The Annals of Mathematical Statistics
1969, Vol. 40, No. 6, 2183–2188

SUFFICIENT CONDITIONS FOR A MIXTURE OF EXPONENTIALS TO BE A PROBABILITY DENSITY FUNCTION

BY D. J. BARTHOLOMEW

University of Kent at Canterbury

1. Introduction. We shall consider the function

$$f(x) = \sum_{i=1}^{k} p_i \lambda_i e^{-\lambda_i x} \qquad (x \geqq 0) \tag{1}$$

where the λ's are positive and $\sum_{i=1}^{k} p_i = 1$. Without loss of generality we may suppose that $\lambda_1 < \lambda_2 < \cdots < \lambda_k$. If all of the p's are positive then it is obvious that (1) represents a probability density function. If some of the p's are negative, $f(x)$ could be negative for some values of x and so may not be a density function. Steutel, in [3], remarked that there appear to be no simple conditions for determining whether or not $f(x)$ is a density. It is the main purpose of this note to provide some simple sufficient conditions. These all have the form of inequalities involving linear functions of the p's; the principal results are given in Theorems 1 and 2 and their Corollaries.

Mixed exponential distributions with negative p's arise in several statistical contexts. One of the best known members of the family is the so-called Erlang distribution. This plays a central role in the derivation of our conditions. It arises as follows. Let $y_1, y_2, \cdots y_k$ be independently and exponentially distributed random variables with scale parameters $\lambda_1, \lambda_2, \cdots \lambda_k$ respectively, then $x = \sum_{i=1}^{k} y_i$ has an Erlang distribution with density given by (1) with

$$p_i = \prod_{j=1, j\neq i}^{k} (\lambda_j/(\lambda_j - \lambda_i)), \qquad (i = 1, 2, \cdots k).$$

(see [1], page 17). In this case the signs of the p's alternate.

The mixed exponential distribution has many attractive properties. The fact that its Laplace transform is a rational algebraic fraction offers many advantages in renewal theory and other branches of stochastic processes. Kingman in [2] has shown that one can approximate arbitrarily closely to any density on $(0, \infty)$ by a function of the form (1), (although this might require a very large value of k in any particular instance). The many advantages which this function offers are off-set to some extent by the difficulty of determining whether a given $f(x)$ is in fact a density function. Our results provide a partial answer to this problem.

A further property, which has a direct application to testing for positivity, is that a mixed exponential distribution, truncated on the left, is also mixed exponential in form. Thus if the point of truncation is $x = X$ and if $y = x - X$ then

$$f(y \mid y \geqq 0) = \sum_{i=1}^{k} p_i' \lambda_i e^{-\lambda_i y} \tag{2}$$

where $p_i' = p_i e^{-\lambda_i X} / \sum_{i=1}^{k} p_i e^{-\lambda_i X}$.

2. Two necessary conditions. By considering the two cases $x = 0$ and $x \to \infty$,

Received 15 April 1968; revised 16 December 1968.

Steutel showed in [3] that the following two conditions are necessary for $f(x)$ to be a density.

(a) $\sum_{i=1}^{k} p_i\lambda_i \geqq 0$, and (b) $p_1 > 0$.

A slightly more general form of (a) is

(a′) $f^{(r)}(0) \geqq 0$ where $f^{(r)}(0)$ is the first non-zero derivative of $f(x)$ at $x = 0$ and $f^{(0)}(0) \equiv f(0) = \sum_{i=1}^{k} p_i\lambda_i$. Except for the case $k = 2$, conditions (a′) and (b) are not sufficient for $f(x)$ to be a density function.

3. Some sufficient conditions. Summation of (1) by parts gives the alternative expression

$$f(x) = e^{-\lambda_k x} \sum_{i=1}^{k} p_i\lambda_i + \sum_{r=1}^{k-1} (e^{-\lambda_r x} - e^{-\lambda_{r+1} x}) \sum_{i=1}^{r} p_i\lambda_i . \tag{3}$$

Since $\lambda_1 < \lambda_2 < \cdots < \lambda_k$ it follows that

$$e^{-\lambda_r x} - e^{-\lambda_{r+1} x} \geqq 0 \qquad (r = 1, 2, \cdots k - 1).$$

Hence every term in (3) is positive for all x if

$$\sum_{i=1}^{r} p_i\lambda_i \geqq 0, \qquad (r = 1, 2, \cdots k). \tag{4}$$

This ensures that $f(x)$ is everywhere positive; we state the result as:

THEOREM 1. *The following conditions are sufficient to ensure that* $f(x)$ *in* (1) *is a probability density function*:

$$\sum_{i=1}^{r} p_i\lambda_i \geqq 0, \qquad (r = 1, 2, \cdots k).$$

The conditions in the theorem are extremely simple but they are not the best that can be obtained. To demonstrate this we proceed as follows.

Let

$$p_r^{(2)} = (\lambda_{r+1} - \lambda_r) \sum_{i=1}^{r} p_i\lambda_i/\lambda_r^{(2)}, \tag{5}$$

where $\lambda_r^{(2)} = \lambda_r\lambda_{r+1} (r = 1, 2, \cdots k - 1)$, $\quad p_k^{(2)} = \sum_{i=1}^{k} p_i\lambda_i/\lambda_k$,

and

$$f_r^{(2)}(x) = \lambda_r^{(2)} (\lambda_{r+1} - \lambda_r)^{-1} (e^{-\lambda_r x} - e^{-\lambda_{r+1} x}).$$

The function $f_r^{(2)}(x)$ is the Erlang density function of the sum of two exponential variates with parameters λ_r and λ_{r+1}. Hence it is necessarily non-negative for all x. In this notation (3) becomes

$$f(x) = p_k^{(2)} f_k^{(1)}(x) + \sum_{r=1}^{k-1} p_r^{(2)} f_r^{(2)}(x) \tag{6}$$

where $f_k^{(1)}(x) = \lambda_k e^{-\lambda_k x}$.

We have thus expressed (3) as a mixture of one exponential function and $k - 1$ Erlang densities of the second order. If the conditions of Theorem 1 are satisfied, all of the coefficients in (6) are non-negative. If some of the partial sums, $\sum_{i=1}^{r} p_i\lambda_i$, are negative the corresponding $p_r^{(2)}$'s will also be negative. We may then hope to treat the second sum on the right hand side of (6) in exactly the same way as we dealt with (1) since it too is a mixture of densities.

Let us write

$$f_r^{(2)}(x)/\lambda_r^{(2)} = f_{r+1}^{(2)}(x)/\lambda_{r+1}^{(2)} + (\lambda_{r+2} - \lambda_r)f_r^{(3)}(x)/\lambda_r^{(3)} \tag{7}$$

where $\lambda_r^{(3)} = \lambda_{r+3}\lambda_r^{(2)}$, $(r = 1, 2, \cdots k - 2)$, and $f_r^{(3)}(x)$ is the Erlang density for independent exponential variates with parameters λ_r, λ_{r+1}, λ_{r+2}. A further summation by parts applied to (6) now gives

$$f(x) = p_k^{(2)}f^{(1)}(x) + p_{k-1}^{(3)}f_{k-1}^{(2)}(x) + \sum_{r=1}^{k-2} p_r^{(3)}f_r^{(3)}(x) \tag{8}$$

where $\lambda_r^{(3)}p_r^{(3)} = (\lambda_{r+2} - \lambda_r)\sum_{i=1}^{r} p_i^{(2)}\lambda_i^{(2)}$, $(r = 1, 2, \cdots k - 2)$, and $p_{k-1}^{(3)} = \sum_{i=1}^{k-1} p_i^{(2)}\lambda_i^{(2)}$. Since $f_r^{(3)}(x)$ is positive for all r and x another set of conditions for $f(x)$ to be a density is thus

$$p_k^{(2)} \geqq 0, \qquad \sum_{i=1}^{r} p_i^{(2)}\lambda_i^{(2)} \geqq 0, \quad (r = 1, 2, \cdots k - 1). \tag{9}$$

If the conditions of Theorem 1 are satisfied then those of (9) are obviously satisfied also. However, the converse is not necessarily true. The conditions (9) are thus better than those of Theorem 1.

The same method can now be applied repeatedly. At each stage the Erlang densities in the mixture are expressed in terms of those of higher order. Thus for any $j \leqq k - r$ we have

$$f_r^{(j)}(x)/\lambda_r^{(j)} = f_{r+1}^{(j)}(x)/\lambda_r^{(j)} + (\lambda_{r+j} - \lambda_r)f_r^{(j+1)}(x)/\lambda_r^{(j+1)}$$

where $f_r^{(j+1)}(x)$ is the $(j + 1)$th order Erlang density with parameters λ_r, $\lambda_{r+1}, \cdots \lambda_{r+j}$ and where $\lambda_r^{(j+1)} = \prod_{i=0}^{j} \lambda_{r+i}$. After j repetitions of the procedure we have

$$f(x) = p_k^{(2)}f_k^{(1)}(x) + p_{k-1}^{(3)}f_{k-1}^{(2)}(x) + \cdots + p_{k-j+1}^{(j+1)}f_{k-j+1}^{(j)}(x) + \sum_{i=1}^{k-j} p_i^{(j+1)}f_i^{(j+1)}(x) \tag{10}$$

where

$$\begin{aligned} \lambda_r^{(h)}p_r^{(h)} &= (\lambda_{r+h-1} - \lambda_r)\sum_{i=1}^{r} p_i^{(h-1)}\lambda_i^{(h-1)}, \quad (r + h - 1 < k) \\ &= \sum_{i=1}^{r} p_i^{(h-1)}\lambda_i^{(h-1)} \qquad (r + h - 1 = k). \end{aligned} \tag{11}$$

Sufficient conditions for $f(x)$ to be a density are thus that

$$\begin{aligned} &p_k^{(2)} \geqq 0, \quad p_{k-r}^{(r+2)} \geqq 0 \qquad (r = 1, 2, \cdots j - 1) \\ &p_{k-r}^{(j+1)} \geqq 0, \qquad (r = j, j + 1, \cdots k - 1). \end{aligned} \tag{12}$$

If these conditions hold for $j = m$ they will certainly hold for $j > m$ so the best result obtainable by this method is found by putting $j = k - 1$. We then have:

THEOREM 2. *The following conditions are sufficient to ensure that $f(x)$ in (1) is a probability density function.*

$$p_k^{(2)} \geqq 0, \qquad p_{k-r}^{(r+2)} \geqq 0, \ (r = 1, 2, \cdots k - 2), \qquad p_1^{(k)} \geqq 0.$$

The conditions of Theorem 2 can be expressed in various forms. Two such, which facilitate calculation and give some insight into the nature of the conditions, are given in the following corollaries.

COROLLARY 1. *The conditions of Theorem 2 are equivalent to the following*

$$\sum_{i=1}^{k} p_i\lambda_i \geqq 0, \qquad \sum_{i=1}^{k-r} p_i\lambda_i \prod_{m=k-r+1}^{k} (\lambda_m - \lambda_i) \geqq 0 \quad (r = 1, 2, \cdots k - 1).$$

PROOF. It is clear from the definition of $p_k^{(2)}$ in (5) that $p_k^{(2)} \geqq 0$ implies $\sum_{i=1}^{k} p_i\lambda_i \geqq 0$ and conversely.
In order to establish the equivalence of the remaining conditions we proceed as follows. We have expanded $f(x)$ in the form

$$f(x) = \sum_{j=1}^{k-1} p_{k-j+1}^{(j+1)} f_{k-j+1}^{(j)}(x) + p_1^{(k)} f_1^{(k)}(x). \tag{13}$$

Now

$$f_{k-j+1}^{(j)}(x) = \sum_{h=k-j+1}^{k} \lambda_h e^{-\lambda_h x} \prod_{m=k-j+1, m\neq h}^{k} \{\lambda_m/(\lambda_m - \lambda_h)\}.$$

Hence, substituting this expression in (13) and reversing the order of summation we have

$$f(x) = \sum_{i=1}^{k-1} \lambda_i e^{-\lambda_i x} \sum_{j=k-i+1}^{k} p_{k-j+1}^{(j+1)} \prod_{m=k-j+1, m\neq i}^{k} \{\lambda_m/(\lambda_m - \lambda_i)\} \tag{14}$$

where $p_1^{(k+1)} \equiv p_1^{(k)}$. Comparing this with (1) we see that

$$\sum_{j=k-i+1}^{k} p_{k-j+1}^{(j+1)} \prod_{m=k-j+1, m\neq i}^{k} \{\lambda_m/(\lambda_m - \lambda_i)\} = p_i \quad (i = 1, 2, \cdots k). \tag{15}$$

On solving this system of equations for the $p_{k-j+1}^{(j+1)}$ we find

$$p_{k-r}^{(r+2)} = \sum_{i=1}^{k-r} p_i\lambda_i \prod_{m=k-r+1}^{k} (\lambda_h - \lambda_i) \Big/ \prod_{m=k-r}^{k} \lambda_i \quad (r = 1, 2, \cdots k - 1). \tag{16}$$

The equivalence of the conditions of Theorem 2 and those of the Corollary follows at once.

COROLLARY 2. *The conditions of Theorem 2 are equivalent to the following*:

$$\sum_{i=0}^{r} S_{ri} f^{(i)}(0) \geqq 0, \qquad (r = 1, 2, \cdots k)$$

where S_{ri} *is the coefficient of* x^i *in the product*

$$\prod_{m=k-r+1}^{k} (\lambda_m + x).$$

This result follows at once from Corollary 1. It is interesting to see from this corollary that the conditions can be expressed in terms of the derivatives of $f(x)$ at the origin.

Our derivation of the condition of Theorem 1 required that the terms $p_i\lambda_i$ be taken in increasing order of their λ's. The argument leading to Theorem 2 assumed that this initial ordering would continue to yield the best conditions obtainable by this method. It will be proved in the Appendix that if the conditions of Corollary 1 (and hence of Theorem 2 and Corollary 2) are satisfied for some other ordering of the λ's then they are satisfied for $\lambda_1 < \lambda_2 < \cdots < \lambda_k$. The conditions stated there are thus the best of their kind.

4. Concluding Remarks. All of the conditions obtained are sufficient to ensure that $f(x)$ is a probability density function. Except when $k = 2$, when they coincide with conditions (a) and (b), they are not necessary. A simple example will illustrate the point. It is obvious that

$$f(x) = 3\lambda e^{-\lambda x}(1 - 2e^{-\lambda x})^2 = 3\lambda e^{-\lambda x} - 12\lambda e^{-2\lambda x} + 12\lambda e^{-3\lambda x}$$

is a density function. However, to satisfy the conditions of Corollary 2 we require that $f(0) \geqq (0)$ and $f^{(1)}(0) + \lambda_3 f(0) \geqq 0$. For this example $f^{(1)}(0) + 3\lambda f(0) = -15\lambda^2 + 9\lambda^2 < 0$. This curve has a zero ordinate when $e^{-\lambda x} = \frac{1}{2}$ and we may use our conditions to show that $f(x)$ is positive to the right of this point. If we let $y = x - (\ln 2)/\lambda$ then, from (2),

$$f(y \mid y > 0) = 3\lambda e^{-\lambda y} - 6\lambda e^{-2\lambda y} + 3\lambda e^{-3\lambda y}. \tag{17}$$

Since $f(0) = 0$ and $f^{(1)}(0) = 0$, the conditions are now satisfied.

Our method can also be used to find conditions for $f(x)$ to be a monotonic decreasing (J-shaped) density. For this to be the case we require

$$-f^{(1)}(x) = \sum_{i=1}^{k} p_i\lambda_i^2 e^{-\lambda_i x} \geqq 0$$

for all x. The methods of this paper can obviously be applied with only minor modifications. For example, the conditions of Corollary 2 become

$$\sum_{i=0}^{r} S_{ri} f^{(i+1)}(0) \geqq 0, \qquad (r = 1, 2, \cdots k) \tag{18}$$

The condition $-f^{(1)}(x) \geqq 0$ for all x together with $f(0) \geqq 0$ is also sufficient to ensure that $f(x)$ is a density function. Conditions (18) are thus sufficient for $f(x)$ to be a density although they are not as general as those of Theorem 2.

It would be very useful if some simple conditions could be found which were necessary as well as sufficient. Research towards this end has so far proved fruitless. One possibility, suggested by a referee, is as follows. The Laplace transform of $f(x)$ is

$$\phi(s) = \sum_{i=1}^{k} p_i\lambda_i(\lambda_i + s)^{-1}$$

and $\phi(s)$ belongs to a non-negative function if and only if

$$(-1)^m \phi^{(m)}(s) \geqq 0, \qquad s > 0, \qquad (m = 0, 1, 2, \cdots).$$

Thus a set of necessary and sufficient conditions for positivity is

$$\sum_{i=1}^{k} p_i\lambda_i/(\lambda_i + s)^m \geqq 0, \qquad s > 0, \qquad (m = 0, 1, 2, \cdots).$$

This device converts the problem into one concerning the zeros of an infinite set of polynomials but does not appear to yield simple and easily applied conditions.

APPENDIX

Here we shall show that if the conditions of Corollary 2 are satisfied for some permutation of the λ's then they are satisfied for the permutation $\lambda_1 < \lambda_2 < \cdots < \lambda_k$. Let S_{ri} now denote the coefficient of x^i in

$$\prod_{h=k-r+1}^{k} (\nu_h + x)$$

where $\nu_1, \nu_2 \cdots \nu_k$ is a permutation of $\lambda_1, \lambda_2 \cdots \lambda_k$. We are given that

$$\text{(A1)} \qquad \sum_{i=0}^{r} S_{ri} f^{(i)}(0) \geqq 0, \qquad (r = 1, 2, \cdots k).$$

Let G denote the set $\{\nu_k, \nu_{k-1} \cdots \nu_{k-r+1}\}$ and $\bar{G}$ the set $\{\nu_{k-r}, \nu_{k-r-1}, \cdots \nu_1\}$. Our method of proof is to show that the value of the sum in (A1), for given r, is not decreased if the smallest member of G is replaced by the largest member of the union of $\bar{G}$ and min $\{\nu_k, \nu_{k-1} \cdots \nu_{k-r+1}\}$. By repeating the argument for the set G thus obtained we shall arrive at the conclusion that if (A1) holds then it holds also when the set G becomes $\{\lambda_k, \lambda_{k-1}, \cdots \lambda_{k-r+1}\}$.

First we observe that permuting the ν's does not affect the derivatives $f^{(i)}(0)$. The proof is by induction on r. The result is certainly true when $r = 1$ because then

$$\sum_{i=0}^{1} S_{1i} f^{(i)}(0) = \nu_k f^{(0)}(0) + f^{(1)}(0).$$

Since $f^{(0)}(0) \geqq 0$ it follows that replacing ν_k by max $\{\nu_k, \nu_{k-1}, \cdots \nu_1\}$ cannot decrease the sum. Assume now that the result holds for $r = m - 1$. It follows from the definition of S_{ri} that

$$\text{(A2)} \quad \sum_{i=0}^{m} S_{mi} f^{(i)}(0) = \nu^* \sum_{i=0}^{m-1} S^*_{m-1,i} f^{(i)}(0) + \sum_{i=0}^{m-1} S^*_{i-1,i} f^{(i+1)}(0)$$

where $\nu^* = \min \{\nu_k, \nu_{k-1} \cdots \nu_{k-r+1}\}$ and $S^*_{m-1,i}$ is the coefficient formed from $\{\nu_k \cdots \nu_{k-r+1}\}$ with ν^* deleted. Replacing ν^* by max $\{\nu^*, \nu_{k-r}, \cdots \nu_1\}$ will either increase the expression, or leave it the same if the first sum on the righthand side of (A2) is positive. We are given that

$$\sum_{i=0}^{m-1} S_{m-1\, i} f^{(i)}(0) \geqq 0.$$

The first sum on the right hand side of (A2) is identical with this if $\nu^* = \nu_{k-r+1}$. If this is not the case it differs from it only in that min $\{\nu_k, \nu_{k-1} \cdots \nu_{k-r+2}\} = \nu^*$ has been replaced by ν_{k-r+1} which is greater than ν^*. But, by hypothesis, this has the effect of not decreasing the sum. Hence

$$\sum_{i=0}^{m-1} S^*_{m-1\, i} f^{(i)}(0) \geqq 0$$

also. The argument can be repeated until there is no member of $\bar{G}$ which is greater than any member of G in which case $G = \{\lambda_k \lambda_{k-1}, \cdots \lambda_{k-m+1}\}$. At each stage the sum either increases or remains the same so that if (A1) is satisfied with $G = \{\nu_k, \nu_{k-1}, \cdots \nu_{k-m+1}\}$ it must also be satisfied when $G = \{\lambda_k, \lambda_{k-1}, \cdots \lambda_{k-m+1}\}$. Our hypothesis is true for $r = 1$; hence, by induction it is true for all r.

REFERENCES

[1] Cox, D. R. (1962). *Renewal Theory*. Methuen, London.
[2] Kingman, J. F. C. (1966). The algebra of queues. *J. Appl. Probability* **3** 285–326.
[3] Steutel, F. W. (1967). Note on the infinite divisibility of exponential mixtures. *Ann. Math. Statist.* **38** 1303–1305.