CATEGORICAL DATA ANALOGS OF SOME MULTIVARIATE TESTS

by

V. P. Bhapkar
University of North Carolina
and
University of Poona

Institute of Statistics Mimeo Series No. 450

October 1965

(To be published in Roy Memorial Volume)

This research was supported by the National Institutes of Health Institute of General Medical Sciences Grant No. GM-12868-01.

DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA
Chapel Hill, N. C.

1. INTRODUCTION

Problems of association between several continuous variables are generally handled under the assumption of joint normality of the variables concerned. Similary in the general multifactor multiresponse situation with observations on several characters of each experimental unit for various factor-combinations, data are generally analysed under the assumption of normality; the situation is referred to as one involving analysis of variance (ANOVA) or multivariate analysis of variance (MANOVA) according as whether we have observations on a single character or several characters, respectively, of each experimental unit. These methods are fairly well-developed now and form a part of the classical statistical theory.

The position is not so satisfactory when the assumption of normality is discarded. This is particularly true when the experimental data are categorical in nature, i.e., are given in the form of frequencies in cells determined by a finitely multi-way cross-classification with predefined categories along each way of classification. Even though analysis of categorical data goes back to the pioneering work of Karl Pearson (1900) and has been developed at subsequent stages among others, by Fisher (1922), Cramer (1946) and, in particular, Neyman (1949), a lot remained to be done.

Barnard (1947) and Pearson (1947) pointed out by considering the simple 2 x 2 table that the same 2 x 2 table could be the outcome of different sampling schemes which makes it necessary to assume appropriate probability models, which may lead to different statistical procedures with obviously different interpretations appropriate to each experimental situation. This line of thought was developed extensively by Roy and his students; refer for example to Mitra (1955), Roy and Mitra (1956), Bhapkar (1959), Roy and

Bhapkar (1960). Formulations of some of these categorical data problems as analogs of ANOVA, and MANOVA and "Normal association" problems were offered first by Roy and Mitra (1956) and later on by Roy and Bhapkar (1960); statistical procedures to handle these were given in Roy and Mitra (1956) and Bhapkar (1961) respectively. The present paper is in the same line and develops further some such methods. It deals exclusively with suitable formulations of hypotheses and appropriate tests for multi-response situations.

Section 2 gives notation and preliminaries, section 3 includes some further results and section 4 gives specific test criteria for some hypotheses; the test statistic mentioned is usually the Pearson -χ^2 statistic if the maximum likelihood estimates are easy to obtain, and otherwise is the Neyman-χ^2_1 statistic computed from results in section 3.

2. Notation and Preliminaries

Suppose that s independent random samples of experimental units are taken from s populations, n_{oj} is the size of the sample from the j-th population and n_{ij} is the observed frequency in the i-th category of the j-th sample, i=1,2,...,r; j=1,2,...s. We assume that p_{ij} is the probability that an experimental unit drawn at random from the j-th population belongs to the i-th category and that either the sampling is with replacement or, if without replacement, sampling fractions are negligible, so that the probability distribution of the observed frequenceies is given by

$$(2.1) \qquad \phi = \prod_{j=1}^{s} \left(\frac{n_{oj}!}{\prod_{i=1}^{r} n_{ij}!} \prod_{i=1}^{r} p_{ij}^{n_{ij}} \right) \quad ,$$

where $\sum_i n_{ij} = n_{oj}$ a given integer and $\sum_i p_{ij} = 1$ for each j = 1,2,...,s;

zero in place of a suffix will indicate sum over that suffix. Let $N = \Sigma_j n_{oj}$, $q_{ij} = n_{ij}/n_{oj}$, $Q_j = n_{oj}/N$, $\underset{\sim}{p}' = [p_{11},\ldots,p_{(r-1)1};\ldots;p_{1s},\ldots,p_{(r-1)s}]$ and $\underset{\sim}{q}' = [q_{11},\ldots,q_{(r-1)1};\ldots;q_{1s},\ldots,q_{(r-1)s}]$.

According as the marginal frequencies along any dimension or way of classification are held fixed or left free by the experimental scheme that dimension will be said to be a "factor" or a "response". Thus in the above model (2.1) $\underline{i}$ refers to response categories while $\underline{j}$ refers to factor categories; n_{io} is a random variable while n_{oj} is a fixed integer. $\underline{i}$ may be a multiple subscript, say, $(i_1 i_2 \ldots i_k)$ with $i_\alpha = 1,2,\ldots,r_\alpha$, $\alpha = 1,2,\ldots,k$ so that $r = r_1 r_2 \ldots r_k$; similarly, j also might be a multiple subscript, say $j_1 j_2 \ldots j_\ell$ with $j_\beta = 1,2,\ldots,s_\beta$, $\beta = 1,2,\ldots,\ell$ but with the distinction that all combinations may not be selected for the experiment. This will be called a k-response (or k-variate) and ℓ-factor problem where i_α refer to a category of the $\underline{\alpha}$-th response while j_β to that of the $\underline{\beta}$th factor.

If a set of real values (scroes) is associated with the categories along any way of classification (factor or response), that way of classification will be said to be structured. These may be, for example, the mid-points of class-intervals for a response (or factor) or the values themselves if the response (or factor) is discrete or may be any scores assigned on some other considerations even for a way of classificat ion without any implied ranking, to start with, for its categories.

Suppose that we have to test the hypothesis

$$(2.2) \qquad H_o: \; F_m(\underset{\sim}{p}) = 0 \qquad , \; m = 1,2,\ldots = t \; (t \leq rs-s)$$

where F's are t independent given functions of $\underset{\sim}{p}$. It is assumed that F's

possess continuous partial derivatives up to the second order and that the rank of the $tx(rs-s)$ matrix $[\partial F_m(\underset{\sim}{p})/\partial p_{ij}]$ is t. It is assumed that there is at least one solution such that $p_{ij} > 0$ for all i,j. It is then well known (e.g. refer to [11]) that H_o can be tested in various ways by using either the χ^2, χ_1^2 or the likelihood-ratio statistic λ defined, respectively, by

$$\chi^2 = \sum_{j=1}^{s} \sum_{i=1}^{r} (n_{ij}-n_{oj}\hat{p}_{ij})^2/n_{oj}\hat{p}_{ij}$$

$$\chi_1^2 = \sum_{j=1}^{s} \sum_{i=1}^{r} (n_{ij}-n_{oj}\hat{p}_{ij})^2/n_{ij} \tag{2.3}$$

$$-2 \log \lambda = 2 \sum_{j=1}^{s} \sum_{i=1}^{r} n_{ij} \{\log n_{ij} - \log n_{oj}\hat{p}_{ij}\} ,$$

where $\hat{p}$'s are any BAN [11] estimates of p's obtained subject to constraints (2.2). Neyman (1949) has shown that, in particular, estimates minimizing χ^2 or χ_1^2 or those maximizing ϕ are BAN estimates. If the p's are subject to constraints (2.2), the equations giving these estimates are, in general, complicated to solve; the minimum - χ_1^2 estimates, though, can be obtained by solving only linear equations whenever the constraints (2.2) are linear in p's. If the functions F_m are not linear, Neyman (1949) has proposed the technique of 'linearization' to reduce the problem to the linear case whereby minimum - χ_1^2 estimates are obtained subject to constraints

$$F_m^*(\underset{\sim}{p}) \equiv F_m(\underset{\sim}{q}) + \sum_{j=1}^{s} \sum_{i=1}^{r-1} \left(\frac{\partial F_m(\underset{\sim}{p})}{\partial p_{ij}}\right)_{\underset{\sim}{p}=\underset{\sim}{q}} (p_{ij}-q_{ij}) = 0 , \quad m = 1,2,\ldots,t .$$

Neyman has proved that each of the statistics in (2.3), using any system of BAN estimates (using linearization if necessary), has a limiting chi-square distribution with t degrees of freedom as $N \to \infty$ with Q's fixed if H_o holds; he has also proved that these tests are asymptotically equivalent in the sense that the probability of any two of them contradicting each other tends to 0

as $N \to \infty$ irrespective of whether H_o is true or false. The author (1965) has shown recently that the χ_1^2 statistic, whenever it is defined, is identical to Wald's statistic (1943) as adapted to the categorical situation and, hence, possesses the same asymptotic optimality properties as those possessed by Wald's statistic and the likelihood-ratio statistic as shown by Wald for the case of sampling from one population (i.e. for $s = 1$) and conjectured for the general case (i.e. for $s \geq 2$).

3. Some Further Results

Theorem 3.1. Let H_o be the hypothesis specified by t independent constraints

$$(3.1) \qquad F_m(p) = \sum_{j=1}^{s} \sum_{i=1}^{r} f_{mij} p_{ij} + f_m \equiv \sum_{j=1}^{s} \sum_{i=1}^{r-1} f^*_{mij} p_{ij} + f^*_m = 0$$

$$m = 1, \ldots, t,$$

where $f^*_{mij} = f_{mij} - f_{mrj}$ and $f^*_m = f_m + \sum_{j=1}^{s} f_{mrj}$; we assume that these are independent of the basic constraints $\sum_{i=1}^{r} p_{ij} = 1$. Let

$$\underset{\sim}{c}' = (c_1 \ldots c_t) \quad \text{with} \quad c_m = F_m(q)$$

$$\underset{\sim}{G} = [g_{mm'}] \qquad m, m' = 1, 2, \ldots, t$$

with

$$g_{mm'} = \sum_{j=1}^{s} n_{oj}^{-1} \{ \sum_{i=1}^{r} f_{mij} f_{m'ij} q_{ij} - (\sum_{i=1}^{r} f_{mij} q_{ij})(\sum_{i=1}^{r} f_{m'ij} q_{ij}) \}$$

$$= \sum_{j=1}^{s} n_{oj}^{-1} \{ \sum_{i=1}^{r-1} f^*_{mij} f^*_{m'ij} q_{ij} - (\sum_{i=1}^{r-1} f^*_{mij} q_{ij})(\sum_{i=1}^{r-1} f^*_{m'ij} q_{ij}) \} .$$

Then if H_o holds, the statistic

$$(3.2) \qquad \underset{\sim}{c}' \underset{\sim}{G}^{-1} \underset{\sim}{c}$$

has a limiting chi-square distribution with t d.f. as $N \to \infty$ with Q's

remaining fixed.

Proof: It can be shown easily (e.g. refer to [2]) that (3.2) is the χ^2_1-statistic to test H_o and the theorem then follows from Neyman's results.

It may be noted that $\underset{\sim}{G}$ is the matrix obtained after replacing $\underset{\sim}{p}$ by $\underset{\sim}{q}$ in the covariance matrix of $\underset{\sim}{c}$ and hence will be referred to as the 'sample covariance matrix' of $\underset{\sim}{c}$; it is nonsingular almost everywhere in view of the assumed conditions.

Theorem 3.2 Let a linear hypothesis be defined by

$$(3.3)\qquad \sum_{i=1}^{r} a_{\beta i}p_{ij} = d_{j1}\theta_{\beta 1} + d_{j2}\theta_{\beta 2} + \ldots + d_{ju}\theta_{\beta u}\ , \quad \beta = 1,2,\ldots,k\ ,$$

where a's and d's are known constants, θ's are unknown parameters, $\underset{\sim}{D} = [d_{j\gamma}]_{sxu}$ with rank $\underset{\sim}{D} = v < s$, and the linear functions on the left in (3.3) are linearly independent and also linearly independent of $\Sigma_i p_{ij}$ ($\equiv 1$, of course). Suppose

$$\alpha_{\beta j} = \Sigma_i a_{\beta i} q_{ij}\ , \qquad \underset{\sim}{\alpha}_j' = [\alpha_{1j}, \ldots, \alpha_{kj}]\ ,$$

$$\underset{\sim}{\theta}_\gamma' = [\theta_{1\gamma}, \ldots, \theta_{k\gamma}]$$

(3.4) and

$$\underset{\sim}{\Lambda}_j = [\lambda_{\beta\beta' j}]_{kxk}\ , \qquad \beta,\beta' = 1,2,\ldots,k$$

with

$$\lambda_{\beta\beta' j} = n_{oj}^{-1}(\Sigma_i a_{\beta i}a_{\beta' i}q_{ij} - \alpha_{\beta j}\alpha_{\beta' j})\ .$$

Then the χ^2_1-statistic to test (3.3) is equal to the minimum value of

$$(3.5)\qquad S^2 = \sum_{j=1}^{s} (\underset{\sim}{\alpha}_j - d_{j1}\underset{\sim}{\theta}_1 - \ldots - d_{ju}\underset{\sim}{\theta}_u)'\underset{\sim}{\Lambda}_j^{-1}(\underset{\sim}{\alpha}_j - \ldots - d_{ju}\underset{\sim}{\theta}_u)$$

with respect to the $\underset{\sim}{\theta}$'s and has $k(s-v)$ degrees of freedom.

Proof: Let $\underset{\sim}{B} = [b_{\delta j}]$ be a (s-v)xs matrix of rank s-v such that $\underset{\sim}{BD} = \underset{\sim}{0}$, i.e.,

$$\sum_j b_{\delta j} d_{j\gamma} = 0 , \qquad \delta = 1,2,\ldots,s-v; \quad \gamma=1,2,\ldots,u .$$

(3.3), then, implies that

$$\sum_j b_{\delta j} \sum_i a_{\beta i} p_{ij} = \sum_j b_{\delta j} \sum_\gamma d_{j\gamma} \theta_{\beta\gamma} = \sum_\gamma \theta_{\beta\gamma} \sum_j b_{\delta j} d_{j\gamma} = 0 .$$

Conversely, $\sum_j b_{\delta j} \sum_i a_{\beta i} p_{ij} = 0$ implies that the vector $(\Sigma_i a_{\ i} p_{ij}$, $j = 1,2,\ldots,s)$ belongs to the vectorspace orthogonal to that generated by the rows of $\underset{\sim}{B}$ and, hence, belongs to the vectorspace generated by the columns of $\underset{\sim}{D}$; thus there exist θ's such that $\Sigma_i a_{\beta i} p_{ij} = \Sigma_\gamma \theta_{\beta\gamma} d_{j\gamma}$ which means (3.3) holds. Thus (3.3) is equivalent to k(s-v) independent linear constraints

$$(3.6) \qquad \sum_j \sum_i b_{\delta j} a_{\beta i} p_{ij} = 0 \qquad \begin{matrix} \beta = 1,2,\ldots,k \\ \delta = 1,2,\ldots,s-v . \end{matrix}$$

The χ^2_1-statistic with k(s-v) degrees of freedom to test (3.6) can be derived, then, from theorem 3.1. Let

$$c_{\delta\beta} = \sum_j \sum_i b_{\delta j} a_{\beta i} q_{ij} = \sum_j b_{\delta j} \alpha_{\beta j}$$

$$\underset{\sim}{c}_\delta' = [c_{\delta 1}, c_{\delta 2}, \ldots, c_{\delta k}], \qquad \underset{\sim}{c}' = [\underset{\sim}{c}_1', \ldots, \underset{\sim}{c}_{s-v}'] ,$$

so that

$$\underset{\sim}{c}_\delta = \sum_j b_{\delta j} \underset{\sim}{\alpha}_j$$

and, hence,

$$(3.7) \qquad \underset{\sim}{c} = \underset{\sim}{B}^* \underset{\sim}{\alpha} ,$$

where

$$\underset{\sim}{B}^* = \begin{bmatrix} b_{11}\underset{\sim}{I}_k & \cdots & b_{1s}\underset{\sim}{I}_k \\ & \cdots & \\ b_{s-v,1}\underset{\sim}{I}_k & \cdots & b_{s-v,s}\underset{\sim}{I}_k \end{bmatrix} = \underset{\sim}{B} \otimes \underset{\sim}{I}_k ,$$

the Kronecker product(sometimes also called the Direct product) of $\underset{\sim}{B}$ and $\underset{\sim}{I}_k$. Also

$$\text{Cov}\ (\underset{\sim}{c}) = \underset{\sim}{B}^{*}\ \text{Cov}(\underset{\sim}{\alpha})\ \underset{\sim}{B}^{*'} = \underset{\sim}{B}^{*}\underset{\sim}{\Sigma}\ \underset{\sim}{B}^{*'} \quad ,$$

where

$$\underset{\sim}{\Sigma} = \begin{bmatrix} \underset{\sim}{\Sigma}_1 & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{\Sigma}_2 & \cdots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \cdots & \underset{\sim}{\Sigma}_s \end{bmatrix} \quad ,$$

and $\underset{\sim}{\Sigma}_j$ is the covariance matrix of $\underset{\sim}{\alpha}_j$. Note that replacing p by q in $\underset{\sim}{\Sigma}_j$ gives $\underset{\sim}{\Lambda}_j$ defined by (3.4). Thus, the sample covariance matrix' of $\underset{\sim}{c}$ is $\underset{\sim}{B}^{*}\ \underset{\sim}{\Lambda}\ \underset{\sim}{B}^{*'}$, where

$$\underset{\sim}{\Lambda} = \begin{bmatrix} \underset{\sim}{\Lambda}_1 & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{\Lambda}_2 & \cdots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \cdots & \underset{\sim}{\Lambda}_s \end{bmatrix} \quad ,$$

and the χ_1^2-statistic, in view of (3.2), is

$$(3.8) \qquad \underset{\sim}{\alpha}'\underset{\sim}{B}^{*'}(\underset{\sim}{B}^{*}\underset{\sim}{\Lambda}\ \underset{\sim}{B}^{*'})^{-1}\underset{\sim}{B}^{*}\underset{\sim}{\alpha} \quad .$$

On the other hand minimum value of S^2 with respect to $\underset{\sim}{\theta}$'s is seen to be

$$(3.9) \qquad \sum_j \underset{\sim}{\alpha}_j' \underset{\sim}{\Lambda}_j^{-1}\underset{\sim}{\alpha}_j - \sum_{\gamma=1}^{u} \hat{\underset{\sim}{\theta}}_\gamma' \underset{\sim}{\psi}_\gamma \quad ,$$

where

$$\underset{\sim}{\psi}_\gamma = \sum_j d_{j\gamma}\underset{\sim}{\Lambda}_j^{-1}\underset{\sim}{\alpha}_j \quad ,$$

and $\hat{\underset{\sim}{\theta}}' \equiv [\hat{\underset{\sim}{\theta}}_1', \ldots, \hat{\underset{\sim}{\theta}}_u']$ is a solution of

$$\underset{\sim}{\psi}_\gamma = \sum_{\epsilon=1}^{u} \underset{\sim}{\Phi}_{\gamma\epsilon}\hat{\underset{\sim}{\theta}}_\epsilon \quad ,$$

where

$$\underset{\sim}{\Phi}_{\gamma\epsilon} = \sum_j d_{j\gamma} d_{j\epsilon} \underset{\sim}{\Lambda}_j^{-1} \quad .$$

If we let

$$\underset{\sim}{\Psi}' = [\underset{\sim}{\Psi}_1', \ldots, \underset{\sim}{\Psi}_u'] \quad , \qquad \underset{\sim}{D}^* = \underset{\sim}{D} \otimes \underset{\sim}{I}_k$$

and

$$\underset{\sim}{\Phi} = \begin{bmatrix} \underset{\sim}{\Phi}_{11}, \ldots, \underset{\sim}{\Phi}_{1u} \\ \cdots \\ \underset{\sim}{\Phi}_{u1} \quad \ldots, \underset{\sim}{\Phi}_{uu} \end{bmatrix} \quad ,$$

then it follows that

$$\underset{\sim}{\Psi} = \underset{\sim}{D}^{*'} \underset{\sim}{\Lambda}^{-1} \underset{\sim}{\alpha} \quad , \qquad \underset{\sim}{\Phi} = \underset{\sim}{D}^{*'} \underset{\sim}{\Lambda}^{-1} \underset{\sim}{D}^*$$

(3.10) and

$$\underset{\sim}{\Psi} = \underset{\sim}{\Phi} \hat{\underset{\sim}{\theta}} \quad ;$$

hence (3.9) reduces to

$$(3.11) \qquad \underset{\sim}{\alpha}' \underset{\sim}{\Lambda}^{-1} \underset{\sim}{\alpha} - \hat{\underset{\sim}{\theta}}' \underset{\sim}{D}^{*'} \underset{\sim}{\Lambda}^{-1} \underset{\sim}{\alpha} \quad .$$

We have to show that (3.8) is equal to (3.11), taking into account (3.10) and the fact that $\underset{\sim}{B}\,\underset{\sim}{D} = \underset{\sim}{0}$, i.e., $\underset{\sim}{B}^* \underset{\sim}{D}^* = \underset{\sim}{0}$.

Now $\underset{\sim}{D}$ is a s x u matrix of rank v so that $\underset{\sim}{D}^*$ is a ks x ku matrix of rank kv; there exist then a ks x kv matrix $\underset{\sim}{M}$ and a kv x ku matrix $\underset{\sim}{E}$, both of rank kv, such that

$$\underset{\sim}{D}^* = \underset{\sim}{M}\,\underset{\sim}{E} \quad .$$

From (3.10) then we have

$$\underset{\sim}{E}'\,\underset{\sim}{M}'\,\underset{\sim}{\Lambda}^{-1}\,\underset{\sim}{\alpha} = \underset{\sim}{E}'\,\underset{\sim}{M}'\,\underset{\sim}{\Lambda}^{-1}\,\underset{\sim}{M}\,\underset{\sim}{E}\,\hat{\underset{\sim}{\theta}} \quad ,$$

which gives

$$\underset{\sim}{\eta} = (\underset{\sim}{M}'\underset{\sim}{\Lambda}^{-1}\underset{\sim}{M})^{-1}\underset{\sim}{M}'\underset{\sim}{\Lambda}^{-1}\underset{\sim}{\alpha} \quad ,$$

where $\underset{\sim}{\eta} = \underset{\sim}{E}\,\hat{\underset{\sim}{\theta}}$.

The second term in (3.11) is then $\underset{\sim}{\eta}'\underset{\sim}{M}'\underset{\sim}{\Lambda}^{-1}\underset{\sim}{\alpha}$.

The theorem follows if we show that

$$(3.12) \qquad \underset{\sim}{B}^{*'}(\underset{\sim}{B}^{*}\underset{\sim}{\Lambda}\,\underset{\sim}{B}^{*'})^{-1}\underset{\sim}{B}^{*} = \underset{\sim}{\Lambda}^{-1} - \underset{\sim}{\Lambda}^{-1}\underset{\sim}{M}(\underset{\sim}{M}'\underset{\sim}{\Lambda}^{-1}\underset{\sim}{M})^{-1}\underset{\sim}{M}'\underset{\sim}{\Lambda}^{-1},$$

noting that $\underset{\sim}{B}^{*}\underset{\sim}{M} = \underset{\sim}{0}$ since $\underset{\sim}{B}^{*}\underset{\sim}{D}^{*} = \underset{\sim}{0}$. Let $\underset{\sim}{R}$, $\underset{\sim}{S}$, $\underset{\sim}{T}$ be nonsingular matrices such that

$$\underset{\sim}{R}\underset{\sim}{\Lambda}\underset{\sim}{R}' = \underset{\sim}{I}, \quad \underset{\sim}{S}(\underset{\sim}{B}^{*}\underset{\sim}{\Lambda}\,\underset{\sim}{B}^{*'})\underset{\sim}{S}' = \underset{\sim}{I} \ , \quad \underset{\sim}{T}\,\underset{\sim}{M}'\,\underset{\sim}{\Lambda}^{-1}\underset{\sim}{M}\,\underset{\sim}{T}' = \underset{\sim}{I}$$

so that

$$\underset{\sim}{\Lambda}^{-1} = \underset{\sim}{R}'\underset{\sim}{R} \ , \quad \underset{\sim}{B}^{*}\underset{\sim}{\Lambda}\,\underset{\sim}{B}^{*'} = (\underset{\sim}{S}'\underset{\sim}{S})^{-1} \quad \text{and}$$

$$\underset{\sim}{M}'\,\underset{\sim}{\Lambda}^{-1}\,\underset{\sim}{M} = (\underset{\sim}{T}'\underset{\sim}{T})^{-1} \ .$$

Then

$$\begin{bmatrix} \underset{\sim}{S} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{T} \end{bmatrix}\begin{bmatrix} \underset{\sim}{B}^{*}\underset{\sim}{R}^{-1} \\ \underset{\sim}{M}'\underset{\sim}{R}' \end{bmatrix}\left[\, \underset{\sim}{R}'^{-1}\underset{\sim}{B}^{*'} \mid \underset{\sim}{R}\,\underset{\sim}{M} \,\right]\begin{bmatrix} \underset{\sim}{S}' & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{T}' \end{bmatrix} = \underset{\sim}{I}$$

which implies

$$\left[\, \underset{\sim}{R}'^{-1}\underset{\sim}{B}^{*'} \mid \underset{\sim}{R}\,\underset{\sim}{M} \,\right]\begin{bmatrix} \underset{\sim}{S}' & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{T}' \end{bmatrix}\begin{bmatrix} \underset{\sim}{S} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{T} \end{bmatrix}\begin{bmatrix} \underset{\sim}{B}^{*}\underset{\sim}{R}^{-1} \\ \underset{\sim}{M}'\underset{\sim}{R}' \end{bmatrix} = \underset{\sim}{I} \ ,$$

that is,

$$\underset{\sim}{R}'^{-1}\underset{\sim}{B}^{*'}\underset{\sim}{S}'\underset{\sim}{S}\underset{\sim}{B}^{*}\underset{\sim}{R}^{-1} + \underset{\sim}{R}\,\underset{\sim}{M}\,\underset{\sim}{T}'\underset{\sim}{T}\,\underset{\sim}{M}'\underset{\sim}{R}^{-1} = \underset{\sim}{I} \ ,$$

which leads to (3.12). Q.E.D.

Note that α's are the natural unbiased estimates of the quantities on the left in (3.3) while $\underset{\sim}{\Lambda}_j$ is the 'sample covariance matrix' of $\underset{\sim}{\alpha}_j$.

4. Applications.

A. Two-dimensional tables with both dimensions responses.

Here we have a single population with $\underline{i}$ a double subscript $(i_1 i_2)$, $i_\alpha = 1,2,\ldots,r_\alpha$; $\alpha = 1,2$ and $r = r_1 r_2$. The probabilities $p_{i_1 i_2}$ are subject to the constraint $\Sigma_{i_1,i_2} p_{i_1 i_2} = 1$.

(i) Hypothesis of independence: This hypothesis is expressed by the condition

$$H_1: \quad p_{i_1 i_2} = p_{i_1 0} p_{0 i_2} \ , \tag{4.A.1}$$

where, of course, $p_{i_1 0} = \Sigma_{i_2} p_{i_1 i_2}$ and $p_{0 i_2} = \Sigma_{i_1} p_{i_1 i_2}$. The well known χ^2-statistic to test H_1 is

$$\sum_{i_1,i_2} \left(n_{i_1 i_2} - \frac{n_{i_1 0} n_{0 i_2}}{N} \right)^2 \Big/ \left(\frac{n_{i_1 0} n_{0 i_2}}{N} \right) \ , \quad \text{d.f.} = (r_1 - 1)(r_2 - 1). \tag{4.A.2}$$

H_1 has been already seen ([15]) to be an obvious analog of the hypothesis of independence (i.e. no correlation) in the bivariate normal analysis.

(ii) Hypothesis of equality of two marginal distributions:

Assuming $r_1 = r_2$, this hypothesis is expressed by the condition

$$H_2: \quad p_{i_1 0} = p_{0 i_1} \quad ; \tag{4.A.3}$$

this may be seen to be an analog of the hypothesis $\mu_1 = \mu_2$ and $\sigma_{11} = \sigma_{22}$, in the usual notation, in the normal analysis. Let

$$h_{i_1}(\underset{\sim}{q}) = q_{i_1 0} - q_{0 i_1} \qquad i_1 = 1,2,..,r_1 - 1 \ ,$$

and

$$N\,\underset{\sim}{G} = \left[\delta_{i_1 i_2}(q_{i_1 0} + q_{0 i_2}) - q_{i_1 i_2} - q_{i_2 i_1} - h_{i_1}(\underset{\sim}{q}) h_{i_2}(\underset{\sim}{q})\right]$$

$$i_1, i_2 = 1, 2, \ldots, r_1 - 1 \ ,$$

where $\delta_{i_1 i_2} = 1$ if $i_1 = i_2$ and 0 otherwise. Then the χ^2_1-statistic is seen to be

(4.A.4) $$\underset{\sim}{h}'(\underset{\sim}{q})\underset{\sim}{G}^{-1}\underset{\sim}{h}(\underset{\sim}{q}) \quad , \quad \text{d.f.} = r_1 - 1 \ .$$

The same expression had been obtained by Sathe (1962) for Wald's criterion and the two statisticsmust be identical as observed by the author (1965). The statistic (4.A.4) differs from the one proposed by Stuart (1955) in that the latter deletes the lastterm in the rectangular brackets for $\underset{\sim}{G}$; our statistic should be preferred since $N\underset{\sim}{G}$ is a consistant estimator of the covariance matrix of $\sqrt{N}\,\underset{\sim}{h}(\underset{\sim}{q})$ even when H_2 is false while the matrix used by Stuart is consistent only when H_2 holds.

(iii) Hypothesis of equality of "means" of two variables:

Let us now suppose that the two responses are "structured", i.e., have an implied ranking, and we have a system of scores $\{a_{i_1}\}$ associated with the categories of the first response with a similar system $\{b_{i_2}\}$ for the second response. Then the above hypothesis may be expressed in the form

(4.A.5) $$H_3: \quad \sum_{i_1} a_{i_1} p_{i_1 0} = \sum_{i_2} b_{i_2} p_{0 i_2} \ .$$

Let

$$c = \sum_{i_1} a_{i_1} q_{i_1 0} - \sum_{i_2} b_{i_2} q_{0 i_2}$$

and

$$N g = \sum_{i_1} a_{i_1}^2 q_{i_1 0} + \sum_{i_2} b_{i_2}^2 q_{0 i_2} - 2 \sum_{i_1} \sum_{i_2} a_{i_1} b_{i_2} q_{i_1 i_2} - c^2 \ ;$$

then the χ^2_1-statistic is seen to be

(4.A.6) $\qquad c^2/g \qquad\qquad$ d.f. = 1 .

H_3 may be considered as an analog of the hypothesis of equality of means of two possibly correlated variables in the normal analysis. Note that in the special case $r_1 = r_2$ with $a_{i_1} = b_{i_1}$, $H_2 \Longrightarrow H_3$ and thus H_3 is a weaker hypothesis than H_2 and may be of interest if H_2 does not hold.

B. Three-dimensional tables with all dimensions responses.

Here again we have a single population, but now $\underline{i}$ is a triple subscript $(i_1 i_2 i_3)$, $i_\alpha = 1,2,\ldots,r_\alpha$, $\alpha = 1,2,3$ and $r = r_1 r_2 r_3$, the probabilities $p_{i_1 i_2 i_3}$ are subject to the constraint $\Sigma_{i_1 i_2 i_3} p_{i_1 i_2 i_3} = 1$.

(i) Hypothesis of complete independence:

(4.B.1) $\qquad H_4: \; p_{i_1 i_2 i_3} = p_{i_1 00} p_{0 i_2 0} p_{00 i_3}$

(ii) Hypothesis of independence of the first response with the last two:

(4.B.2) $\qquad H_5: \; p_{i_1 i_2 i_3} = p_{i_1 00} p_{0 i_2 i_3}$

(iii) Hypothesis of independence of the first two responses given the third response:

(4.B.3) $\qquad H_6: \; p_{i_1 i_2 i_3} = \dfrac{p_{i_1 0 i_3} p_{0 i_2 i_3}}{p_{00 i_3}}$

It has been pointed out by Roy and Mitra (1956) that H_5, H_6 can be considered analogs of the hypotheses of no multiple correlation and no partial correlation, respectively, in the normal analysis while H_4 is that of the hypothesis of zero correlations. The χ^2-statistics are knownto be

$$N^2 \sum_{i_1,i_2,i_3} \left(n_{i_1i_2i_3} - \frac{n_{i_100}n_{0i_20}n_{00i_3}}{N^2}\right)^2/n_{i_100}n_{0i_20}n_{00i_3},$$

$$\text{d.f.} = r_1r_2r_3 - r_1 - r_2 - r_3 + 2$$

(4.B.4)
$$N \sum_{i_1,i_2,i_3} \left(n_{i_1i_2i_3} - \frac{n_{i_100}n_{0i_2i_3}}{N}\right)^2/n_{i_100}n_{0i_2i_3}, \quad \text{d.f.} = (r_1-1)(r_2r_3-1)$$

$$\sum_{i_1,i_2,i_3} \left(n_{i_1i_2i_3} - \frac{n_{i_10i_3}n_{0i_2i_3}}{n_{00i_3}}\right)^2/\left(\frac{n_{i_10i_3}n_{0i_2i_3}}{n_{00i_3}}\right), \quad \text{d.f.} = (r_1-1)(r_2-1)r_3$$

for testing H_4, H_5 and H_6 respectively.

It may be pointed out here that the hypothesis of pairwise independence of three responses

$$H_7: p_{i_1i_20} = p_{i_100}p_{0i_20}, \quad p_{0i_2i_3} = p_{0i_20}p_{00i_3}, \quad p_{i_10i_3} = p_{i_100}p_{00i_3},$$

that of pairwise independence of the first two (separately) with the third, viz.,

$$H_8 = p_{i_10i_3} = p_{i_100}p_{00i_3}, \quad p_{0i_2i_3} = p_{0i_20}p_{00i_3},$$

and the hypothesis of "no interaction" between three responses (see, for example, [4] or [9]) are hypotheses which have no formal analogs in the normal analysis where pairwise independence between two sets of variables is equivalent to the complete independence of the two sets of variables.

(iv) Hypothesis of equality of three marginal distributions is specified by the conditions

(4.B.5)
$$H_9: p_{i_100} = p_{0i_10} = p_{00i_1},$$

assuming, of course, that $r_1=r_2=r_3$. Let

$$\underset{\sim}{q}_1' = [q_{100}, q_{200}, \cdots, q_{(r_1-1)00}] ,$$

$$\underset{\sim}{q}_2' = [q_{010}, q_{020}, \cdots, q_{0(r_1-10}] ,$$

$$\underset{\sim}{q}_3' = [q_{001}, q_{002}, \cdots, q_{00(r_1-1)}] , \qquad q' = [q_1', q_2', q_3'] ,$$

$$\underset{\sim}{Q}_{11} = \text{diagonal } (q_{i_1 00},\ i_1 = 1,2,..,r_1-1) \text{ etc.}$$

$$\underset{\sim}{Q}_{12} = [q_{i_1 i_2 0}] \quad i_1, i_2 = 1,2,\ldots,r_1-1 \text{ etc.}$$

(4.B.6)
$$N\underset{\sim}{\Lambda} = [\underset{\sim}{Q}_{\alpha\beta} - \underset{\sim}{q}_\alpha \underset{\sim}{q}_\beta'] \qquad \alpha,\beta = 1,2,3 ,$$

$$\underset{\sim}{\Lambda}^{-1} = \underset{\sim}{M} \equiv [\underset{\sim}{M}_{\alpha\beta}] \qquad \alpha,\beta = 1,2,3 ,$$

$$\underset{\sim}{M}_\beta = \Sigma_\alpha \underset{\sim}{M}_{\alpha\beta} , \quad \underset{\sim}{M}_0 = \Sigma_\beta \underset{\sim}{M}_\beta ,$$

and finally

$$\underset{\sim}{m} = \Sigma_\beta \underset{\sim}{M}_\beta \underset{\sim}{q}_\beta .$$

Note that $\underset{\sim}{\Lambda}$ is the 'sample covariance matrix' of $\underset{\sim}{q}$ and, hence, is nonsingular almost everywhere excluding, of course, the degenerate case where some variables (or rather the associated probabilities) are linear functions of the remaining variables (i.e., their corresponding probabilities). By applying theorem 3.2 it can be shown that the χ_1^2-statistic to test H_9 is given by

(4.B.7)
$$\sum_{\alpha,\beta} \underset{\sim}{q}_\alpha' \underset{\sim}{M}_{\alpha\beta} \underset{\sim}{q}_\beta - \underset{\sim}{m}' \underset{\sim}{M}_0^{-1} \underset{\sim}{m} , \quad \text{d.f.} = 2(r_1-1) .$$

The method can be immediately extended to the case of k variables; the statistic then has $(k-1)(r_1-1)$ degrees of freedom. The case $k = 2$ leads to the statistic (4.A.4). Cochran (1950) has considered the k-variate problem only for the special case $r_1 = r_2 = \ldots = r_k = 2$; even for this special case the statistic (4.B.7) is expected to be more efficient.

H_9 is easily seen to be an analog of the hypothesis $\mu_1 = \mu_2 = \mu_3$ and $\sigma_{11} = \sigma_{22} = \sigma_{33}$ in the normal analysis.

(v) Hypothesis of equality of "means" of three variables:

Assume now that the responses are "structured" with $\{a_{i_1}\}$, $\{b_{i_2}\}$ and $\{c_{i_3}\}$ as the scores associated with the respective categories. Then the above hypothesis is expressed by

$$(4.B.8) \qquad H_{10}: \sum_{i_1} a_{i_1} p_{i_1 00} = \sum_{i_2} b_{i_2} p_{0 i_2 0} = \sum_{i_3} c_{i_3} p_{00 i_3} .$$

Let

$$\gamma_1 = \sum_{i_1} a_{i_1} q_{i_1 00} \quad , \quad \gamma_2 = \sum_{i_2} b_{i_2} q_{0 i_2 0} \quad , \quad \gamma_3 = \sum_{i_3} c_{i_3} q_{00 i_3}$$

$$N\lambda_{11} = \left(\sum_{i_1} a_{i_1}^2 q_{i_1 00} \right) - \gamma_1^2 \qquad \text{etc.}$$

$$(4.B.9) \qquad N\lambda_{12} = \left(\sum_{i_1, i_2} a_{i_1} b_{i_2} q_{i_1 i_2 0} \right) - \gamma_1 \gamma_2 \qquad \text{etc.}$$

$$\underset{\sim}{\Lambda} = [\lambda_{\alpha\beta}] , \qquad [m_{\alpha\beta}] \equiv \underset{\sim}{M} = \underset{\sim}{\Lambda}^{-1} ,$$

and finally

$$m_\beta = \Sigma_\alpha m_{\alpha\beta} \quad , \qquad m_0 = \Sigma_\beta m_\beta \quad , \quad m = \Sigma_\beta m_\beta \gamma_\beta \quad ,$$

with $\alpha, \beta = 1,2,3$. By applying Theorem 3.2 it can be shown that the χ_1^2-statistic is

$$(4.B.10) \qquad \sum_{\alpha,\beta} m_{\alpha\beta} \gamma_\alpha \gamma_\beta - (m^2/m_o) \quad , \text{ d.f.} = 2 .$$

This can be immediately extended to the case of k variables where the statistic would have $k-1$ degrees of freedom.

H_{10} is seen to be an analog of the hypothesis of equality of means of three (possibly correlated) variables in the normal analysis.

C. Three-dimensional tables with two responses and one factor

With the basic set up (2.1) we have now s populations indicated by the subscript j; also i is a double subscript $(i_1 i_2)$, $i_\alpha = 1,2,\ldots,r_\alpha$, $\alpha = 1,2$ so that $r = r_1 r_2$ and the probabilities $p_{i_1 i_2 j}$ are subject to the constraints $\Sigma_{i_1,i_2} p_{i_1 i_2 j} = 1$.

(i) Hypothesis of independence of the two responses:

(4.C.1) $$H_{11}: \quad p_{i_1 i_2 j} = p_{i_1 0 j} p_{0 i_2 j} \ .$$

This implies that H_1 holds for each of the s populations and it is known that we have a χ^2-statistic

(4.C.2) $$\sum_j \sum_{i_1,i_2} (n_{i_1 i_2 j} - \frac{n_{i_1 0 j} n_{0 i_2 j}}{n_{00 j}})^2 / (\frac{n_{i_1 0 j} n_{0 i_2 j}}{n_{00 j}}) \ ,$$

$$\text{d.f.} = (r_1 - 1)(r_2 - 1)s$$

which follows immediately from (4.A.2). H_{11} is seen to be an obvious analog of the hypothesis of independence of two variables in each of s (bivariate) normal populations.

(ii) Hypothesis of "no interaction" between the two responses and one factor essentially means that the nature of association between the two responses is the same over all factor categories, i.e., for all populations; the formulation depends on what measure of association is chosen for comparison.

(4.C.3)

$$H_{12}: \quad \frac{p_{i_1 i_2 j} p_{r_1 r_2 j}}{p_{i_1 r_2 j} p_{r_1 i_2 j}} \quad \text{is independent of } j \ .$$

$$H_{13}: \quad \frac{p_{i_1 i_2 j}}{p_{i_1 0 j} p_{0 i_2 j}} \quad \text{is independent of } j.$$

The formulation H_{12} is due to Goodman (1964) while H_{13} is due to Bhapkar and Koch (1965). These hypotheses are nonlinear and the Wald statistics can be obtained by the "linearization" technique. For details the reader is referred to [9] and [4].

This hypothesis can be seen to be the analog of the hypothesis of equality of s correlations given samples from s bivariate normal populations. Note that the hypothesis H_{11} is a very special case of the hypotheses H_{12}, H_{13}.

(iii) Hypothesis of homogeneity of marginal distributions:

$$(4.C.4) \qquad H_{14}: \quad \begin{matrix} p_{i_1 0j} \text{ is independent of } j \\ p_{0 i_2 j} \text{ also is independent of } j \, . \end{matrix}$$

This is seen to be an analog of the hypothesis that $\mu_{\alpha j}$, $\sigma_{\alpha\alpha j}$, $\alpha = 1,2$, are independent of j, using the usual notation, in the normal analysis.

Let

$$\underset{\sim}{q}'_{1j} = [q_{10j}, q_{20j}, \ldots, q_{(r_1-1)0j}] \, ,$$

$$\underset{\sim}{q}'_{2j} = [q_{01j}, q_{02j}, \ldots, q_{0(r_2-1)j}], \quad \underset{\sim}{q}'_{j} = [\underset{\sim}{q}'_{1j}, \underset{\sim}{q}'_{2j}] \, ,$$

$$\underset{\sim}{Q}_{11j} = \text{diagonal } (q_{i_1 0j}, \; i_1 = 1,2,\ldots,r_1-1) \text{ etc.}$$

$$(4.C.5) \qquad \underset{\sim}{Q}_{12j} = [q_{i_1 i_2 j}] \qquad i_1 = 1,\ldots,r_1-1 \text{ and } i_2 = 1,\ldots,r_2-1 \, ,$$

$$n_{00j} \underset{\sim}{\Delta}_j = [\underset{\sim}{Q}_{\alpha\beta j} - q_{\alpha j} \underset{\sim}{q}'_{\beta j}] \qquad \alpha,\beta = 1,2,$$

$$\underset{\sim}{\Delta}_j^{-1} = \underset{\sim}{M}_j \, , \quad \underset{\sim}{M} = \Sigma_j \underset{\sim}{M}_j \quad \text{and} \quad \underset{\sim}{t} = \Sigma_j \underset{\sim}{M}_j \underset{\sim}{q}_j \, .$$

Note that $\underset{\sim}{\Delta}_j$ is the 'sample covariance matrix' of $\underset{\sim}{q}_j$ and is nonsingular almost everywhere (excluding the degenerate case). Theorem 3.2 immediately gives the χ^2_1-statistic

(4.C.6) $$\sum_j \underset{\sim}{q}_j' \underset{\sim}{M}_j \underset{\sim}{q}_j - \underset{\sim}{t}'\underset{\sim}{M}^{-1}\underset{\sim}{t} \quad , \qquad \text{d.f.} = (r_1+r_2-2)(s-1)$$

to test H_{14}. The extension to the general k-response case is quite obvious giving a statistic of the same form with $\{\Sigma_\alpha(r_\alpha-1)\}(s-1)$ degrees of freedom.

(iv) Hypothesis of equality of 'means': Suppose now that the two responses are 'structured' with $\{a_{i_1}\}$ and $\{b_{i_2}\}$ as the scores. Consider

(4.C.7) $$H_{15}: \quad \begin{matrix} \sum_{i_1} a_{i_1} p_{i_1 0 j} \text{ is independent of } j \\ \sum_{i_2} b_{i_2} p_{0 i_2 j} \text{ also is independent of } j \ ; \end{matrix}$$

this is an obvious analog of the hypothesis of equality of mean-vectors of several populations in MANOVA. Let

$$\gamma_{ij} = \sum_{i_1} a_{i_1} q_{i_1 0 j} \quad , \quad \gamma_{2j} = \sum_{i_2} b_{i_2} q_{0 i_2 j} \quad , \quad \underset{\sim}{\gamma}_j' = [\gamma_{ij}, \gamma_{2j}] \ ,$$

$$n_{00j}\lambda_{11j} = \left(\sum_{i_1} a_{i_1}^2 q_{i_1 0 j} \right) - \gamma_{ij}^2 \quad \text{etc.},$$

$$n_{00j}\lambda_{12j} = \left(\sum_{i_1, i_2} a_{i_1} b_{i_2} q_{i_1 i_2 j} \right) - \gamma_{ij}\gamma_{2j} \ ,$$

(4.C.8) $$\underset{\sim}{\Lambda}_j = [\lambda_{\alpha\beta j}] \ , \quad \alpha,\beta = 1,2, \qquad \underset{\sim}{M}_j = \underset{\sim}{\Lambda}_j^{-1} \ ,$$

and finally

$$\underset{\sim}{M} = \Sigma_j \underset{\sim}{M}_j \qquad \text{and} \qquad \underset{\sim}{t} = \Sigma_j \underset{\sim}{M}_j \underset{\sim}{\gamma}_j \ .$$

The theorem 3.2 gives

(4.C.9) $$\sum_j \underset{\sim}{\gamma}_j' \underset{\sim}{M}_j \underset{\sim}{\gamma}_j - \underset{\sim}{t}'\underset{\sim}{M}^{-1}\underset{\sim}{t} \quad , \qquad \text{d.f.} = 2(s-1) \ ,$$

as the χ_1^2-statistic to test H_{15} . For the k-response problem we have a statistic of the same type but with k(s-1) degrees of freedom.

Note that $H_{14} \Longrightarrow H_{15}$ and the weaker hypothesis H_{15} may be of interest when H_{14} does not hold.

(v) Hypothesis of linearity of regression: Assume now that the factor is also 'structured' with d_j as the score associated with the jth factor-category (or the jth level of the factor). If H_{15} does not hold, we may test the hypothesis

$$(4.C.10) \qquad H_{16}: \begin{aligned} \sum_{i_1} a_{i_1} p_{i_1 0j} &= \xi^{(1)} + \eta^{(1)} d_j \\ \sum_{i_2} b_{i_2} p_{0 i_2 j} &= \xi^{(2)} + \eta^{(2)} d_j \end{aligned} ,$$

where η's are in the nature of regression coefficients, η's and ξ's are unknown. This is an analog of the hypothesis: $\underset{\sim}{\mu}_j = \underset{\sim}{\xi} + \underset{\sim}{\eta}\, d_j$ in MANOVA. In the notation of (4.C.8) and with

$$\underset{\sim}{W} = \Sigma_j \underset{\sim}{M}_j \underset{\sim}{\gamma}_j d_j \ , \quad \underset{\sim}{R} = \Sigma_j \underset{\sim}{M}_j d_j \ , \quad \underset{\sim}{S} = \Sigma_j \underset{\sim}{M}_j d_j^2 \ ,$$

the χ_1^2-statistic is seen to be

$$(4.C.11) \qquad \sum_j \underset{\sim}{\gamma}_j' \underset{\sim}{M}_j \underset{\sim}{\gamma}_j - [\underset{\sim}{t}', \underset{\sim}{W}'] \begin{bmatrix} \underset{\sim}{M} & \underset{\sim}{R} \\ \underset{\sim}{R} & \underset{\sim}{S} \end{bmatrix}^{-1} \begin{bmatrix} \underset{\sim}{t} \\ \underset{\sim}{W} \end{bmatrix} , \quad \text{d.f.} = 2(s-2) .$$

Again the k-response extension is immediate where the statistic would have k(s-2) degrees of freedom.

It may be pointed out here that the hypotheses of the type H_{14}, H_{15} and H_{16} were proposed earlier by Roy and Bhapkar (1960). The test of the hypothesis

$$\underset{\sim}{\eta} = \underset{\sim}{0}$$

with $\underset{\sim}{\eta}$ given in H_{16} , or in other words, of the hypothesis H_{15} with H_{16}

as the model is provided by the statistic

$$(4.C.9) - (4.C.11) \qquad \text{d.f.} = 2 \ ;$$

a statistic of this type would have k degrees of freedom in the general k-response case.

D. Four-dimensional tables with two responses and two factors.

In the basic set up (2.1) now we have $\underset{\sim}{i}$ a double subscript $(i_1 i_2)$, $i_\alpha = 1,2,\ldots,r_\alpha$, $\alpha = 1,2$ so that $r = r_1 r_2$ as before and $\underset{\sim}{j}$ also a double subscript $(j_1 j_2)$, $j_\beta = 1,2,\ldots,s_\beta$, $\beta = 1,2$ with not all combinations selected necessarily for the experiment; let s be the number of $(j_1 j_2)$ combinations selected so that $s \leq s_1 s_2$. For convenience we shall denote by j_1 the categories of the 'block-factor' (i.e., the 'blocks') and by j_2 the 'treatments' (i.e., the categories of the 'treatment-factor').

(i) Hypothesis of no 'treatment-effect' on the marginal distributions:

$$(4.D.1) \qquad H_{17}: \quad \begin{matrix} p_{i_1 0 j_1 j_2} \text{ is independent of } j_2 \\ p_{0 i_2 j_1 j_2} \text{ is also independent of } j_2 \ . \end{matrix}$$

This may be seen to be an analog of the hypothesis that $\mu_{\alpha j_1 j_2}$, $\sigma_{\alpha\alpha\, j_1 j_2}$, $\alpha = 1,2$, are independent of j_2 in MANOVA . In the notation of (4.C.5) with $\underset{\sim}{j}$ a double subscript $(j_1 j_2)$ let further

$$(4.D.2) \qquad \underset{\sim}{M}_{j_1} = \sum_{j_2} \underset{\sim}{M}_{j_1 j_2} \quad , \quad \underset{\sim}{t}_{j_1} = \sum_{j_2} \underset{\sim}{M}_{j_1 j_2} \underset{\sim}{q}_{j_1 j_2} \quad ,$$

the summation being over those j_2 only which occur in conjunction with j_1 . From theorem 3.2 we have the χ^2_1-statistic

(4.D.3) $$\sum_{j_1,j_2} \underset{\sim}{q}'_{j_1j_2} \underset{\sim}{M}_{j_1j_2} \underset{\sim}{q}_{j_1j_2} - \sum_{j_1} \underset{\sim}{t}'_{j_1} \underset{\sim}{M}^{-1}_{j_1} \underset{\sim}{t}_{j_1} \quad , \quad \text{d.f.} = (r_1+r_2-2)(s-s_1)$$

to test H_{17}. The method can be immediately generalized to the case of k response giving a statistic of the same type with $\{\Sigma_\alpha(r_\alpha-1)\}(s-s_1)$ degrees of freedom.

(ii) Hypothesis of no 'treatment-effect' on the 'means':

If the two responses are 'structured' with $\{a_{i_1}\}$ and $\{b_{i_2}\}$ as the associated scores, consider

(4.D.4) H_{18}:

$$\sum_{i_1} a_{i_1} p_{i_1 0 j_1 j_2} \quad \text{is independent of } j_2$$

$$\sum_{i_2} b_{i_2} p_{0 i_2 j_1 j_2} \quad \text{is also independent of } j_2 .$$

This is the analog of the usual hypothesis of no 'treatment-effects' in MANOVA. In the notation of (4.C.8) but with j a double subscript $(j_1 j_2)$ let further

(4.D.5) $$\underset{\sim}{M}_{j_1} = \sum_{j_2} \underset{\sim}{M}_{j_1 j_2} \qquad \underset{\sim}{t}_{j_1} = \sum_{j_2} \underset{\sim}{M}_{j_1 j_2} \underset{\sim}{\chi}_{j_1 j_2} .$$

For testing H_{18} we have the χ^2_1-statistic

(4.D.6) $$\sum_{j_1 j_2} \underset{\sim}{\chi}'_{j_1 j_2} \underset{\sim}{M}_{j_1 j_2} \underset{\sim}{\chi}_{j_1 j_2} - \sum_{j_1} \underset{\sim}{t}'_{j_1} \underset{\sim}{M}^{-1}_{j_1} \underset{\sim}{t}_{j_1} \quad , \quad \text{d.f.} = 2(s-s_1) .$$

For the k-response situation the statistic would have $k(s-s_1)$ degrees of freedom. Here again we note that $H_{17} \Longrightarrow H_{18}$ and the weaker hypothesis H_{18} would be of interest when H_{17} does not hold.

(iii) Hypothesis of linearity of regression on treatment level:

We assume now that the second factor is 'structured' and d_{j_2} is the score associated with the category j_2; in other words, d_{j_2} is the 'level' of the j_2th treatment. If H_{18} does not hold, we may consider

$$(4.D.7) \qquad H_{19}: \quad \begin{aligned} \sum_{i_1} a_{i_1} p_{i_1 0 j_1 j_2} &= \xi^{(1)}_{j_1} + \eta^{(1)}_{j_1} d_{j_2} \\ \sum_{i_2} b_{i_2} p_{0 i_2 j_1 j_2} &= \xi^{(2)}_{j_1} + \eta^{(2)}_{j_1} d_{j_2} \end{aligned} ,$$

where ξ's and η's are unknown. This hypothesis is the analog of the hypothesis that

$$\underset{\sim}{\mu}_{j_1 j_2} = \underset{\sim}{\xi}_{j_1} + \underset{\sim}{\eta}_{j_1} d_{j_2}$$

in MANOVA. In the notation of (4.C.8) with $j = (j_1 j_2)$ and of (4.D.5) let further

$$(4.D.8) \qquad \underset{\sim}{W}_{j_1} = \sum_{j_2} \underset{\sim}{M}_{j_1 j_2} \underset{\sim}{\chi}_{j_1 j_2} d_{j_2} , \quad \underset{\sim}{R}_{j_1} = \sum_{j_2} \underset{\sim}{M}_{j_1 j_2} d_{j_2} , \quad \underset{\sim}{S}_{j_1} = \sum_{j_2} \underset{\sim}{M}_{j_1 j_2} d^2_{j_2} .$$

Then for testing H_{19} from theorem 3.2 we have the χ^2_1-statistic

$$(4.D.9) \qquad \sum_{d_1, d_2} \underset{\sim}{\chi}'_{j_1 j_2} \underset{\sim}{M}_{j_1 j_2} \underset{\sim}{\chi}_{j_1 j_2} - \sum_{j_1} [\underset{\sim}{t}'_{j_1} , \underset{\sim}{w}'_{j_1}] \begin{bmatrix} \underset{\sim}{M}_{j_1} & \underset{\sim}{R}_{j_1} \\ \underset{\sim}{R}_{j_1} & \underset{\sim}{S}_{j_1} \end{bmatrix} \begin{bmatrix} \underset{\sim}{t}_{j_1} \\ \underset{\sim}{w}_{j_1} \end{bmatrix}$$

$$\text{d.f.} = 2(s-2s_1)$$

For the k-response problem, the statistic would have $k(s-2s_1)$ degrees of freedom.

It may be mentioned here that the hypotheses H_{17}, H_{18}, H_{19} were offered as analogs earlier by Roy and Bhapkar (1960) . To test the hypothesis

$$\underset{\sim}{\eta}_{j_1} = \underset{\sim}{0} ,$$

with $\underset{\sim}{\eta}$'s given by (4.D.7), i.e., the hypothesis H_{18} with H_{19} as the model, we have the statistic

$$(4.D.6) - (4.D.9) \qquad \text{d.f.} = 2s_1 \ ;$$

a statistic of the same type would have ks_1 degrees of freedom in the k-response case.

Similarly we can test the hypotheses $\underset{\sim}{\eta}_{j_1} = \underset{\sim}{\eta}$ (known or unknown as the case may be) and/or $\underset{\sim}{\xi}_{j_1} = \underset{\sim}{\xi}$ under the model H_{19}; these details are omitted.

(iv) Hypotheses of "no interaction":

This raises a number of problems indeed depending on whether we are interested in the nature of association between the two responses over the categories of the two factors or in the effects of factors on the marginal distributions of the responses. We are then faced with defining 'interactions' of different orders and testing hypotheses about these. Hence this is omitted in further discussion and the reader is referred to Bhapkar and Koch (1965) for further details.

5. Remarks on higher dimensional tables

Most of these problems can be handled in much the same way as the simpler problems discussed earlier. Thus, as mentioned earlier, the hypothesis of equality of marginal distributions of several responses, or of equality of 'means' of several responses can be tested by methods discussed in 4.B; the hypothesis of homogeneity of marginal distributions or the one of equality of 'means' etc., for several populations and the general multi-response problem can be tested by methods in 4.C; the various hypotheses of no

'treatment-effect' or of the linearity of the treatment-effect in the general multi-response two-factor situation (with one factor in the nature of 'blocks') can be handled by methods in 4.D. Problems of this nature for the general k-response ℓ-factor situation present no further difficulties and can be handled in precisely the same way. It is when we are considering problems of interactions of various types and of various orders that further difficulties arise; some of these are discussed in [4] and [5].

Problems of associations similar to those in H_4, H_5 and H_6 can be handled in a fairly straightforward manner for the general k-response single population situation as follows:

(i) Hypothesis of complete independence of k responses;

$$(5.1) \qquad H_{20}: \; p_{i_1 i_2 \ldots i_k} = p_{i_1 0 0 \ldots 0} p_{0 i_2 0 \ldots 0} \cdots p_{0 0 0 \ldots 0 i_k} \; .$$

It is easy to check that, as in H_4, we immediately get maximum likelihood estimate $\hat{p}_{i_1 0 \ldots 0} = n_{i_1 0 \ldots 0}/N$ etc. giving the χ^2-statistic

$$(5.2) \qquad N^{k-1} \sum_{i_1 \ldots, i_k} \left(n_{i_1 \ldots i_k} - \frac{n_{i_1 0 \ldots 0} \cdots n_{0 0 \ldots i_k}}{N^{k-1}} \right)^2 \Big/ \left(n_{i_1 0 \ldots 0} \cdots n_{0 \ldots 0 i_k} \right)$$

$$\text{d.f.} = r_1 r_2 \ldots r_k - (r_1 + r_2 + \ldots + r_k) + (k-1) \; .$$

(ii) Hypothesis of independence of two sets of responses:

$$(5.3) \qquad H_{21}: \; p_{i_1 i_2 \ldots i_k} = p_{i_1 i_2 \ldots i_{k_1} 0 \ldots 0} \, p_{0 0 \ldots 0 i_{k_1+1} \ldots i_k} \; ,$$

where we assume, without loss of generality, that the first k_1 responses form the first set. As in H_5, we get immediately the maximum likelihood estimates

$\hat{p}_{i_1\ldots i_{k_1}0\ldots0} = n_{i_1\ldots i_{k_1}0\ldots0}/N$ etc. giving the χ^2-statistic

$$(5.4)\quad \sum_{i_1,\ldots,i_k} N\left(n_{i_1 i_2\ldots i_k} - \frac{n_{i_1\ldots i_{k_1}0\ldots0}\, n_{0\ldots0 i_{k_1+1}\ldots i_k}}{N}\right)^2 \Big/ \left(n_{i_1\ldots i_{k_1}0\ldots0}\, n_{0\ldots0 i_{k_1+1}\ldots i_k}\right),$$

$$\text{d.f.} = (r_1\ldots r_{k_1}-1)(r_{k_1+1}\ldots r_k-1)\ .$$

Actually the statistic follows directly from (4.A.2) noting that $(i_1\ldots i_{k_1})$ can be regarded as one subscript i_1^* and $(i_{k_1+1}\ldots i_k)$ as i_2^*.

(iii) Hypothesis of independence of several sets of responses:

$$(5.5)\quad H_{22}:\ p_{i_1 i_2\ldots i_k} = p_{i_1\ldots i_{k_1}0\ldots0}\, p_{0\ldots0 i_{k_1+1}\ldots i_{k_1+k_2}0\ldots0}\cdots p_{0\ldots0 i_{k_1+\ldots+k_{t-1}+1}\ldots i_k}$$

where we assume that there are t sets of responses, the first containing the first k_1 responses and so on. The statistic for H_{22} follows immediately from (5.2) regarding $i_1^* = (i_1\ldots i_{k_1})$ and so on so that $r_1^* = r_1\ldots r_{k_1}$ etc., and replacing k in (5.2) by t with $r_1^*\ldots r_t^* - (r_1^*+\ldots+r_t^*) + (t-1)$, i.e., $r_1 r_2\ldots r_k - (r_1 r_2\ldots r_{k_1}+\ldots+r_{k_1+\ldots+k_{t-1}+1}\ldots r_k) + t - 1$ degrees of freedom.

(iv) Hypothesis of independence of two sets of responses given a third set:

$$(5.6)\quad H_{23}:\ p_{i_1 i_2\ldots i_k} = \frac{p_{i_1\ldots i_{k_1}0\ldots0 i_{k_1+k_2+1}\ldots i_k}\, p_{0\ldots0 i_{k_1+1}\ldots i_k}}{p_{0\ldots0 i_{k_1+k_2+1}\ldots i_k}}\ ;$$

here we want to test the conditional independence of the first two sets given the third set of responses. The χ^2-statistic is immediately seen to be

$$(5.7)\qquad \sum_{i_1,\ldots,i_k} \frac{\left(n_{i_1\ldots i_k} - \dfrac{n_{i_1\ldots i_{k_1}0\ldots 0 i_{k_1+k_2+1}\ldots i_k}\, n_{0\ldots 0 i_{k_1+1}\ldots i_k}}{n_{0\ldots 0 i_{k_1+k_2+1}\ldots i_k}}\right)^2}{\left(\dfrac{n_{i_1\ldots i_{k_1}0\ldots 0 i_{k_1+k_2+1}\ldots i_k}\, n_{00\ldots 0 i_{k_1+1}\ldots i_k}}{n_{0\ldots 0\, i_{k_1+k_2+1}\,\ldots\, i_k}}\right)},$$

$$\text{d.f.} = (r_1\ldots r_{k_1}-1)(r_{k_1+1}\ldots r_{k_1+k_2}-1)r_{k_1+k_2+1}\ldots r_k .$$

Note that H_{20} is an analog of the hypothesis that the correlation matrix is the unit matrix, while H_{21}, H_{22} and H_{23} are analogs of the hypothesis of no canonical correlations, Wilks' hypothesis of independence of sets of variates and the hypothesis of no partial canonical corrleations, respectively, in the normal multivariate analysis.

REFERENCES

[1] Barnard, G. A. (1947), "Significance tests for 2 x 2 tables," Biometrika, 34, pp. 123-138.

[2] Bhapkar, V. P. (1961), "Some tests for categorical data," Ann. Math. Stat., 32, pp. 72-83.

[3] ----- (1965), "A note on the equivalence of two test criteria for hypotheses in categorical data," to appear in the March 1966 issue of Jour. Amer. Stat. Assoc.

[4] ----- and Koch, Gary G. (1965), "On the hypothesis of no interaction in three-dimensional contingency tables," Institute of Statistics Mimeo Series No. 440, University of North Carolina.

[5] ----- and -----(1965), "The hypotheses of no interaction in four-dimensional contingency tables," Institute of Statistics Mimeo Series No. 447, University of North Carolina.

[6] Cochran, W. G. (1950), "The comparison of percentages in matched samples," Biometrika, 37, pp. 256-266.

[7] Cramer, H. (1946), Mathematical Methods of Statistics, Princeton University Press, Princeton.

[8] Fisher, R. A. (1922), "On the interpretation of chi-square from contingency tables and the calculation of p," Jou. Roy. Stat. Soc., 85, pp. 87-94.

[9] Goodman, L. A. (1964), "Simple methods for analyzing three-factor interaction in contingency tables," Jou. Amer. Stat. Assoc., 59, pp. 319-352.

[10] Mitra, S. K. (1955), "Contributions to the statistical analysis of categorical data," Institute of Statistics Mimeo Series No. 142, University of North Carolina.

[11] Neyman, J. (1949), "Contribution to the theory of the χ^2 test, " Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, pp. 239-273.

[12] Pearson, E. S. (1947), "The choice of statistical tests illustrated on the interpretation of data classed in a 2 x 2 table," Biometrika, 34, pp. 139-167.

[13] Pearson, K. (1900), "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," Philosophical Magazine, Series 5, 50, pp. 157-172.

[14] Roy, S. N. and Bhapkar, V. P. (1960), "Some nonparametric analogs of normal ANOVA, MANOVA, and of studies in normal association," Contributions to Probability and Statistics, Stanford University Press, Stanford, pp. 371-387.

[15] ------ and Mitra, S. K. (1956), "An introduction to some nonparametric generalizations of analysis of variance and multivariate analysis," Biometrika, 43, pp. 361-376.

[16] Sathe, Y. S. (1962), "Studies of some problems in nonparametric inference," Institute of Statistics Mimeo Series No. 325, University of North Carolina.

[17] Stuart, A. (1955), "A test for homogeneity of the marginal distributions in a two-way classification," Biometrika, 42, 412-416.

[18] Wald, A. (1943), "Tests of statistical hypotheses concerning several parameters when the number of observations is large," Trans. Amer. Math. Soci., 54, 426-482.