# DISTRIBUTIONS WITH SUFFICIENT STATISTICS FOR MULTIVARIATE LOCATION PARAMETER AND TRANSFORMATION PARAMETER*

by

Andrew L. Rukhin
Purdue University

Mimeograph Series #80-12

Department of Statistics
Division of Mathematical Statistics

June 1980

*Research was supported by the National Science Foundation under grant MCS 78-02300.

# Distributions with Sufficient Statistics For Multivariate Location Parameter and Transformation Parameter

Andrew L. Rukhin

Department of Statistics
Purdue University
W. Lafayette, IN 47907 USA

## Abstract

The form of distributions with sufficient statistics for multivariate location parameter is discussed, and a formula for the rank of such family is obtained. We also give generalizations of these formulas for the case of general transformation parameter.

Keywords: Sufficient statistics, rank of an exponential family, multivariate location parameter, transformation parameter, exponential polynomials.

## 1. Introduction and Summary

The important role of exponential families in mathematical statistics is well known. This special role suggests a closer examination of such families in the multivariate case where reasonable alternatives to multidimensional normal distribution are still scarce. A study of more general exponential distributions, which are generated from a single member of the family by a group of transformations, is also of interest.

In this paper we discuss the form of exponential families with multivariate location parameter. Logarithms of densities of these distributions can be represented as exponential polynomials. This result in one-dimensional case was obtained by Dynkin (1951). We obtain another analogue of a Dynkin's one-dimensional result: an inequality for the rank of such family is derived in Section 2. It is seen that in the case of a real location parameter these bounds reduce to Dynkin's formula.

In Section 3 we obtain similar results for abstract transformation parameter. A general formula, which gives the logarithm of a density as a matrix element of a finite dimensional representation of the group, is investigated. In the case of commutative group this formula is used to show that exponential families again correspond to exponential polynomials, i.e., the logarithm of densities belongs to the algebra of functions generated by complex valued additive and multiplicative homomorphisms of the group. In the case of a locally compact group a formula for the rank similar to that of Section 2 follows.

## 2. Sufficient statistics for location parameter and their rank

Let P be a probability distribution given on Euclidean space $\mathbb{R}^m$, and let $P_\theta$, $\theta \in \mathbb{R}^m$, denote a translate of P by a vector $\theta$, i.e. $P_\theta(A) = P(A-\theta)$ for all Borel sets A. It is known (cf. for example Rukhin (1976) or the next section) that the family $\{p_\theta, \theta \in \mathbb{R}^m\}$, where $p_\theta$ is the density of $P_\theta$ with respect to Lebesque measure, is exponential one if and only if the density p of measure P admits a representation

$$\log p(u) = \sum_{i=0}^{b} Q_i(u) e^{<a_i, u>} . \tag{1}$$

Here $Q_0(u),\ldots,Q_b(u)$ are polynomials in m-dimensional vector u with complex coefficients, $a_0,\ldots,a_b$ are different m-dimensional vectors of complex numbers and $<a,u>$ signifies the dot product of vectors a and u.

In the case m=1 Dynkin (1951) established a useful formula for the rank r of the corresponding exponential family. Namely he proved that the density (1) forms an r-parameter exponential family, where

$$r = \max(k_0 - 1, 0) + \sum_{i=1}^{b} (k_i + 1). \tag{2}$$

Here $k_i$ denotes the degree of polynomial $P_i$ and it is assumed that in (1) $a_0=0$.

To generalize this formula to the case of arbitrary m we need the following notation.

Let

$$Q(u) = \sum q_{j_1 \ldots j_m} u_1^{j_1} \ldots u_m^{j_m}$$

be a non-zero polynomial in vector $u = (u_1,\dots,u_m)$ of degree k, i.e. there exists a coefficient $q_{i_1\dots i_m} \neq 0$, $i_1+\dots+i_m = k$ and $q_{j_1\dots j_m} = 0$ for $j_1+\dots+j_m > k$. Define two sets F and G in the following way:

$$F = \{(i_1,\dots,i_m): \sum_{s=1}^{m} i_s \leq k,\ q_{j_1\dots j_m} = 0 \text{ for all } j_1 \geq i_1,\dots j_m \geq i_m\}$$

and

$$G = \{(i_1,\dots,i_m): \sum_{s=1}^{m} i_s \leq k,\ q_{j_1\dots j_m} = 0 \text{ for all } j_1 \geq i_1,\dots,j_m \geq i_m \text{ and } j_s > i_s \text{ for some } s\}.$$

Denote $f = f(Q) = \text{card } F$ and $g = g(Q) = \text{card } G$. (Here card A is just the number of elements of a set A).

In the case m=1 clearly $f(Q) = 0$ and $g(Q) = 1$ for any polynomial Q. In the case m=2 if, for example, $Q(u) = u_1^2u_2$, so that k=3, then $F = \{(3,0), (0,2), (1,2), (0,3)\}$ and $f(Q) = 4$. Also $G = \{(3,0), (2,1), (0,2), (1,2), (0,3)\}$, so that $g(Q) = 5$.

It is easy to establish inequalities

$$\binom{m+k}{m} > g(Q) > f(Q)$$

and

$$g(Q) \geq \binom{m+k-1}{m-1}.$$

Theorem 1. Let $p_\theta(u) = p(u-\theta)$, $u, \theta \in \mathbb{R}^m$, be a location parameter density which constitutes an r-parameter exponential family. Then

$$\log p(u) = Q_0(u) + \sum_{i=1}^{b} Q_i(u)e^{\langle a_i,u\rangle},$$

where $Q_i$ are polynomials of degree $k_i$, $i = 0,1,\dots,b$, and $a_i$, $i = 1,\dots,b$ are non-zero complex vectors. Also

$$r \leq \max\,[\binom{m+k_0}{m} - g_0 - 1,\ 0] + \sum_{i=1}^{b} [\binom{m+k_i}{m} - f_i]. \qquad (3)$$

Here $f_i = f(Q_i)$, $i = 1,\ldots b$; $g_0 = g(Q_0)$.

Proof. Let $\mathcal{K}$ denote the linear space spanned by functions $\log p_\theta(u) - \log p(u)$, $\theta \in \mathbb{R}^m$ and constants. By the definition $r = \dim \mathcal{K} - 1$.

Let

$$Q_i(u) = Q(u) = \sum q_{j_1\ldots j_m} u_1^{j_1}\ldots u_m^{j_m}$$

and $a_i = a \neq 0$. Then

$$e^{\langle a,u+t\rangle}Q(u+t) - e^{\langle a,u\rangle}Q(u) = e^{\langle a,u\rangle} \sum_{i_1,\ldots,i_m} u_1^{i_1}\ldots u_m^{i_m} p_{i_1\ldots i_m}(t)$$

where

$$p_{i_1\ldots i_m}(t) = e^{\langle a,t\rangle} \sum_{j_1\geq i_1,\ldots,j_m\geq i_m} q_{j_1\ldots j_m} \prod_{s=1}^{m} \binom{j_s}{i_s} t_s^{j_s-i_s} - q_{i_1\ldots i_m}$$

Let $k_i$ be the number of linearly independent exponential polynomials $p_{i_1\ldots i_m}(t)$. Clearly $k_i$ does not exceed the number of these polynomials which are not identically zero. The number of the polynomials, which are equal to zero, conincides with the number of coefficients $q_{i_1\ldots i_m}$ such that $q_{j_1\ldots j_m} = 0$ for all $j_1 \geq i_1,\ldots,j_m \geq i_m$, i.e. with $f(Q_i)$.

Since the number of all monomials $u_1^{i_1}\ldots u_m^{i_m}$ of degree at most $k_i$ is equal to

$$\sum_{\ell=0}^{k}\binom{\ell+m-1}{m-1} = \binom{m+k}{m},$$

we see that the impact in the dimension of $\mathcal{K}$ contributed by the term $e^{\langle a_i,u\rangle} Q_i-(u)$, $a_1 \neq 0$ does not exceed

$$\binom{m+k_i}{m} - f_i.$$

One also has

$$Q_0(u+t) - Q_0(u) = \sum_{i_1,\ldots,i_m} u_1^{i_1}\ldots u_m^{i_m}\ q_{i_1\ldots i_m}(t),$$

where

$$q_{i_1\ldots i_m}(t) = \sum_{\substack{j_1\geq i_1,\ldots,j_m\geq i_m\\(j_1,\ldots,j_m)\neq(i_1,\ldots,i_m)}} q^0_{j_1\ldots j_m} \prod_{s=1}^{m} \cdot \binom{j_s}{i_s} t_s^{j_s-i_s}.$$

Thus if $k_0 \geq 1$, the polynomial $Q_0(u)$ contributes at most $\binom{m+k_0}{m} - g(Q_0)$ basis functions of the form $u_1^{i_1}\ldots u_m^{i_m}$. Therefore in this case

$$\dim \mathcal{K} \leq \binom{m+k_0}{m} - g_0 + \sum_{i=1}^{b}\left[\binom{m+k_i}{m} - f_i\right].$$

If $k_0 = 0$, the basis of $\mathcal{K}$ is formed just by linear combinations of functions $e^{\langle a,u\rangle}\ u_1^{i_1}\ldots u_m^{i_m}$ and constants. Hence

$$\dim \mathcal{K} \leq \sum_{i=1}^{b}\left[\binom{m+k_i}{m} - f_i\right] + 1 .$$

Combining these formulae we obtain (3) and complete the proof.

Corollary. Under notations of Theorem 1

$$r \leq \max\left[\binom{m+k_0-1}{m} - 1,\ 0\right] + \sum_{i=1}^{k}\binom{m+k_i}{m} \tag{4}$$

Indeed for $k_0 \geq 1$

$$\binom{m+k_0}{m} - g_0 \leq \binom{m+k_0}{m} - \binom{m+k_0-1}{m-1} = \binom{m+k_0-1}{m} .$$

In the case m = 1 inequality (4) reduces to Dynkin's formula (2).

Notice that the numbers f (Q) and g (Q) are not invariant under nonsingular linear transforms of the argument, but the rank r is invariant. Thus by taking a particular linear transformation one can obtain a sharper bound for r.

## 3. Exponential families with transformation parameter

In this section we consider the case of general transformation parameter. Assume that a topological group $\mathcal{G}$ of transformations acts transitively on a space $\mathcal{X}$. Thus we can suppose that $\mathcal{X}$ is the left cosets space $\mathcal{G}/\mathcal{H}$ where $\mathcal{H}$ is a subgroup of $\mathcal{G}$. Define the transformation parameter family $\{P_g, g \in \mathcal{G}\}$ as $\{P_g(A) = P(g^{-1}A)$ for all measurable A$\}$. We shall be interested in the form of measures which form an exponential family. This is a natural generalization of the location parameter families we have considered in Section 2.

Let $\mathcal{C} = \{g: P_g = P\}$. Then $\mathcal{C}$ is a compact subgroup of $\mathcal{G}$ and the parameter space $\Theta$ should be identified with $\mathcal{G}/\mathcal{C}$.

Assume that there exists a relatively invariant measure $\mu$ on $\mathcal{X}$ i.e. $\mu(gA) = \chi(g)\mu(A)$, where $\chi$ is a positive function. If $p(u) = \frac{dP}{d\mu}(u)$, is the density of P with respect to $\mu$, which we assume to be continuous, then $\chi(g^{-1})p(g^{-1}u) = \frac{dP_g}{d\mu}(u)$, $u \in \mathcal{X}$. Since $\chi(g_1g_2) = \chi(g_1)\chi(g_2)$, so that $\chi(c) = 1$ for all $c \in \mathcal{C}$, we see that $\mathcal{C} = \{g: p(gu) = p(u)$ for all $u\}$.

The next result gives the form of densities p from exponential families. Formula (5) was proved under different assumptions by several authors (see Maksimov (1967), Roy (1975), Rukhin (1975)). The uniqueness up to equivalence of a cyclic finite dimensional representation in this formula apparently was not noticed earlier. By M' we will denote the transpose of a matrix M.

Theorem 2. If the family of densities $\{\chi(g^{-1})p(g^{-1}u),\ u \in \mathcal{X},\ g \in \mathcal{G}\}$ constitutes an exponential family, then there exists a matrix homomorphism M of the group $\mathcal{G}$, i.e. $M(g_1 g_2) = M(g_1)M(g_2)$, $g_1, g_2 \in \mathcal{G}$ such that

$$\log p(u) = \langle M(u^{-1})\ell, \Delta \rangle. \tag{5}$$

Here $\ell, \Delta$ are fixed vectors of dimension equal to that of $M(g)$, $M(c)\ell = \ell$ for all $c \in \mathcal{C}$, $M'(h)\Delta = \Delta$ for all $h \in \mathcal{H}$. The vectors $M(g)\ell$, $g \in \mathcal{G}$, span the space $\mathcal{L}$ and the vectors $M'(g)\Delta$, $g \in \mathcal{G}$, span the dual space $\mathcal{L}'$. The representation (5) is unique in the following sense: if (5) holds and

$$\log p(u) = \langle M_1(u^{-1})\ell_1, \Delta_1 \rangle$$

for some matrix homomorphism $M_1$ and vectors $\ell_1$ and $\Delta_1$ with properties specified above, then there exists a nonsingular matrix C such that

$$M_1(u) = C^{-1}M(u)C \text{ and } C\ell_1 = \ell, (C^{-1})'\Delta_1 = \Delta.$$

Proof. Under our assumption the linear space $\mathcal{L}$ spanned by functions $\log p(g^{-1}u)$, $g \in \mathcal{G}$ is finite-dimensional. We can consider the function $\log p(u)$ as defined on $\mathcal{G}$. Then $\log p(gh) = \log p(g)$ for all $h \in \mathcal{H}$, $g \in \mathcal{G}$. Under this agreement all functions from $\mathcal{L}$ are right invariant under multiplication by elements of $\mathcal{H}$. The space $\mathcal{L}$ is invariant under all operators $L(g)$, $g \in \mathcal{G}$, $L(g)f(u) = f(g^{-1}u)$. Let $M(g)$

denote the restriction of the operator $L(g)$ onto $\mathcal{L}$. Clearly for all $f \in \mathcal{L}$

$$M(g_1)M(g_2)f = M(g_1)f(g_2^{-1}\cdot) = f(g_2^{-1}g_1^{-1}\cdot) = f((g_1g_2)^{-1}\cdot) = M(g_1g_2)f,$$

so that

$$M(g_1)M(g_2) = M(g_1g_2).$$

Let $\Delta$ be a linear functional such that $< f, \Delta > = f(e)$, where e is the identity element of the group $\mathcal{G}$. (Note that $f(h) = f(e)$ for $h \in \mathcal{H}$, $f \in \mathcal{L}$). Then for all $h \in \mathcal{H}$ and all $f \in \mathcal{L}$

$$< M(h)f, \Delta > = < f, \Delta >,$$

so that $M'(h)\Delta = \Delta$, $h \in \mathcal{H}$. Also

$$\log p(g^{-1}) = < M(g)\ell, \Delta >,$$

where $\ell$ denotes the function $\log p(\cdot)$ considered as an element of $\mathcal{L}$. Since by the definition of $\mathcal{L}$, $M(c)\ell = \ell$ the formula (5) is proved.

It follows from the definition of the operator $M(g)$ that the vectors $M(g)\ell$, $g \in \mathcal{G}$, span $\mathcal{L}$ and that the vectors $M'(g)\Delta$, $g \in \mathcal{G}$, span $\mathcal{L}'$. If $\log p(u)$ admits another representation of the form (5) then the space $\mathcal{L}$ contains all functions of the form $< M_1(u^{-1})\ell_1, M'(g)\Delta_1 >$ so that with some matrix C

$$M_1(u)\ell_1 = CM(u)\ell .$$

It follows immediately that $\ell_1 = C\ell$ and $M_1(u) = C^{-1}M(u)C$.

As an example to Theorem 2 notice that in the case of multivariate location parameter every matrix homomorphism of $\mathbb{R}^m$ has the form

$$M(u) = \exp\{u_1 N_1 + \ldots + u_m N_m\},$$

where $N_i$, $i = 1,\ldots,n$ are commuting matrices. Therefore the family $\{p(u-\theta)\}$ is exponential one if and only if formula (1) holds.

We generalize this example to the case of a commutative group $\mathcal{G}$. In this situation all matrices $M(g)$ commute, $M(g_1)M(g_2) = M(g_2g_1) = M(g_2)M(g_1)$. Therefore (see Suprunenko and Tyshkevich (1968) p. 16) the whole space $\mathcal{L}$ can be represented as direct sum of subspaces $W_n$, $n = 1,\ldots,N$, which are invariant with respect to all operators $M(g)$. The irreducible parts of restrictions of $M(g)$ onto $W_n$ are equivalent, while for $n \neq s$ the irreducible parts of restrictions of $M(g)$ onto $W_n$ and $W_s$ are not equivalent. Shur's lemma shows that all irreducible parts of restriction of $M(g)$ on $W_n$, $n = 1,\ldots,N$, are one-dimensional operators. Thus all matrices $M(g)$ have the form $M(g) = T^{-1}U(g^{-1})T$, where $U(g)$ is a quasi-diagonal complex matrix with blocks $U_1(g),\ldots,U_N(g)$ on the principal diagonal, and $U_n(g)$ is lower triangular matrix of dimension $w_n = \dim W_n$, $n = 1,\ldots,N$ with the same diagonal elements $d_n(g)$, $d_n(g) \neq d_s(g)$, $n \neq s$.

It is clear that

$$d_n(g_1g_2) = d_n(g_1)d_n(g_2),$$

so that all $d_n$, $n = 1,\ldots,N$ are different non-zero multiplicative continuous homomorphisms of $\mathcal{G}$ into complex numbers. Also if a density $p$ forms an exponential family with transformation parameter, i.e. formula (5) holds, then

$$\log p(g) = \langle U(g)\lambda,\delta\rangle = \sum_{n=1}^{N} \langle U_n(g)\lambda_n,\delta_n\rangle.$$

Here $\lambda = T\ell$, $\delta = (T^{-1})'\Delta$, and $\lambda_n(\delta_n)$ is the projection of $\lambda(\delta)$ onto $W_n(W_n')$, $n = 1,\ldots,N$.

One has $U_n(g) = d_n(g)Y_n(g) = d_n(g)(\, I + V_n(g))$, where $Y_n(g_1g_2) = Y_n(g_1)Y_n(g_2)$ and all eigenvalues of $Y_n$ are equal to one. Also $V_n$ is a nilpotent matrix, $V_n^{w_n} = 0$. Therefore the function $S_n(g) = \langle V_n(g)\lambda_n, \delta_n\rangle$ is a polynomial of degree at most $w_n - 1$, i.e.

$$[L(g) - I]^{w_n} S_n(\cdot) = 0.$$

Indeed

$$0 = \langle V_n^{w_n}(g)\lambda_n, \delta_n\rangle = \langle [d_n(g^{-1})U_n(g) - I]^{w_n}\lambda_n, \delta_n\rangle = [L(g) - I]^{w_n} Q_n(\cdot)$$

Also $Q_n(g) = \langle Y_n(g)\lambda_n, \delta_n\rangle = S_n(g) + \langle \lambda_n, \delta_n\rangle$ is a polynomial of the same degree, and we have established the following result.

Theorem 3. Let $\mathcal{G}$ be a commutative group. If a density p generates an exponential family with transformation parameter from $\mathcal{G}$, then

$$\log p(g) = \sum_{n=1}^{N} d_n(g)Q_n(g),$$

where $d_n(g)$ are different complex-valued continuous multiplicative homomorphisms of $\mathcal{G}$, and $Q_n(g)$, $n = 1,\ldots,N$ are polynomials on $\mathcal{G}$ of the form

$$Q_n(g) = \langle Y_n(g)\lambda_n, \delta_n\rangle, \tag{6}$$

where $Y_n$ is a matrix representation of the group $\mathcal{G}$ with all eigenvalues being identically equal to one.

If $\mathcal{G}$ is locally compact Abelian group then there is a finite number, say m, of different linearly independent additive homomorphisms $\chi$, i.e. $\chi(g_1g_2) = \chi(g_1) + \chi(g_2)$, and every polynomial Q over $\mathcal{G}$ admits a representation

$$Q(g) = \sum q_{j_1\ldots j_m} \chi_1^{j_1}(g)\ldots\chi_m^{j_m}(g)$$

with complex coefficients $q_{j_1\ldots j_m}$. Thus by using the proof of Theorem 1 one easily obtains an inequality for the rank of the corresponding exponential family. Note, however, that if $\mathcal{G}$ is a compact group, then necessarily all additive homomorphisms of $G$ are identically zero, so that every polynomial is a constant. Thus in the case of a compact group log p(u) is just a linear combination of multiplicative homomorphisms, or characters, which are homomorphisms of $G$ into the unit circle.

In the case of a locally compact Abelian group every polynomial has the form (6). However if, say, $G$ is a Hilbert space considered as an additive group and $Q(g) = \|g\|^2$, then Q is a polynomial of degree two, but Q cannot be represented with the form (6).

## References

Dynkin, E.B. (1951). Necessary and sufficient statistics for a family of probability distributions, *Uspehi Matematičeskih Nauk.* 6, 68-90. English translation in *Selected Translations in Mathematical Statistics and Probability*, 1, 23-41.

Maksimov, V.M. (1967). Necessary and sufficient statistics for the family of shifts of probability distributions on continuous bicompact groups. *Theory of Probability and its Applications* (English Translation), 12(2), 267-280.

Roy, K.K. (1975). Exponential families of densities on an analytic group and sufficient statistics. *Sankhyā*, 37 ser A, 82-92.

Rukhin, A.L. (1975). Characterizations of distributions by statistical properties on groups. In *Statistical Distributions for Scientific Work*, G.P. Patil editor, vol 3 Boston, D. Reidel, 173-184.

Rukhin, A.L. (1976). Remark on distributions with sufficient statistics for a multidimensional location parameter. *Theory of Probability and Mathematical Statistics*, 11, 155-160.

Suprunenko, D.A. and Tyshkevich, R.I. (1968). *Commutative Matrices*, Academic Press, New York, London.