A CHARACTERIZATION OF THE GAMMA DISTRIBUTION FROM A RANDOM DIFFERENCE EQUATION

by

Eric S. Tollar

FSU Statistics Report M705
USARO Technical Report No. D-82

July, 1985

The Florida State University
Department of Statistics
Tallahassee, Florida 32306-3033

Research supported in part by the U.S. Army Research Office Grant Number DAAG 29-82-K-0168.
The United States Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

Key words: random difference equations, gamma distribution, independence.
A CHARACTERIZATION OF THE GAMMA DISTRIBUTION FROM A RANDOM DIFFERENCE EQUATION

by

Eric S. Tollar

ABSTRACT

A characterization of the gamma distribution is considered which arises from a random difference equation. A proof without characteristic functions is given that if $V$ and $Y$ are independent random variables, then the independence of $V \cdot Y$ and $(1-V) \cdot Y$ results in a characterization of the gamma distribution (after excluding the trivial cases).
1. Introduction

The general random difference equation is defined recursively by

\[(1.1) \quad Y_n = M_n Y_{n-1} + Q_n, \quad n \geq 1,\]

where \(M_n\) are random \(d \times d\) matrices, and \(Q_n\) and \(Y_n\) are random \(d\)-vectors. This equation has proven useful as a model for physical phenomena (see Bernard, Shenton and Uppuluri (1967), and Cavalli-Sforza and Feldman (1973)), and as a useful mathematical tool (see, for example, Solomon (1975)).

Kesten (1973) has established some general conditions under which \(Y_n\) converges in distribution for \(d \geq 1\) as \(n\) approaches infinity. Paulson and Uppuluri (1972), and Verwaat (1979) have some partial results on the characterization of the limiting distribution of (1.1) for \(d = 1\).

A very specific version of (1.1) when \(d = 2\) will be examined. Under reasonable conditions it is shown that asymptotic independence of \(Y_{1,n}\) and \(Y_{2,n}\) results in a characterization of the gamma distribution. More specifically, it is shown that often asymptotic independence of \(Y_{1,n}\) and \(Y_{2,n}\) implies that there are independent random variables \(V\) and \(Y\) where \(V \cdot Y\) and \((1 - V) \cdot Y\) are also independent. In the non-trivial cases this implies that \(Y\) has a gamma distribution, and \(V\) has a beta distribution.

As can be easily shown, this is yet another generalization of the celebrated characterization of the gamma distribution of Lukacs (1955), (for an example of other generalizations, see Marsaglia (1974)). A simple proof of this (i.e., one without characteristic functions) is given in section 2 of this paper. Section 3 is devoted to the difference equation from which the characterization arose.
2. A Characterization of the Gamma Distribution

In the following, it is said \( X \sim \Gamma(\lambda, \beta) \) if

\[
(2.1) \quad P(X \leq x) = I(x > 0) \Gamma(\beta)^{-1} \int_0^x \lambda \gamma^{\beta-1} e^{-\lambda y} dy,
\]

and \( X \sim \beta(\alpha, \beta) \) if

\[
(2.2) \quad P(X \leq x) = I(0 < x < 1) \beta(\alpha, \beta)^{-1} \int_0^x \gamma^{\alpha-1} (1 - y)^{\beta-1} dy.
\]

Also, for any random variable \( X \) and arbitrary set \( A \), we define the random variable \( X_A \) by the restriction of \( X \) to the set \( A \). That is,

\[
(2.3) \quad P(X_A \leq x) = P(X \in A)^{-1} P(X \leq x, X \in A).
\]

The following lemma proves to be useful in this section.

**Lemma 3.1.** Let \( U, W \) be independent random variables, where \( U > W > 0 \). Then \( U(U - W)^{-1} \) and \( U - W \) are independent if and only if \( U \geq c, W \geq d, c > d > 0 \).

**Proof.** It is clear there must be a constant \( e \) where

\[
(2.4) \quad P(U \geq e) = P(W \leq e) = 1.
\]

Let

\[
(2.5) \quad \begin{cases} 
 b_1 = \inf \{x: P(W \leq x) = 1\}, \\
 b_2 = \sup \{x: P(U \geq x) = 1\}.
\end{cases}
\]

Then, since \( U - W \geq b_2 - b_1, W \leq b_1 \),

\[
(2.6) \quad U(U - W)^{-1} = 1 + W(U - W)^{-1} \leq b_2 (b_2 - b_1)^{-1}.
\]

From the definition of \( b_1 \) and \( b_2 \), it is equally clear that for all \( e > 0 \),

\( P(U - W < b_2 - b_1 + e) > 0 \). Because \( U - W \) and \( U(U - W)^{-1} \) are independent,
(2.7) \[ P(U(U-W)^{-1} > b_2(b_2 - b_1 + \varepsilon)^{-1}) \]
\[ = P(U(U-W)^{-1} > b_2(b_2 - b_1 + \varepsilon)^{-1} | U-W < b_2 - b_1 + \varepsilon). \]

Since

(2.8) \{ \omega: U-W < b_2 - b_1 + \varepsilon \} \subset \{ \omega: (U < b_2 + \varepsilon) \cap (W > b_1 - \varepsilon) \},

\[ U-W < b_2 - b_1 + \varepsilon \] therefore implies that \( U(U-W)^{-1} > 1 + (b_1 - \varepsilon)(b_2 - b_1 + \varepsilon)^{-1} = b_2(b_2 - b_1 + \varepsilon)^{-1}. \) Therefore, for all \( \varepsilon > 0, \)

(2.9) \[ P(U(U-W)^{-1} > b_2(b_2 - b_1 + \varepsilon)^{-1} | U-W < b_2 - b_1 + \varepsilon) = 1. \]

Coupled with (2.7), (2.9) implies that

(2.10) \[ P(U(U-W)^{-1} \geq b_2(b_2 - b_1)^{-1}) = 1. \]

Combined with (2.6), it is established that \( P(U(U-W)^{-1} = b_2(b_2 - b_1)^{-1}) = 1, \) which in turn establishes the lemma, where \( c = b_2 \) and \( d = b_1. \)

\[ \square \]

As a first step in showing that for independent random variables \( V \) and \( Y, \)
\( V \cdot Y \) being independent of \( (1-V) \cdot Y \) leads to a characterization of the Gamma distribution, the following theorem is established.

**Theorem 2.2.** If \( Y \geq 0, \) and if \( V \) and \( Y \) are independent, then \( VY \) and \( (1-V)Y \) are independent if and only if one of the five conditions below are true:

1) \( Y \equiv 0, \)
2) \( V \equiv 0, \)
3) \( V \equiv 1, \)
4) \( Y \equiv c, V \equiv d, \)
5) \( Y \sim \Gamma(\lambda, \alpha + \beta), V \sim B(\alpha, \beta). \)
PROOF. Since sufficiency is obvious, only the necessity of the conditions need be established. For convenience, let

\[(2.11) \quad X = VY \text{ and } Z = (1 - V)Y.\]

Further, let

\begin{align*}
A_1 &= \{Y = 0\} = \{X = 0, Z = 0\} \\
A_2 &= \{Y > 0, V < 0\} = \{X < 0, Z > 0\} \\
A_3 &= \{Y > 0, V = 0\} = \{X = 0, Z > 0\} \\
A_4 &= \{Y > 0, 0 < V < 1\} = \{X > 0, Z > 0\} \\
A_5 &= \{Y > 0, V = 1\} = \{X > 0, Z = 0\} \\
A_6 &= \{Y > 0, V > 1\} = \{X > 0, Z < 0\}.
\end{align*}

From (2.3), it is easily seen that since \(V\) and \(Y\) are independent, so are \(V_{A_i}\) and \(Y_{A_i}\) for all \(i\). Similarly \(X\) and \(Z\) being independent implies \(X_{A_i}\) and \(Z_{A_i}\) are also.

The proof requires little other than the observation that since \(Y\) and \(V\) are independent, and \(X\) and \(Z\) are also assumed to be independent, then since \(Y_{\{Y > 0\}} \sim Y_{\{Y > 0, V \in B\}}\), we have for \(i \neq 1, j \neq 1\)

\[(2.13) \quad Y_{A_i} \sim Y_{A_j}.
\]

It can be similarly seen that for \(i \geq 4, j \geq 4\)

\[(2.14) \quad X_{A_i} \sim X_{A_j},
\]

and for \(2 \leq i \leq 4, 2 \leq j \leq 4,\)

\[(2.15) \quad Z_{A_i} \sim Z_{A_j}.
\]

It can then be shown that only the 5 conditions of the theorem do not violate one of (2.13), (2.14), or (2.15).
Since \( P(Y = 0) = 1 \) is condition 1 of the theorem, it is assumed that \( P(Y = 0) < 1 \) throughout. Two cases will be considered; when \( P(0 < V < 1) = 0 \) and when \( P(0 < V < 1) > 0 \).

Case I: \( P(0 < V < 1) = 0 \).

\( P(0 < V < 1) = 0 \) implies \( P(X > 0, Z > 0) = 0 \). Therefore, either \( P(X > 0) = 0 \), or \( P(Z > 0) = 0 \). Assume that \( P(Z > 0) = 0 \). As such, \( P(A_2 \cup A_3 \cup A_4) = 0 \).

If \( P(A_6) > 0 \), then \( X_{A_6} > 0, -Z_{A_6} > 0 \), and \( X_{A_6} + Z_{A_6} = Y_{A_6} > 0 \). Since \( X_{A_6}, -Z_{A_6} \) are independent, and \( X_{A_6} + Z_{A_6} = Y_{A_6} \) and \( X_{A_6}(X_{A_6} + Z_{A_6})^{-1} = V_{A_6} \) are also independent, the application of lemma 2.1 where \( U = X_{A_6}, W = -Z_{A_6} \) yields \( X_{A_6} = c, Z_{A_6} = -d \), and \( Y_{A_6} = c - d \), for \( c > d > 0 \). If in addition \( P(A_5 > 0) \), then \( X_{A_5} = Y_{A_5} \). However, from (2.13) and (2.14) it must be true that \( X_{A_5} \sim X_{A_6} \) and \( Y_{A_5} \sim Y_{A_6} \) if both \( P(A_5) > 0 \) and \( P(A_6) > 0 \). Thus at most one of \( A_5 \) and \( A_6 \) can have probability greater than 0.

Assume that \( P(A_5) > 0 \) (and therefore \( P(A_2 \cup A_3 \cup A_4 \cup A_6) = 0 \)). Then from the independence of \( Y \) and \( V \) it can be seen that

\[
(2.16) \quad P(Y = 0, V \neq 1) = P(Y = 0)P(V \neq 1) = P(Y = 0)P(Y = 0, V \neq 1).
\]

As such, \( P(Y = 0) = 1 \), which is condition 1, or \( P(Y = 0, V \neq 1) = 0 \), which yields \( P(V = 1) = 1 \) (condition 3).

If \( P(A_6) > 0 \) is assumed instead, a similar condition based on the independence of \( X \) and \( Z \) will yield that \( P(A_6) = 0 \) or \( P(A_6) = 1 \) which are conditions 1 and 4, respectively.

Finally, if it is instead assumed that \( P(X > 0) = 0 \), then by similar arguments it can be shown that one of condition 1, condition 2 or condition 4 must apply.
Case II: \( P(0 < V < 1) > 0 \).

Since it is assumed that \( P(Y = 0) < 1 \), it follows that \( P(A_4) = P(Y > 0, 0 < V < 1) > 0 \). Clearly \( X_{A_4} > 0 \), \( Z_{A_4} > 0 \), \( X_{A_4} \) and \( Z_{A_4} \) are independent. Also, \( V_{A_4} \) and \( Y_{A_4} \) are independent and as can be seen from (2.11),

\[
Y_{A_4} = X_{A_4} + Z_{A_4},
\]

\[
V_{A_4} = X_{A_4} (X_{A_4} + Z_{A_4})^{-1}.
\]

Therefore Lukacs' characterization of the gamma distribution yields,

\[
X_{A_4} \sim \Gamma(\lambda, \alpha), \ Z_{A_4} \sim \Gamma(\lambda, \beta),
\]

which implies

\[
Y_{A_4} \sim \Gamma(\lambda, \alpha + \beta), \ V_{A_4} \sim B(\alpha, \beta).
\]

If \( P(A_2) > 0 \), then letting \( U = Z_{A_2}, W = -X_{A_2} \), appealing to lemma 2.1 again yields that \( Z_{A_2} = c, X_{A_2} = -d \). However, from (2.15), \( Z_{A_2} \sim Z_{A_4} \), which implies that \( P(A_2) = 0 \). Assuming that \( P(A_6) > 0 \) yields a similar contradiction for the distribution of \( X \).

If \( P(A_3) > 0 \), then \( Y_{A_3} = Z_{A_3} \), since \( X_{A_3} = 0 \). But (2.13) and (2.15) imply that \( Y_{A_3} \sim \Gamma(\lambda, \alpha + \beta) \) and \( Z_{A_3} \sim \Gamma(\lambda, \beta) \). As such, \( P(A_3) = 0 \). A similar argument yields that \( P(A_5) = 0 \).

These observations collectively imply \( P(0 < V < 1) = 1 \). Finally, the observation that the independence of \( X \) and \( Z \) implies

\[
(2.17) \quad P(A_1) = P(X = 0, Z = 0) = P(X = 0)P(Z = 0) = P(A_1)^2
\]
yields that $P(A_1) = 1$ which is condition 1, or $P(A_1) = 0$ which yields $P(A_4) = 1$, and in turn implies condition 5. □

In the next theorem, the condition that $Y \geq 0$ in theorem 2.2 is removed.

**THEOREM 2.3.** If $V$ and $Y$ are independent random variables, then $VY$ and $(1 - V)Y$ are independent if and only if one of the conditions below are true:

1) $Y \equiv 0$,
2) $V \equiv 0$,
3) $V \equiv 1$,
4) $Y \equiv c$, $V \equiv d$,
5) $Y \sim \Gamma(\lambda, \alpha + \beta)$, $V \sim B(\alpha, \beta)$,
6) $-Y \sim \Gamma(\lambda, \alpha + \beta)$, $V \sim B(\alpha, \beta)$.

**PROOF.** Again the sufficiency is obvious, so only the necessity of the conditions need be established. Clearly if either $P(Y \geq 0) = 1$ or $P(Y \leq 0) = 1$, then by theorem 2.2, the proof is complete. It will therefore be assumed throughout that $P(Y > 0) > 0$ and $P(Y < 0) > 0$, which will be shown to generate a contradiction.

As in theorem 2.2, again sets are defined and restricted random variables are examined. The sets are:

$$
\begin{align*}
A_1 &= \{Y = 0\} = \{X = 0, Z = 0\} \\
A_2 &= \{Y > 0, V < 0\} = \{X < 0, Z > 0, X + Z > 0\} \\
A_3 &= \{Y > 0, V = 0\} = \{X = 0, Z > 0\} \\
A_4 &= \{Y > 0, 0 < V < 1\} = \{X > 0, Z = 0\} \\
A_5 &= \{Y > 0, V = 1\} = \{X > 0, Z = 0\} \\
A_6 &= \{Y > 0, V > 1\} = \{X > 0, Z < 0, X + Z > 0\} \\
A_7 &= \{Y < 0, V < 0\} = \{X > 0, Z < 0, X + Z < 0\} \\
A_8 &= \{Y < 0, V = 0\} = \{X = 0, Z < 0\} \\
A_9 &= \{Y < 0, 0 < V < 1\} = \{X < 0, Z < 0\} \\
A_{10} &= \{Y < 0, V = 1\} = \{X < 0, Z = 0\} \\
A_{11} &= \{Y < 0, V > 1\} = \{X < 0, Z > 0, X + Z < 0\}.
\end{align*}
$$
Note that while all $A_i$ are of the form $A_i = \{Y \in A, V \in B\}$, the sets $A_2, A_6, A_7$ and $A_{11}$ cannot be expressed as $A_i = \{X \in A, Z \in B\}$, for $A \subset R, B \subset R$.

If it is temporarily assumed that $P(0 < V < 1) > 0$, then $P(X > 0, Z > 0) > 0, P(X < 0, Z < 0) > 0$. Since $X$ and $Z$ are independent, this condition also implies that $P(X > 0, Z < 0) > 0$ and $P(X < 0, Z > 0) > 0$. Thus $P(A_2 \cup A_{11}) > 0, P(A_4) > 0, P(A_6 \cup A_7) > 0, P(A_9) > 0$.

By appealing to Lukacs' characterization as in theorem 2.2, it follows that

\[
\begin{cases}
    X_{A_4} \sim \Gamma(\lambda_1, \alpha_1), Z_{A_4} \sim \Gamma(\lambda_1, \beta_1), Y_{A_4} \sim \Gamma(\lambda_1, \alpha_1 + \beta_1) \\
    -X_{A_9} \sim \Gamma(\lambda_2, \alpha_2), -Z_{A_9} \sim \Gamma(\lambda_2, \beta_2), -Y_{A_9} \sim \Gamma(\lambda_2, \alpha_2 + \beta_2).
\end{cases}
\] (2.19)

By arguments similar to those of (2.13), (2.14) and (2.15), it follows that

\[
\begin{cases}
    X_{A_6 \cup A_7}, -Z_{A_6 \cup A_7} \sim \Gamma(\lambda_1, \alpha_1) \times \Gamma(\lambda_2, \beta_2) \\
    -X_{A_2 \cup A_{11}}, Z_{A_2 \cup A_{11}} \sim \Gamma(\lambda_2, \alpha_2) \times \Gamma(\lambda_1, \beta_1).
\end{cases}
\] (2.20)

Since $P(A_6 \cup A_7) > 0$, it will be temporarily assumed that $P(A_6) > 0$. Clearly, it follows from (2.13) that

\[
Y_{A_6} \sim Y_{A_4} \sim \Gamma(\lambda_1, \alpha_1 + \beta_1).
\] (2.21)

It can be trivially verified from (2.20) for $x_0 > 0, z_0 < 0$ that

\[
P(X_{A_6} \leq x_0, Z_{A_6} \leq z_0) = \iint_A \frac{\alpha_1 \alpha_1^{-1} \lambda_1^{-1} \frac{\beta_2}{\beta_2 - 1} \lambda_2}{\Gamma(\alpha_1) \Gamma(\beta_2) P(A_6)} e^{\lambda_1 x} (-z)^{\beta_2 - 1} \lambda_2 z dx dz
\] (2.22)

where $A = \{(x, z): x \leq x_0, z \leq z_0, x + z > 0\}$. Because $Y_{A_6} = X_{A_6} + Z_{A_6}$, from a
transformation of variables in (2.22), for (2.21) to be also true it must follow that

\[ (2.23) \quad \lambda_{1}^{\beta_{1}} y^{\alpha_{1} + \beta_{1} - 1} \frac{\Gamma(\alpha_{1} + \beta_{1})}{\Gamma(\alpha_{1}) \Gamma(\beta_{1})} \frac{\lambda_{2}^{\beta_{2}}}{\lambda_{2}^{\alpha_{2}}} \int_{0}^{\infty} (y + w)^{\alpha_{1} - 1} (\lambda_{1} + \lambda_{2})^{-1} w^{\beta_{2} - 1} dw. \]

Examination of the behavior of the two sides of this equation (particularly as \( y \) approaches 0) yields that equality is impossible for \( \alpha_{1} > 0, \beta_{1} > 0 \). As such, \( P(A_{0}) = 0 \). It is clear an identical argument will yield that \( P(A_{0}) = 0 \). As such, it must be true that the assumption of \( P(0 < V < 1) > 0 \) leads to a contradiction, so \( P(0 < V < 1) = 0 \), and therefore \( P(X > 0, Z > 0) = P(X < 0, Z < 0) = 0 \).

If \( P(V < 0) > 0 \), then since both \( P(Y < 0) > 0 \) and \( P(Y > 0) > 0 \), we have that \( P(0 < Y < 0) > 0 \) and \( P(0 < Y < 0) > 0 \). This contradicts \( P(X > 0, Z > 0) = 0 \), since \( X \) and \( Z \) are assumed to be independent. As such, \( P(V < 0) = 0 \). Similarly \( P(V > 1) > 0 \) contradicts \( P(X < 0, Z < 0) = 0 \), and therefore \( P(V > 1) = 0 \). If both \( P(V = 1) > 0 \) and \( P(V = 0) > 0 \), then again \( P(X > 0) > 0 \) and \( P(Z > 0) > 0 \), respectively, which is again a contradiction. Therefore we finally have that if \( P(Y > 0) > 0 \) and \( P(Y < 0) > 0 \), then \( P(V = 0) = 1 \) (condition 2), or \( P(V = 1) = 1 \) (condition 3).

In the next section, the random difference equation which motivated this characterization of the gamma distribution is given.

3. Concluding Comments

Paulson and Uppuluri (1972) and Verwaat (1979), have characterized some of the limiting distributions for a random difference equation with one dimension. For a very specific two-dimensional model, the above characterization of the gamma arises when the asymptotic independence of the two compartments is considered.
The equation to be considered is defined for $n > 0$ by the recursive equation

$$
\begin{pmatrix}
Y_{1,n} \\
Y_{2,n}
\end{pmatrix} =
\begin{pmatrix}
V_n & 1-V_n \\
0 & W_n
\end{pmatrix}
\begin{pmatrix}
Y_{1,n-1} \\
Y_{2,n-1}
\end{pmatrix} +
\begin{pmatrix}
U_n \\
0
\end{pmatrix}
$$

where $\{V_n, W_n, U_n\}$ is an i.i.d. sequence, independent of each other and $Y_0$.

Kesten (1973) has given some conditions under which the convergences in distribution of $Y_n$ is assured. Should $Y_n$ converge in distribution to $\mathcal{Y} = (Y_1, Y_2)$, then for

$$
\begin{cases}
\phi(s, t) = E(\exp(isY_1 + itY_2)) \\
\psi(s) = E(\exp(isU_1))
\end{cases}
$$

it is easily verified that

$$
\phi(s, t) = \psi(s) E(\phi(sV_1 + t(1-V_1), tW_1)).
$$

From the assumed independence of $Y_1$ and $Y_2$ it follows that

$$
\phi(s, t) = \psi(s) E(\phi(sV_1 + t(1-V_1), 0) E(0, tW_1))
$$

and that

$$
\phi(s, t) = \psi(s) E(\phi(sV_1, 0) E(t(1-V_1), 0) E(0, tW_1)).
$$

Define the set $A \subset \mathbb{R}$ by

$$
A = \{s: (\psi(s) = 0) \cup (E(0, sW_1) = 0)\}.
$$

If $A^c$ is dense in $\mathbb{R}$ then by equating (3.4) and (3.5) it is clear that for all $s, t \in \mathbb{R}^2$

$$
E(\phi(sV_1 + t(1-V_1), 0) E(0, sV_1, 0) E(0, t(1-V_1), 0)).
$$
From (3.7) it is clear that should $A^c$ be dense, then $Y_1$ and $Y_2$ are independent if and only if $V_1Y_1$ and $(1-V_1)Y_1$ are independent. This is the concern of section 2.

The most reasonable and realistic condition on $\{U_n, V_n, W_n\}$ in order for $A^c$ to be dense is given in the following corollary.

**COROLLARY 3.1.** If $U_n \geq 0$, $0 \leq V_n \leq 1$, and $W_n \geq 0$ for all $n$, then $Y_{1,n}$ and $Y_{2,n}$ are asymptotically independent if and only if $U_n$ has a Gamma distribution and $V_n$ has a beta distribution, or one of the four trivial conditions of theorem 2.2 are true.

The proof follows immediately from theorem 2.2, the conditions for convergence of (3.1) given in Kesten (1973), and the fact that the characteristic functions of non-negative random variables have dense support (see Smith (1962)).

Clearly, other restrictions of $U_n$, $V_n$, $W_n$ will yield that $A^c$ is dense, however, the more interesting problem of characterizing the asymptotic independence of $Y_{1,n}$ and $Y_{2,n}$ appears to be an open question.
REFERENCES


A characterization of the gamma distribution from a random difference equation.

A characterization of the gamma distribution is considered which arises from a random difference equation. A proof without characteristic functions is given that if V and Y are independent random variables, then the independence of V·Y and (1-V)·Y results in a characterization of the gamma distribution (after excluding the trivial cases).