Sankhyā : The Indian Journal of Statistics
1987, Volume 49, Series A, Pt. 2, pp. 199-200

CHARACTERIZATION OF PROBABILITY MEASURES BY LINEAR FUNCTIONS DEFINED ON A HOMOGENEOUS MARKOV CHAIN

By B. L. S. PRAKASA RAO
Indian Statistical Institute

SUMMARY. Let $\zeta_1, \zeta_2, \zeta_3$ be three independent random variables and $Z_1 = \zeta_1-\zeta_2$ and $Z_2 = \zeta_2-\zeta_3$. It is known that if the characteristic function of (Z_1, Z_2) does not vanish, then the distribution of (Z_1, Z_2) determines those of $\zeta_1, \zeta_2, \zeta_3$ up to a possible change in location. Generalizations of this result, to random variables $\zeta_1, ..., \zeta_n$ defined on a homogeneous Markov chain in the sense of Gyires, are obtained.

1. Introduction

Let ξ_1, ξ_2, ξ_3 be three independent real-valued random variables and let

$$Z_1 = \xi_1-\xi_2, \;\; Z_2 = \xi_2-\xi_3. \qquad \dots \quad (1.1)$$

If the characteristic function of (Z_1, Z_2) does not vanish, then it was shown by Kotlarski (1967) that the distribution of (Z_1, Z_2) determines those of ξ_1, ξ_2, ξ_3 upto a change in location. This result was generalized to random elements taking values in a locally compact abelian group in Prakasa Rao (1968).

We now consider the same problem but in the case when the random variables ξ_1, ξ_2, ξ_3 are defined on a homogeneous Markov chain as introduced in Gyires (1981a). Related results, generalizing some characterizations of probability laws by linear function of independent random variables, as given in Rao (1971), are presented.

2. Preliminaries

Suppose that

$$\theta_{hj}^{(k)}, \; 1 \leqslant h, j \leqslant p, \; 1 \leqslant k \leqslant n \qquad \dots \quad (2.1)$$

are independent real-valued random variables. Let $\{\eta_j, j \geqslant 0\}$ be a homogeneous Markov chain with state space $\{1, ..., p\}$ and with a non-singular transition matrix $A = ((a_{hj}))$.

We denote this homogeneous Markov chain by $\{A\}$.

AMS (1980) *subject classification* : 62E10.

Key words and phrases : Characterization, Markov chain, Matrix valued distribution function, Matrix valued characteristic function, Linear form.

Definition 2.1 : The random variables $\{\xi_k, 1 \leqslant k \leqslant n\}$ are said to be *defined on the homogeneous Markov chain* $\{A\}$ if

$$\xi_k = O^{(k)}_{\eta_{k-1}, \eta_k}, \qquad 1 \leqslant k \leqslant n \qquad \dots \quad (2.2)$$

i.e., $\xi_k = O^{(k)}_{hj}$ if $\eta_{k-1} = h,\ \eta_k = j,\ 1 \leqslant k \leqslant n.$

Let

$$a^{(k)}_{hj}(x) = P(\xi_k \leqslant x,\ \eta_k = j \mid \eta_{k-1} = h),$$

$$A_k(x) = ((a^{(k)}_{hj}(x))),$$

and

$$\phi_k(t) = \int_{-\infty}^{\infty} e^{itx} dA_k(x), \quad 1 \leqslant k \leqslant n,\ t \in R. \qquad \dots \quad (2.3)$$

Observe that $\phi_k(0) = A$ and $\phi_k(t)$ is continuous in $t \in R$.

$A_k(x)$ is called the *matrix-valued distribution function* of ξ_k and $\phi_k(t)$ is called the *matrix-valued characteristic function* of ξ_k defined on the homogeneous Markov chain $\{A\}$. It is easy to see that

$$a^{(k)}_{hj}(x) = a_{hj} F^{(k)}_{hj}(x)$$

where $F^{(k)}_{hj}(x)$ is the distribution function of $O^{(k)}_{hj}$. Further the matrix-valued characteristic function of the linear form

$$a_1\xi_1 + a_2\xi_2 + a_3\xi_3 \qquad \dots \quad (2.4)$$

is

$$\phi_1(a_1 t)\phi_2(a_2 t)\phi_3(a_3 t) \qquad \dots \quad (2.5)$$

(cf. Gyires, 1981a, b).

Given a nonsingular matrix M, there always exists a matrix L such that

$$M = \sum_{\nu=0}^{\infty} \frac{1}{\nu\,!} L^{\nu}$$

(Hille, 1948, p. 125). The matrix L is called the *logarithm of the matrix* M and is denoted by $L = \log M$. Since A is non-singular matrix, it can be seen that the matrix-valued characteristic function ϕ_k of ξ_k given by (2.3) is non-singular in a neighbourhood of the origin and there exists $\Phi_k(t) = \log \phi_k(t)$ in this neighbourhood of origin. We choose that version of $\log \phi_k(t)$ for which $\Phi_k(0) = \log A$. Note that, if two non-singular matrices M and N commute, then $\log MN = \log M + \log N$.

It is easy to check that the matrix-valued characteristic function of

$$Z = a_1\xi_{t_1} + \dots + a_j\xi_{t_j} \qquad \dots \quad (2.6)$$

is

$$A^{t_1-1}\phi_{t_1}(a_1 t)A^{t_2-t_1-1}\phi_{t_2}(a_2 t) \dots A^{t_j-t_{j-1}-1}\phi_{t_j}(a_j t)A^{n-t_j} \qquad \dots \quad (2.7)$$

whenever $1 \leqslant i_1 < \ldots < i_j \leqslant n$ (cf. Gyires, 1981a, b). In particular if $\phi_{i_r}(t)$, $1 \leqslant r \leqslant j$ commute with A, then the matrix-valued characteristic function of Z can be written in the form

$$A^{n-j}\phi_{i_1}(a_1 t) \ldots \phi_{i_j}(a_j t). \qquad \ldots \quad (2.8)$$

3. Characterizations

Theorem 3.1 : *Let ξ_1, ξ_2, ξ_3 be random variables defined on a homogeneous Markov chain $\{A\}$. Define*

$$Z_1 = \xi_1 - \xi_2, \; Z_2 = \xi_2 - \xi_3. \qquad \ldots \quad (3.0)$$

If the matrix-valued characteristic function of (Z_1, Z_2) is non-singular, then the matrix-valued distribution function of (Z_1, Z_2) determines the matrix-valued distribution functions of ξ_1, ξ_2, ξ_3 upto change in location.

Proof : Let t, u be real. Note that

$$E[e^{it(\xi_1-\xi_2)+iu(\xi_2-\xi_3)}\, I[\eta_3 = j] \,|\, \eta_0 = h]$$

$$= E[e^{it\xi_1+i(-t+u)\xi_2-iu\xi_3}\, I[\eta_3 = j] \,|\, \eta_0 = h] \qquad \ldots \quad (3.1)$$

where $I(A)$ denotes the indicator function of event A. Hence the matrix-valued characteristic function of (Z_1, Z_2) is

$$\phi_1(t)\phi_2(u-t)\phi_3(-u) \qquad \ldots \quad (3.2)$$

from (2.5). Suppose that $\{\gamma_1, \gamma_2, \gamma_3\}$ is another set of random variables defined on the homogeneous Markov chain $\{A\}$ such that the matrix-valued characteristic function of $(\gamma_1-\gamma_2, \gamma_2-\gamma_3)$ is the same as that of $(\xi_1-\xi_2, \xi_2-\xi_3)$. Let ψ_i, $1 \leqslant i \leqslant 3$ be the matrix-valued characteristic functions of γ_i, $1 \leqslant i \leqslant 3$ respectively. It is clear that

$$\phi_1(t)\phi_2(u-t)\phi_3(-u) = \psi_1(t)\psi_2(u-t)\psi_3(-u) \qquad \ldots \quad (3.3)$$

for all t, u in R. Note that ϕ_i's and ψ_i's are non-singular matrices for all t and u since the joint matrix-valued characteristic function of $(\xi_1-\xi_2, \xi_2-\xi_3)$ is non-singular by hypothesis.

Substituting $t = 0$ in (3.3), we have

$$A\phi_2(u)\phi_3(-u) = A\psi_2(u)\psi_3(-u), \qquad u \in R$$

i.e.,

$$\psi_2^{-1}(u)\phi_2(u) = \psi_3(-u)\phi_3^{-1}(-u), \quad u \in R \qquad \ldots \quad (3.4)$$

since A is non-singular by hypothesis. Similarly substituting $u = 0$ in (3.3) we have

$$\phi_1(t)\phi_2(-t)A = \psi_1(t)\psi_2(-t)A, \qquad t \in R$$

i.e.,

$$\psi_1^{-1}(-t)\phi_1(t) = \psi_2(-t)\phi_2^{-1}(-t), \quad t \in R. \qquad \ldots \quad (3.5)$$

But

$$\psi_1^{-1}(t)\phi_1(t)\phi_2(u-t)\phi_3(-u)\psi_3^{-1}(-u) = \psi_2(u-t), \quad t, u \in R \quad \dots \quad (3.6)$$

from (3.3). Applying (3.4) and (3.5), we have

$$\psi_2(-t)\phi_2^{-1}(-t)\phi_2(u-t)\phi_2^{-1}(u)\psi_2(u) = \psi_2(u-t), \quad t, u \in R. \quad \dots \quad (3.7)$$

Therefore

$$\psi_2(-t)\phi_2^{-1}(-t)\phi_2(u-t) = \psi_2(u-t)\psi_2^{-1}(u)\phi_2(u), \quad t, u \in R. \quad \dots \quad (3.8)$$

Let $\zeta_2 = \psi_2\phi_2^{-1}$. It is easy to obtain from (3.8) that

$$\zeta_2(-t)\phi_2(u-t) = \psi_2(u-t)\psi_2^{-1}(u)\phi_2(u), \quad t, u \in R, \quad \dots \quad (3.9)$$

Substituting $t = u$ in (3.9), we have

$$\zeta_2(-u)A = A\psi_2^{-1}(u)\phi_2(u), \quad u \in R. \quad \dots \quad (3.10)$$

Hence, from (3.9) again, it follows that

$$\zeta_2(-t)\phi_2(u-t) = \psi_2(u-t)A^{-1}\zeta_2(-u)A, \; t, u \in R$$

or

$$A^{-1}\zeta_2^{-1}(-u)A\zeta_2(-t) = \zeta_2(u-t), \quad t, u \in R$$

or

$$A\zeta_2(-t) = \zeta_2(-u)A\zeta_2(u-t), \quad t, u \in R.$$

Hence

$$A\zeta_2(x+y) = \zeta_2(x)A\zeta_2(y), \quad x, y \in R. \quad \dots \quad (3.11)$$

In particular, let $y = 0$ in (3.11). Then $\zeta_2(0) = I$ and

$$A\zeta_2(x) = \zeta_2(x)A, \qquad x \in R. \quad \dots \quad (3.12)$$

Hence A commutes with $\zeta_2(x)$ for all $x \in R$ and we have

$$A\zeta_2(x+y) = A\zeta_2(x)\zeta_2(y), \qquad x, y \in R.$$

Since A is non-singular, it follows that

$$\zeta_2(x+y) = \zeta_2(x)\zeta_2(y), \qquad x, y \in R. \quad \dots \quad (3.13)$$

Since ζ_2 is continuous at 0 with $\zeta_2(0) = I$, it follows from Theorem 9.6.1, p. 287 of Hille and Phillips (1957) that there exists a matrix D_2 such that

$$\zeta_2(x) = e^{xD_2}, \quad x \in R.$$

In particular, it follows that

$$\psi_2(u)e^{-uD_2} = \phi_2(u), \; u \in R.$$

It can be checked that similar relations hold for $i = 1$ and $i = 3$. By the uniqueness theorem for the characteristic functions, the above relation proves that the matrix-valued distribution functions of ξ_1, ξ_2, ξ_3 are determined upto changes in location. From (3.4) and (3.5), it is easy to check that $D_1 = D_2 = D_3$. This completes the proof of the theorem.

Theorem 3.2 : *Let* $\{\xi_k, 1 \leqslant k \leqslant n\}$ *be random variables defined on a homogeneous Markov chain* $\{A\}$. *Suppose* $1 \leqslant i_1 < i_2 < i_3 \leqslant n$. *Define*

$$Z_1 = a_1\xi_{i_1} + a_2\xi_{i_2} + a_3\xi_{i_3},$$

and

$$Z_2 = b_1\xi_{i_1} + b_2\xi_{i_2} + b_3\xi_{i_3}.$$

Further suppose that the matrix-valued characteristic functions $\phi_{i_j}(t)$, $1 \leqslant j \leqslant 3$ *of* ξ_{i_j}, $1 \leqslant j \leqslant 3$ *commute with each other and with* A. *Let* $\{\zeta_k, 1 \leqslant k \leqslant n\}$ *be another set of random variables defined on the homogeneous Markov chain* $\{A\}$ *such that the matrix-valued characteristic functions* $\psi_{i_j}(t)$, $1 \leqslant j \leqslant 3$ *of* ζ_{i_j}, $1 \leqslant j \leqslant 3$ *commute with each other and with* A. *Define*

$$W_1 = a_1\zeta_{i_1} + a_2\zeta_{i_2} + a_3\zeta_{i_3}$$

and

$$W_2 = b_1\zeta_{i_1} + b_2\zeta_{i_2} + b_3\zeta_{i_3}.$$

Assume that the joint matrix-valued characteristic function of (Z_1, Z_2) *is the same as that of* (W_1, W_2) *and is non-singular. Suppose that* $a_i : b_i \neq a_j : b_j$ *for* $i \neq j$, $1 \leqslant i, j \leqslant 3$.

Then $\xi_{i_j} = \zeta_{i_j} + D_j$ *where* D_j *is a constant depending on the state* $(\eta_{i_{j-1}}, \eta_{i_j})$ *only of the Markov chain* $\{A\}$. *In other words, the matrix valued distribution functions of* ξ_{i_j}, $1 \leqslant j \leqslant 3$ *are determined upto change of location.*

Before we give a proof of this theorem we first state few lemmas.

Lemma 3.1 : *Let* ψ_i, $1 \leqslant i \leqslant n$ *be continuous matrix-valued functions such that*

$$\sum_{i=1}^{n} \psi_i(t + c_i u) = A(t\,|\,u) + B(u\,|\,t), \quad t, u \in R \qquad \ldots \quad (3.14)$$

where $c_i \neq 0$, $1 \leqslant i \leqslant n$, $A(x\,|\,y)$ *and* $B(x\,|\,y)$ *are, for any fixed* $y \in R$, *matrix-valued polynomials in* x *of degree* $\leqslant a$ *and* b *respectively. Then* $\psi_i(t)$, $1 \leqslant i \leqslant n$. *are matrix-valued polynomials of degree* $\leqslant a + b + n$.

Lemma 3.2 : *If in Lemma* 3.1

$$A(t\,|\,u) = A(u) \text{ and } B(u\,|\,t) = B(t) \qquad \ldots \quad (3.15)$$

where $A(t)$ *and* $B(u)$ *are matrix-valued continuous functions, then* $\psi_i(t)$, $1 \leqslant i \leqslant n$ *and* $A(t)$, $B(t)$ *are all matrix-valued polynomials of degree* $\leqslant n$.

Lemma 3.3 : *If the R.H.S. of* (3.14) *is of the form*

$$A(t) + B(u) + P_k(t, u) \qquad \ldots \quad (3.16)$$

where $A(t)$ and $B(u)$ are matrix-valued continuous functions and $P_k(t, u)$ is a matrix-valued polynomial of degree k, then $\psi_i(t)$, $1 \leqslant i \leqslant n$, $A(t)$ and $B(t)$ are all matrix-valued polynomials of degree $\leqslant max(n, k)$.

Lemma 3.4 : *If the R.H.S of* (3.14) *consists of only $P_k(t, u)$ as defined, then $\psi_i(t)$, $1 \leqslant i \leqslant n$ are polynomials of degree $\leqslant max(n-2, k)$.*

All of the lemmas follow as consequences of the corresponding results for complex-valued functions by considering the equations component wise (cf. Rao, 1971). Note that the results in Lemmas 3.1 to 3.4 hold if the equations are satisfied for t and u in a neighbourhood of the origin (cf. Rao, 1971).

We now come back to the proof of Theorem 3.2.

Proof of Theorem 3.2 : Note that the matrix-valued joint characteristic function of (Z_1, Z_2) is

$$((E[e^{it(a_1\xi_{i_1}+a_2\xi_{i_2}+a_3\xi_{i_3})+iu(b_1\xi_{i_1}+b_2\xi_{i_2}+b_3\xi_{i_3})}I(\eta_n = j) \mid \eta_0 = h]))$$
$$= ((E[e^{i(a_1t+b_1u)\xi_{i_1}+i(a_2t+b_2u)\xi_{i_2}+i(a_3t+b_3u)\xi_{i_3}}I(\eta_n = j) \mid \eta_0 = h]))$$
$$= A^{n-3}\phi_{i_1}(a_1t+b_1u)\phi_{i_2}(a_2t+b_2u)\phi_{i_3}(a_3t+b_3u)$$

in view of the hypothesis that ϕ_{i_j}, $1 \leqslant j \leqslant 3$ commute with A. Since the joint characteristic function of (Z_1, Z_2) and (W_1, W_2) agree, it follows that

$$A^{n-3}\phi_{i_1}(a_1t+b_1u)\phi_{i_2}(a_2t+b_2u)\phi_{i_3}(a_3t+b_3u)$$
$$= A^{n-3}\psi_{i_1}(a_1t+b_1u)\psi_{i_2}(a_2t+b_2u)\psi_{i_3}(a_3t+b_3u).$$

Note that $\phi_{i_j}(t)$ and $\psi_{i_j}(t)$ are non-singular for all t since the product on either side is non-singular by hypothesis. Since ϕ_i's commute and ψ_i's commute by hypothesis, it follows that

$$\sum_{j=1}^{3} [\log \phi_{i_j}(a_jt+b_ju) - \log \psi_{i_j}(a_jt+b_ju)] = 0$$

for all t, u in a neighbourhood of the origin.

Let

$$f_j(t) = \log \phi_{i_j}(t) - \log \psi_{i_j}(t).$$

Then it follows that

$$\sum_{j=1}^{3} f_j(a_jt+b_ju) = 0.$$

Applying Lemmas 3.3 and 3.4 depending on the coefficients a_j, b_j, $1 \leqslant j \leqslant 3$, it can be checked that

$$f_j(t) = D_jt+E_j$$

where D_j and E_j are matrices. Since $f_j(0) = 0$, it follows that $E_j = 0$ and hence

$$\phi_{i_j}(t) = \psi_{i_j}(t)e^{D_j t}.$$

In other words ξ_{i_j} and η_{i_j} differ by the matrix D_j for $1 \leqslant j \leqslant 3$ almost surely.

Remarks: Note that no conditions on commutativity of the matrices $\phi_i(t)$ and A are assumed in Theorem 3.1 whereas these conditions are part of the hypothesis in Theorem 3.2. On the other hand, in Theorem 3.1, we have ξ_1, ξ_2, ξ_3 in succession but in Theorem 3.2 the linear forms involve $\xi_{i_1}, \xi_{i_2}, \xi_{i_3}$ where $1 \leqslant i_1 < i_2 < i_3 \leqslant n$. Rao (1971) proved Theorem 3.2 for independent random variables ξ_1, ξ_2, ξ_3 and it is extended to locally compact abelian groups in Prakasa Rao (1975) and to generalized random fields in Prakasa Rao (1976).

Theorem 3.3: *Let $\{\xi_k, 1 \leqslant k \leqslant n\}$ be random variables defined on a homogeneous Markov chain $\{A\}$. Suppose $1 \leqslant i_1 < i_2 < \ldots < i_j \leqslant n$. Define*

$$Z_1 = a_1\xi_{i_1} + \ldots + a_j\xi_{i_j},$$

and

$$Z_2 = b_1\xi_{i_1} + \ldots + b_j\xi_{i_j}$$

where $a_i : b_i \neq a_l : b_l$ for $i \neq l$. Suppose that the joint matrix-valued characteristic function of (Z_1, Z_2) is non-singular. If $\phi_r(t)$ and $\psi_r(t)$ are two alternative possible matrix-valued characteristic functions of ξ_{i_r}, $1 \leqslant r \leqslant j$ and if they satisfy the commutativity conditions as in Theorem 3.2, then

$$\phi_r(t) = \psi_r(t)e^{P_{j-2}(t)}$$

where $P_k(t)$ is a matrix-valued polynomial in t of degree $\leqslant k$.

Proof: This result follows from Lemma 3.4 by arguments analogous to those given in Theorem 3.2.

In particular, it follows that, if $j = 4$, then

$$\phi_r(t) = \psi_r(t)e^{P_2(t)}$$

where $P_2(t)$ is a matrix-valued polynomial in t of degree at most 2 with $P_2(0) = 0$.

Remarks : One can derive results similar to others in Rao (1971) under the commutativity condition. The problem of obtaining the solutions of the functional equation of the type

$$A^{t_1-1}\phi_{t_1}(a_1t+b_1u)A^{t_2-t_1-1}\phi_{t_2}(a_2t+b_2u)\ldots A^{t_n-t_{n-1}-1}\phi_{t_n}(a_nt+b_nu)A^{n-t_n}$$

$$=A^{t_1-1}\psi_{t_1}(a_1t+b_1u)A^{t_2-t_1-1}\psi_{t_2}(a_2t+b_2u)\ldots A^{t_n-t_{n-1}-1}\psi_{t_n}(a_nt+b_nu)A^{n-t_n}$$

where ϕ's, ψ's and A need not satisfy commutativity condition is an interesting mathematical problem in itself. A solution to this will lead to general results on characterization of probability measures by linear functions defined on a homogeneous Markov chain.

References

Gyires, B. (1981a): Linear forms in random variables defined on a homogeneous Markov chain. In *The First Pannonian Symp. on Math. Stat.* (Ed. P. Revesz et. al.) Lectures Notes in Statistics 8, Springer-Verlag, New York, 110-121.

——— (1981b): Constant regression of quadratic statistics on the sum of random variables defined on a Markov chain. In *Contributions to Probability* (Ed. J. Gani and V. K. Rohatgi) Academic Press, New York, 255-266.

Hille, E. (1948): *Functional Analysis and Semigroups*, Amer. Math. Soc.

Hille, E. and Phillips (1957): *Functional Analysis and Semigroups*, Amer. Math. Soc.

Kotlarski, I. (1967): On characterizing the gamma and the normal distribution. *Pacific J. Math.* 20, 69-76.

Prakasa Rao, B. L. S. (1968): On a characterization of probability distributions on locally compact abelian groups. *Z. Wahrscheinlichkeitstheorie verw. Geb.* 9, 98-100.

——— (1975): On a characterization of probability distributions on locally compact abelian groups II. In *Statistical Distributions in Scientific Work. Vol.* 3 *Characterizations and Applications* (Ed. G. P. Patil *et al*) D. Reidel Publishing Company, Dordrecht, Holland, 231-236.

——— (1976): On a property of generalized random fields. *Studia Sci. Math. Hung.*, 11, 277-282.

Rao, C. R. (1971): Characterizations of probability laws by linear functions. *Sankhyā*, Ser. A, 33, 265-270.

Paper received : January, 1986.