J. Appl. Prob. **47,** 293–299 (2010)
Printed in England

A CHARACTERIZATION RELATED TO THE EQUILIBRIUM DISTRIBUTION ASSOCIATED WITH A POLYNOMIAL STRUCTURE

SHAUL K. BAR-LEV,* *University of Haifa*
ONNO BOXMA,** *EURANDOM and Eindhoven University of Technology*
GÉRARD LETAC,*** *Université Paul Sabatier*

Abstract

Let f be a probability density function on $(a, b) \subset (0, \infty)$, and consider the class $\mathcal{C}_f$ of all probability density functions of the form Pf, where P is a polynomial. Assume that if X has its density in $\mathcal{C}_f$ then the equilibrium probability density $x \mapsto \mathrm{P}(X > x)/\mathrm{E}(X)$ also belongs to $\mathcal{C}_f$: this happens, for instance, when $f(x) = C\mathrm{e}^{-\lambda x}$ or $f(x) = C(b-x)^{\lambda-1}$. We show in the present paper that these two cases are the only possibilities. This surprising result is achieved with an unusual tool in renewal theory, by using ideals of polynomials.

Keywords: Renewal theory; excess lifetime; polynomial density; ideals of polynomials

2000 Mathematics Subject Classification: Primary 60K05; 13P10

1. Introduction: equilibrium distribution

Let $X_1, X_2, \ldots$ be a sequence of nonnegative independent random variables with a common distribution F, with probability density function (PDF) f and Laplace–Stieltjes transform (LST) ϕ. Letting $\mu = \mathrm{E}(X_i)$, it is assumed that $0 < \mu < \infty$. The random variable X_i denotes the interoccurrence time between the $(i-1)$th and ith events in some probability problem. The counting process $\{N(t),\ t \geq 0\}$, where $N(t) = \max\{n \geq 0 \colon X_1 + \cdots + X_n \leq t\}$, is called the renewal process generated by the interoccurrence times $X_1, X_2, \ldots$ (cf. the classical textbooks [3], [5], and [6]). An important role in renewal theory is played by the backward recurrence time A_t (the time since the last renewal before t) and the forward recurrence time B_t (the time until the first renewal after t). If the X_i are interpreted as lifetimes then A_t is the past lifetime at t and B_t is the residual or excess lifetime at t. It is well known that the limiting distributions of A_t and B_t for $t \to \infty$ are given by (with X a generic random variable with distribution F)

$$\lim_{t\to\infty} \mathrm{P}(A_t \leq x) = \lim_{t\to\infty} \mathrm{P}(B_t \leq x) = \int_{y=0}^{x} \frac{\mathrm{P}(X > y)}{\mu}\,\mathrm{d}y. \tag{1.1}$$

Denote this limiting or equilibrium excess lifetime distribution by F_{e}, and its PDF by $f_{\mathrm{e}}(x) = \mathrm{P}(X > x)/\mu = \int_x^\infty (f(y)/\mu)\,\mathrm{d}y$. Its LST is given by $\varphi_{\mathrm{e}}(s) = (1 - \varphi(s))/s\mu$. Excess lifetimes play an extremely important role in applied probability. They arise in a host of real-life problems,

Received 20 July 2009; revision received 2 November 2009.
* Postal address: Department of Statistics, University of Haifa, Haifa 31905, Israel.
Email address: barlev@stat.haifa.ac.il
** Postal address: Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands. Email address: boxma@win.tue.nl
*** Postal address: Laboratoire de Statistique et Probabilité, Université Paul Sabatier, 31062 Toulouse, France.
Email address: gerard.letac@alsatis.net

ranging from reliability theory to inventory and queueing theory. In many queueing problems we need to know the time until completion of the ongoing service (residual service time); e.g. the residual service time plays a key role in the celebrated Pollaczek–Khintchine formula for the steady-state waiting time distribution in the M/G/1 first-come–first-served queue (see [1, Chapter VIII]). In reliability and maintenance problems, we need to know the time until breakdown of a machine, or until an ongoing repair is completed, etc. We refer the reader to Chapter 1 of [6] for a host of other examples, which confirm the importance of obtaining insight into the characteristics of the distribution of the residual lifetime.

A related important random variable is $X_{N(t)+1}$, the length of the renewal interval seen by an outside observer at t. Denote by $\hat{X}$ a random variable with distribution the limiting distribution of $X_{N(t)+1}$. Its steady-state PDF is $yf(y)/\mu$ and $\mathrm{P}(\hat{X} > x) = \int_x^\infty (yf(y)/\mu)\,\mathrm{d}y$.

In a recent report [2], it is shown that the class of distributions on the positive reals with a rational LST, also known as matrix-exponential distributions, is closed under formation of moment distributions (distributions with density $y^i f(y)/\int_0^\infty x^i f(x)\,\mathrm{d}x$). In Section 2 we also observe that, for the classes of exponential, Erlang, and hyperexponential distributions, the PDF f_{e}, and also the PDF of $\hat{X}$, are again exponential, Erlang, hyperexponential, or mixtures of those. The beta distribution has a similar closure property. This has led us to study a much more general question: which PDFs have the property that, for any polynomial P, $\int_x^b P(t)f(t)\,\mathrm{d}t$ can be written in the form of a product of another polynomial and $f(x)$? This question is answered in our main result, Proposition 3.1 in Section 3. But first, in Section 2, we provide several examples where we demonstrate the property of Proposition 3.1. In considering these examples, it should be realized that $P(t)f(t)$ is also a PDF, up to a multiplicative constant.

2. Examples

In this section we consider two examples. One is related to the exponential distribution, the other to the beta distribution.

Example 2.1. (i) If $X \sim \exp(\lambda)$, i.e. $\varphi(s) = \lambda/(\lambda + s)$, then $\varphi_{\mathrm{e}}(s) = \varphi(s)$, and $f_{\mathrm{e}}(x) = f(x)$: the residual lifetime is again exponential. Of course, this is the familiar memoryless property.

(ii) If F is a hyperexponential distribution, i.e. a mixture of exponential distributions, then F_{e} is also hyperexponential. If F is Erlang(n) then F_{e} is a mixture of Erlang(i) with weights $1/n$. If F is a mixture of Erlang(i) with weights p_i, $i = 1, \dots, n$, then F_{e} is also a mixture of Erlang(i), $i = 1, \dots, n$, with different weights

$$p_i^* = \sum_{j=i}^n \frac{p_j}{\sum_{k=1}^n kp_k}, \qquad i = 1, \dots, n.$$

In the above example F_{e} is either a mixture of exponential distributions or a mixture of convolutions of exponential distributions; or, equivalently, the related PDF f_{e} has the form

$$f_{\mathrm{e}}(x) = \sum_{i=1}^n P_i(x)\mathrm{e}^{-\lambda_i x},$$

where $n \in \mathbb{N}$, $P_i(x)$ is a polynomial in x, and $\lambda_i > 0$. A similar statement holds for

$$\mathrm{P}(\hat{X} > x) = \frac{1}{\mu}\int_x^\infty tf(t)\,\mathrm{d}t.$$

https://doi.org/10.1239/jap/1269610832 Published online by Cambridge University Press

It should be further noted that in Example 2.1(i) and (ii), we have a PDF of the form $P(t)f(t)$ with P a polynomial and f an exponential; furthermore, $\int_x^b P(t)f(t)\,\mathrm{d}t$ has the form of the product of another polynomial and f.

Example 2.2. Now consider the beta PDF

$$f(x) = \left(\frac{x-a}{b-a}\right)^{\zeta-1}\left(\frac{b-x}{b-a}\right)^{\lambda-1}\frac{1}{(b-a)B(\zeta,\lambda)},$$

where $B(\zeta,\lambda) = \int_0^1 x^{\zeta-1}(1-x)^{\lambda-1}\,\mathrm{d}x$ is the beta function. If $\zeta = 1$ then

$$f_{\mathrm{e}}(x) = \frac{(b-x)^{\lambda}}{(b-a)^{\lambda}},$$

which is again a (special) beta PDF with $\zeta = 1$. We see here a similar closure property as in the previous example. We could also have taken a weighted sum of special beta PDFs multiplied by polynomials, and it is easily seen that taking the integration $\int_x^b$ with respect to such a sum results in other polynomials multiplied by special beta PDFs.

This raises the following question. For which PDFs f (or, equivalently, LSTs φ) is the equilibrium PDF in (1.1) a PDF in the same 'class' of PDFs as f, or a polynomial multiplied with f? In the next section we introduce such a closure property in a more general setting, and we prove a characterization result. If f is concentrated on $0 \le a < b \le \infty$ then which PDFs f have the property that, for any polynomial P, $\int_x^b P(t)f(t)\,\mathrm{d}t$ can be written in the form of a product of another polynomial and $f(x)$? We show in Proposition 3.1, below, that a necessary and sufficient condition for this to hold is that either $b = \infty$ and $f(x) = C\mathrm{e}^{-\lambda x}$, where $\lambda > 0$ and $1/C = \int_a^b (t-a)f(t)\,\mathrm{d}t$, or b is finite and $f(x) = C(b-x)^{\lambda-1}$, i.e. f is either exponential or of beta type.

3. The main result

Let f be a PDF on (a,b) with $0 \le a < b \le \infty$ such that $1/C = \int_a^b (t-a)f(t)\,\mathrm{d}t < \infty$. Consider the new PDF on (a,b) defined by $T(f)(x) = C\int_x^b f(t)\,\mathrm{d}t$. Note that, up to a multiplicative constant, this is $f_{\mathrm{e}}(x)$. For instance, if $(a,b) = (0,\infty)$, consider the class $\mathcal{F}$ of the PDFs of the form

$$f(x) = \sum_{i=1}^{n} P_i(x)\mathrm{e}^{-\lambda_i x},$$

where $P_i(x)$ is a polynomial and $\lambda_i > 0$. Because of the formula

$$\int_x^{\infty} \lambda^n \frac{t^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda t}\,\mathrm{d}t = \sum_{k=0}^{n-1}\lambda^k\frac{x^k}{k!}\mathrm{e}^{-\lambda x}, \tag{3.1}$$

clearly $T(f)$ is also in $\mathcal{F}$. A similar situation occurs when considering a bounded interval (a,b) and the class $\mathcal{G}$ of PDFs on (a,b) which are polynomials P multiplied by the function $f(x) = (b-x)^{\lambda-1}$, where $\lambda > 0$. Here, $\mathcal{G}$ is stable by T, meaning that $T(\mathcal{G}) \subset \mathcal{G}$ (write $P(x)f(x)$ in the form $\sum_{k=0}^{n} p_k(b-x)^{k+\lambda-1}$ to be convinced of this fact). Of course, choosing a class $\mathcal{C}$ of PDFs on (a,b) having all their moments implies that the class of PDFs defined by

$$\mathcal{C}_1 = \bigcup_{n=0}^{\infty} T^n(\mathcal{C})$$

is stable by T. But we are going to show that the classes $\mathcal{F}$ and $\mathcal{G}$ above are unique in the following sense.

Proposition 3.1. *Let f be a positive measurable function on (a, b) with $0 \leq a < b \leq \infty$ such that $\int_a^b t^n f(t)\,\mathrm{d}t < \infty$ for any nonnegative integer n. Suppose that, for any polynomial P, there exists a polynomial $A(P)$ such that, for all $x \in (a, b)$, we have*

$$\int_x^b P(t) f(t) \mathrm{d}t = A(P)(x) f(x). \tag{3.2}$$

Then there exist $C, \lambda > 0$ such that either b is infinite and $f(x) = C\mathrm{e}^{-\lambda x}$, or b is finite and $f(x) = C(b - x)^{\lambda - 1}$.

Remarks. The statement of Proposition 3.1 describes the few functions f on (a, b) such that the class $\mathcal{C}_f$ of PDFs of the form $P(x) f(x)$ is stable by the operation T described above, with $T(Pf) = A(P)f$. Note that in both cases a is not necessarily 0. For instance, if $f(x) = \mathrm{e}^{-\lambda x}$ on $(a, b) = (a, \infty)$ and $P(x) = \lambda^n x^{n-1}/(n-1)!$, we have (cf. (3.1))

$$A(P)(x) = \sum_{k=0}^{n-1} \frac{\lambda^k x^k}{k!}. \tag{3.3}$$

Note that $A(1) = 1/\lambda$. Since A is a linear operator, these formulae describe A completely. Similarly, if f is $(b - x)^{\lambda - 1}$ on the bounded interval (a, b) and if $P(x) = (b - x)^n$, we have

$$A(P)(x) = \frac{(b - x)^{n+1}}{n + \lambda}. \tag{3.4}$$

For instance, $A(1) = (b - x)/\lambda$.

Let us also insist on the fact that the proposition describes the only two possibilities. We could be tempted if f satisfies (3.2) to coin the new function $f_1(x) = R(x) f(x)$, where R is a nonconstant polynomial which is positive on (a, b), and to observe that, for all polynomials P, we have

$$\int_x^b P(t) f_1(t)\,\mathrm{d}t = \frac{A(PR)(x)}{R(x)} f_1(x).$$

A consequence of the proposition is that it is impossible that R divides $A(PR)$ for all polynomials P.

3.1. Proof of Proposition 3.1

For $P \equiv 1$, we define $Q(x) = A(1)(x)$. Writing $\int_x^b f(t)\,\mathrm{d}t = Q(x) f(x)$ shows that the polynomial Q must be positive on (a, b). Since f is integrable, writing

$$f(x) = \frac{1}{Q(x)} \int_x^b f(t)\,\mathrm{d}t$$

shows that f must be continuous and differentiable, and, thus, infinitely differentiable. Now taking the derivative in x of $\int_x^b P(t) f(t)\,\mathrm{d}t = A(P)(x) f(x)$ gives the differential equation

$$-P(x) f(x) = A(P)'(x) f(x) + A(P)(x) f'(x),$$

which we rewrite as

$$\frac{f'(x)}{f(x)} = -\frac{P(x) + A(P)'(x)}{A(P)(x)}.$$

Note that, since the left-hand side of this equation does not depend on P, we can get information on $A(P)$ by replacing P with 1, giving the following differential equation in $A(P)$:

$$\frac{P(x) + A(P)'(x)}{A(P)(x)} = \frac{1 + Q'(x)}{Q(x)}.$$

As a consequence, all information on f and $A(P)$ is actually given by the polynomial Q.

Case 1: Q of degree 0. If Q is the nonzero constant $1/\lambda$, the equation $f'/f = -(1+Q')/Q$ gives $f(x) = \mathrm{e}^{-\lambda x}$ on (a, b). If $b = \infty$, we have already seen that, if $\lambda > 0$, the identity $\int_x^b P(t) f(t)\,\mathrm{d}t = A(P)(x) f(x)$ holds for a suitable operator A defined by (3.3). If $\lambda \leq 0$, the condition $\int_a^b t^n f(t)\,\mathrm{d}t < \infty$ is not fulfilled. If $b < \infty$ then $\int_x^b P(t) f(t)\,\mathrm{d}t = A(P)(x) f(x)$ does not hold since, for $P = \lambda$, we obtain

$$\int_x^b \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t = \mathrm{e}^{-\lambda x} - \mathrm{e}^{-\lambda b},$$

which is not of the desired form of a polynomial multiplied by $\mathrm{e}^{-\lambda x}$.

Case 2: Q of degree 1. If Q is a first-degree polynomial, we write it as $Q(x) = (b_1 - x)/\lambda$, where b_1 is a real number and λ is a nonzero number. From the equation $f'/f = -(1+Q')/Q$ on (a, b) and the fact that $(\mathrm{d}/\mathrm{d}t)\log|t| = 1/t$, we find that $f(x) = C|b_1 - x|^{\lambda - 1}$ for some positive number C. Suppose that $b = \infty$. Clearly, $\int_a^\infty t^n f(t)\,\mathrm{d}t < \infty$ is impossible if n is large enough. Thus, $b < \infty$. Now, for all x in (a, b), we have

$$\int_x^b |b_1 - t|^{\lambda - 1}\,\mathrm{d}t = |b_1 - x|^{\lambda - 1}\frac{b_1 - x}{\lambda} = \frac{|b_1 - x|^{\lambda}}{|\lambda|}.$$

Since the left-hand side must converge to 0 when $x \to b$, this would imply that $b = b_1$ and that $\lambda > 0$.

Case 3: Q of degree greater than or equal to 2. We now claim that Q has necessarily degree less than or equal to 1, a more difficult part of the proof. Suppose that Q has degree greater than or equal to 2, and suppose that the differential equation $QP = (1 + Q')Y - QY'$ always has a polynomial solution $Y = A(P)$ for any polynomial P.

To reach a contradiction, we introduce the following notation. We denote by $\mathcal{A}$ the algebra of polynomials with real coefficients. If $A \in \mathcal{A}$, we denote by $\mathcal{I}_A$ the ideal generated by A, that is, the set of polynomials divisible by A:

$$\mathcal{I}_A = \{AP;\ P \in \mathcal{A}\}.$$

Recall that in general an ideal of $\mathcal{A}$ is a linear subspace $\mathcal{I}$ of $\mathcal{A}$ such that PB is in $\mathcal{I}$ for any $B \in \mathcal{I}$ and any $P \in \mathcal{A}$. The important result is that in this algebra $\mathcal{A}$ of polynomials, for any ideal $\mathcal{I}$, there exists an $A \in \mathcal{A}$ such that $\mathcal{I} = \mathcal{I}_A$: this is called the *principal ideal* property; see, e.g. [4, p. 105].

Finally, we introduce the notation for the linear application φ of $\mathcal{A}$ into itself, defined by

$$Y \mapsto \varphi(Y) = (1 + Q')Y - QY'.$$

Assuming that $QP = \varphi(Y)$ has a solution Y in $\mathcal{A}$ for each $P \in \mathcal{A}$ is equivalent to saying that the image $\varphi(\mathcal{A})$ of φ contains the ideal $\mathcal{I}_Q$.

Lemma 3.1. *Let B_0 and C_0 be two polynomials, and consider the linear application φ_0 of $\mathcal{A}$ into itself, defined by*

$$\varphi_0(Y) = B_0 Y - C_0 Y'.$$

We assume that $\varphi_0(\mathcal{A}) \supset \mathcal{I}_{C_0}$. Then there exist $A_1, B_1, C_1 \in \mathcal{A}$ such that $\varphi_0(\mathcal{A}) = \mathcal{I}_{A_1}$, $B_0 = A_1 B_1$, and $C_0 = A_1 C_1$. Furthermore, if $\varphi_1(Y) = B_1 Y - C_1 Y'$, we have $\varphi_1(\mathcal{A}) \supset \mathcal{I}_{C_1}$.

Proof. We show that $\varphi_0(\mathcal{A})$ is an ideal. Since φ_0 is linear, the set $\varphi_0(\mathcal{A})$ is a linear subspace of $\mathcal{A}$. Thus, we want to show that if $B = \varphi(Y_0)$ and P are arbitrary elements of $\varphi_0(\mathcal{A})$ and $\mathcal{A}$, respectively, then the polynomial $P\varphi_0(Y_0)$ is in $\varphi_0(\mathcal{A})$. Since $\varphi_0(\mathcal{A}) \supset \mathcal{I}_{C_0}$, there exists $Y_1 \in \mathcal{A}$ such that $\varphi_0(Y_1) = C_0 P' Y_0$. Thus,

$$\begin{aligned}
\varphi_0(PY_0 + Y_1) &= \varphi_0(PY_0) + \varphi_0(Y_1) \\
&= B_0 P Y_0 - C_0 P Y_0' - C_0 P' Y_0 + \varphi_0(Y_1) \\
&= B_0 P Y_0 - C_0 P Y_0' \\
&= P\varphi_0(Y_0).
\end{aligned}$$

Equality $P\varphi_0(Y_0) = \varphi_0(PY_0 + Y_1)$ shows that $\varphi_0(\mathcal{A})$ is an ideal of $\mathcal{A}$. From the principal ideal property, there exists A_1 such that $\varphi_0(\mathcal{A}) = \mathcal{I}_{A_1}$. Since $\mathcal{I}_{A_1} \supset \mathcal{I}_{C_0}$, A_1 divides C_0. Thus, there exists C_1 such that $C_0 = A_1 C_1$. Since $\varphi_0(Y) = B_0 Y - C_0 Y' = B_0 Y - A_1 C_1 Y'$ is a multiple of A_1 for any Y, then $B_0 = \varphi_0(1) = A_1 B_1$ is also a multiple of A_1. Finally, since, for each P, there exists Y such that

$$\varphi_0(Y) = B_0 Y - C_0 Y' = C_0 P = A_1 B_1 Y - A_1 C_1 Y' = A_1 C_1 P,$$

this implies that the same Y satisfies $\varphi_1(Y) = B_1 Y - C_1 Y' = C_1 P$, showing that $\varphi_1(\mathcal{A}) \supset \mathcal{I}_{C_1}$. This completes the proof of Lemma 3.1.

We now iterate Lemma 3.1. For each $n = 1, 2, \ldots$, there exist A_n, B_n, and C_n such that

$$B_0 = A_1 A_2 \cdots A_n B_n, \qquad C_0 = A_1 A_2 \cdots A_n C_n,$$

and such that if we define $\varphi_n(Y) = B_n Y - C_n Y'$, we have $\varphi_n(\mathcal{A}) = \mathcal{I}_{A_{n+1}} \supset \mathcal{I}_{C_n}$. In particular, $\sum_{k=1}^{n} \deg A_k \le \deg C_0$ implies that, for large enough n, $\deg A_{n+1} = 0$. As a consequence, A_{n+1} must be a constant polynomial, which is equivalent to saying that φ_n is *surjective.*

We now apply the above considerations to the particular case where $B_0 = 1 + Q'$ and $C_0 = Q$, where Q is a polynomial of degree $d_0 \ge 2$. Thus, $\varphi = \varphi_0$ in the lemma. With this choice of (B_0, C_0), we show that whatever n is, the map φ_n cannot be surjective when the degree of Q is greater than or equal to 2. Write

$$Q(x) = C_0(x) = c_0 x^{d_0} + \text{lower-degree terms},$$

and, more generally,

$$C_n(x) = c_n x^{d_n} + \text{lower-degree terms}, \qquad B_n(x) = b_n x^{d_n - 1} + \text{lower-degree terms}.$$

We show by induction on n that $b_n = d_0 c_n$. This is obvious for $n = 0$ since $B_0 = 1 + Q'$ and $d_0 \ge 2$. Suppose that it is true for $n - 1$. Since $B_{n-1} = A_n B_n$ and $C_{n-1} = A_n C_n$, and if the term of maximum degree of A_n is $a_n x^m$, then $d_{n-1} = d_n + m$, $b_{n-1} = a_n b_n$, and $c_{n-1} = a_n c_n$. Since, by definition, $a_n \ne 0$, the equality $b_n = d_0 c_n$ holds.

We finally use this fact to prove that φ_n cannot be surjective, by showing that there is no Y such that

$$\varphi_n(Y)(x) = B_n(x)Y(x) - C_n(x)Y'(x) = x^{d_0+d_n-1}$$

holds. Suppose that there exists such a Y, with highest degree term αx^m. The highest degree term of $B_nY - C_nY'$ is $(d_0 - m)\alpha c_n x^{d_n+m-1}$ if $m \neq d_0$, which cannot be equal to $x^{d_0+d_n-1}$. If $m = d_0$, the highest degree term of $B_nY - C_nY'$ has degree less than $d_0 + d_n - 1$. We obtain the desired contradiction. This completes the proof of Proposition 3.1.

Remark. As observed by the referee about the statement of Proposition 3.1, positivity of f can be replaced by a slightly weaker hypothesis: (i) nonnegativity of f; (ii) if $G(x_0)$ is the Lebesgue measure of the set $\{x_0 \leq x < b;\ f(x) > 0\}$ then $G(x_0) > 0$ for all $x_0 \in (a, b)$. The equality $\int_x^b f(t)\,\mathrm{d}t = Q(x)f(x)$ shows that this weaker hypothesis implies the positivity of f.

References

[1] Asmussen, S. (2003). *Applied Probability and Queues*, 2nd. edn. Springer, New York.
[2] Bladt, M. and Nielsen, B. F. (2009). Moment distributions of phase type. Submitted.
[3] Cox, D. R. (1962). *Renewal Theory*. Methuen, London.
[4] Herstein, I. N. (1964). *Topics in Algebra*. Blaisdell Publishing, New York.
[5] Ross, S. M. (1996). *Stochastic Processes*, 2nd edn. John Wiley, New York.
[6] Tijms, H. C. (1994). *Stochastic Models. An Algorithmic Approach.* John Wiley, Chichester.