) &= P(X(t)=n)\\ &= P \text{ (the process is in state } n \text{ at time } t)\\ &= F_{n-1}(t)-F_n(t)n=0,1,\ldots \end{aligned}$$

where

$$F_{-1}(x)=I(x)=\begin{cases}0 & x<0\\ 1 & x\geq 0\end{cases}$$

The forward Kolmogorov equations of this process are (Feller (1958) p. 402)

$$\begin{aligned} P_n'(t) &= -\gamma_n P_n(t)+\gamma_{n-1}P_{n-1}(t) \quad n\geq 1\\ P_0'(t) &= -\gamma_0 P_0(t) \end{aligned}$$

Summing from 1 up to n and using

$$\sum_{j=0}^{n} P_j(t)=I(t)-F_n(t),$$

https://doi.org/10.4153/CMB-1981-032-5 Published online by Cambridge University Press

for $t>0$

$$-DF_n(t)=-\gamma_n P_n(t) \quad \text{or} \quad (1+D/\gamma_n)F_n(t)=F_{n-1}(t)$$

and F_n is solution of

$$\prod_{j=0}^{n}(1+D/\gamma_j)F(t)=I(t). \tag{2.1}$$

Using Corollary 2 and a subsequent remark, there is a unique solution to equation (2.1) which is the distribution of $\sum_{j=0}^{n} X_j/\gamma_j$. To obtain an algebraic expression for $F_n(x)$ when the γ_j's are distinct, first obtain a partial fraction representation of $\psi(t)=\prod_{j=0}^{n}(1-it/\gamma_j)^{-1}$, the characteristic function of F_n

$$\psi(t)=\sum_{j=0}^{n}\left(\prod_{k\neq j}(1-\gamma_j/\gamma_k)\right)^{-1}(1-it/\gamma_j)^{-1}$$

using the linearity of the inversion operator for characteristic function:

$$F_n(x)=1-\sum_{j=0}^{n} e^{-\gamma_j x}\left(\prod_{k\neq j}(1-\gamma_j/\gamma_k)\right)^{-1}$$

Now $P_n(t)=DF_n(t)/\gamma_n$ is easily obtained. This derivation is simpler than the usual one (Feller (1971) p. 489, Prabhu (1965) p. 135) which is using the Laplace transform.

References

Feller, W. (1959). *An Introduction to Probability Theory and its Applications.* Volume I. Second edition. John Wiley and Sons.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications.* Volume II. Second edition. John Wiley and Sons.

Gnedenko, B. V. and Kolmogorov, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables.* Addison-Wesley, Cambridge, Massachusetts.

Hirschman, I. I. and Widder, D. V. (1955). *The Convolution Transform.* Princeton University Press.

Karlin, S. (1968). *Total Positivity.* Stanford University Press.

Khintchine, A. Y. (1938). On Unimodal Distributions. (In Russian). *Izv Nauchno-Issled. Inst. Mat. Mekh., Tomsk. Cos. Univ.*, **2,** 1–7.

Levy, P. (1962). Extensions d'un théorème de D. Dugué et M. Girault, Z. *Wahrscheinlichkeitstheorie*, **1,** 159–173.

Olshen, R. A. and Savage, L. J. (1970). A Generalized Unimodality. *J. Appl. Prob.*, **7,** 21–34.

Prabhu, N. V. (1965). *Stochastic processes.* Macmillan, New York.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications.* Second edition. John Wiley and Sons.

Simmons, G. F. (1972). *Differential Equations.* McGraw-Hill Book Company.

Widder, D. V. (1971). *An Introduction to transform theory.* Academic Press.

Williamson, R. T. (1956). Multiply Monotone Functions and Their Laplace Transform, *Duke Math. J.*, **23,** 189–207.

Department of Statistics
University of Toronto
Toronto, Ontario
M5S 1A1