COMPLETENESS FOR A FAMILY OF MULTIVARIATE NORMAL DISTRIBUTIONS

By Justus Seely

Oregon State University

A necessary and sufficient condition is given for a family of multivariate normal distributions with zero mean vectors to admit a complete sufficient statistic.

1. Introduction and summary. Suppose Z is a $q \times 1$ random vector distributed according to a multivariate normal distribution with mean vector zero and covariance matrix $\sum_{i=1}^{m} \nu_i V_i$ where $V_1, V_2, \cdots, V_m$ are known real symmetric matrices and $\nu' = (\nu_1, \nu_2, \cdots, \nu_m)$ is a vector of unknown parameters ranging over a non-void open set Ω of R^m. Suppose further $V_m = I$, the matrices V_1, $V_2, \cdots, V_m$ are linearly independent, and for each $\nu \in \Omega$ the matrix $\sum_i \nu_i V_i$ is positive definite. In [3] it was shown that this family of distributions admits a complete sufficient statistic if the set $\mathscr{V}$ consisting of all real linear combinations of the matrices $V_1, V_2, \cdots, V_m$ is a quadratic subspace of the linear space $\mathscr{A}$ composed of all $q \times q$ real symmetric matrics. It is the purpose of the present note to establish the converse of this statement. In so doing explicit representations are obtained for locally best unbiased estimators (Barankin [1], Stein [5]) for parametric functions of the form $\sum_i \lambda_i \nu_i = \lambda' \nu$ and the information matrix is given for each $\nu \in \Omega$. Additionally it is indicated how the results are applicable through the principle of invariance when the mean vector of Z is not necessarily zero but interest is in parametric functions of the form $g(\nu)$.

2. Locally best unbiased estimators and completeness. For each $a \in R^m$ let W_a denote the matrix $\sum_{i=1}^{m} a_i V_i$ and for each $\nu \in \Omega$ let $p(-|\nu)$ denote the density function of a $N_q(0, W_\nu)$ distribution, i.e., a q-variate normal distribution with mean vector zero and covariance matrix W_ν. Further, for each $a, b \in \Omega$ let $C(a \mid b)$ denote the $m \times m$ matrix whose (k, h) element is

$$\operatorname{tr}\left(W_a^{-1} V_k W_a^{-1} W_b W_a^{-1} V_h W_a^{-1} W_b\right),$$

and for $a \in \Omega$ let $Q(a)$ denote the random vector composed of $Q_1(a), Q_2(a), \cdots,$ $Q_m(a)$ where $Q_i(a) = Z' W_a^{-1} V_i W_a^{-1} Z$ for $i = 1, 2, \cdots, m$.

For vectors $a, b \in \Omega$ several points concerning the matrix $C(a \mid b)$ and the random vector $Q(a)$ should be noted. First observe that $2C(a \mid b)$ is the covariance matrix of the random vector $Q(a)$ when $\nu = b$, i.e., with respect to the distribution $N_q(0, W_b)$. Also, $C(a \mid b)$ is a covariance matrix and hence is positive semidefinite, and in fact is positive definite. To see the positive definiteness observe for $\rho \in R^m$ that $\rho' C(a \mid b) \rho$ is equal to $\operatorname{tr}(HW_b)$ where

$$H = W_a^{-1} W_\rho W_a^{-1} W_b W_a^{-1} W_\rho W_a^{-1}.$$

Thus, if $\rho'C(a\,|\,b)\rho = 0$ then $H = 0$ since H is positive semidefinite and W_b is positive definite; however, $H = 0$ implies $W_\rho = 0$ and hence $\rho = 0$ from the linear independence of the matrices $V_1, V_2, \cdots, V_m$. It follows that $C(a\,|\,b)$ is positive definite.

PROPOSITION 1. *For each $\nu \in \Omega$ the information matrix is $\frac{1}{2}C(\nu\,|\,\nu)$.*

PROOF. For each $k = 1, 2, \cdots, m$ let S_k denote the random variable whose value at $Z = z$ is given by

$$\frac{\partial \ln p(z\,|\,\nu)}{\partial \nu_k} = -\tfrac{1}{2}\,\mathrm{tr}\,(W_\nu^{-1}V_k) + \tfrac{1}{2}z'W_\nu^{-1}V_kW_\nu^{-1}z \;.$$

The result follows immediately upon noting that the (k, h) element of $\frac{1}{2}C(\nu\,|\,\nu)$ is $\mathrm{Cov}\,(S_k, S_h\,|\,\nu)$.

PROPOSITION 2. *Let a be a fixed element in Ω. If $\hat\nu = C(a\,|\,a)^{-1}Q(a)$ and $\lambda \in R^m$, then the random variable $\lambda'\hat\nu$ is an unbiased estimator for the parametric function $\lambda'\nu$ and*

$$\mathrm{Var}\,(\lambda'\hat\nu\,|\,\nu) = 2\lambda'C(a\,|\,a)^{-1}C(a\,|\,v)C(a\,|\,a)^{-1}\lambda$$

for all $\nu \in \Omega$.

PROOF. The variance formula follows from the observations in the paragraph preceding Proposition 1, and the unbiasedness follows upon noting that the expectation of $Q(a)$ at $\nu \in \Omega$ is $C(a\,|\,a)\nu$.

Let $a \in \Omega$ and $\lambda \in R^m$ be fixed and let $\hat\nu$ be defined as in Proposition 2. Observe that

$$\mathrm{Var}\,(\lambda'\hat\nu\,|\,\nu = a) = \lambda'[\tfrac{1}{2}C(a\,|\,a)]^{-1}\lambda \;.$$

Thus, the variance of $\lambda'\hat\nu$ is equal to its Cramér–Rao lower bound at $\nu = a$ and hence is a locally best unbiased estimator for $\lambda'\nu$ at the parameter point $\nu = a$. Further, let $\rho = C(a\,|\,a)^{-1}\lambda$ and note that $\lambda'\hat\nu$ may be written in the form $Z'GZ$ where

$$G = \textstyle\sum_{i=1}^m \rho_i W_a^{-1}V_iW_a^{-1} = W_a^{-1}W_\rho W_a^{-1} \;.$$

The matrix G just defined is somewhat interesting. For each $A \in \mathscr{A}$ let $\|A\|_a = [\mathrm{tr}\,(W_a AW_a A)]^{\frac{1}{2}}$, then $\|-\|_a$ is a norm on $\mathscr{A}$ and it is easily established that the matrix G is the unique solution to the following minimization problem:

$$\min \|A\|_a$$

subject to   (i) $A \in \mathscr{A}$

           (ii) $\mathrm{tr}\,(AV_i) = \lambda_i$ for $i = 1, 2, \cdots, m$.

Thus, in the recent terminology of Rao [2] it follows that $\lambda'\hat\nu$ is a MINQUE estimator for $\lambda'\nu$ with respect to the norm $\|-\|_a$. Alternatively, in the terminology of [4] if $\mathscr{A}$ is viewed as a finite-dimensional inner product space with the trace inner product, then $\lambda'\hat\nu$ is a Σ_a-min estimator within the class of quadratic estimators where Σ_a is the positive definite linear operator on $\mathscr{A}$ which induces the norm $\|-\|_a$, i.e., $\Sigma_a A = W_a AW_a$ for all $A \in \mathscr{A}$.

THEOREM. *The family of distributions* $\{N_q(0, W_\nu) : \nu \in \Omega\}$ *admits a complete sufficient statistic if and only if* $\mathscr{V}$ *is a quadratic subspace of* $\mathscr{A}$, *i.e., if and only if* $V \in \mathscr{V}$ *implies* $V^2 \in \mathscr{V}$.

PROOF. Theorem 2 in [3] implies the sufficiency of the condition. Conversely, suppose the family of distributions admits a complete sufficient statistic. For each $\lambda \in R^m$ let T_λ denote the uniformly minimum variance unbiased estimator for $\lambda'\nu$. Such an estimator must exist since the family admits a complete sufficient statistic and since there exists an unbiased estimator for $\lambda'\nu$ for every $\lambda \in R^m$. Let an $a \in \Omega$ be fixed and let $\hat{\nu}$ be defined as in Proposition 2. Since $\lambda'\hat{\nu}$ is a locally best unbiased estimator for $\lambda'\nu$ at $\nu = a$ and since T_λ is the uniformly minimum variance unbiased estimator for $\lambda'\nu$ it follows easily that

$$\Pr\{\lambda'\hat{\nu} = T_\lambda \,|\, \nu = a\} = 1 \;.$$

However, since W_ν is positive definite for every $\nu \in \Omega$ it follows that $\Pr\{\lambda'\hat{\nu} = T_\lambda \,|\, \nu\} = 1$ for all $\nu \in \Omega$ and hence $\lambda'\hat{\nu}$ is a uniformly minimum variance unbiased estimator for $\lambda'\nu$ for each $\lambda \in R^m$. The result now follows from Theorem 1 in [3] upon noting that $\lambda'\hat{\nu}$ is a quadratic form in Z.

With regard to the above theorem several additional points may be mentioned. In the event that $\mathscr{V}$ is a quadratic subspace, it then follows from Theorem 2 in [3] that the random vector $(Z'V_1 Z, \cdots, Z'V_m Z)'$ is a complete sufficient statistic. The assumption that the matrices $V_1, \cdots, V_m$ are linearly independent is not necessary for the validity of the theorem since the family of distributions could always be reparametrized to meet the linear independence assumption. Further, the assumption that Ω is an open subset of R^m may be relaxed to any subset of R^m which contains a non-void open set. This follows because Theorem 2 in [3] remains true under this assumption and since the above proof requires only for each parametric function $\lambda'\nu$ the existence of a quadratic form which is a locally best unbiased estimator for $\lambda'\nu$.

3. The case of a nonzero mean vector. Suppose $Y \sim N_n(X\beta, \sum_{i=1}^m \nu_i V_i)$ where X is a known $n \times p$ matrix of rank p, β is an unknown vector of parameters ranging over R^p, and ν, $V_1, \cdots, V_m$ are as stated in Section 1. Additionally, assume β and ν are not functionally related, i.e., the entire parameter space is $R^p \times \Omega$. Let $\mathscr{P}$ denote this family of distributions and for a matrix A let $\mathbf{R}(A)$ and $\mathbf{N}(A)$ denote the range and null space respectively. Observe that the family $\mathscr{P}$ is invariant under the group of transformations

$$\rho(y) = y + \rho , \qquad\qquad \rho \in \mathbf{R}(X) .$$

Let $\mathscr{G}$ denote this group of transformations and let $Z = Q'Y$ where Q is an $n \times q$ $(q = n - p)$ matrix such that $Q'Q = I$ and $\mathbf{R}(Q) = \mathbf{N}(X')$. Observe that Z is invariant under the group $\mathscr{G}$ and that $Q'y_1 = Q'y_2$ implies $y_1 - y_2 \in \mathbf{R}(X)$ so that $y_1 = y_2 + \rho$ for some $\rho \in \mathbf{R}(X)$. It follows that Z is a maximal invariant with respect to the group $\mathscr{G}$. Furthermore, a transformation in the group G applied to the vector Y leaves the parameter ν of the resulting distribution

unchanged, and thus the invariance principle of estimation for any parametric function of the form $g(\nu)$ indicates that we may restrict attention to the maximal invariant Z. However, the family of distributions corresponding to the maximal invariant Z is $\{N_q(0, \sum_{i=1}^{m} \nu_i Q'V_i Q) : \nu \in \Omega\}$ which corresponds exactly to the family of distributions considered in the first two sections except possibly that the matrices $Q'V_1 Q, \cdots, Q'V_m Q$ may no longer be linearly independent. Thus, for estimating parametric functions of the form $g(\nu)$ with estimators which are invariant under the group $\mathscr{G}$, all statements and results (possibly after a reparametrization) in Section 2 may be utilized.

REFERENCES

[1] BARANKIN, E. W. (1949). Locally best unbiased estimates. *Ann. Math. Statist.* **20** 477–501.
[2] RAO, C. R. (1970). Estimation of variance and covariance components in linear models. Technical Report No. Math-Stat/44/70, Research and Training School, Indian Statistical Institute.
[3] SEELY, J. (1971). Quadratic subspaces and completeness. *Ann. Math. Statist.* **42** 710–721.
[4] SEELY, J. and ZYSKIND, G. (1971). Linear spaces and minimum variance unbiased estimation. *Ann. Math. Statist.* **42** 691–703.
[5] STEIN, C. (1950). Unbiased estimates with minimum variance. *Ann. Math. Statist.* **21** 406–415.

DEPARTMENT OF STATISTICS
OREGON STATE UNIVERSITY
CORVALLIS, OREGON 97331