A NOTE ON WILKS' INTERNAL SCATTER[1]

By H. Robert van der Vaart

North Carolina State University, Raleigh, N. C.

0. Introduction and summary. Let Y_1 and Y_2 be two real valued, independently and identically distributed random variables with variance σ^2. Then $\sigma^2 = \mathcal{E}(Y_1 - \mathcal{E}Y_1)^2 = 2^{-1} \cdot \mathcal{E}(Y_1 - Y_2)^2$. Note that

$$(0.1) \qquad (Y_1 - Y_2)^2 = \begin{vmatrix} 1 & 1 \\ Y_1 & Y_2 \end{vmatrix}^2$$

is the square of the length of the interval $[Y_1, Y_2]$. Given a sample of size n, a well known unbiased estimator for σ^2 is given by

$$(0.2) \qquad [n(n-1)]^{-1} \textstyle\sum_{1 \leq i_1 < i_2 \leq n} (Y_{i_1} - Y_{i_2})^2 = (n-1)^{-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2.$$

The present note will discuss a k-dimensional generalization of this situation. Throughout our discussion we will assume that the following condition is satisfied.

Condition $\mathcal{I}$. *$X_1, X_2, \cdots, X_n$ are $n(>k)$ independently and identically distributed, k-vector valued random variables with (unknown) expected vector μ and covariance matrix Σ.*

One natural generalization of the above parameter σ^2 then is (the X_i in the following formulae being understood as one-column matrices with k components) the parameter θ defined as follows:

$$(0.3) \qquad \theta = [(k+1)!]^{-1} \mathcal{E} \left(\begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ X_1 & X_2 & \cdots & X_k & X_{k+1} \end{vmatrix}^2 \right).$$

Theorem 1 will show that $\theta = \det \Sigma$. Note that the absolute value of the determinant in the second member of (0.3) is known to be $k!$ times the k-dimensional content of the simplex (for a definition e.g. see [4], p. 10) which has the $(k+1)$ points (i.e., the $(k+1)$ k-tuples) $X_1, X_2, \cdots, X_{k+1}$ for its vertices (an enlightening discussion of determinants as connected with volumes can be found in [3], pp. 152–162). This furnishes a geometric interpretation to the parameter $\theta = \det \Sigma$.

By an argument well known from the theory of U-statistics and using the equality $\binom{n}{k+1} \cdot (k+1)! = n(n-1) \cdots (n-k) \equiv n_{k+1}$, say, one finds (see Corollary 1.1) that an unbiased estimator for $\theta = \det \Sigma$ is given by $\hat{\theta}$, where $\hat{\theta}$ is defined by:

$$(0.4) \qquad \hat{\theta} = n_{k+1}^{-1} \cdot \sum \begin{vmatrix} 1 & 1 & \cdots & 1 \\ X_{i_1} & X_{i_2} & \cdots & X_{i_{k+1}} \end{vmatrix}^2.$$

Received 31 December 1964.

[1] Work done under Public Health Service Grant GM-678 from the National Institute of General Medical Sciences to the Biomathematics Training Program, Institute of Statistics, Raleigh Section.

where *summation is over all* $(i_1, i_2, \cdots, i_{k+1})$ *with* $1 \leqq i_1 < i_2 < \cdots < i_{k+1} \leqq n$.

Theorem 2 will point out a simple relation between $\hat{\theta}$ and *Wilks' internal scatter* $S_{k,\bar{X},n} = \det U$ (cf. Wilks [5], equations (4.8), (4.2) and (4.3); Wilks [6], equation (18.1.23); and Section 2 below):

$$(0.5) \qquad (n-1)(n-2)\cdots(n-k)\hat{\theta} = S_{k,\bar{X},n} = \det U.$$

This solves a question, left open by Wilks [5], p. 493, namely: under which conditions is it true that

$$(0.6) \qquad \mathcal{E}S_{k,\bar{X},n} = (n-1)(n-2)\cdots(n-k)\det \mathbf{\Sigma}\ ?$$

Corollary 2.1 shows that equality (0.6) is true as soon as the trivial Condition $\mathcal{I}$ is satisfied. This clearly constitutes a much stronger result than Wilks' preliminary statement. In fact, equation (0.6) is a full generalization of the equality $\mathcal{E}\sum_i (Y_i - \bar{Y})^2 = (n-1)\sigma^2$, and valid under conditions of the same scope.

1. A generalization of the parameter $\frac{1}{2}\mathcal{E}(Y_1 - Y_2)^2 = \sigma^2$ to k-variate distributions.

LEMMA 1. *If the random vectors* X_i $(i = 1, 2, \cdots, k+1)$ *satisfy condition* $\mathcal{I}$, *then*

$$(1.1a) \quad \mathcal{E}(|X_1 - \mu \quad X_2 - \mu \quad \cdots \quad X_k - \mu|^2) = k! \det \mathbf{\Sigma};$$

$$(1.1b) \quad \mathcal{E}(|X_1 - \mu \quad X_2 - \mu \quad \cdots \quad X_{k-1} - \mu \quad X_k - \mu| \cdot |X_1 - \mu \quad X_2 - \mu \quad \cdots \quad X_{k-1} - \mu \quad X_{k+1} - \mu|) = 0.$$

PROOF. Define a function ϵ on the set of all permutations

$$\begin{pmatrix} 1 & 2 & \cdots & k \\ r_1 & r_2 & \cdots & r_k \end{pmatrix}:$$

$\epsilon = +1$ for even permutations, $\epsilon = -1$ for odd permutations. Then (see [2], Theorem 3.6 in Chapter 10):

$$(1.2) \quad \begin{aligned} \epsilon\begin{pmatrix} 1 & 2 & \cdots & k \\ r_1 & r_2 & \cdots & r_k \end{pmatrix} \cdot \epsilon\begin{pmatrix} 1 & 2 & \cdots & k \\ s_1 & s_2 & \cdots & s_k \end{pmatrix} &= \epsilon\begin{pmatrix} r_1 & r_2 & \cdots & r_k \\ 1 & 2 & \cdots & k \end{pmatrix} \\ \cdot\, \epsilon\begin{pmatrix} 1 & 2 & \cdots & k \\ s_1 & s_2 & \cdots & s_k \end{pmatrix} &= \epsilon\begin{pmatrix} r_1 & r_2 & \cdots & r_k \\ s_1 & s_2 & \cdots & s_k \end{pmatrix}. \end{aligned}$$

In the course of this proof $\sum_r$ will indicate summation over all permutations $\begin{pmatrix} 1 & 2 & \cdots & k \\ r_1 & r_2 & \cdots & r_k \end{pmatrix}$ and $\sum_s$ over all permutations $\begin{pmatrix} 1 & 2 & \cdots & k \\ s_1 & s_2 & \cdots & s_k \end{pmatrix}$; X_{1r} and μ_r will represent the rth component of vectors X_1 and μ. Now the first member of equation (1.1a) can be written as

$$\mathcal{E}\sum_r \sum_s \epsilon\begin{pmatrix} 1 & \cdots & k \\ r_1 & \cdots & r_k \end{pmatrix} \epsilon\begin{pmatrix} 1 & \cdots & k \\ s_1 & \cdots & s_k \end{pmatrix} (X_{1r_1} - \mu_{r_1}) \cdots (X_{kr_k} - \mu_{r_k}) \cdot (X_{1s_1} - \mu_{s_1}) \cdots (X_{ks_k} - \mu_{s_k}),$$

which by equation (1.2) and by the definition of covariance is equal to

$$\sum_r \sum_s \epsilon \begin{pmatrix} r_1 & \cdots & r_k \\ s_1 & \cdots & s_k \end{pmatrix} \sigma_{r_1 s_1} \cdots \sigma_{r_k s_k} = \sum_r \det \boldsymbol{\Sigma} = k! \det \boldsymbol{\Sigma}.$$

For the first member of equation (1.1b) we find in a similar way:

$$\mathcal{E} \sum_r \sum_s \epsilon \begin{pmatrix} 1 & \cdots & k \\ r_1 & \cdots & r_k \end{pmatrix} \epsilon \begin{pmatrix} 1 & \cdots & k \\ s_1 & \cdots & s_k \end{pmatrix} (X_{1r_1} - \mu_{r_1}) \cdots (X_{k-1,r_{k-1}} - \mu_{r_{k-1}})$$
$$\cdot (X_{kr_k} - \mu_{r_k})(X_{1s_1} - \mu_{s_1}) \cdots (X_{k-1,s_{k-1}} - \mu_{s_{k-1}})(X_{k+1,s_k} - \mu_{s_k}).$$

Since the vectors X_k and X_{k+1} are independent, it is obvious that $\mathcal{E}[(X_{kr_k} - \mu_{r_k})(X_{k+1,s_k} - \mu_{s_k})] = 0$ for all pairs (r_k, s_k). This proves equation (1.1b).

Now we can prove

THEOREM 1. *If the random vectors X_i $(i = 1, 2, \cdots k + 1)$ satisfy Condition $\mathcal{I}$, then*

$$\mathcal{E}\left(\begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ X_1 & X_2 & \cdots & X_k & X_{k+1} \end{vmatrix}^2\right) = (k + 1)! \det \boldsymbol{\Sigma}. \tag{1.3}$$

PROOF. By elementary results from the theory of determinants we have

$$\begin{vmatrix} 1 & \cdots & 1 & 1 \\ X_1 & \cdots & X_k & X_{k+1} \end{vmatrix} = \begin{vmatrix} 1 & \cdots & 1 & 1 \\ X_1 - \mu & \cdots & X_k - \mu & X_{k+1} - \mu \end{vmatrix} = \sum_{j=1}^{k+1} (-1)^{j-1} D_j,$$

where D_j is the determinant of the matrix obtained from the matrix $(X_1 - \mu \quad \cdots \quad X_k - \mu \quad X_{k+1} - \mu)$ by deleting the jth column. So

$$\begin{vmatrix} 1 & \cdots & 1 & 1 \\ X_1 & \cdots & X_k & X_{k+1} \end{vmatrix}^2 = \sum_{j=1}^{k+1} D_j^2 + 2 \sum_{j=1}^{k+1} \sum_{l=j+1}^{k+1} (-1)^{j+l} D_j D_l. \tag{1.4}$$

Since the vectors X_i are identically distributed, $\mathcal{E}D_j^2 = k! \det \boldsymbol{\Sigma}$ for all $j = 1, \cdots, k + 1$, as a consequence of equation (1.1a), and (since in addition D_j and D_l have all columns in common but one) $\mathcal{E}D_jD_l = 0$ for all $j \neq l$, as a consequence of equation (1.1b). Taking expected values in equation (1.4), Theorem 1 follows immediately.

Theorem 1 establishes $\det \boldsymbol{\Sigma}$ as an estimable parameter. Since $(k + 1)$ numbers $i_1 < i_2 < \cdots < i_{k+1}$ can be chosen from n numbers in $\binom{n}{k+1}$ different ways, the sum in equation (0.4) contains $\binom{n}{k+1}$ terms, all of which have expected value $(k + 1)! \det \boldsymbol{\Sigma}$. Hence concerning the statistic $\hat{\theta}$ defined in equation (0.4) we have:

COROLLARY 1.1. *If the random vectors X_i $(i = 1, \cdots, n)$ satisfy Condition $\mathcal{I}$, then*

$$\mathcal{E}\hat{\theta} = \det \boldsymbol{\Sigma}. \tag{1.5}$$

2. Wilks' internal scatter. Wilks [5], equation (4.8), introduced a statistic $S_{k,\bar{X},n}$, which he called the *internal scatter of the sample* $X_1, X_2, \cdots, X_n$. In the following formulae the X_i and $\bar{X}$ are still understood to be one-column matrices with k components, the X_i' and $\bar{X}'$ one-row matrices with k components; $\bar{X} = n^{-1} \cdot \sum_{i=1}^{n} X_i$. Wilks' definition is:

$$(2.1) \qquad S_{k,\bar{X},n} = \sum (|X_{i_1} - \bar{X} \quad X_{i_2} - \bar{X} \quad \cdots \quad X_{i_k} - \bar{X}|^2),$$

where *summation is over all* $(i_1, i_2, \cdots, i_k)$ *with* $1 \leqq i_1 < i_2 < \cdots < i_k \leqq n$; $n > k$. By the theorem of Binet-Cauchy (e.g. see [1], p. 9) $S_{k,\bar{X},n}$ is equal to the determinant of the matrix

$$(2.2) \quad (X_1 - \bar{X} \cdots X_n - \bar{X}) \begin{pmatrix} X_1' - \bar{X}' \\ \cdots \\ X_n' - \bar{X}' \end{pmatrix} = \sum_{i=1}^{n} (X_i - \bar{X})(X_i' - \bar{X}') \equiv U,$$

say, where, except for a multiplicative constant, U is the sample covariance matrix; $u_{pq} = \sum_{i=1}^{n} (X_{ip} - \bar{X}_p)(X_{iq} - \bar{X}_q)$. It follows that

$$(2.3) \qquad S_{k,\bar{X},n} = \det U,$$

a result cited by Wilks, but proved in a different manner, namely as a corollary of the minimum property of the internal scatter among a set of sample scatters (with other pivotal points than $\bar{X}$). Now we can prove

THEOREM 2. *The following identity is true for all values of the random vectors*:

$$(2.4) \qquad \sum_{1 \leqq i_1 < \cdots < i_{k+1} \leqq n} \begin{vmatrix} 1 & \cdots & 1 & 1 \\ X_{i_1} & \cdots & X_{i_k} & X_{i_{k+1}} \end{vmatrix}^2 = n \cdot S_{k,\bar{X},n} = n \det U.$$

PROOF. By elementary results from the theory of determinants and by the Binet-Cauchy theorem the first member of equation (2.4) equals

$$\sum \begin{vmatrix} 1 & \cdots & 1 & 1 \\ X_{i_1} - \bar{X} & \cdots & X_{i_k} - \bar{X} & X_{i_{k+1}} - \bar{X} \end{vmatrix}^2$$
$$= \det \left\{ \begin{pmatrix} 1 & \cdots & 1 \\ X_1 - \bar{X} & \cdots & X_n - \bar{X} \end{pmatrix} \begin{pmatrix} 1 & X_1' - \bar{X}' \\ \cdots & \cdots \\ 1 & X_n' - \bar{X}' \end{pmatrix} \right\}$$
$$= \det \begin{pmatrix} n & 0' \\ 0 & U \end{pmatrix} = n \det U.$$

COROLLARY 2.1. *If the random vectors* X_i $(i = 1, \cdots, n)$ *satisfy Condition* $\mathcal{J}$, *then* $\mathcal{E} S_{k,\bar{X},n} = \mathcal{E} \det U = (n-1)(n-2) \cdots (n-k) \det \Sigma$.

PROOF. By Theorem 2 and equations (2.3) and (0.4),

$$S_{k,\bar{X},n} = \det U = (n-1)(n-2) \cdots (n-k)\hat{\theta}.$$

Corollary 2.1 then follows by applying Corollary 1.1.

REFERENCES

[1] GANTMACHER, F. R. (1959). *The Theory of Matrices* **1** Chelsea, New York. English translation.

[2] MOSTOW, GEORGE D., SAMPSON, JOSEPH H. and MEYER, JEAN-PIERRE (1963). *Fundamental Structures of Algebra*. McGraw-Hill, New York.

[3] PISOT, CHARLES and ZAMANSKY, MARC (1959). *Mathématiques Générales; Algèbre-Analyse*. Dunod, Paris.

[4] PONTRYAGIN, L. S. (1952). *Foundations of Combinatorial Topology*. Graylock, Rochester. English translation.

[5] WILKS, S. S. (1960). Multidimensional statistical scatter. *Contributions to Probability and Statistics* (*Essays in Honor of Harold Hotelling*) (Olkin, Ingram *et al.*, eds.), Stanford Univ. Press.

[6] WILKS, S. S. (1962). *Mathematical Statistics*. Wiley, New York.