Marginal and conditional distributions of singular distributions

By

W. A. Harris, Jr.[1] and T. N. Helvig[2]

This note derives marginal and conditional means and covariances when the joint distribution may be singular and discusses the resulting invariants.

1. Introduction.

For convenience our discussion will be in terms of the linear operator expected value, written $\mathcal{E}(x)$, and multivariate normal distributions. However, our proofs are algebraic and apply to any multivariate distribution for which zero correlation and independence are equivalent.

A normal distribution is characterized by its mean μ, and its covariance matrix $\sum$, defined by $\mu=\mathcal{E}(X)$, $\sum=\mathcal{E}([X-\mu][X-\mu]')$ and we write X is distributed according to $N(\mu, \sum)$. If X is distributed according to $N(\mu, \sum)$ and D is a real matrix, then $Z=DX$ is distributed according to $N(D\mu, D\sum D')$.

Our first result is

Theorem 1.

Let X be distributed according to $N(\mu, \sum)$ (possibly singular) and let X, μ, $\sum$ have the compatible partitionings:

$$X=\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad \mu=\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \sum=\begin{pmatrix} \sum_{11} & \sum_{12} \\ \sum_{21} & \sum_{22} \end{pmatrix}.$$

Then the equation $\sum_{11}M+\sum_{12}=0$ has at least one solution and the

Received November 17, 1965.

1) School of Mathematics, University of Minnesota, U. S. A., and Research Institute for Mathematical Sciences, Kyoto University, Japan.

2) Ordnance Division, Honeywell, Inc., U. S. A.

random vectors X_1 and $X_2+M'X_1$ are independent. Hence

(i) *X_1 has a marginal distribution which is distributed according to $N(\mu_1, \Sigma_{11})$, and*

(ii) *the conditional distribution of X_2 given that $X_1=\xi$ is: $\mathcal{E}(X_2|X_1=\xi)=\mu_2-M'(\xi-\mu_1)$, with covariance matrix: $cov(X_2|X_1=\xi)=\Sigma_{22}+M'\Sigma_{12}$.*

G. Marsaglia [2] has shown that (with A^+ denoting the pseudo-inverse of A in the sense of Penrose [3]) $\Sigma_{11}(-\Sigma_{11}^+\Sigma_{12})+\Sigma_{12}=0$ (i.e. that one choice for M is $-\Sigma_{11}^+\Sigma_{12}$) and hence that X_1 and $X_2-\Sigma_{21}\Sigma_{11}^+X_1$ are independent and concluded from this (corresponding to (ii) above) that $\mathcal{E}(X_2|X_1=\xi)=\mu_2+\Sigma_{21}\Sigma_{11}^+(\xi-\mu_1)$ and $\text{cov}(X_2|X_1=\xi)=\Sigma_{22}-\Sigma_{21}\Sigma_{11}^+\Sigma_{12}$.

If Σ_{11} is nonsingular, then $\Sigma_{11}^+=\Sigma_{11}^{-1}$, M is unique, and these are the familiar formulas. However, if Σ_{11} is singular, then if the equation $\Sigma_{11}M+\Sigma_{12}=0$ has one solution, it has many solutions and it appears worthwhile to determine the invariants of the conditional means and covariances resulting from different choices of the solution M. To this end we prove the following theorems.

Theorem 2.

Let the hypotheses of Theorem 1 hold. If M is a solution of the equation $\Sigma_{11}M+\Sigma_{12}=0$, then $\Sigma_{22}+M'\Sigma_{12}$ is an invariant and hence

$$\Sigma_{22}+M'\Sigma_{12}\equiv\Sigma_{22}-\Sigma_{21}\Sigma_{11}^+\Sigma_{12}.$$

Theorem 3.

Let the hypotheses of Theorem 1 hold. If ξ is compatible(3) with Σ_{11} and M is a solution of $\Sigma_{11}M+\Sigma_{12}=0$, then $M'(\xi-\mu_1)$ is an invariant and hence

$$\mu_2-M'(\xi-\mu_1)=\mu_2+\Sigma_{21}\Sigma_{11}^+(\xi-\mu_1).$$

(3) The definition of compatible is a natural one and occurs in Section 5.

2. Preliminaries.

Solutions of linear equations of the form $AX=C$ can be completely characterized in terms of the pseudoinverse, A^+, of the matrix A which is defined as the unique solution of the equations

$$AA^+A=A,\ A^+AA^+=A^+,\ AA^+=(AA^+)^*,\ A^+A=(A^+A)^*.$$

R. Penrose [3] has shown that $AX=C$ has a solution if and only if $AA^+C=C$. Further, if a solution exists, the general solution is given by

$$X=A^+C+(I-A^+A)Y \tag{2.1}$$

where Y is arbitrary.

If A is real, A^+ is real, and the following properties are easily verified from the definition of A^+:

$$(A'A)^+=A^+(A^+)',\ A'AA^+=A',\ A'(A^+)'A'=A'; \tag{2.2}$$

and $A=A'$ implies $A^+=(A^+)'$ and $A^+A=AA^+$.

Further, if T is a real rxm matrix of rank r, then TT' is non-singular and it is easily verified that if $A=T'T$, then $A^+=T'(TT')^{-2}T$ and $AA^+=A^+A=T'(TT')^{-1}T$.

3. Proof of Theorem 1.

Our proof of this theorem is not substantially different from that of Marsaglia [2] but is given for completeness. In fact, it is the classical proof [1] for nonsingular distributions except that we must show that the equation $\Sigma_{11}M+\Sigma_{12}=0$ has a solution even if Σ_{11} is singular.

Now Σ is a positive, semi-definite matrix, and hence can be written in the form, with real S and U,

$$\Sigma=\begin{pmatrix} S'S & S'U \\ U'S & U'U \end{pmatrix}. \tag{3.1}$$

Thus the solvability of $\Sigma_{11}M+\Sigma_{12}=0$ becomes $S'S(S'S)^+S'U=S'U$, which will be satisfied for all U only if $S'S(S'S)^+S'=S'$. But, using

(2. 2) we have $S'S(S'S)^{+}S'=S'SS^{+}(S^{+})'S'=S'(S^{+})'S'=S'$. Hence $\Sigma_{11}M+\Sigma_{12}=0$ has a solution.

Let M be a solution of $\Sigma_{11}M+\Sigma_{12}=0$ and consider the random vector

$$Y=\begin{pmatrix}Y_1\\ Y_2\end{pmatrix}=\begin{pmatrix}I & 0\\ M' & I\end{pmatrix}\begin{pmatrix}X_1\\ X_2\end{pmatrix}=DX,$$

then Y is distributed $N(D\mu, D\Sigma D')$.

A simple computation yields that

$$D\Sigma D'=\begin{pmatrix}\Sigma_{11} & 0\\ 0 & \Sigma_{22}+M'\Sigma_{12}\end{pmatrix}.$$

Hence Y_1 and Y_2 are independent and Theorem 1 is proved.

4. Proof of Theorem 2.

Consider the equation $\Sigma_{11}M=-\Sigma_{12}$. By Theorem 1 this equation has a solution and hence by (2. 1) the general solution is

$$M=-\Sigma_{11}^{+}\Sigma_{12}+(I-\Sigma_{11}^{+}\Sigma_{11})Y \tag{4.1}$$

or $M'=-\Sigma_{21}\Sigma_{11}^{+}+Y'(I-\Sigma_{11}^{+}\Sigma_{11})$. Thus

$$M'\Sigma_{12}\equiv-\Sigma_{21}\Sigma_{11}^{+}\Sigma_{12},$$

since $Y'(I-\Sigma_{11}^{+}\Sigma_{11})\Sigma_{12}=Y'(\Sigma_{12}-\Sigma_{11}^{+}\Sigma_{11}\Sigma_{12})=0$ by the solvability condition. Hence $M'\Sigma_{12}$ is an invariant and the theorem is proved.

5. Proof of Theorem 3.

If Σ_{11} is singular, then X_1 is a distribution in a p-dimensional space which is concentrated in a lower dimensional subspace and if ξ is compatible with this lower dimensional subspace we shall prove that $M'(\xi-\mu_1)$ is an invariant.

Assume, without loss of generality that Σ_{11} and μ_1 have the partitioning

$$\Sigma_{11}=\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix},\quad \mu_1=\begin{pmatrix}\nu_1\\ \nu_2\end{pmatrix}$$

where rank $\sum_{11}=\text{rank}\,\sigma_{11}$ (a simple reordering of the components of X_1 will achieve this). Make the transformation $Y_1=CX_1$, where

$$C=\begin{pmatrix} I & 0 \\ -\sigma_{21}\sigma_{11}^{-1} & I \end{pmatrix},\quad Y_1=\begin{pmatrix} y_1 \\ y_2 \end{pmatrix},\quad X_1=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

and the partitioning is induced by the partitioning of $\sum_{11}$ and μ_1.

Hence Y_1 is distributed according to $N(C\mu_1, C\sum_{11}C')$. But

$$C\sum_{11}C'=\begin{pmatrix} \sigma_{11} & 0 \\ 0 & 0 \end{pmatrix},$$

hence, with probability one, $y_2=\mathcal{E}(y_2)$, or

$$x_2=\sigma_{21}\sigma_{11}^{-1}(x_1-\nu_1)+\nu_2.$$

Definition.

The vector ξ is said to be compatible with $\sum_{11}$ if it has the form

$$\xi=\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \textit{ with } \xi_2=\sigma_{21}\sigma_{11}^{-1}\xi_1.$$

Thus the proof of Theorem 3 reduces to showing that

$$(I-\textstyle\sum_{11}^{+}\sum_{11})\begin{pmatrix} I \\ \sigma_{21}\sigma_{11}^{-1} \end{pmatrix}=0. \tag{5.1}$$

Since $\sum_{11}$ is a positive, semi-definite matrix, it may be written in the form $\sum_{11}=T'T$ where $T=(s\ u)$ is a rectangular matrix with the same number of rows as the rank of $\sum_{11}$. Thus by Section 2, we have

$$\textstyle\sum_{11}^{+}\sum_{11}=\begin{pmatrix} s'(ss'+uu')^{-1}s & s'(ss'+uu')u \\ u'(ss'+uu')^{-1}s & u'(ss'+uu')u \end{pmatrix}$$

and $\sigma_{21}\sigma_{11}^{-1}=u's(s's)^{-1}$. Using this formulation, a simple straight forward calculation shows that (5.1) is valid and the theorem is proved.

REFERENCES

[1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley, New York (1958).

[2] G. Marsaglia, Conditional means and covariances of normal variables with singular covariance matrix, J. Amer. Statist. Assoc. (1964), 1203-1204.

[3] R. Penrose, A generalized inverse for matrices, Proc. Cambridge Philos. Soc. **51** (1955), 406-413.