THE DEGREES OF FREEDOM OF THE X^2-TEST OF DIMENSIONALITY

by

A. M. Kshirsagar

Technical Report No. 100
Department of Statistics ONR Contract

May 1, 1971

Research sponsored by the Office of Naval Research
Contract N00014-68-A-0515
Project NR 042-260

DEPARTMENT OF STATISTICS
Southern Methodist University

The Degrees of Freedom of the χ^2-test of Dimensionality

A. M. Kshirsagar
Southern Methodist University
Dallas, Texas 75222

1. INTRODUCTION: Wani and Kabe (abbreviated as W & K hereafter) [2] have recently given an elegant derivation of the likelihood ratio criterion for testing the hypothesis H_0 that the dimensionality of the space of the means of k p-variate normal populations is s. The main difference between their derivation and the one given in Rao [1] is that Rao uses geometrical terminology while W & K's derivation is completely analytical. However, their proof is incomplete without the degrees of freedom of the χ^2-test. It will be a pity to be required to go to Rao's geometrical terminology just for the degrees of freedom (d.f.) of the test. The object of this note is therefore to derive the number of d.f. here analytically and complete the W & K derivation.

2. Degrees of Freedom of the Likelihood Ratio Criterion:

We shall use the same notation as W & K and shall not reproduce it here to economize space. The number of d.f. of the likelihood ratio criterion is the difference between the number of parameters with respect to (w.r.t.) which the likelihood L is maximized in the entire parameter space Ω and the space ω, restricted by the hypothesis

$$H_0; \quad H\mu_i = \xi, \quad (i = 1, 2, \ldots, k)$$

given by equation (1) of W & K. The number of parameters in Ω is pk, the p means of each of the k populations. Let us now count the number of parameters estimated while deriving max. L in ω. W & K's first step in maximizing L is equivalent to making a transformation from

$\mu_i (i = 1, 2, \ldots k)$ to

$$\left[\frac{H}{K}\right] \mu_i \tag{1}$$

where K is an $r \times p$ matrix of rank $r = p - s$, such that $H\Sigma K' = 0$, so that the exponent (apart from the factor ½) in L given by (2) of W & K, reduces to (on using H_0),

$$\sum_{i=1}^{k} N \left(\xi - H \bar{U}_i\right)' \left(H\Sigma H'\right)^{-1} \left(\xi - H \bar{U}_i\right)$$

$$+ \sum_{i=1}^{k} N \left(K\mu_i - K \bar{U}_i\right)' \left(K\Sigma K'\right)^{-1} \left(K\mu_i - K \bar{U}_i\right) . \tag{2}$$

This is then minimized w.r.t. $K\mu_i (i = 1, \ldots, k)$ first. In other words, we estimate the rk parameters, $K\mu_i (i = 1, \ldots k)$ here. The second term in (2) therefore vanishes, when the minimum value is taken. This step is hidden in W & K. Next they minimize (2) w.r.t. the unknown ξ ie they estimate a further s parameters. Finally, they minimize

$$\mathrm{tr}\{(H\Sigma H')^{-1} HBH'\}, \tag{3}$$

where $B = N\left(\sum_{i=1}^{k} \bar{U}_i \bar{U}_i' - k \bar{U} \bar{U}'\right)$, w.r.t. the unknown H, the only condition being that rank H is s. One may think here that the additional number of parameter estimated in this is p, the number of elements of H, but it is not true because, the quantity in (3) is also equal to

$$\mathrm{tr}\ H^{*\prime} \ (H^* \Sigma H^{*\prime})^{-1} H^* B, \tag{4}$$

where $H^* = AH$ and A is any arbitrary non-singular $s \times s$ matrix, and we can choose A to be H_1^{-1}, where

$$H = [H_1 | H_2], \tag{5}$$

H_1 being $s \times s$, H_2 being $s \times (p - s)$. H_1 can be assumed to be non-singular without loss of generality, as rank $H = s$. This reduces H^* to

$$[I_s \,|\, H_1^{-1} H_2] \tag{6}$$

which has only $s(p - s)$ unknown elements. Thus, the number of unknown parameters estimated in minimizing (3) is only $s(p - s)$. The total number of parameters in ω is therefore

$$rk + s + s(p - s) \tag{7}$$

and the degrees of freedom of the x^2 test are

$$pk - (rk + s + ps - s^2) = (p - r)(k - 1 - r). \tag{8}$$

3. Equation of the r-dimensional Flat:

Rao bases his derivation on the fact that the hypothesis H_0 is geometrically equivalent to the fact that the k points (representing the means of the k populations) collapse on an r-dimensional flat and he then proceeds to write its vectorial equation. Perhaps, it will be instructive to demonstrate this analytically. If H_0 is true,

$$HM = \xi E_{1k}, \tag{9}$$

where
$$M = [\mu_1, \mu_2, \ldots, \mu_k] \tag{10}$$

and E_{ab} denotes an $a \times b$ matrix, with all unit elements. Hence $HM^* = 0$, where $M^* = M(I - \frac{1}{k} E_{kk})$. So that M^* is of rank $p - s = r$, as H is of rank s, and that its rank cannot be improved upon. M^* has therefore r linearly independent column vectors. Let us denote them by $\mu^*_{j_i}$ $(i = 1, 2, \ldots r)$. But, it is easy to see from the relationship between M and M^* that the difference between any two columns of M is the same as the difference between the corresponding columns of M^* and so,

$$\mu_i = \mu_1 + (i^{th} \text{ column of } M^* - \text{1st column of } M^*)$$

$$= \mu_1 + \text{a linear combination of } \mu^*_{j_i} \ (I = 1, 2, \ldots r)$$

This is the vectorial equation of the r-dimensional flat which Rao uses and is determined by the (r + 1) independent points μ_1 and $\mu^*_{j_i}$ (i = 1, 2, ..., r).

I am indebted to Dr. J. T. Webster, for his help and discussion.

References

[1] Rao, C. Radhakrishna. Linear Statistical Inference and its applications. John Wiley, N.Y. 1965.

[2] Wani, J. K. and Kabe, D. G. (1970) "On a certain minimization problem", The American Statistician, 24, p. 29.