THE INSTITUTE OF STATISTICS

THE CONSOLIDATED UNIVERSITY
OF NORTH CAROLINA

CHARACTERIZATIONS OF THE EXPONENTIAL DISTRIBUTION BY CONDITIONAL MOMENTS

by

Steel T. Huang

Institute of Statistics Mimeo Series #1214

February, 1979

DEPARTMENT OF STATISTICS
Chapel Hill, North Carolina

CHARACTERIZATIONS OF THE EXPONENTIAL DISTRIBUTION BY CONDITIONAL MOMENTS

by Steel T. Huang*

Department of Mathematics, University of Cincinnati and
Department of Statistics, University of North Carolina.

ABSTRACT

In this note, two characterizations of the exponential distribution are established. Theorem 1: Let $\alpha > -1$, $\alpha \neq 0$. A positive random variable X has an exponential distribution if and only if $E[(X-y)^{\alpha}|X > y]$ is a finite constant for all $y \geq 0$. Theorem 2: Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables having a continuous distribution F(x), and $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ their order statistics. Let $\alpha > 0$ and $m = 1,2,\ldots,n-1$ be given. Then F(x) is an exponential distribution up to a location parameter if and only if $E[(X_{(m+1)} - X_{(m)})^{\alpha}|X_{(m)} = y]$ is a finite constant for almost all y with respect to F(y).

*This research was supported by the Air Force Office of Scientific Research under contract AFOSR-75-2796 and the Office of Naval Research under contract N00014-75-C-0809.

CHARACTERIZATIONS OF THE EXPONENTIAL DISTRIBUTION BY CONDITIONAL MOMENTS

1. INTRODUCTION

In recent years there has been much interest and development in the theory of characterizations of the exponential distribution. A comprehensive account of this theory and its significance in model building is given by Galambos and Kotz [3]. In this note, two known characterization theorems are generalized. A functional equation is solved first (Theorem 3) using Mellin transforms. As a consequence, the following characterization theorems are established.

THEOREM 1: Let $\alpha > -1$, $\alpha \neq 0$ be given. A positive random variable X has an exponential distribution if and only if

$$E[(X - y)^{\alpha} | x > y] = c \quad \text{for all } y \geq 0, \tag{1.1}$$

where c is a (finite) constant.

If $P(X > y) = 0$, the conditional expectation in (1.1) is chosen to satisfy the equality.

THEOREM 2: Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables having a continuous distribution F(x), and $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ their order statistics. Let $\alpha > 0$ and $m = 1,2,\ldots,n-1$ be given. Then F(x) is an exponential distribution with some location parameter a (i.e. $F(x) = 1 - e^{-\lambda(x-a)}$, $x \geq a$, for some $\lambda > 0$) if and only if

$$E[(X_{(m+1)} - X_{(m)})^{\alpha} | X_{(m)} = y] = c \quad \text{a.e. } [dF(y)] \tag{1.2}$$

where c is a (finite) constant.

REMARK: If $F(x)$ is absolutely continuous and has a bounded density, then Theorem 2 holds for $\alpha > -1$, $\alpha \neq 0$.

Shanbhag [5] showed that (1.1) is a characteristic property of the exponential distribution for the case $\alpha = 1$. Later, this result was generalized to abitrary $\alpha = 1,2,\ldots$ by Sahobov and Geshev (see [3, p.33]). A special case of Theorem 2 ($\alpha = 1$) was obtained by Ferguson [2] (also see [3, p. 61]).

In the context of applied probability, X represents the life time of a component and $X - y$ given $X > y$ represents its residual life time; $X_1, X_2, \ldots, X_n$ represent the life times of n identical components working independently. Then Theorem 1 asserts that the life time is exponentially distributed if and only if the α-th moment ($\alpha > -1$, $\alpha \neq 0$) of the residual life time is a finite constant, and Theorem 2 asserts that the life time has an exponential distribution up to a location parameter if and only if the conditional α-th moment ($\alpha > 0$) of the waiting time between the m-th and (m+1)-st failures given the waiting time to the m-th failure is a finite constant.

2. DERIVATION OF CHARACTERIZATION THEOREMS

Some preliminaries are needed for the proof of Theorem 3. The *Mellin transforms* of a distribution function $F(x)$ with $F(0^+) = 0$ and a nonnegative measurable function $\phi(x)$, $0 < x < \infty$, are given by

$$m(s) = \int_0^\infty x^{s-1}\, dF(x) \text{ and } n(s) = \int_0^\infty x^{s-1} \phi(x) dx$$

respectively, where s is a complex number. The set of all s for which $m(s)$ converges is either the single line $\mathcal{Re}\, s = 1$ or an open strip $(-\infty \leq)\ c_1 < \mathcal{Re}\, s < c_2\ (\leq \infty)$; the set of all s for which $n(s)$ converges is empty or a single vertical line or an open strip. When a Mellin transform converges for all s in an open strip, it is analytic in that region. It follows from the uniqueness property of the characteristic functions that if two Mellin transforms coincide on the line $\mathcal{Re}\ s = 1$ then their corresponding distribution functions are identical. For detailed properties of Mellin transforms, see [4].

LEMMA 1: Suppose F(x) is a distribution function with $F(0^+) = 0$ and m(s), $c_1 < \mathcal{Re}\ s < c_2$, its Mellin transform. Then m(s) is log convex (i.e. log m(s) is convex) on the interval $c_1 < s < c_2$.

<u>Proof</u>. The lemma is a slight generalization of Theorem 1.9 in [1]; the proof is essentially the same.

Consider the Riemann-Stieltjes sum of the integral $\int_{a+}^{b} x^{s-1}\, dF(x)$ for $c_1 < s < c_2$, where $0 < a < b < \infty$:

$$S_n(s) = \sum_{j=0}^{n-1} (a + jh)^{s-1} [F(a + (j+1)h) - F(a + jh)] ,$$

where $h = (b-a)/n$. Each summand is either identically zero or a log convex function. For large b, $S_n(s)$ is positive. Being the sum of log convex functions, $S_n(s)$ is itself log convex [1, Theorem 1.8]. As $n \to \infty$, $S_n(s)$ converges to $\int_{a+}^{b} x^{s-1}\, dF(x)$ for $c_1 < s < c_2$ by the Bounded convergence theorem. Hence this integral, as the limit of log convex functions, is also log convex [1, Theorem 1.6]. Since m(s), $c_1 < s < c_2$, is the limit of $\int_{a+}^{b} x^{s-1}\, dF(x)$ as $a \to 0$ and $b \to \infty$, it is log convex. □

LEMMA 2: If a real-valued function m(s), $s \geq 1$, satisfies the following conditions, then it is identical with the Gamma function $\Gamma(s)$ for $s \geq 1$:

i) For some constant $\alpha > 0$,

$$m(s + \alpha) = \frac{\Gamma(s + \alpha)}{\Gamma(s)} m(s) , \quad s \geq 1.$$

ii) $m(s)$ is log convex.

iii) $m(1) = 1$.

<u>Proof</u>. Suppose $m(s)$ is a function satisfying these three conditions. Applying (i) recursively, we have

$$m(s + k\alpha) = \frac{\Gamma(s + k\alpha)}{\Gamma(s)} m(s), \quad k = 0,1,2,\ldots \quad . \tag{2.1}$$

Setting $s = 1$, (2.1) becomes by (iii)

$$m(1 + k\alpha) = \Gamma(1 + k\alpha), \quad k = 0,1,2,\ldots \tag{2.2}$$

It suffices to show that $m(s)$ agrees with $\Gamma(s)$ on the interval $1 \leq s \leq 1+\alpha$ because of (i). Fix s between 1 and $1+\alpha$. It follows from (ii) that for $k = 1,2,\ldots$

$$\frac{\log m(1 + k\alpha) - \log m(1 + k\alpha - \alpha)}{(1 + k\alpha) - (1 + k\alpha - \alpha)} \leq \frac{\log m(s + k\alpha) - \log m(1 + k\alpha)}{(s + k\alpha) - (1 + k\alpha)}$$

$$\leq \frac{\log m(1 + k\alpha + \alpha) - \log m(1 + k\alpha)}{(1 + k\alpha + \alpha) - (1 + k\alpha)}$$

which reduces to

$$\left[\frac{m(1 + k\alpha)}{m(1 + k\alpha - \alpha)}\right]^{\frac{s-1}{\alpha}} \leq \frac{m(s + k\alpha)}{\Gamma(1 + k\alpha)} \leq \left[\frac{m(1 + k\alpha + \alpha)}{m(1 + k\alpha)}\right]^{\frac{s-1}{\alpha}} \tag{2.3}$$

substituting (2.1) and (2.2) into (2.3), we get

$$\frac{\Gamma(1 + k\alpha)}{\Gamma(s + k\alpha)} \left[\frac{\Gamma(1 + k\alpha)}{\Gamma(1 + k\alpha - \alpha)}\right]^{\frac{s-1}{\alpha}} \leq \frac{m(s)}{\Gamma(s)}$$

$$\leq \frac{\Gamma(1 + k\alpha)}{\Gamma(s + k\alpha)} \left[\frac{\Gamma(1 + k\alpha + \alpha)}{\Gamma(1 + k\alpha)}\right]^{\frac{s-1}{\alpha}} \tag{2.4}$$

A trite calculation employing Stirling's formula shows that, for any $a \geq 0$, the ratio of $\Gamma(t + a)/\Gamma(t)$ and t^a tends to 1 as $t \to \infty$. Applying this result, we deduce that the quantities at both ends of the inequality (2.4) tend 1 as $k \to \infty$. Thus it follows that $m(s) = \Gamma(s)$ for $1 \leq s \leq 1+\alpha$ as desired. □

THEOREM 3: Let a, $\alpha > -1$ and $\alpha \neq 0$, $c > 0$ be given (finite) constants. Suppose $F_a(x)$ is a distribution function satisfying $F_a(a^+) = 0$ and

$$\int_y^\infty (x-y)^\alpha \, dF_a(x) = c\,(1 - F_a(y)), \quad y \geq a . \tag{2.5}$$

Then $F_a(x) = 1 - e^{-\lambda(x-a)}$ for $x \geq a$, where λ is given by $\lambda^\alpha c = \Gamma(1 + \alpha)$.

Proof. Let $F(x) = F_a(a + \frac{x}{\lambda})$. Then $F(x)$ is a distribution function satisfying $F(0^+) = 0$. Also it follows from (2.5) by a change of variable that $F(x)$ satisfies

$$\int_y^\infty (x-y)^\alpha \, dF(x) = \Gamma(1 + \alpha)(1 - F(y)), \quad y \geq 0. \tag{2.6}$$

We shall show that $F(x) = 1 - e^{-\lambda x}$ for $x \geq 0$, and the Theorem then follows. For definiteness let us assume $\alpha > 0$, the case $-1 < \alpha < 0$ being similar.

Since (2.6) implies $\int_0^\infty x^\alpha \, dF(x) > \infty$, the Mellin transform $m(s) = \int_0^\infty x^{s-1} \, dF(x)$ of $F(x)$ converges at $s = 1$ and $1 + \alpha$. Therefore $m(s)$ converges on a strip $c_1 < \mathrm{Re}\, s < c_2$ where $c_1 < 1 < 1 + \alpha < c_2$. Now consider the function $\phi(y)$, $y > 0$, defined by either side of (2.6), and its Mellin transform $n(s)$. On the one hand

$$\begin{aligned} n(s) &= \int_0^\infty y^{s-1} \phi(y)\,dy \\ &= \int_0^\infty \int_y^\infty y^{s-1} (x-y)^\alpha \, dF(x)\,dy \\ &= \int_0^\infty [\int_0^x y^{s-1} (x-y)^\alpha \, dy]\,dF(x) \\ &= \int_0^\infty [\int_0^1 z^{s-1} (1-z)^\alpha \, dz]\,x^{s+\alpha} \, dF(x) \\ &= \frac{\Gamma(s)\Gamma(1+\alpha)}{\Gamma(s+1+\alpha)} \; m(s+1+\alpha) \;, \end{aligned} \tag{2.7}$$

where the set of s for which n(s) converges is seen to be

$$\{s \mid \mathcal{Re}\ s > 0\} \cap \{s \mid c_1 - 1 - \alpha < \mathcal{Re}\ s < c_2 - 1 - \alpha\}$$
$$= \{s \mid 0 < s < c_2 - 1 - \alpha\} \ ; \tag{2.8}$$

on the other hand

$$n(s) = \Gamma(1+\alpha) \int_0^\infty \int_{y+}^\infty y^{s-1}\, dF(x)dy$$
$$= \Gamma(1+\alpha) \int_0^\infty \int_0^x y^{s-1}\, dy\, dF(x) \tag{2.9}$$
$$= \frac{\Gamma(1+\alpha)}{s}\ m(s+1) \ ,$$

where the set of s for which n(s) converges is seen to be

$$\{s \mid \mathcal{Re}\ s > 0\} \cap \{c_1 - 1 < \mathcal{Re}\ s < c_2 - 1\} = \{s \mid 0 < \mathcal{Re}\ s < c_2 - 1\}. \tag{2.10}$$

The interchange of the order of integration above is justified by Fubini's theorem because the integrands are nonnegative. Comparing the convergence regions in (2.8) and (2.10), we have that $c_2 = \infty$ and n(s) converges for all s, $\mathcal{Re}\ s > 0$. Thus it follows from (2.7) and (2.9) that

$$m(s + 1 + \alpha) = \frac{\Gamma(s + 1 + \alpha)}{\Gamma(s + 1)}\ m(s + 1) \ , \quad \mathcal{Re}\ s > 0.$$

Note that m(s) restricted to $s \geq 1$ is a real-valued function which clearly satisfies conditions (i) and (ii) in Lemma 2 and which is log convex by Lemma 1. Therefore, by Lemma 2, $m(s) = \Gamma(s)$ for $s \geq 1$. Since m(s) is analytic in $\mathcal{Re}\ s > c_1$ and $\Gamma(s)$ is analytic in $\mathcal{Re}\ s > 0$, we have $m(s) = \Gamma(s)$ on $\mathcal{Re}\ s > \max(0,c_1)$. In particular $m(s) = \Gamma(s)$ on $\mathcal{Re}\ s = 1$. Finally notice that $\Gamma(s)$, $\mathcal{Re}\ s > 0$, is the Mellin transform of the exponential distribution $1 - e^{-x}$, $x \geq 0$. Thus by the uniqueness property of the Mellin transform, we conclude $F(x) = 1 - e^{-x}$, $x \geq 0$, as was to be proved. □

Proof of Theorem 1. If X has an exponential distribution then it has the lack of memory property (i.e. X given X > y ($y \geq 0$) has the same distribution as X) and hence (1.1) holds. Conversely, suppose (1.1) holds. It is easy to show that (1.1) is equivalent to

$$\int_y^\infty (x-y)^\alpha \, dF(x) = c(1 - F(y)), \quad y \geq 0 ,$$

where F(x) is the distribution function of X. Since X is positive, we have $F(0^+) = 0$ and $c = E(X^\alpha) > 0$.

Thus, by Theorem 3, $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$, for some λ. The proof is complete. □

Proof of Theorem 2 "only if" part. Suppose $F(x) = 1 - e^{-\lambda(x-a)}$, $x \geq a$. We may assume a = 0, for otherwise we can argue with $X_1-a, \ldots, X_n-a$. Then $X_{(m+1)} - X_{(m)}$ is independent of $X_{(m)}$ and it has an exponential distribution (see for example, [3, Theorem 3.1.1]). Hence (1.2) holds.

"if" part. From Theorem 3.1.3 in [3], it follows that $X_{(m+1)}$ given $X_{(m)} = y$ has the distribution function

$$F_y^*(x) = 1 - \left(1 - \frac{F(x) - F(y)}{1 - F(y)}\right)^{n-m}, \quad x \geq y$$

for almost all y with respect to F(y). Therefore, (1.2) implies

$$\int_y^\infty (x-y)^\alpha \, dF_y^*(x) = c \quad \text{a.e. } [dF(y)] ,$$

or equivalently

$$\int_y^\infty (x-y)^\alpha \, dG(x) = c(1 - G(y)) \quad \text{a.e. } [dF(y)] \tag{2.11}$$

where

$$G(x) = 1 - (1-F(x))^{n-m} . \tag{2.12}$$

It is easily seen that $G(x)$ is a continuous distribution function and $\sup\{x|G(x) = 0\} = \sup\{x|F(x) = 0\} = a$ (say), $\inf\{x|G(x) = 1\} = \inf \{x|F(x) = 1\} = b$ (say). We shall show that (2.11) in fact holds for all $y \geq a$ and a is finite. Then, by Theorem 3, $G(x)$ is an exponential distribution with location parameter a. By (2.12), $F(x)$ is also an exponential distribution with location parameter a as asserted.

Denote by $A(y)$ and $B(y)$, $-\infty < y < \infty$, the functions defined by the left hand side and the right hand side of (2.11) respectively. $B(y)$ is continuous since $G(y)$ is. Also $A(y)$ can be shown to be continuous by the Monotone convergence theorem (here the assumption $\alpha > 0$ is used). We claim that $F(y)$ is strictly increasing on $a < y < b$. Otherwise, there exist $a < y_1 < y_2 < b$ such that $F(y_1) = F(y_2)$ and $F(y)$ is strictly increasing on the intervals $(y_1 - \epsilon, y_1)$ and $(y_2, y_2 + \epsilon)$ for sufficiently small $\epsilon > 0$. Now we have $B(y_1) = B(y_2)$, and

$$A(y_1) - A(y_2) = \int_{y_2}^{\infty} [(x-y_1)^{\alpha} - (x-y_2)^{\alpha}]\, dG(x) > 0$$

since the integrand is strictly positive on the set (y_2,∞) which has positive dF-measure. Thus, by the continuity of $A(y)$ and $B(y)$, $A(y) \neq B(y)$ for every y in $(y_1 - \epsilon, y_1)$ or for every y in $(y_2, y_2 + \epsilon)$ for sufficiently small ϵ. This contradicts (2.11) since both $(y_1 - \epsilon, y_1)$ and $(y_2, y_2 + \epsilon)$ have positive dF-measure. Therefore $F(y)$ is strictly increasing on $a < y < b$. This implies (2.11) holds for a dense subset of (a,b). By the continuity of $A(y)$ and $B(y)$ again, (2.11) holds for all y, $a < y < b$. A further inspection reveals that (2.11) in fact holds for all $y \geq a$. Finally, we need to show that $a > -\infty$. Assume the contrary. Then there exists a sequence $a_n < b$ decreasing to $-\infty$. Consider the distribution function

$$F_{a_n}(x) = \frac{F(x) - F(a_n)}{1 - F(a_n)}, \qquad x \geq a_n .$$

Note that $F_{a_n}(a_n^+) = 0$ and, by the improved version of (2.11),

$$\int_y^\infty (x-y)^\alpha \, dF_{a_n}(x) = \frac{1}{1 - F(a_n)} \int_y^\infty (x-y)^\alpha \, dF(x)$$

$$= \frac{c}{1 - F(a_n)} (1 - F(y))$$

$$= c(1 - F_{a_n}(y)), \; y \geq a_n.$$

Now Theorem 3 implies that $F_{a_n}(x) = 1 - e^{-\lambda(x-a_n)}$, $x \geq a_n$ where $\lambda^\alpha c = \Gamma(1 + \alpha)$. Hence

$$F(x) = (1 - F(a_n))(1 - e^{-\lambda(x-a_n)}) + F(a_n), \quad x \geq a_n .$$

Letting $a_n \downarrow -\infty$, we have $F(x) = 1$ for all x. This is a contradiction, and thus a must be finite. The proof is now complete. □

To substantiate the remark after Theorem 2, we note that if F(x) has a bounded density then the function A(y) can be shown to be continuous for $\alpha > -1$ and the same proof of Theorem 3 will carry over.

REFERENCES

[1] Artin, E., The Gamma Function, New York: Holt, Rinehart and Winston, 1964.

[2] Ferguson, T.S., "A Characterization of the Exponential Distribution," *Ann. Math. Statist.*, 35 (1964), 1199-1207.

[3] Galambos, J. and Kotz, S., Characterizations of Probability Distributions, Berlin, Springer-Verlag, 1978.

[4] Kawata, T., Fourier Analysis in Probability Theory, New York: Academic Press, 1972.

[5] Shanbhag, D.N., "The Characterizations for Exponential and Geometric Distributions," *J. Amer. Statist. Assoc.*, 65 (1970), 1256-1259.