STATISTICAL RESEARCH REPORT No 6

Institute of Mathematics september 1975

University of Oslo

INFERENCE IN RANDOM COEFFICIENT REGRESSION MODELS WITH ONE-WAY AND NESTED CLASSIFICATION.

Variations in absorption and excretion of drugs.

by

Ragnar Norberg

Summary.

We consider IJ linear regressions

$$\underset{\sim}{X}_{ij} = \underset{\sim}{T}\underset{\sim}{B}_{ij} + \underset{\sim}{U}_{ij} \quad , \qquad i=1,\ldots,I; \quad j=1,\ldots,J \ ,$$

where $\underset{\sim}{X}_{ij}$ is the vector of observations of the dependent variable, $\underset{\sim}{B}_{ij}$ the vector of regression coefficients and $\underset{\sim}{U}_{ij}$ the vector of disturbances in the regression labelled (i,j). The design matrix $\underset{\sim}{T}$ is common for all regressions, the situation being a typical experimental one. The vectors $\underset{\sim}{B}_{ij}$ are of the form

$$\underset{\sim}{B}_{ij} = \underset{\sim}{\beta} + \underset{\sim}{C}_i + \underset{\sim}{D}_{ij} \ ,$$

where the $\underset{\sim}{C}_i$'s are i.i.d. with zero mean and independent of the $\underset{\sim}{D}_{ij}$'s which are also i.i.d. with zero mean and $\underset{\sim}{\beta} = E\underset{\sim}{B}_{ij}$. Typically the subscript i represents individuals and subscript j represents repetitions of the experiment, and the $\underset{\sim}{C}_i$'s and $\underset{\sim}{D}_{ij}$'s represent inter- and intraindividual variations of the regression coefficients. Estimation and test procedures are derived under normality assumptions. Asymptotic minimum variance Fisherconsistent estimates are found for the parameters. For testing hypotheses concerning $\underset{\sim}{\beta}$ a simultaneous test procedure analogous to Scheffé's S-method is employed. Hypotheses concerning the covariance matrices are tested under various assumptions about the covariance structure. The derived tests are shown to be UMP unbiased when $\underset{\sim}{B}_{ij}$ is one-dimensional.

0. Notations and terminology.

In the following μ denotes Lebesgue measure on the Borel class of sets in a euclidean space, the dimension of which will be clear from the context. Otherwise greek letters denote subsets, points or components of points in a parameter space Ω .

As a rule random variables are denoted by capital latin letters. However, an estimator of a parameter α , say, is denoted by α^* or $\hat{\alpha}$.

Matrices are underlined with a tilde, and their order may be indicated by a top-script. Example: $\underset{\sim}{G}^{P\times Q}$ is a $P \times Q$-matrix. If $Q=1$, i.e. $\underset{\sim}{G}$ is a column vector, we may write $\underset{\sim}{G}^{P}$ instead of $\underset{\sim}{G}^{P\times 1}$. $\|\underset{\sim}{G}\| = (\underset{\sim}{G}'\underset{\sim}{G})^{\frac{1}{2}}$ is the euclidean norm of $\underset{\sim}{G}^{P}$. We write G_{pq} , or occasionally $(\underset{\sim}{G})_{pq}$, for the element in the p'th row and q'th column of $\underset{\sim}{G}$. If $\underset{\sim}{G}$ is a vector, ($P=1$ or $Q=1$), we number its elements by a single subscript. If the matrix or vector itself is numbered by a subscript, we place the element subscript after the matrix subscript. Example: The h'th element of the column vector $\underset{\sim}{B}_{ij}^{H}$ is denoted B_{ijh} or $(\underset{\sim}{B}_{ij})_h$. $\underset{\sim}{I}$ is the identity matrix, and $\underset{\sim}{0}$ is the null matrix.

We introduce special symbols for some distribution laws: $N_P(\underset{\sim}{\xi},\underset{\sim}{\Sigma})$ is the P-variate multinormal distribution with mean vector $\underset{\sim}{\xi}$ and covariance matrix $\underset{\sim}{\Sigma}$, χ_m^2 is the chi-square distribution with m degrees of freedom, $F_{m,n}$ is the F-distribution with m and n degrees of freedom and T_m is the t-distribution with m degrees of freedom. We do not distinguish notationally between a law and its distribution function. The (lower) ϵ- points in χ_m^2, $F_{m,n}$ and T_m are denoted by

$\chi^2_{\epsilon;m}$, $f_{\epsilon;m}$ and $t_{\epsilon;m}$ respectively. That X is distributed according to the law F is written $X \sim F$.

A dot in the place of a subscript denotes averaging with respect to that subscript.

Examples: $\underset{\sim}{X}_{\cdot} = N^{-1} \sum_{i=1}^{N} \underset{\sim}{X}_i$, $\underset{\sim}{X}_{i\cdot} = M^{-1} \sum_{j=1}^{M} \underset{\sim}{X}_{ij}$ and $\underset{\sim}{X}_{\cdot\cdot} = (MN)^{-1} \sum_{i=1}^{N} \sum_{j=1}^{M} \underset{\sim}{X}_{ij}$.

Uniformly most powerful is written UMP.

1. Introductory examples.

1 A. In standard regression theory the regression coefficients are unknown, fixed parameters. A random coefficient regression (RCR) model was first studied by Wald (1947). Swamy (1971) gives a survey of the work done so far on RCR models. In this paper we consider the situation where the vector of regression coefficients is of the form

$$\underset{\sim}{B} = \underset{\sim}{\beta} + \underset{\sim}{C} + \underset{\sim}{D} , \tag{1.1}$$

where $\underset{\sim}{\beta} = E\underset{\sim}{B}$, $\underset{\sim}{C}$ and $\underset{\sim}{D}$ have zero expectations, are uncorrelated and typically represent inter- and intraindividual variations respectively. We give two examples of this situation.

1 B. Fluor washing. As a caries-preventing measure childrens teeth are washed with dissoluted natriumfluorid. The concentration CN of natriumfluorid in the plaque, (i.e. the coating on the teeth), after washing is a measure of the effect of the washing. CN is assumed to depend proportionally on the concentration of natriumfluorid in the concentration, t ;

$$CN = Bt .$$

The coefficient B depends on nutritional factors, oral hygiene and strains of bacteria in the mouth, which vary between persons

and also from day to day for each person. Consider a person who is chosen at random from the population on a random day. By the line of argument in Scheffé (1959), pp. 221-223 and pp. 238-242, his B-value is composed of β, C and D as in (1.1). β is the average of all B-values taken over all persons and days. $\beta + C$ is the individual mean B-value of the chosen person, and his mean deviation C from the total mean thus represents inter- (or between-) individual variations. On the selected day his B-value differs from the individual mean by a random term D, which represents intra- (or within-) individual variations.
CN cannot be observed directly, but is measured by a chemical method with a random error U. The measured CN-value corresponding to concentration t is then

$$X = Bt + U \quad .$$

Knowledge of the distribution of B is important in connection with large scale production and sale of the preparation. For the purpose of drawing inference concerning β and the distributions of C and D the X-values corresponding to the concentrations $t_1,\ldots,t_K$ are measured on I persons on J different days. The resulting observations are

$$X_{ijk} = (\beta+C_i+D_{ij})t_k+U_{ijk} \;, \quad \begin{array}{l} i=1,\ldots,I \;;\; j=1,\ldots,J \;; \\ k=1,\ldots,K \;. \end{array} \tag{1.2}$$

β, C_i, D_{ij} and U_{ijk} are unobservable. C_i is the C-value of the i'th person, D_{ij} his D-value on the j'th day and U_{ijk} the measurement error in his k'th washing on that day.

1 C. Variations in absorption and excretion of drugs.

When a medical preparation is taken non-intravenously the active drug is absorbed and spread in the body. For many drugs the plasma concentration (i.e. the concentration of active drug in the blood) at time t after ingestion of the dose, $C(t)$, is a reliable measure of the effect of the dose at time t. It may be shown theoretically, (see e.g. Wagner, (1961)), that for certain simple preparations $C(t)$ is determined by the relation

$$C(t) = \frac{dR}{V(R-E)} \{\exp(-Et) - \exp(-Rt)\} , \qquad t \geq 0 ,$$

where d is the dose of active drug, V is the "distribution volume" in the body which is accessible for the drug, R is the rate of absorption and E is the rate of excretion of the drug. By the mean value theorem for derivatives the expression on the right side is equal to $(dRtV^{-1})\exp\{B(t)\}$, where $-B(t)$ is a number between Et and Rt for each t. We consider a preparation for which E and R are not too much different. Then $B(t)$ is approximately linear, $B(t) \approx B_2 t$, and putting $B_1 = \ln(RV^{-1})$ we get (approximately)

$$C(t) = dt \exp(B_1 + B_2 t) . \tag{1.3}$$

At time t a blood sample is taken from the patient. The measured value of the plasma concentration is

$$C^*(t) = C(t) + U^*(t) ,$$

where the random error $U^*(t)$ is due to the fact that the drug is not ideally distributed in the volume V.

It seems reasonable to assume that $U^*(t)$ is of the form $U^*(t) = C(t)U$, where U has expectation zero and variance σ^2, say. Then we have

$$C^*(t) = C(t)(1+U) . \tag{1.4}$$

We now substitute (1.3) in (1.4), divide by dt and take logarithms on both sides and get (approximately)

$$X(t) = \ln\{(dt)^{-1}C^*(t)\} = B_1 + B_2t + U , \tag{1.5}$$

having assumed that $\sigma \ll 1$ so that $\ln(1+U) \approx U$ with large probability. Pharmaceutists speak of R, E and V as "pharmacokinetic constants". However, Frislid et.al. (1973) and many others have pointed at the possible existence of great inter- and intraindividual variations in the ability to absorb and excrete drugs. Thus R, E and V, and hence also B_1 and B_2, should rather be considered as random variables. We now proceed as in example 1 B and decompose the vector $\underset{\sim}{B} = (B_1, B_2)'$ as in (1.1). Before mass production and marketing of a preparation its physiological availability is investigated in clinical trials. The following observational plan is standard. Equal doses of the preparation are given to I persons and blood samples are taken from each person at times $t_1, t_2, \ldots, t_K$ after ingestion. This experiment is performed with the same I persons on J different days. By (1.5) the resulting observations can be written in the form

$$X_{ijk} = B_{ij1} + B_{ij2} t_k + U_{ijk} \ , \ i=1,\ldots,I \ ; \ j=1,\ldots,J \ ; \ k=1,\ldots,K \ ,$$

where

$$\underset{\sim}{B}_{ij} = \begin{pmatrix} B_{ij1} \\ B_{ij2} \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} C_{i1} \\ C_{i2} \end{pmatrix} + \begin{pmatrix} D_{ij1} \\ D_{ij2} \end{pmatrix}$$

in accordance with (1.1). On the basis of these observations we want to draw inference concerning $\underset{\sim}{\beta}$ and the distributions of $\underset{\sim}{C}$ and $\underset{\sim}{D}$.

2. The model.

2 A. Nested (or hierarchical) classification. Our observations are of the form

$$\underset{\sim}{X}_{ij}^{K\times 1} = \underset{\sim}{T}^{K\times H} \underset{\sim}{B}_{ij}^{H\times 1} + \underset{\sim}{U}_{ij}^{K\times 1} \ , \ i=1,\ldots,I \ ; \ j=1,\ldots,J \ . \tag{2.1}$$

The "design matrix" $\underset{\sim}{T}$ is known and has full rank H $(<K)$. The vector of regression coefficients $\underset{\sim}{B}_{ij}$ is an unobservable random vector of the form

$$\underset{\sim}{B}_{ij} = \underset{\sim}{\beta} + \underset{\sim}{C}_i + \underset{\sim}{D}_{ij} \ , \tag{2.2}$$

where $\underset{\sim}{\beta}^H$ is constant , $\underset{\sim}{C}_i \sim N_H(\underset{\sim}{0},\underset{\sim}{\Sigma}_C)$ and $D_{ij} \sim N_H(\underset{\sim}{0},\underset{\sim}{\Sigma}_D)$. The vector of disturbances $\underset{\sim}{U}_{ij}$ is distributed according to $N_K(\underset{\sim}{0},\sigma^2 \underset{\sim}{I})$. All $\underset{\sim}{C}_i$, $\underset{\sim}{D}_{ij}$ and $\underset{\sim}{U}_{ij}$ are unobservable, and they are uncorrelated and hence independent.

2 B. One-way classification. If $J=1$ in the model above, we can drop the subscript j and put $\underset{\sim}{X}_{i1} = \underset{\sim}{X}_i$, $\underset{\sim}{B}_{i1} = \underset{\sim}{B}_i$ and so forth. We introduce

$$\underset{\sim}{\Sigma}_B = \underset{\sim}{\Sigma}_C + \underset{\sim}{\Sigma}_D .$$

Then our observations are of the form

$$\underset{\sim}{X}_i = \underset{\sim}{T}\underset{\sim}{B}_i + \underset{\sim}{U}_i , \qquad i=1,\ldots,I ,$$

where $\underset{\sim}{B}_i \sim N_H(\underset{\sim}{\beta},\underset{\sim}{\Sigma}_B)$ and $\underset{\sim}{U}_i \sim N_K(\underset{\sim}{0},\sigma^2\underset{\sim}{I})$. This model has been studied by Rao (1965) .

2 C. In the special case when $\underset{\sim}{B}$ is scalar ($H=1$) and $T=(1,1,\ldots,1)'$, the models 2 A and 2 B reduce to the wellknown models II with nested and one-way classification, which are treated by Lehmann (1959), pp. 286-293.

3. A canonical form of the observations.

3 A. Let the colums of $\underset{\sim}{E}^{K\times(K-H)}$ form an orthonormal basis for the orthocomplement of the H-dimensional linear space spanned by the colums of $\underset{\sim}{T}^{K\times H}$. For each i and j we transform $\underset{\sim}{X}_{ij}$ to

$$\hat{\underset{\sim}{B}}{}^H_{ij} = (\underset{\sim}{T}'\underset{\sim}{T})^{-1}\underset{\sim}{T}'\underset{\sim}{X}_{ij} , \tag{3.1}$$

$$\underset{\sim}{V}^{K-H} = \underset{\sim}{E}'\underset{\sim}{X}_{ij} .$$

$\hat{\underset{\sim}{B}}_{ij}$ is the ordinary least squares estimator of $\underset{\sim}{B}_{ij}$.

We substitute $\underset{\sim}{X}_{ij}$ from (2.1) in these expressions and get

$$\hat{\underset{\sim}{B}}_{ij} = (\underset{\sim}{T}'\underset{\sim}{T})^{-1}\underset{\sim}{T}'(\underset{\sim}{T}\underset{\sim}{B}_{ij}+\underset{\sim}{U}_{ij}) = \underset{\sim}{B}_{ij} + (\underset{\sim}{T}'\underset{\sim}{T})^{-1}\underset{\sim}{T}'\underset{\sim}{U}_{ij} \, ,$$

$$\underset{\sim}{V}_{ij} = \underset{\sim}{E}'(\underset{\sim}{T}\underset{\sim}{B}_{ij}+\underset{\sim}{U}_{ij}) = \underset{\sim}{E}'\underset{\sim}{U}_{ij} \, . \tag{3.2}$$

Substituting

$$\underset{\sim}{W}_{ij} = (\underset{\sim}{T}'\underset{\sim}{T})^{-1}\underset{\sim}{T}'\underset{\sim}{U}_{ij}$$

and $\underset{\sim}{B}_{ij}$ from (2.2) in (3.2) , we arrive at the following form of our transformed observations.

$$\hat{\underset{\sim}{B}}{}^{H}_{ij} = \underset{\sim}{\beta} + \underset{\sim}{C}_i + \underset{\sim}{D}_{ij} + \underset{\sim}{W}_{ij} \, ,$$

$$\underset{\sim}{V}{}^{K-H}_{ij} \, , \qquad\qquad i=1,\dots,I \, ; \; j=1,\dots,J \, .$$

$\underset{\sim}{\beta}$, $\underset{\sim}{C}_i$, $\underset{\sim}{D}_{ij}$ are as explained in subsection 2 A. Straightforward calculations show that $\underset{\sim}{W}_{ij} \sim N_H(\underset{\sim}{O},\sigma^2\underset{\sim}{M}_W)$, with

$$\underset{\sim}{M}_W = (\underset{\sim}{T}'\underset{\sim}{T})^{-1} \, , \tag{3.3}$$

and $\underset{\sim}{V}_{ij} \sim N_{K-H}(\underset{\sim}{O},\sigma^2\underset{\sim}{I})$. All $\underset{\sim}{C}_i$, $\underset{\sim}{D}_{ij}$, $\underset{\sim}{V}_{ij}$, $\underset{\sim}{W}_{ij}$ are independent and unobservable.

3 B. Now let $\underset{\sim}{G}^{J\times J} = (G_{pq})$ be an orthogonal matrix with the first column equal to $J^{-\frac{1}{2}}(1,\dots,1)'$, and for each i transform the $\hat{\underset{\sim}{B}}_{ij}$'s to $\underset{\sim}{Y}_{ij}$'s given by

$$(\underset{\sim}{Y}_{i1},\ldots,\underset{\sim}{Y}_{iJ}) = (\hat{\underset{\sim}{B}}_{i1},\ldots,\hat{\underset{\sim}{B}}_{iJ})\underset{\sim}{G} \quad . \tag{3.4}$$

The h'th component in $\underset{\sim}{Y}_{ij}$ is

$$Y_{ijh} = \sum_{p=1}^{J} \hat{B}_{iph}G_{pj} = \sum_{p=1}^{J}(\beta_h+C_{ih})G_{pj} + \sum_{p=1}^{J}(D_{iph} + W_{iph})G_{pj}$$

$$= J^{\frac{1}{2}}(\beta_h+C_{ih})\sum_{p=1}^{J} G_{p1}G_{pj} + \sum_{p=1}^{J}(D_{iph} + W_{iph})G_{pj}$$

$$= \delta_{j1}\; J^{\frac{1}{2}}(\beta_h+C_{ih}) + \sum_{p=1}^{J}(D_{iph} + W_{iph})G_{pj} \quad ,$$

(δ_{j1} is the Kroenecker delta). Thus we get

$$(\underset{\sim}{Y}_{i1},\ldots,\underset{\sim}{Y}_{iJ}) = (J^{\frac{1}{2}}(\underset{\sim}{\beta} + \underset{\sim}{C}_i),\underset{\sim}{0},\ldots,\underset{\sim}{0}) + (\underset{\sim}{Y}^*_{i1},\ldots,\underset{\sim}{Y}^*_{iJ}) \ ,$$

where $(\underset{\sim}{Y}^*_{i1},\ldots,\underset{\sim}{Y}^*_{iJ}) = (\underset{\sim}{D}_{i1}+\underset{\sim}{W}_{i1},\ldots,\underset{\sim}{D}_{iJ}+\underset{\sim}{W}_{iJ})\underset{\sim}{G}$ is distributed as $(\underset{\sim}{D}_{i1}+\underset{\sim}{W}_{i1},\ldots,\underset{\sim}{D}_{iJ}+\underset{\sim}{W}_{iJ})$. (For a proof of the last assertion, see Anderson (1958), pp. 51-52). Our observations are now on the form

$$\begin{aligned}
&\underset{\sim}{Y}_{i1} \sim N_H(J^{\frac{1}{2}}\underset{\sim}{\beta},\underset{\sim}{\Sigma}_Z) \quad , \quad i=1,\ldots,I \ , \\
&\underset{\sim}{Y}_{ij} \sim N_H(\underset{\sim}{0},\underset{\sim}{\Sigma}_Y) \quad , \quad i=1,\ldots,I \ ; \ j=2,\ldots,J \ , \\
&V_{ijk} \sim N_1(\underset{\sim}{0},\sigma^2) \ , \quad i=1,\ldots,I \ ; \ j=1,\ldots,J \ ; \ k=1,\ldots,K-H,
\end{aligned} \tag{3.5}$$

where

$$\underset{\sim}{\Sigma}_Y = \underset{\sim}{\Sigma}_D+\sigma^2\underset{\sim}{M}_W \quad \text{and} \quad \underset{\sim}{\Sigma}_Z = J\Sigma_C+\underset{\sim}{\Sigma}_D+\sigma^2\underset{\sim}{M}_W \quad . \tag{3.6}$$

All variables in (3.5) are independent.

3 C. Let $\underset{\sim}{H}^{I\times I}$ be an orthogonal matrix with the first column equal to $I^{-\frac{1}{2}}(1,\ldots,1)'$, and transform the $\underset{\sim}{Y}_{i1}$'s to $\underset{\sim}{Z}_i$'s by

$$(\underset{\sim}{Z}_1,\ldots,\underset{\sim}{Z}_I) = (\underset{\sim}{Y}_{i1},\ldots,\underset{\sim}{Y}_{I1})\underset{\sim}{H} .$$

By the line of argument in last subsection we find that

$$(\underset{\sim}{Z}_1,\ldots,\underset{\sim}{Z}_I) = \{(IJ)^{\frac{1}{2}}\underset{\sim}{\beta},\underset{\sim}{0},\ldots,\underset{\sim}{0}\} + (\underset{\sim}{Z}_1^*,\ldots,\underset{\sim}{Z}_I^*) ,$$

where the $\underset{\sim}{Z}_i^*$'s are independent and identically distributed according to $N_H(\underset{\sim}{0},\Sigma_Z)$. Thus we have arrived at the following form of our observations.

$$\begin{aligned}
&\underset{\sim}{Z}_1 \sim N_H\{(IJ)^{\frac{1}{2}}\underset{\sim}{\beta},\underset{\sim}{\Sigma}_Z\} , \\
&\underset{\sim}{Z}_i \sim N_H(\underset{\sim}{0},\underset{\sim}{\Sigma}_Z) , \quad i=2,\ldots,I, \\
&\underset{\sim}{Y}_{ij} \sim N_H(\underset{\sim}{0},\underset{\sim}{\Sigma}_Y) , \quad i=1,\ldots,I ; \ j=2,\ldots,J , \\
&V_{ijk} \sim N_1(0,\sigma^2) , \quad i=1,\ldots,I ; \ j=1,\ldots,J ; \ k= 1,\ldots,K-H.
\end{aligned} \tag{3.7}$$

All these variables are independent.
The joint density of the transformed observations in (3.7) is given by

$$\begin{aligned}
dP_{\underset{\sim}{\omega}} = {} & (2\pi)^{-\frac{1}{2}IH}|\underset{\sim}{\Sigma}_Z|^{-I/2}\exp[-\tfrac{1}{2}\{\underset{\sim}{z}_1-(IJ)^{\frac{1}{2}}\underset{\sim}{\beta}\}'\underset{\sim}{\Sigma}_Z^{-1}\{\underset{\sim}{z}_1-(IJ)^{\frac{1}{2}}\underset{\sim}{\beta}\} \\
& -\tfrac{1}{2}\sum_{i=2}^{I}\underset{\sim}{z}_i'\underset{\sim}{\Sigma}_Z^{-1}z_i] \\
& (2\pi)^{-\frac{1}{2}I(J-1)H}|\underset{\sim}{\Sigma}_Y|^{-I(J-1)/2}\exp(-\tfrac{1}{2}\sum_{i=1}^{I}\sum_{j=2}^{J}\underset{\sim}{y}_{ij}'\underset{\sim}{\Sigma}_Y^{-1}\underset{\sim}{y}_{ij}) \\
& (2\pi)^{-\frac{1}{2}IJ(K-H)}\sigma^{-IJ(K-H)}\exp\{-(2\sigma^2)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K-H}v_{ijk}^2\}\,d\mu .
\end{aligned} \tag{3.8}$$

The components of the parameter $\underset{\sim}{\omega}$ are

$$
\begin{aligned}
&\beta_h \;, \qquad h=1,\dots,H \;, \\
&(\underset{\sim}{\Sigma}_Z)_{hh'} \;, \; 1 \le h \le h' \le H \;, \\
&(\underset{\sim}{\Sigma}_Y)_{hh'} \;, \; 1 \le h \le h' \le H \;, \\
&\sigma^2 \;.
\end{aligned}
\tag{3.9}
$$

The domain of variation of $\underset{\sim}{\omega}$ is the parameter space Ω given by

$$
\begin{aligned}
&\beta_h \in (-\infty,\infty) \;, \quad h=1,\dots,H \;, \\
&\underset{\sim}{\Sigma}_Z-\underset{\sim}{\Sigma}_Y(=J\Sigma_C) \text{ and } \underset{\sim}{\Sigma}_Y - \sigma^2\underset{\sim}{M}_W(=\Sigma_D) \text{ are positive} \\
&\text{definite,} \\
&\sigma^2 > 0 \;.
\end{aligned}
\tag{3.10}
$$

$\underset{\sim}{\omega}$ (and Ω) is of dimension $H+1+\frac{1}{2}H(H+1)+\frac{1}{2}H(H+1) = (H+1)^2$.

3 D. <u>One-way classification.</u> In subsection 2 B we regarded one-way classification as a special case of nested classification. Putting $J=1$ we get the following version of the canonical form (3.5).

$$
\begin{aligned}
&\underset{\sim}{Y}_i \sim N_H(\underset{\sim}{\beta} \;,\; \underset{\sim}{\Sigma}_B + \sigma^2\underset{\sim}{M}_W) \;, \\
&V_{ik} \sim N_1(0,\sigma^2) \;, \qquad i=1,\dots,I \;;\; k=1,\dots,K-H \;.
\end{aligned}
\tag{3.11}
$$

Corresponding to the canonical form (3.7) we now get

$$
\begin{aligned}
&\underset{\sim}{Z}_1 \sim N_H(I^{\frac{1}{2}}\underset{\sim}{\beta} \;,\; \underset{\sim}{\Sigma}_B+\sigma^2\underset{\sim}{M}_W) \;, \\
&\underset{\sim}{Z}_i \sim N_H(\underset{\sim}{0}, \underset{\sim}{\Sigma}_B+\sigma^2\underset{\sim}{M}_W) \;,\; i=2,\dots,I \;, \\
&V_{ik} \sim N_1(0,\sigma^2) \;, \qquad i=1,\dots,I \;,\; k=1,\dots,K-H \;.
\end{aligned}
\tag{3.12}
$$

All variables in (3.11) are independent, and so are the variables in (3.12) .

4. Point estimates of the parameters.

4 A. The likelihood function of the observations (3.7) is given by (3.8) and is seen to be of the form

$$L(\underset{\sim}{\omega};\underset{\sim}{Z},\underset{\sim}{Y},\underset{\sim}{V}) = L_1(\underset{\sim}{\beta},\underset{\sim}{\Sigma}_Z;\underset{\sim}{Z})L_2(\underset{\sim}{\Sigma}_Y;\underset{\sim}{Y})L_3(\sigma^2;\underset{\sim}{V}) \ , \tag{4.1}$$

with $\underset{\sim}{Z} = (\underset{\sim}{Z}_1,\ldots,\underset{\sim}{Z}_I)$, $\underset{\sim}{Y} = (Y_{12},\ldots,Y_{IJ})$ and $\underset{\sim}{V}=(V_{111},\ldots,V_{IJ,K-H})$.

The maximum likelihood estimator $\hat{\underset{\sim}{\omega}}$ is the point in Ω which maximizes L . $\hat{\omega}$ may be difficult to find because Ω is restricted by the positive definiteness of $\underset{\sim}{\Sigma}_C$ and $\underset{\sim}{\Sigma}_D$. We easily obtain a modified maximum likelihood estimator by maximizing L in the wider region Ω'
defined by

$$\begin{aligned} &\beta_h \in (-\infty,\infty) \ , \quad h=1,\ldots,H \ , \\ &\underset{\sim}{\Sigma}_Z \text{ and } \underset{\sim}{\Sigma}_Y \text{ are positive definite,} \\ &\sigma^2 > 0 \ . \end{aligned} \tag{4.2}$$

$\Omega \subset \Omega'$ since the sum of any two positive definite matrices is itself positive definite. Ω' is the direct product of the $(\underset{\sim}{\beta},\underset{\sim}{\Sigma}_Z)$-space, the $\underset{\sim}{\Sigma}_Y$-space and the σ^2-space, and hence we find the maximum of L as $\underset{\sim}{\omega}$ varies in Ω' by maximizing each of the factors L_1,L_2 and L_3 in (4.1) separately with respect to the parameter occuring in it. Let $\underset{\sim}{\omega}^*$ be the point in Ω' defined by

$$\underset{\sim}{\beta}^* = (IJ)^{-\frac{1}{2}}\underset{\sim}{Z}_1 = \underset{\sim}{\hat{B}}_{\bullet\bullet} \ ,$$

$$\underset{\sim}{\Sigma}^*_Z = I^{-1}\sum_{i=2}^{I}\underset{\sim}{Z}_i\underset{\sim}{Z}_i' = JI^{-1}\sum_{i=1}^{I}(\underset{\sim}{\hat{B}}_{i\bullet}-\underset{\sim}{\hat{B}}_{\bullet\bullet})(\underset{\sim}{\hat{B}}_{i\bullet}-\underset{\sim}{\hat{B}}_{\bullet\bullet})' \ ,$$

$$\underset{\sim}{\Sigma}^*_Y = \{I(J-1)\}^{-1}\sum_{i=1}^{I}\sum_{j=2}^{J}\underset{\sim}{Y}_{ij}\underset{\sim}{Y}_{ij}' \tag{4.3}$$

$$= \{I(J-1)\}^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}(\underset{\sim}{\hat{B}}_{ij}-\underset{\sim}{\hat{B}}_{i\bullet})(\underset{\sim}{\hat{B}}_{ij}-\underset{\sim}{\hat{B}}_{i\bullet})' \ ,$$

$$\sigma^{2*} = \{IJ(K-H)\}^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K-H}V_{ijk}^2 = \{IJ(K-H)\}^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}S_{ij}^2 \ ,$$

where S_{ij}^2 is the sum of squared residuals in the regression labelled (i,j) , i.e.

$$S_{ij}^2 = \sum_{k=1}^{K}(X_{ijk}-\sum_{h=1}^{H}t_{kh}\hat{B}_{ijh})^2 \ . \tag{4.4}$$

The relations stated in (4.3) between original variables and canonical variables will be proved in subsection 4 B .

From theorem (3.2.1) and its proof in Anderson (1958), we obtain

Lemma (4.5). (Almost surely) the point $\underset{\sim}{\omega}^*$ defined by (4.3) maximizes L in Ω' and is the unique solution of the first order equations of a local extremum, i.e.

$$\partial \ln L/\partial \omega_p \big|_{\underset{\sim}{\omega}=\underset{\sim}{\omega}^*} = 0 \ , \quad p=1,\ldots,(H+1)^2 \ .$$

$\underset{\sim}{\omega}^*$ possesses the following optimum property.

Theorem (4.6). (a) $\underset{\sim}{\omega}^*$ given by (4.3) is a Fisherconsistent estimator of $\underset{\sim}{\omega}$, (see Sverdrup (1965)) . For each $p=1,\ldots,(H+1)^2$, ω_p^* has uniformly minimum asymptotic variance among all Fisherconsistent estimators of ω_p as $I \to \infty$.

(b) $\underset{\sim}{\omega}^*$ is asymptotic maximum likelihood estimator in the sense that $\lim_{I\to\infty} P_{\underset{\sim}{\omega}}(\underset{\sim}{\omega}^*$ maximizes L in $\Omega) = 1$ for all $\underset{\sim}{\omega} \in \Omega$.

Proof. The proof of (a) goes in several steps. We first prove that

Ω and Ω' are open sets.

To prove that Ω' is open we need only prove that the space M of symmetric and positive definite $n\times n$-matrices is open, since then $\underset{\sim}{\Sigma}_Y$-space and $\underset{\sim}{\Sigma}_Z$-space are open. A $n\times n$-matrix $\underset{\sim}{A} = \{a_{ij}\}$, which is symmetric (i.e. $a_{ij} = a_{ji}$), may be regarded as the point $(a_{11}, a_{21}, a_{22}, a_{31},\dots,a_{nn})$ in the euclidean $\frac{1}{2}n(n+1)$-space . Assume $\underset{\sim}{A} \in M$. That $\underset{\sim}{A}$ is positive definite means that $\underset{\sim}{x}'\underset{\sim}{A}\underset{\sim}{x} > 0$ for all $\underset{\sim}{x}^{n\times 1} \neq \underset{\sim}{0}^{n\times 1}$, or equivalently that $\underset{\sim}{x}'\underset{\sim}{A}\underset{\sim}{x} > 0$ for all $\underset{\sim}{x} \in S_0 = \{\underset{\sim}{x}\,|\,\|\underset{\sim}{x}\|=1\}$, which is seen by replacing $\underset{\sim}{x}$ with $\|\underset{\sim}{x}\|^{-\frac{1}{2}}\underset{\sim}{x}$. Since $\underset{\sim}{x}'\underset{\sim}{A}\underset{\sim}{x}$ is continuous regarded as a function of $\underset{\sim}{x}$ and S_0 is compact, there is a $\underset{\sim}{x}_0 \in S_0$ such that $\underset{\sim}{x}_0'\underset{\sim}{A}\underset{\sim}{x}_0 = \inf_{\underset{\sim}{x} \in S_0} \underset{\sim}{x}'\underset{\sim}{A}\underset{\sim}{x}$. $\underset{\sim}{x}_0'\underset{\sim}{A}\underset{\sim}{x}_0 > 0$ since $x_0 \neq 0$. For any $\underset{\sim}{C}^{n\times n}$ and any $\underset{\sim}{x} \in S_0$ we have $\underset{\sim}{x}'\underset{\sim}{C}\underset{\sim}{x} = \underset{\sim}{x}'\underset{\sim}{A}\underset{\sim}{x} + \underset{\sim}{x}'(\underset{\sim}{C}-\underset{\sim}{A})\underset{\sim}{x}$ $\geq \underset{\sim}{x}_0'\underset{\sim}{A}\underset{\sim}{x}_0 + \sum_i \sum_j (c_{ij}-a_{ij})x_i x_j$. Let $\underset{\sim}{C}$ be symmetric and assume that the distance between $\underset{\sim}{C}$ and $\underset{\sim}{A}$ in the euclidean $\frac{1}{2}n(n+1)$ - space is $< \epsilon$. Then $|c_{ij}-a_{ij}| < \epsilon$ for all i and j , and noting that $|x_i| \leq 1$ when $\underset{\sim}{x} \in S_0$ we get $\underset{\sim}{x}'\underset{\sim}{C}\underset{\sim}{x} \geq \underset{\sim}{x}_0'\underset{\sim}{A}\underset{\sim}{x}_0 - n^2\epsilon$. By choosing $\epsilon \leq (2n^2)^{-1}\underset{\sim}{x}_0'\underset{\sim}{A}\underset{\sim}{x}_0$ we get $\underset{\sim}{x}'\underset{\sim}{C}\,\underset{\sim}{x} \geq \frac{1}{2}\underset{\sim}{x}_0'\underset{\sim}{A}\underset{\sim}{x}_0 > 0$, and hence $\underset{\sim}{C} \in M$ which proves that M is open. We conclude that Ω' is open. Consider the mapping f which takes $(\underset{\sim}{\beta}, \underset{\sim}{\Sigma}_Z, \underset{\sim}{\Sigma}_Y, \sigma^2)$ into $(\underset{\sim}{\beta}, J\underset{\sim}{\Sigma}_C, \underset{\sim}{\Sigma}_D, \sigma^2)$. Let Γ be the subset of the range of f defined by the conditions in (3.10). Ω is the inverse image of Γ under f. By the above reasoning we know that Γ is open, and the continuity of f then implies that Ω is open. (Note that in lemma (4.5) it was tacitly assumed that Ω' is open).

We now return to the form (3.5) of our observations. The simultaneous density of the observations corresponding to a fixed i is

$$(2\pi)^{-\frac{1}{2}H}|\underset{\sim}{\Sigma}_Z|^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(\underset{\sim}{y}_{i1}-J^{\frac{1}{2}}\underset{\sim}{\beta})'\underset{\sim}{\Sigma}_Z^{-1}(\underset{\sim}{y}_{i1}-J^{\frac{1}{2}}\underset{\sim}{\beta})\}$$

$$(2\pi)^{-\frac{1}{2}(J-1)H}|\underset{\sim}{\Sigma}_Y|^{-\frac{1}{2}(J-1)}\exp\{-\tfrac{1}{2}\sum_{j=2}^{J}\underset{\sim}{y}_{ij}'\underset{\sim}{\Sigma}_Y^{-1}\underset{\sim}{y}_{ij}\}$$

$$(2\pi)^{-\frac{1}{2}J(K-H)}\sigma^{-J(K-H)}\exp\{-(2\sigma^2)^{-1}\sum_{j=1}^{J}\sum_{k=1}^{K-H}v_{ijk}^2\}$$

$$= \quad c(\underset{\sim}{\omega})\exp\{J^{\frac{1}{2}}\underset{\sim}{\beta}'\underset{\sim}{\Sigma}_Z^{-1}\underset{\sim}{y}_{i1} - \tfrac{1}{2}\underset{\sim}{y}'_{i1}\underset{\sim}{\Sigma}_Z^{-1}\underset{\sim}{y}_{i1} - \tfrac{1}{2}\sum_{j=2}^{J}\underset{\sim}{y}_{ij}'\underset{\sim}{\Sigma}_Y^{-1}\underset{\sim}{y}_{ij}$$

$$-(2\sigma^2)^{-1}\sum_{j=1}^{J}\sum_{k=1}^{K-H}v_{ijk}^2\},$$

where $c(\underset{\sim}{\omega})$ is independent of the observations. This density can be written on Darmois-Koopman form (Sverdrup (1965), p.206) as

$$\exp\{\tau_0(\underset{\sim}{\omega}) + \sum_{p=1}^{(H+1)^2}\tau_p(\underset{\sim}{\omega})\;U_{ip}(x)\}\frac{dP_0}{d\mu}(x)\;, \tag{4.7}$$

where P_0 corresponds to $\underset{\sim}{\beta} = \underset{\sim}{0}$, $\underset{\sim}{\Sigma}_Z = \underset{\sim}{\Sigma}_Y = \underset{\sim}{I}$, $\sigma^2 = 1$ and $\tau_0(\underset{\sim}{\omega}) = \ln c(\underset{\sim}{\omega})$, $\tau_1(\underset{\sim}{\omega}) = J^{\frac{1}{2}}\sum_{h=1}^{H}\beta_h(\underset{\sim}{\Sigma}_Z^{-1})_{h1}$, $U_{i1} = Y_{i11},\ldots,\tau_{H+1}(\underset{\sim}{\omega})$

$= -\frac{1}{2}(\underset{\sim}{\Sigma}_Z^{-1} - \underset{\sim}{I})_{11}$, $U_{i,H+1} = Y_{i11}^2,\ldots,$etc.

We see that

the functions $\tau_p(\underset{\sim}{\omega})$ have continuous second order derivatives. (4.8)

Since the functions $\tau_p(\underset{\sim}{\omega})$ are bicontinuous and Ω' is open, we also conclude that

the region of convergence of $\int \exp\{\sum_{p=1}^{(H+1)^2} \tau_p U_{ip}(x)\} dP_o(x)$

contains all the points of the range of (4.9)

$\{\tau_1(\underset{\sim}{\omega}),\ldots,\tau_{(H+1)^2}(\underset{\sim}{\omega})\}$ as inner points.

The results (4.8) and (4.9) are exactly the assumptions (i) and (ii) in Sverdrup (1965), appendix B . Assumption (iii) there follows from lemma (4.5) and the fact that $\underset{I\to\infty}{\text{plim}}\ \underset{\sim}{\omega}^* = \underset{\sim}{\omega}$, which is easy to check. Point (a) of theorem (4.6) follows from the conclusion in Sverdrup (1965), p. 211.

Point (b) is now easy to prove. If $\underset{\sim}{\omega}^* \in \Omega$, then $\underset{\sim}{\omega}^*$ maximizes L in Ω since it also maximizes L in the larger region Ω' . Thus we need only prove that $\lim_{I\to\infty} P_{\underset{\sim}{\omega}}(\underset{\sim}{\omega}^* \in \Omega) = 1$.

Since Ω is open, there is an $\epsilon > 0$ such that $\{\underset{\sim}{\omega}' \mid \|\underset{\sim}{\omega}'-\underset{\sim}{\omega}\| < \epsilon\} \subset \Omega$, and hence $P_{\underset{\sim}{\omega}}(\underset{\sim}{\omega}^* \in \Omega) \geq P_{\underset{\sim}{\omega}}(\|\underset{\sim}{\omega}^*-\underset{\sim}{\omega}\| < \epsilon)$. By the Fisherconsistency of $\underset{\sim}{\omega}^*$, the right side and hence also the left side of this inequality tend to 1 as $I \to \infty$. This completes the proof of theorem (4.6) .

4 B. <u>The estimator</u> $\underset{\sim}{\omega}^*$ <u>expressed by the original observations.</u> There exists a matrix $F^{H\times H}$ such that the columns of

$\underset{\sim}{T}(\underset{\sim}{T}'\underset{\sim}{T})^{-1}\underset{\sim}{F}$ form an orthonormal basis for the space spanned by the columns of $\underset{\sim}{T}$. By (3.1) and the orthogonality of the matrix

$$\begin{pmatrix} \underset{\sim}{F}'(\underset{\sim}{T}'\underset{\sim}{T})^{-1}\underset{\sim}{T}' \\ \underset{\sim}{E}' \end{pmatrix} ,$$

where $\underset{\sim}{E}$ is defined in subsection 3 A, we get

$$\|\underset{\sim}{X}_{ij}\|^2 = \left\|\begin{pmatrix} \underset{\sim}{F}'\underset{\sim}{\hat{B}}_{ij} \\ \underset{\sim}{V}_{ij} \end{pmatrix}\right\|^2 = \underset{\sim}{\hat{B}}'_{ij}\underset{\sim}{F}\underset{\sim}{F}'\underset{\sim}{\hat{B}}_{ij} + \underset{\sim}{V}'_{ij}\underset{\sim}{V}_{ij}. \tag{4.10}$$

Again due to orthonormality we have $\underset{\sim}{I} = \{\underset{\sim}{T}(\underset{\sim}{T}'\underset{\sim}{T})^{-1}\underset{\sim}{F}\}'\{\underset{\sim}{T}(\underset{\sim}{T}'\underset{\sim}{T})^{-1}\underset{\sim}{F}\}$ $= \underset{\sim}{F}'(\underset{\sim}{T}'\underset{\sim}{T})^{-1}\underset{\sim}{F}$, or equivalently that $\underset{\sim}{T}'\underset{\sim}{T} = \underset{\sim}{F}\underset{\sim}{F}'$. Hence

$$\underset{\sim}{\hat{B}}'_{ij}\underset{\sim}{F}\underset{\sim}{F}'\underset{\sim}{\hat{B}}_{ij} = \underset{\sim}{\hat{B}}'_{ij}\underset{\sim}{T}'\underset{\sim}{T}\underset{\sim}{\hat{B}}_{ij} . \tag{4.11}$$

We also have

$$\underset{\sim}{\hat{B}}'_{ij}\underset{\sim}{T}'\underset{\sim}{T}\underset{\sim}{\hat{B}}_{ij} = \underset{\sim}{X}'_{ij}\underset{\sim}{T}(\underset{\sim}{T}'\underset{\sim}{T})^{-1}(\underset{\sim}{T}'\underset{\sim}{T})\underset{\sim}{\hat{B}}_{ij} = \underset{\sim}{X}'_{ij}\underset{\sim}{T}\underset{\sim}{\hat{B}}_{ij} . \tag{4.12}$$

Combining (4.10), (4.11) and (4.12) we get

$$\begin{aligned} \underset{\sim}{V}'_{ij}\underset{\sim}{V}_{ij} = \underset{\sim}{X}'_{ij}\underset{\sim}{X}_{ij} - \underset{\sim}{\hat{B}}'_{ij}\underset{\sim}{F}\underset{\sim}{F}'\underset{\sim}{\hat{B}}_{ij} &= \underset{\sim}{X}'_{ij}\underset{\sim}{X}_{ij} - 2\underset{\sim}{X}'_{ij}\underset{\sim}{T}\underset{\sim}{\hat{B}}_{ij} \\ &\quad + (\underset{\sim}{T}\underset{\sim}{\hat{B}}_{ij})'(\underset{\sim}{T}\underset{\sim}{\hat{B}}_{ij}) \\ &= \|\underset{\sim}{X}_{ij} - \underset{\sim}{T}\underset{\sim}{\hat{B}}_{ij}\|^2 , \end{aligned}$$

which by (4.4) may be written

$$\sum_{k=1}^{K-H} V_{ijk}^2 = S_{ij}^2 \quad . \tag{4.13}$$

The first column of $\underset{\sim}{G}$ in subsection 3 B is $J^{-\frac{1}{2}}(1,\ldots,1)'$, and by (3.4) we have

$$\underset{\sim}{Y}_{i1} = J^{\frac{1}{2}}\,\hat{\underset{\sim}{B}}_{i\cdot} \quad . \tag{4.14}$$

By lemma 3.3.1 on p. 52 in Anderson (1958) we have

$$\sum_{j=1}^{J} \underset{\sim}{Y}_{ij}\underset{\sim}{Y}'_{ij} = \sum_{j=1}^{J} \hat{\underset{\sim}{B}}_{ij}\hat{\underset{\sim}{B}}'_{ij} \quad . \tag{4.15}$$

Combining (4.14) and (4.15) we get

$$\begin{aligned}\sum_{j=2}^{J} \underset{\sim}{Y}_{ij}\underset{\sim}{Y}'_{ij} &= \sum_{j=1}^{J} \underset{\sim}{Y}_{ij}\underset{\sim}{Y}'_{ij} - \underset{\sim}{Y}_{i1}\underset{\sim}{Y}'_{i1} = \sum_{j=1}^{J} \hat{\underset{\sim}{B}}_{ij}\hat{\underset{\sim}{B}}'_{ij} - J\hat{\underset{\sim}{B}}_{i\cdot}\hat{\underset{\sim}{B}}'_{i\cdot} \ , \\ &= \sum_{j=1}^{J} (\hat{\underset{\sim}{B}}_{ij}-\hat{\underset{\sim}{B}}_{i\cdot})(\hat{\underset{\sim}{B}}_{ij}-\hat{\underset{\sim}{B}}_{i\cdot})' \quad .\end{aligned} \tag{4.16}$$

The first column of $\underset{\sim}{H}$ in subsection 3 C is $I^{-\frac{1}{2}}(1,\ldots,1)$, and hence

$$\underset{\sim}{Z}_1 = I^{\frac{1}{2}}\underset{\sim}{Y}_{\cdot 1} = (IJ)^{\frac{1}{2}}\hat{\underset{\sim}{B}}_{\cdot\cdot} \quad . \tag{4.17}$$

By repeating the arguments leading to (4.16) we also find

$$\sum_{i=2}^{I} \underset{\sim}{Z}_i\underset{\sim}{Z}'_i = J\sum_{i=1}^{I} (\hat{\underset{\sim}{B}}_{i\cdot}-\hat{\underset{\sim}{B}}_{\cdot\cdot})(\hat{\underset{\sim}{B}}_{i\cdot}-\hat{\underset{\sim}{B}}_{\cdot\cdot})' \quad . \tag{4.18}$$

(4.13), (4.16), (4.17) and (4.18) establishes the relations given in (4.3) between canonical and original variables.

The modified maximum likelihood estimators $\underset{\sim}{\Sigma}_D^*$ and $\underset{\sim}{\Sigma}_C^*$ are now determined by the relation (3.6).

4 C . By application of theorem 3.3.2 on p. 53 in Anderson (1958) we find that $E\underset{\sim}{\Sigma}_Y^* = \underset{\sim}{\Sigma}_Y$ and $E\underset{\sim}{\Sigma}^*_Z = (I-1)I^{-1}\underset{\sim}{\Sigma}_Z$. We also have $E\underset{\sim}{\beta}^* = \underset{\sim}{\beta}$ and $E\sigma^{2*} = \sigma^2$. Hence $\underset{\sim}{\beta}^*$,σ^{2*} and $\underset{\sim}{\Sigma}_D^*$ are unbiased estimators. An unbiased estimator of $\underset{\sim}{\Sigma}_C$ is $I\{(I-1)J\}^{-1}\underset{\sim}{\Sigma}_Z^* - \underset{\sim}{\Sigma}_Y^*$. It follows from elementary limit theorems that even if we drop the normality assumptions, $\underset{\sim}{\omega}^*$ is a consistent estimator in the sense that $\text{plim}(\beta^* , \underset{\sim}{\Sigma}_Z^*) = (\underset{\sim}{\beta} , \underset{\sim}{\Sigma}_Z)$ as $I \to \infty$, $\text{plim}\, \underset{\sim}{\Sigma}_Y^* = \underset{\sim}{\Sigma}_Y$ as $I(J-1) \to \infty$ and $\text{plim}\, \sigma^{2*} = \sigma^2$ as $IJ(K-H) \to \infty$. In the one-way classification situation $J=1$ and $\underset{\sim}{\Sigma}_Y^*$ is not defined. $\underset{\sim}{\Sigma}_Z^*$, with $J=1$ and $\hat{\underset{\sim}{B}}_{i\cdot} = \hat{\underset{\sim}{B}}_i$, is now the estimator of $\underset{\sim}{\Sigma}_B + \sigma^2\underset{\sim}{M}_W$, and the estimators of $\underset{\sim}{\beta}$ and σ^2 remain unchanged. Rao (1965) studied the case $J=1$ and found that $\hat{\underset{\sim}{B}}_i$ is a BLU estimator of $\underset{\sim}{\beta}$ based on $\underset{\sim}{X}_i$ and a BLU predictor of $\underset{\sim}{B}_i$.

5. Testing hypotheses in the case of one regressor.

5 A. The observations are now

$$X_{ijk} = (\beta+C_i+D_{ij})t_k + U_{ijk}, \quad \begin{matrix} i=1,\dots,I \ ; \ j=1,\dots,J \ ; \\ k=1,\dots,K \ , \end{matrix} \qquad (5.1)$$

where β is a constant, $C_i \sim N_1(0,\sigma_C^2)$, $D_{ij} \sim N_1(0,\sigma_D^2)$ and $U_{ijk} \sim N_1(0,\sigma^2)$. All C_i , D_{ij} and U_{ijk} are independent. The design matrix is $\underset{\sim}{T} = (t_1,\dots t_k)$, and the matrix $\underset{\sim}{M}_W$ in (3.3) reduces to the scalar

$$m = (\sum_{k=1}^{K} t_k^2)^{-1} .$$

The canonical form (3.7) becomes

$$
\begin{aligned}
&Z_1 \sim N_1\{(IJ)^{\frac{1}{2}}\beta,\ \sigma_Z^2\} , \\
&Z_i \sim N_1(0,\ \sigma_Z^2) , \qquad\qquad i=2,\dots,I , \\
&Y_{ij} \sim N_1(0,\ \sigma_Y^2) , \quad i=1,\dots,I ;\ j=2,\dots,J, \\
&V_{ijk} \sim N_1(0,\ \sigma^2) , \quad i=1,\dots,I ;\ j=1,\dots,J ; \\
&\qquad\qquad\qquad\qquad k=1,\dots,K-1 .
\end{aligned}
\tag{5.2}
$$

where, analogous to (3.6),

$$\sigma_Y^2 = \sigma_D^2 + \sigma^2 m \quad \text{and} \quad \sigma_Z^2 = J\sigma_C^2+\sigma_D^2+\sigma^2 m . \tag{5.3}$$

(3.8) reduces to

$$
\begin{aligned}
dP_{\underset{\sim}{\omega}} = A'\exp(&-(2\sigma_Z^2)^{-1}\Big[\{z_1-(IJ)^{\frac{1}{2}}\beta\}^2+\sum_{i=2}^{I} z_i^2\Big] \\
&-(2\sigma_Y^2)^{-1}\sum_{i=1}^{I}\sum_{j=2}^{J} y_{ij}^2 - (2\sigma^2)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K-1} v_{ijk}^2)d\mu ,
\end{aligned}
\tag{5.4}
$$

where A' is an analytic function of $\underset{\sim}{\omega} = (\beta,\sigma_Z^2,\sigma_Y^2,\sigma^2)$. The parameter space is $\Omega : -\infty < \beta < \infty$, $\sigma_Z^2 - \sigma_Y^2\ (=J\sigma_C^2) > 0$, $\sigma_Y^2 - \sigma^2 m\ (=\sigma_D^2) > 0$ and $\sigma^2 > 0$.

5 B. We want to test hypotheses concerning the parameters β , $\varkappa$, Δ and σ^2 , where

$$\varkappa = \sigma_C^2(\sigma_D^2 + \sigma^2 m)^{-1} \quad \text{and} \quad \Delta = \sigma_D^2/\sigma^2 . \tag{5.5}$$

In example 1 B $\varkappa$ expresses interindividual variations, as measured by σ_C^2 , in fractions of intraindividual variations and measurement error variations, as measured by $\sigma_D^2 + \sigma^2 m$.

Δ expresses intraindividual variations in fractions of measurement error variations. From (5.3) and (5.5) we get

$$\sigma_Y^2 = (\Delta + m)\sigma^2 \quad \text{and} \quad \sigma_Z^2 = (J\varkappa+1)(\Delta+m)\sigma^2 \;. \tag{5.6}$$

For testing the hypotheses $\beta \leqq \beta_o$, $\varkappa \leqq \varkappa_o$, $\Delta \leqq \Delta_o$ and $\sigma^2 \leqq \sigma_o^2$ it is convenient to rewrite (5.4) by substituting (5.6) , adding and subtracting $(IJ)^{\frac{1}{2}}\beta_o$ in the term $\{z_1-(IJ)^{\frac{1}{2}}\beta\}^2$ and dividing and multiplying by the density $dP_o/d\mu$ corresponding to the parameter point defined by $\beta = \beta_o$, $\varkappa = \varkappa_o$, $\Delta = \Delta_o$ and $\sigma^2 = \sigma_o^2$. After collecting all constant factors in one factor A and reordering the remaining factors in a straightforward way, we finally get

$$\begin{aligned} dP_{\underset{\sim}{\omega}} = A\exp(& \tau_1 z_1+\tau_2[\{z_1-(IJ)^{\frac{1}{2}}\beta_o\}^2+u] \\ & + \tau_3\{(J\varkappa_o+1)^{-1}[\{z_1-(IJ)^{\frac{1}{2}}\beta_o\}^2+u] +v\} \\ & + \tau_4\{(J\varkappa_o+1)^{-1}(\Delta_o+m)^{-1}[\{z_1-(IJ)^{\frac{1}{2}}\beta_o\}^2+u] \\ & + (\Delta_o+m)^{-1}v+w\})dP_o \quad , \end{aligned} \tag{5.7}$$

where

$$\tau_1 = \frac{(IJ)^{\frac{1}{2}}(\beta-\beta_o)}{(J\varkappa+1)(\Delta+m)\sigma^2} \quad , \quad \tau_2 = \frac{1}{2(\Delta+m)\sigma^2}\left(\frac{1}{J\varkappa_o+1} - \frac{1}{J\varkappa+1}\right) ,$$
$$\tau_3 = \frac{1}{2\sigma^2}\left(\frac{1}{\Delta_o+m} - \frac{1}{\Delta+m}\right) \quad \text{and} \quad \tau_4 = \tfrac{1}{2}\left(\frac{1}{\sigma_o^2} - \frac{1}{\sigma^2}\right) , \tag{5.8}$$

and

$$U = \sum_{i=2}^{I} Z_i^2 \;,\; V = \sum_{i=1}^{I}\sum_{j=2}^{J} Y_{ij}^2 \quad \text{and} \quad W = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K-1} V_{ijk}^2 \tag{5.9}$$

The distributions $P_{\underset{\sim}{\omega}}$ given by (5.7) constitute a multiparameter exponential family. We see that P_0 corresponds to the value $\underset{\sim}{0}$ of the exponential parameter $\underset{\sim}{\tau}$.

5 C. Testing the hypothesis $\beta \leq \beta_0$. This hypothesis may be expressed equivalently as $\tau_1 \leqq 0$. Let

$$\underset{\sim}{R} = (\{Z_1-(IJ)^{\frac{1}{2}}\beta_0\}^2 + U \ , \quad (J\varkappa_0+1)^{-1}[\{Z_1-(IJ)^{\frac{1}{2}}\beta_0\}^2+U] + V \ ,$$

$$\{(J\varkappa_0+1)(\Delta_0+m)\}^{-1}[\{Z_1-(IJ)^{\frac{1}{2}}\beta_0\}^2+\tfrac{1}{2}U] + (\Delta_0+m)^{-1}V+W) \ .$$

According to Lehmann (1959), theorem 3 on p. 136, a UMP unbiased level ϵ test has the critical function φ_β which is equal to 1, $\gamma(\underset{\sim}{r})$ or 0 as $\{Z_1-(IJ)^{\frac{1}{2}}\beta_0\}$ is $>$, $=$ or $<$ $c'(\underset{\sim}{r})$, with γ and c' determined by

$$E_{\tau_1=0}\{\varphi_\beta(Z_1,U,V,W)\,|\,\underset{\sim}{R} = \underset{\sim}{r}\} = \epsilon \quad \text{for all } \underset{\sim}{r} \ . \tag{5.10}$$

(In the referred theorem it is assumed that the domain of variation of $\underset{\sim}{\tau}$ is a convex space of dimension equal to the number of components in $\underset{\sim}{\tau}$. This assumption is used only to establish that the domain of variation of (τ_1,τ_2,τ_3) on the boundary $\{\underset{\sim}{\tau}\ ;\ \tau_1 = 0\}$ contains a non-degenerate rectangle, which is satisfied here).

Since $z_1-(IJ)^{\frac{1}{2}}\beta_0 \gtreqless c'(\underset{\sim}{r}) \iff \{z_1-(IJ)^{\frac{1}{2}}\beta_0\}/\sqrt{\{z_1-(IJ)^{\frac{1}{2}}\beta_0\}^2+u} \gtreqless c''(\underset{\sim}{r})$ $\iff \{z_1-(IJ)^{\frac{1}{2}}\beta_0\}u^{-\frac{1}{2}} \gtreqless c'''(\underset{\sim}{r})$, we have

$$\varphi_\beta = \begin{cases} 1 & \text{when } \dfrac{z_1-(IJ)^{\frac{1}{2}}\beta_0}{u^{\frac{1}{2}}}(I-1)^{\frac{1}{2}} > c(\underset{\sim}{r}) \\ \gamma(\underset{\sim}{r}) & \text{"} \quad \text{-- " --} \quad = \text{-"-} \\ 0 & \text{otherwise} \end{cases} \tag{5.11}$$

Now U and Z_1 are independent, $U/\sigma_Z^2 \sim \chi^2_{I-1}$ and when $\tau_1 = 0$, $\{Z_1-(IJ)^{\frac{1}{2}}\beta_o\}/\sigma_Z \sim N_1(o,1)$. Hence, on the boundary we have $T = \{Z_1-(IJ)^{\frac{1}{2}}\beta_o\}(I-1)^{\frac{1}{2}}U^{-\frac{1}{2}} \sim T_{I-1}$, which is independent of $\underset{\sim}{\tau}$. $\underset{\sim}{R}$ is sufficient on the boundary, and according to Lehmann (1959), theorem 1 on p. 132, $\underset{\sim}{R}$ is also complete on the boundary. Then according to a theorem of Basu (1955) $\underset{\sim}{R}$ and T are independent on the boundary, and hence (5.10) is satisfied by the choice $\gamma(\underset{\sim}{r}) = 0$ and $c(\underset{\sim}{r}) = t_{1-\epsilon, I-1}$. The one-dimensional versions of (4.17) and (4.18) are $Z_1 = (IJ)^{\frac{1}{2}}\hat{B}_{\cdot\cdot}$ and $U = J\sum_{i=1}^{I}(\hat{B}_{i\cdot}-\hat{B}_{\cdot\cdot})^2$, and our result can therefore be expressed as follows.

<u>Theorem (5.12).</u> A UMP unbiased level ϵ test for the hypothesis $\beta \leqq \beta_o$ is given by the critical function

$$\varphi_\beta = \begin{cases} 1 \text{ when } \dfrac{I^{\frac{1}{2}}(\hat{b}_{\cdot\cdot}-\beta_o)}{\sqrt{(I-1)^{-\frac{1}{2}}\sum_{i=1}^{I}(\hat{b}_{i\cdot}-\hat{b}_{\cdot\cdot})^2}} \geqq t_{1-\epsilon;I-1}, \\ 0 \text{ otherwise.} \end{cases}$$

5 D. <u>Testing the hypothesis $\sigma_C^2 \leqq \varkappa_o(\sigma_D^2+\sigma^2 m)$.</u> This hypothesis may be expressed equivalently as $\varkappa \leqq \varkappa_o$ or as $\tau_2 \leqq 0$. Let $\underset{\sim}{S} = (Z_1, (J\varkappa_o+1)^{-1}[\{Z_1-(IJ)^{\frac{1}{2}}\beta_o\}^2+U] + V,$ $\{(J\varkappa_o+1)(\Delta_o+m)\}^{-1}[\{Z_1-(IJ)^{\frac{1}{2}}\beta_o\}^2+U] + (\Delta_o+m)^{-1}V + W)$.

Again according to Lehmann (1959), theorem 3 on p. 136, a UMP unbiased level ϵ test has the critical function φ_C which is equal to 1, $\gamma(\underset{\sim}{s})$ or 0 as $\{z_1-(IJ)^{\frac{1}{2}}\beta_0\}^2+u$ is $>$, $=$ or $<$ $c(\underset{\sim}{s})$, with γ and c' determined by

$$E_{\tau_2=0}\{\varphi_C(Z_1,U,V,W)\,|\,\underset{\sim}{S}=\underset{\sim}{s}\} = \epsilon \quad \text{for all } \underset{\sim}{s} \;. \tag{5.13}$$

Since $\{Z_1-(IJ)^{\frac{1}{2}}\beta_0\}^2$ and V are functions of $\underset{\sim}{S}$, we have $\{z_1-(IJ)^{\frac{1}{2}}\beta_0\}^2+u \gtreqless c'(\underset{\sim}{s}) \iff u \gtreqless c''(\underset{\sim}{s}) \iff u\,v^{-1} \gtreqless c'''(\underset{\sim}{s})$ and hence

$$\varphi_C = \begin{cases} 1 & \text{when } \dfrac{u}{v}\,\dfrac{I(J-1)}{I-1}\,\dfrac{1}{J\kappa_0+1} > c(\underset{\sim}{s}) \,, \\ \gamma(\underset{\sim}{s}) & \text{———} \;\;''\;\; \text{————} = -\;''\; = \,, \\ 0 & \text{otherwise} \,. \end{cases}$$

By (5.2), (5.6) and (5.9) U and V are independent, $U\{(J\kappa+1)(\Delta+m)\sigma^2\}^{-1} \sim \chi^2_{I-1}$ and $V\{(\Delta+m)\sigma^2\}^{-1} \sim \chi^2_{I(J-1)}$.

Therefore

$$\begin{aligned} &\frac{U\{(J\kappa+1)(\Delta+m)\sigma^2\}^{-1}}{V\{(\Delta+m)\;\sigma^2\}^{-1}}\;\frac{I(J-1)}{I-1} \\ &= \frac{U}{V}\;\frac{I(J-1)}{I-1}\;\frac{1}{J\kappa+1} \sim F_{I-1,I(J-1)} \;. \end{aligned} \tag{5.14}$$

Thus the distribution of UV^{-1} is independent of $\underset{\sim}{\tau}$ on the boundary, where $\kappa = \kappa_0$. $\underset{\sim}{S}$ is sufficient and complete on the boundary, and using the theorem of Basu (1955) we conclude that UV^{-1} and $\underset{\sim}{S}$ are independent on the boundary. We may put $\gamma(\underset{\sim}{s}) = 0$ since UV^{-1} has a continuous distribution. (5.13) then becomes

$P_{\varkappa=\varkappa_0}\left[UV^{-1}\{I(J-1)\}\{(I-1)(J\varkappa_0+1)\}^{-1} > c(\underset{\sim}{S})\,|\underset{\sim}{S}=\underset{\sim}{s}\right] = \epsilon$ for all $\underset{\sim}{s}$, and by (5.14) and the independence of UV^{-1} and $\underset{\sim}{S}$ when $\varkappa=\varkappa_0$ we may choose $c(\underset{\sim}{s}) = f_{1-\epsilon;I-1,I(J-1)}$ for all $\underset{\sim}{s}$. The power function is obtained from (5.14). The one-dimensional versions of (4.18) and (4.16) imply $U = J\sum_{i=1}^{I}(\hat{B}_{i\cdot}-\hat{B}_{\cdot\cdot})^2$ and $V = \sum_{i=1}^{I}\sum_{j=1}^{J}(\hat{B}_{ij}-\hat{B}_{i\cdot})^2$, and by substitution of these expressions we arrive at the following result.

<u>Theorem (5.15).</u> A UMP unbiased level ϵ test for the hypothesis $\varkappa \leqq \varkappa_0$ is given by the critical function

$$\varphi_C = \begin{cases} 1 \text{ when } \dfrac{J\sum_{i=1}^{I}(\hat{b}_{i\cdot}-\hat{b}_{\cdot\cdot})^2}{\sum_{i=1}^{I}\sum_{j=1}^{J}(\hat{b}_{ij}-\hat{b}_{i\cdot})^2}\;\dfrac{I(J-1)}{I-1}\;\dfrac{1}{J\varkappa_0+1} \geqq f_{1-\epsilon;I-1,I(J-1)}, \\ 0 \text{ otherwise.} \end{cases}$$

The power function of the test is

$$\beta_C(\varkappa) = 1 - F_{I-1,I(J-1)}\left(\frac{J\varkappa_0+1}{J\varkappa+1}\,f_{1-\epsilon;\,I-1,I(J-1)}\right).$$

5 E. <u>Testing the hypothesis $\sigma_D^2 \leqq \Delta_0\sigma^2$.</u> This hypothesis may be expressed equivalently as $\Delta \leqq \Delta_0$ or as $\tau_3 \leqq 0$. By reasoning along the same lines as in the preceding two subsections we get the following theorem.

Theorem (5.16). A UMP unbiased level ϵ test for the hypothesis $\Delta \leqq \Delta_0$ is given by the critical function

$$\varphi_D = \begin{cases} 1 \text{ when } \dfrac{\sum_{i=1}^{I}\sum_{j=1}^{J}(\hat{b}_{ij}-\hat{b}_{i\cdot})^2}{\sum_{i=1}^{I}\sum_{j=1}^{J} s_{ij}^2}\,\dfrac{J(K-1)}{J-1}\,\dfrac{1}{\Delta_0+m} \geqq f_{1-\epsilon;I(J-1),IJ(K-1)}\ , \\ 0 \text{ otherwise.} \end{cases}$$

The power function of the test is

$$\beta_D(\Delta) = 1-F_{I(J-1),IJ(K-1)}\left(\frac{\Delta_0+m}{\Delta+m}\ f_{1-\epsilon;I(J-1),IJ(K-1)}\right) .$$

5 F. The hypothesis $\sigma^2 \leqq \sigma_0^2$ is equivalent to $\tau_4 \leqq 0$. By the same kind of reasoning as above we find that a UMP unbiased level ϵ test of this hypothesis rejects when $\sigma_0^{-2} \sum_{i=1}^{I} \sum_{j=1}^{J} S_{ij}^2 \geqq \chi^2_{1-\epsilon;IJ(K-1)}$. We easily find UMP unbiased level ϵ tests of the reversed hypotheses $\beta \geqq \beta_0$, $\varkappa \geqq \varkappa_0$, $\Delta \geqq \Delta_0$ and $\sigma^2 \geqq \sigma_0^2$. Their critical functions may be specified by reversing the inequalities and replacing upper ϵ-points by lower ϵ-points in the critical functions defined above. Each test statistic considered in this section has the property of being a strictly monotonic function of the boundary value which defines the corresponding hypothesis. Thus one- and two-sided confidence intervals may be obtained in an obvious way.

5 G. <u>One-way classification.</u> When $J=1$, the canonical observations are given by (3.12) with $H=1$, $\underset{\sim}{\Sigma}_B=\sigma_B^2$ and $\underset{\sim}{M}_W=m$.

Their density (with respect to μ) is given by (5.4) with $\sigma_Z^2 = \sigma_B^2 + \sigma^2 m$ and $J=1$ (Implying that the second last term in the exponent drops out) . Theorem (5.12) is still valid, but the theorems (5.15) and (5.16) are not applicable when $J=1$. However, by the now familiar way of reasoning we easily get the following result, with $\hat{B}_i = \hat{B}_{i1}$ and $S_i^2 = S_{i1}^2$.

Theorem (5.17). A UMP unbiased level ϵ test for the hypothesis $\lambda \leqq \lambda_o$, where $\lambda = \sigma_B^2/\sigma^2$, is given by the critical function

$$\varphi_B = \begin{cases} 1 & \text{when } \dfrac{\sum_{i=1}^{I} (\hat{b}_i - \hat{b}_.)^2}{\sum_{i=1}^{I} s_i^2} \dfrac{I(K-1)}{I-1} \dfrac{1}{\lambda_o + m} \geqq f_{1-\epsilon; I-1, I(K-1)} , \\ 0 & \text{otherwise.} \end{cases}$$

The power function is

$$\beta_B(\lambda) = 1 - F_{I-1, I(K-1)}\left(\frac{\lambda_o + m}{\lambda + m} f_{1-\epsilon;\; I-1, I(K-1)} \right) .$$

5 H. Invariance considerations. Consider the following groups of transformations. G_1 is the translations $gZ_1 = Z_1 + b$, G_2 the orthogonal transformations of $Z_2, \ldots, Z_I$, G_3 the orthogonal transformations of $Y_{12}, \ldots, Y_{IJ}$ (when $J > 1$) , G_4 the orthogonal transformations of the V_{ijk}'s and G_5 the scale transformations which multiply all variables by the same positive constant. The product of any subset of these five groups is itself a group. (Lehmann (1959), theorem 2 on p. 218).

The problem of testing $\beta \leqq \beta_0$ remains invariant under any transformation in $G' = G_2 \times G_3 \times G_4 \times G_5$. The invariance principle claims that a test should depend on the observations only through $[\{Z_1-(LU)^{\frac{1}{2}}\beta_0\}U^{-\frac{1}{2}}, UV^{-1}, VW^{-1}]$, which is a maximal invariant with respect to G' . From (5.11) we see that φ_β is invariant, recalling that $\gamma = 0$ and c is constant. Since φ_β is UMP unbiased it is also admissible, and hence there exists no other invariant test which dominates φ_β . It may be shown that a UMP invariant test does not exist.

The problems of testing $\varkappa \leqq \varkappa_0$ and $\Delta \leqq \Delta_0$ remain invariant under any transformation in $G'' = G_1 \times G_2 \times G_3 \times G_4 \times G_5$. (UV^{-1}, VW^{-1}) is a maximal invariant, and hence φ_C and φ_D are invariant. Being UMP unbiased tests they are admissible, and hence admissible in the class of invariant tests. φ_C and φ_D are not UMP invariant, but we shall see presently that they possess a weaker optimum property. $(\varkappa, \Delta)$ is a maximal invariant with respect to the group induced in the parameter space by G'' . Since Δ appears as a nuisance parameter in the problem of testing $\varkappa \leqq \varkappa_0$ (against $\varkappa > \varkappa_0$) , it seems reasonable to restrict attention to the class Φ of invariant level ϵ tests whose power functions depend only on $\varkappa$. Let Φ' be the class of level ϵ test which are similar on the boundary $\varkappa = \varkappa_0$. Every test in Φ has constant power on the boundary, and so $\Phi \subset \Phi'$. By theorem (5.15) $\varphi_C \in \Phi$. From the proof of theorem 3 on p. 136 in Lehmann (1959) it follows that φ_C is UMP in the larger class Φ', and hence it is UMP in Φ . By the same kind of reasoning we prove an analogous result for φ_D . We summarize the results as follows.

<u>Theorem (5.18).</u> For testing the hypothesis $\varkappa \leqq \varkappa_0$ the test φ_C is UMP in the class of invariant level ϵ tests whose power function depends only on $\varkappa$. For testing the hypothesis $\Delta \leqq \Delta_0$ the test φ_D is UMP in the class of invariant level ϵ tests whose power function depends only on Δ.

Finally we look at the one-way classification. The problem of testing $\lambda \leqq \lambda_0$ remains invariant under any transformation in $G''' = G_1 \times G_2 \times G_4 \times G_5$. UW^{-1} is a maximal invariant with respect to G''' , and its distribution depends on $\underset{\sim}{\omega}$ only through λ . Therefore every invariant test must be similar. φ_B is invariant, and it is UMP in the class of similar tests, and hence we get the following theorem.

<u>Theorem (5.19).</u> For testing the hypothesis $\lambda \leqq \lambda_0$, φ_B is a UMP invariant level ϵ test.

5 I. Herbach (1959) established the analogues of our theorems (5.15) - (5.18) for the classical model II with one- and two-way classification. Lehmann (1959) , pp.286-293, proves the theorems (5.15) and (5.16) in the special case when $\underset{\sim}{T} =$ (1,...,1) , which is just the classical model II with nested classification.

6. Testing hypotheses concerning β in the case of more than one regressor.

6 A. A simultaneous test procedure for linear functions of β.

Suppose we are interested in all linear functions $a'\beta$ with a^H belonging to the Q-dimensional space A spanned by the columns of $A^{H\times Q}$ $(Q \leqq H)$. In other words we are interested in all linear functions of $A'\beta$. The natural thing to do is to look at the corresponding transformations of the individual mean least squares estimators, $A'\hat{B}_{i\cdot}$, $i=1,\ldots,I$, which constitute a sample from a Q-variate normal population with mean vector equal to $A'\beta$. The sample covariance matrix of the $A'\hat{B}_{i\cdot}$'s is

$$(I-1)^{-1} \sum_{i=1}^{I} (A'\hat{B}_{i\cdot}-A'\hat{B}_{i\cdot})(A'\hat{B}_{i\cdot}-A'\hat{B}_{\cdot\cdot})' =$$

$$= (I-1)^{-1}A'\{\sum_{i=1}^{I} (\hat{B}_{i\cdot}-\hat{B}_{\cdot\cdot})(\hat{B}_{i\cdot}-\hat{B}_{\cdot\cdot})'\}A = I\{J(I-1)\}^{-1} A'\Sigma_Z^* A ,$$

the last equality following from (4.3). According to Miller (1966), p. 196, with probability $1-\epsilon$

$$l'A'\beta \in l'A'\hat{B}_{\cdot\cdot} \pm \{Qf_{1-\epsilon;\, Q,I-Q}(I-1)(I-Q)^{-1}\}^{\frac{1}{2}}$$

$$[I^{-1}l'I\{J(I-1)\}^{-1}A'\Sigma_Z^* A l]^{\frac{1}{2}} , \quad \forall l^Q .$$

We state this result and the test procedure derived from it as a theorem.

<u>Theorem (6.1).</u> Let A be a Q-dimensional linear subspace of $\mathbb{R}^H$.

(a) With probability $1-\epsilon$

$$\underset{\sim}{a}'\underset{\sim}{\beta} \in \underset{\sim}{a}'\underset{\sim}{\hat{B}}_{\cdot\cdot} \pm \{f_{1-\epsilon;\, Q,I-Q}Q(I-Q)^{-1}\}^{\frac{1}{2}}(J^{-1}\underset{\sim}{a}'\Sigma_{Z}^{*}\underset{\sim}{a})^{\frac{1}{2}} \;, \quad \forall \underset{\sim}{a} \in A \;.$$

(b) The probability of one or more false statements is at most ϵ if we state that $\underset{\sim}{a}'\underset{\sim}{\beta} > \underset{\sim}{a}'\underset{\sim}{\beta}_o$ for every $\underset{\sim}{a} \in A$ for which

$$\underset{\sim}{a}'\underset{\sim}{\hat{B}}_{\cdot\cdot} > \underset{\sim}{a}'\underset{\sim}{\beta}_o + \{f_{1-\epsilon;Q,I-Q}Q(I-Q)^{-1}\}^{\frac{1}{2}}(J^{-1}\underset{\sim}{a}'\Sigma_{Z}^{*}\underset{\sim}{a})^{\frac{1}{2}} \;.$$

<u>Remark.</u> The constant $C(I,Q) = f_{1-\epsilon;Q,I-Q}Q(I-Q)^{-1}$ increases strictly when Q increases. Thus, in order to obtain short confidence intervals and sensitive tests in theorem (6.1), we should restrict attention to the smallest linear space A which generate all linear functions $\underset{\sim}{a}'\underset{\sim}{\beta}$ of interest. Here is a short proof of the assertion. Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be independent and distributed according to $N_1(0,1)$.

Put $Z_{I,Q} = \sum_{i=1}^{Q} X_i^2 \Big/ \sum_{i=1}^{I-Q} Y_i^2$, $1 \leqq Q < I$. $Z_{I,Q}(I-Q)/Q \sim F_{Q,I-Q}$, and hence

$$P\{Z_{I,Q} > C(I,Q)\} = P\{Z_{I,Q+1} > C(I,Q+1)\} \;. \tag{6.2}$$

(Their common value is ϵ). Since $Z_{I,Q+1} > Z_{I,Q}$ almost surely we also have

$$P\{Z_{I,Q} > C(I,Q)\} < P\{Z_{I,Q+1} > C(I,Q)\} \;. \tag{6.3}$$

Comparison of (6.2) and (6.3) shows that $C(I,Q+1) > C(I,Q)$.

(In the same manner we may prove the obvious thing that $C(I,Q)$ decreases when I increases) .

6 B. If we are interested only in the functions $\underset{\sim}{a}_1'\underset{\sim}{\beta}_1,\ldots,$ $\underset{\sim}{a}_P'\underset{\sim}{\beta}$, and $t_{1-\epsilon(2P)^{-1};I-1}(I-1)^{-\frac{1}{2}} < \{f_{1-\epsilon;Q,I-Q}Q(I-Q)^{-1}\}^{\frac{1}{2}}$, then the Bonferroni-intervals, (see Miller (1966), p. 67) ,

$$\underset{\sim}{a}_p'\underset{\sim}{\beta} \in \underset{\sim}{a}_p'\hat{\underset{\sim}{B}}_{\cdot\cdot} \pm t_{1-\epsilon(2P)^{-1};I-1}[\{J(I-1)\}^{-1}\underset{\sim}{a}_p'\underset{\sim}{\Sigma}_Z^*\underset{\sim}{a}_p]^{\frac{1}{2}} , \qquad (6.4)$$

$$p=1,\ldots,P \quad ,$$

are shorter than those in theorem (6.1) and have a simultaneous confidence level not less than $1-\epsilon$, and the corresponding test procedure is more sensitive than that in theorem (6.1) and has a level not larger than ϵ .

6 C . We apply these results to example 1 C. Suppose we are interested in the concentration of drug at a certain point of time , t_o . From (6.4) we get the level $1-\epsilon$ confidence interval $\beta_1+\beta_2t_o \in \hat{B}_{\cdot\cdot1}+\hat{B}_{\cdot\cdot2}t_o \pm t_{1-\frac{1}{2}\epsilon;I-1}\,\sigma^*_{t_o}\,I^{-\frac{1}{2}}$, where

$$\sigma_t^{*2} = (I-1)^{-1}\{\sum_{i=1}^{I}(\hat{B}_{i\cdot1}-\hat{B}_{\cdot\cdot1})^2 + 2t\sum_{i=1}^{I}(\hat{B}_{i\cdot1}-\hat{B}_{\cdot\cdot1})(\hat{B}_{i\cdot2}-\hat{B}_{\cdot\cdot2})$$
$$+t^2\sum_{i=1}^{I}(\hat{B}_{i\cdot2}-\hat{B}_{\cdot\cdot2})^2\} .$$

From this interval we get an interval for $dt_o\exp(\beta_1+\beta_2t_o)$. The effect at time t_o of the preparation may be compared with that of another preparation with known mean value $\underset{\sim}{\beta}^o$ of $\underset{\sim}{B}$ by means of (6.4) . At level ϵ we state that $\beta_1+\beta_2t_o$ is $>$ or

$< \beta_1^o+\beta_2^o t_o$ as $\hat{B}_{\cdot\cdot 1}+\hat{B}_{\cdot\cdot 2}t_o$ is $> \beta_1^o+\beta_2^o t_o+t_{1-\frac{1}{2}\epsilon;I-1}\sigma^*_{t_o}I^{-\frac{1}{2}}$ or $< \beta_1^o+\beta_2^o t_o-t_{1-\frac{1}{2}\epsilon;I-1}\sigma^*_{t_o}I^{-\frac{1}{2}}$. By point (a) in theorem (6.1) we obtain a Working-Hotelling kind of level $1-\epsilon$ confidence band,

$$\beta_1+\beta_2 t \in \hat{B}_{\cdot\cdot 1}+\hat{B}_{\cdot\cdot 2}t \pm \{f_{1-\epsilon;2,I-2}2(I-2)^{-1}\}^{\frac{1}{2}}\sigma^*_{t,I} \ , \ \forall t > 0 \ .$$

7. Testing hypotheses concerning $\underset{\sim}{\Sigma}_C$ and $\underset{\sim}{\Sigma}_D$ in the general case with more than one regressor.

7 A. Many hypotheses concerning $\underset{\sim}{\Sigma}_C$ and $\underset{\sim}{\Sigma}_D$ may be expressed in terms of $\underset{\sim}{\Sigma}_Z$ and $\underset{\sim}{\Sigma}_Y$ and tested by standard multivariate methods on basis of the transformed observations (3.5). We give a few examples.

The hypothesis of no interindividual variations, $H_1:\underset{\sim}{\Sigma}_C = \underset{\sim}{0}$, may be expressed equivalently as $H_1:\underset{\sim}{\Sigma}_Z = \underset{\sim}{\Sigma}_Y$. This hypothesis may be tested by the method in Anderson (1958), pp. 247-250. (It should be noted that the proposed test is designed for the alternative hypothesis $K_1:\underset{\sim}{\Sigma}_Z \neq \underset{\sim}{\Sigma}_Y$, while we are only interested in the restricted alternative $K_1':\underset{\sim}{\Sigma}_Z-\underset{\sim}{\Sigma}_Y$ is positive definit).

The hypothesis of no intraindividual variations, $H_2:\underset{\sim}{\Sigma}_D = 0$, may be expressed equivalently as $H_2:\underset{\sim}{\Sigma}_Y = \sigma^2\underset{\sim}{M}_W$. Let $\underset{\sim}{A}^{H\times H}$ be a nonsingular matrix such that $\underset{\sim}{A}\underset{\sim}{M}_W\underset{\sim}{A}' = \underset{\sim}{I}$, and hence $\underset{\sim}{A}\underset{\sim}{Y}_{ij} \sim N_H(\underset{\sim}{0},\underset{\sim}{A}\underset{\sim}{\Sigma}_Y\underset{\sim}{A}')$, $i=1,\dots,I$; $j=2,\dots,J$. We see that the hypothesis may be expressed as $H_2:(\underset{\sim}{A}\underset{\sim}{Y}_{ij})_k$, $i=1,\dots,I$; $j=2,\dots,J$; $k=1,\dots,H$, are independent and distributed according to $N_1(o,\sigma^2)$. A possible level ϵ test consists in rejecting when $\sum_{i=1}^{I}\sum_{j=2}^{J}\sum_{k=1}^{H}(\underset{\sim}{A}\underset{\sim}{Y}_{ij})_k^2 / \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K-H}V_{ijk}^2$ is larger than

$f_{1-\epsilon;\ I(J-1)H,IJ(K-H)}(J-1)H\{J(K-H)\}^{-1}$.

The composite hypothesis $H_3 : \Sigma_C = \Sigma_C^o$, $\Sigma_D = \Sigma_D^o$, $\sigma^2 = \sigma_o^2$, where $\underset{\sim}{\Sigma}_C^o$, $\underset{\sim}{\Sigma}_D^o$ and σ_o^2 are known, may be expressed equivalently as $H_3 : \underset{\sim}{\Sigma}_Z = J\underset{\sim}{\Sigma}_C^o + \underset{\sim}{\Sigma}_D^o + \sigma_o^2 \underset{\sim}{M}_W$, $\underset{\sim}{\Sigma}_Y = \underset{\sim}{\Sigma}_D^o + \sigma_o^2 \underset{\sim}{M}_W$, $\sigma^2 = \sigma_o^2$. We may test H_3 at a level not larger than ϵ by testing at level $\epsilon/3$ each component hypothesis by the method described by Anderson (1959) pp. 264-267. In example 1 C, $\underset{\sim}{\Sigma}_C^o$, $\underset{\sim}{\Sigma}_D^o$ and σ_o^2 may belong to an old preparation with known pharmacokinetic properties and with which the new preparation is to be compared.

In the final two subsections we shall consider hypotheses which are analogous to those in section 5.

7 B. <u>Hypotheses concerning inter- and intraindividual variations in a certain factor point.</u> Suppose that in example 1 C we are primarily interested in the concentration $dt_o \exp(B_1 + B_2 t_o)$ of drug at time t_o . It is then of interest to draw inference concerning inter- and intraindividual variations of $B_1 + B_2 t_o$.

In terms of our general model we are now interested in a particular linear function $\underset{\sim}{a}'\underset{\sim}{B}$. $\underset{\sim}{a}'\underset{\sim}{B}$ is normally distributed with expectation $\beta^{(a)} = \underset{\sim}{a}'\underset{\sim}{\beta}$ and variance $\underset{\sim}{a}'\underset{\sim}{\Sigma}_C\underset{\sim}{a} + \underset{\sim}{a}'\underset{\sim}{\Sigma}_D\underset{\sim}{a}$, where $\underset{\sim}{a}'\underset{\sim}{\Sigma}_C\underset{\sim}{a}$ and $\underset{\sim}{a}'\underset{\sim}{\Sigma}_D\underset{\sim}{a}$ represent interindividual and intraindividual variations respectively. We now transform the vectors in the canonical form (3.7) to scalar variables as follows.

$$Z_1^{(a)} = \underset{\sim}{a}'\underset{\sim}{Z}_1 \sim N_1\{(IJ)^{\frac{1}{2}}\beta^{(a)} , \sigma_{Z(a)}^2\} ,$$

$$Z_i^{(a)} = \underset{\sim}{a}'Z_i \sim N_1(0, \sigma_{Z(a)}^2) \quad , \ i=2,\ldots,I ,$$

$$Y_{ij}^{(a)} = \underset{\sim}{a}'\underset{\sim}{Y}_{ij} \sim N_1(0,\sigma^2_{Y(a)}) \ , \quad i=1,\dots,I \ ; \ j=2,\dots,J \ ,$$

$$V_{ijk} \sim N_1(0,\sigma^2) \ , \quad i=1,\dots,I \ ; \ j=1,\dots,J \ ; \ k=1,\dots,K-H \ ,$$

where $\sigma_Y^2 = \underset{\sim}{a}'\underset{\sim}{\Sigma}_D\underset{\sim}{a} + \sigma^2\underset{\sim}{a}'\underset{\sim}{M}_W\underset{\sim}{a}$ and $\sigma_Z^2 = J\underset{\sim}{a}'\underset{\sim}{\Sigma}_C\underset{\sim}{a} + \underset{\sim}{a}'\underset{\sim}{\Sigma}_D\underset{\sim}{a} + \sigma^2\underset{\sim}{a}'\underset{\sim}{M}_W\underset{\sim}{a}$. These variables are exactly on the form (5.2), and hence the hypotheses $\underset{\sim}{a}'\underset{\sim}{\Sigma}_C\underset{\sim}{a} \leqq \varkappa_0(\underset{\sim}{a}'\underset{\sim}{\Sigma}_D\underset{\sim}{a} + \sigma^2\underset{\sim}{a}'\underset{\sim}{M}_W\underset{\sim}{a})$ and $\underset{\sim}{a}'\underset{\sim}{\Sigma}_D\underset{\sim}{a} \leqq \Delta_0\sigma^2$ may be tested by replacing $\hat{b}_{ij}, K-1$ and m with $\underset{\sim}{a}'\hat{\underset{\sim}{b}}_{ij}, K-H$ and $\underset{\sim}{a}'\underset{\sim}{M}_W\underset{\sim}{a}$ in the test criterions in the theorems (5.15) - (5.16). The optimum properties, which these tests possess in the case with one single regressor, are not carried over to the general case.

7 C. <u>The case when the covariance matrices are proportional to known matrices.</u> We consider first the case $J=1$, which was interpreted as one-way classification in subsection 2 B. We assume that $\underset{\sim}{\Sigma}_B = \sigma_B^2\underset{\sim}{M}_B$, where $\underset{\sim}{M}_B$ is a known matrix and σ_B^2 an unknown positive constant.

In example 1 C this assumption is appropriate when the relative significance of- and interdependence between - the different sources of plasma concentration variations is assumed to be the same for the preparation under consideration and a similar preparation, which has been extensively studied in clinical trials, whereas the magnitude of the variations may be different for the two preparations. In that example the case $J=1$ is of special interest due to possible stochastic dependence between the results of clinical trials performed on the same person at different points of time. From the observational scheme with $J=1$ we can draw inference concerning $\underset{\sim}{\Sigma}_B$, which expresses the variation of $\underset{\sim}{B}$ in the patient population and hence is the relevant parameter in connection with large scale production of the preparation.

We now return to the general situation and consider the observations on the form (3.12). In terms of the parameters

$$\lambda = \sigma_B^2/\sigma^2 \quad \text{and} \quad \underset{\sim}{\tau}^{H\times 1} = (\underset{\sim}{\Sigma}_B + \sigma^2 \underset{\sim}{M}_W)^{-1}\underset{\sim}{\beta}' ,$$

their density is given by

$$dP_{\underset{\sim}{\omega}} = A'\exp\{\underset{\sim}{\tau}'\underset{\sim}{z}_1 - (2\sigma^2)^{-1}\sum_{i=1}^{I}\underset{\sim}{z}_i'(\lambda\underset{\sim}{M}_B+\underset{\sim}{M}_W)^{-1}\underset{\sim}{z}_i$$
$$- (2\sigma^2)^{-1}\sum_{i=1}^{I}\sum_{k=1}^{K-H} v_{ik}^2\}\,d\mu .$$

We will test the hypothesis $\lambda \leqq \lambda_o$. Letting P_o be the distribution corresponding to $\underset{\sim}{\tau} = \underset{\sim}{0}$, $\lambda = \lambda_o$ and $\sigma^2 = 1$, we can rewrite $dP_{\underset{\sim}{\omega}}$ in a way analogous to that described in subsection 5 B and get

$$dP_{\underset{\sim}{\omega}} = A\exp[\underset{\sim}{\tau}'\underset{\sim}{z}_1 + (2\sigma^2)^{-1}\sum_{i=1}^{I}\underset{\sim}{z}_i'\{(\lambda_o\underset{\sim}{M}_B+\underset{\sim}{M}_W)^{-1}-(\lambda\underset{\sim}{M}_B+\underset{\sim}{M}_W)^{-1}\}\underset{\sim}{z}_i$$
$$+ \{\tfrac{1}{2}-(2\sigma^2)^{-1}\}\{\sum_{i=1}^{I}\underset{\sim}{z}_i'(\lambda_o\underset{\sim}{M}_B+\underset{\sim}{M}_W)^{-1}\underset{\sim}{z}_i+\sum_{i=1}^{I}\sum_{k=1}^{K-H} v_{ik}^2\}]dP_o .$$

The distributions on the boundary, $\lambda = \lambda_o$, constitute a multiparameter exponential family, whose parameter space contains an open rectangle. Then, according to Lehmann (1959), theorem 1 on p. 132, a sufficient and complete statistic on the boundary is

$\underset{\sim}{R} = \{\underset{\sim}{Z}_1 \ , \sum_{i=1}^{I} Z_i'(\lambda_o \underset{\sim}{M}_B + \underset{\sim}{M}_W)^{-1} \underset{\sim}{Z}_i + W\}$, where $W = \sum_{i=1}^{I} \sum_{k=1}^{K-H} V_{ik}^2$. We consider the alternative $\underset{\sim}{\omega}_1$, defined by $\tau = \tau_1$, $\lambda = \lambda_1 (>\lambda_o)$ and $\sigma^2 = \sigma_1^2$. According to Sverdrup (1953), theorem 3, we have the following result. Among all tests which have similar power ϵ on the boundary, a most powerful test against the alternative $\underset{\sim}{\omega}_1$ has critical function $\varphi_{\underset{\sim}{\omega}_1}$, which is 1 , $\gamma(\underset{\sim}{r})$ or 0 as

$$\underset{\sim}{\Sigma}_1' \underset{\sim}{Z}_1 + (2\sigma_1^2)^{-1} \sum_{i=1}^{I} \underset{\sim}{Z}_i' \{(\lambda_o \underset{\sim}{M}_B + \underset{\sim}{M}_W)^{-1} - (\lambda_1 \underset{\sim}{M}_B + \underset{\sim}{M}_W)^{-1}\} \underset{\sim}{Z}_i + \{\tfrac{1}{2} - (2\sigma_1^2)^{-1}\}$$

$\{\sum_{i=1}^{I} \underset{\sim}{Z}_i'(\Delta_o \underset{\sim}{M}_B + \underset{\sim}{M}_W)^{-1} \underset{\sim}{Z}_i + w\}$ is $>$, $=$ or $<$ $c'(\underset{\sim}{r})$, with γ and c' determined by

$$E_{\lambda=\lambda_o} \{\varphi_{\underset{\sim}{\omega}_1}(Z_1,\ldots,Z_I \ ,W) | \underset{\sim}{R} = \underset{\sim}{r}\} = \epsilon \quad \text{for all } \underset{\sim}{r} \ . \tag{7.1}$$

Define the statistic F_{λ_1} by

$$F_{\lambda_1}(\underset{\sim}{Z}_2,\ldots,\underset{\sim}{Z}_{I1},W) = \frac{\sum_{i=2}^{I} \underset{\sim}{Z}_i'(\lambda_1 \underset{\sim}{M}_B + \underset{\sim}{M}_W)^{-1} \underset{\sim}{Z}_i + W}{\sum_{i=2}^{I} \underset{\sim}{Z}_i'(\lambda_o \underset{\sim}{M}_B + \underset{\sim}{M}_W)^{-1} \underset{\sim}{Z}_i + W} \tag{7.2}$$

After simple rearrangements in the test criterion we find that $\varphi_{\underset{\sim}{\omega}_1}$ is 1 , $\gamma(\underset{\sim}{r})$ or 0 as F_{λ_1} is $<$, $=$ or $>$ $c(\underset{\sim}{r})$, where γ and c is determined by (7.1) . $\underset{\sim}{Z}_2,\ldots,\underset{\sim}{Z}_I$, W are independent, and on the boundary we have $\sigma^{-1}\underset{\sim}{Z}_i \sim N_H(\underset{\sim}{O}, \lambda_o \underset{\sim}{M}_B + \underset{\sim}{M}_W)$ $i=2,\ldots,I$ and $\sigma^{-2}W \sim \chi^2_{I(K-H)}$. Hence, by dividing numerator and denominator in (7.2) by σ^2 , we see that the distribution of

F_{λ_1} is independent of $\underset{\sim}{\omega}$ on the boundary. Then by the earlier mentioned result of Basu (1955), F_{λ_1} is independent of $\underset{\sim}{R}$ on the boundary. Since F_{λ_1} has a continuous distribution, we can put $\gamma(\underset{\sim}{r}) = 0$. Then (7.1) becomes $P_{\lambda=\lambda_0}\{F_{\lambda_1} \leq c(\underset{\sim}{R})|\underset{\sim}{R}=\underset{\sim}{r}\} = \epsilon$ for all $\underset{\sim}{r}$, which by independence reduces to

$$P_0(F_{\lambda_1} \leqq c) = \epsilon \text{ , } c \text{ is a constant.} \tag{7.3}$$

The test depends on the alternative only through the parameter λ_1, and we therefore write φ_{λ_1} instead of $\varphi_{\underset{\sim}{\omega}_1}$. By dividing numerator and denominator in (7.2) by σ^2 we see that the power function β_{λ_1} depends on $\underset{\sim}{\omega}$ only through λ .

According to Lehmann (1959), lemma 1 on p. 126, φ_{λ_1} is most powerful unbiased against the alternative $\lambda=\lambda_1$ if it is a level ϵ test. We shall prove that $\beta_{\lambda_1}(\lambda)$ increases with increasing λ , which ensures that φ_{λ_1} has level ϵ . By theorem 3 on p. 341 in Anderson (1958) there exists a (nonsingular) matrix $\underset{\sim}{L}^{H\times H}$ such that $\underset{\sim}{L}\underset{\sim}{M}_B\underset{\sim}{L}' = \underset{\sim}{P} = \text{diag}(p_1,\dots,p_H)$ and $\underset{\sim}{L}\underset{\sim}{M}_W\underset{\sim}{L}' = \underset{\sim}{I}$. Thus for any $\lambda' > 0$ we have

$$\sum_{i=2}^{I} \underset{\sim}{Z}_i'(\lambda'\underset{\sim}{M}_B+\underset{\sim}{M}_W)^{-1}\underset{\sim}{Z}_i = \sum_{i=2}^{I} (\underset{\sim}{L}\underset{\sim}{Z}_i)'[\underset{\sim}{L}(\lambda'\underset{\sim}{M}_B+\underset{\sim}{M}_W)\underset{\sim}{L}']^{-1}(\underset{\sim}{L}\underset{\sim}{Z}_i)$$

$$= \sum_{i=2}^{I} (\underset{\sim}{L}\underset{\sim}{Z}_i)'(\lambda'\underset{\sim}{P}+\underset{\sim}{I})^{-1}(\underset{\sim}{L}\underset{\sim}{Z}_i) = \sum_{h=1}^{H} (\lambda' p_h+1)^{-1} \sum_{i=2}^{I} (\underset{\sim}{L}\underset{\sim}{Z}_i)_h^2$$

Now $\underset{\sim}{L}\underset{\sim}{Z}_i \sim N_H\{\underset{\sim}{0}, \sigma^2(\lambda\underset{\sim}{P}+\underset{\sim}{I})\}$, $i=2,\dots,I$, and hence we may write

$$\sum_{i=2}^{I} \underset{\sim}{Z}_i'(\lambda'\underset{\sim}{M}_B+\underset{\sim}{M}_W)^{-1}\underset{\sim}{Z}_i = \sum_{h=1}^{H}(\lambda p_h+1)(\lambda' p_h+1)^{-1}R_h\sigma^2 , \tag{7.4}$$

where $R_h \sim \chi^2_{I-1}$, $h=1,\ldots,H$. In the expression on the right side of (7.2) we substitute (7.4) with $\lambda' = \lambda_1$ in the numerator and with $\lambda' = \lambda_o$ in the denominator, put $Q_h = R_h\sigma^2 W^{-1}$ and get

$$F_{\lambda_1} = \frac{\sum_{h=1}^{H}(\lambda p_h+1)(\lambda_1 p_h+1)^{-1}Q_h+1}{\sum_{h=1}^{H}(\lambda p_h+1)(\lambda_o p_h+1)^{-1}Q_h+1} = g(Q_1,\ldots,Q_H;\lambda) .$$

The power function may be written $\beta_{\lambda_1}(\lambda) = \Pr\{g(Q_1,\ldots,Q_H;\lambda) \leq c\}$. Since the distribution of $(Q_1,\ldots,Q_H)$ is independent of λ , we need only prove that $g(Q_1,\ldots,Q_H;\lambda)$ decreases with increasing λ in order to establish that β_{λ_1} is an increasing function of λ . By differentiation of the above expression we find

$$\partial g/\partial\lambda = \frac{\{\sum_{h=1}^{H}(\lambda p_h+1)(\lambda_o p_h+1)^{-1}Q_h+1\}\sum_{k=1}^{H}p_k(\lambda_1 p_k+1)^{-1}Q_k}{\text{denominator}}$$

$$- \frac{\sum_{h=1}^{H}p_h(\lambda_o p_h+1)^{-1}Q_h\{\sum_{k=1}^{H}(\lambda p_k+1)(\lambda_1 p_k+1)Q_k+1\}}{\text{denominator}} .$$

The denominator is a squared number and hence positive. Straightforward calculations show that the numerator is equal to

$$(\lambda_o-\lambda)\sum_{h<k}(p_k-p_h)^2\{(\lambda_o p_h+1)(\lambda_1 p_k+1)(\lambda_o p_k+1)(\lambda_1 p_h+1)\}^{-1}Q_hQ_k$$

$$+ \sum_{k=1}^{H}\{(\lambda_1 p_k+1)^{-1} - (\lambda_o p_k+1)^{-1}\}p_kQ_k .$$

All Q_h are positive, and by the positive definiteness of $\underset{\sim}{P}$ also all p_h are positive. We conclude that the numerator in $\partial g/\partial\lambda$ is negative, and so $\partial g/\partial\lambda$ is negative. We summarize these results as follows.

Theorem (7.5). The most powerful unbiased level ϵ test for the hypothesis $\lambda \leqq \lambda_o$, where $\lambda = \sigma_B^2/\sigma^2$, against the alternatives $\lambda = \lambda_1 (> \lambda_o)$ is given by the critical function

$$\varphi_{\lambda_1} = \begin{cases} 1 & \text{when } F_{\lambda_1} \leqq c , \\ 0 & \text{otherwise} , \end{cases}$$

where F_{λ_1} is defined by (7.2) and c is determined by (7.3) .

Invariance considerations. The problem of testing H is invariant under translations $g\underset{\sim}{Z}_1 = \underset{\sim}{Z}_1 + \underset{\sim}{a}$, orthogonal transformations of all V_{ik}'s and common change of scale of all variables. $(W^{-1}Z_2,\ldots,W^{-1}Z_H)$ is a maximal invariant under the group of these transformations. Its distribution depends only on λ , and hence any invariant test must be similar. We see that the test φ_{λ_1} depends on the observations only through the maximal invariant, and hence it is invariant. Since φ_{λ_1} is most powerful against the alternatives $\lambda=\lambda_1$ among similar level ϵ tests, it is also most powerful against these alternatives in the smaller class of invariant level ϵ tests. We easily show that φ_{λ_1} is a maximin test of the hypothesis $\lambda \leqq \lambda_o$ against $\lambda \geqq \lambda_1$. The problem remains invariant under the group of transformations mentioned above. The conditions of the Hunt-Stein theorem are satisfied, (Lehmann (1959), theorem 2 on p. 336), and hence there exists an invariant maximin test. (Any almost invariant test is equivalent with an invariant test in this case according to Lehmann (1959), theorem 4 on p. 225). Assume ψ is an invariant maximin test. Then

$$\inf_{\lambda\geqq\lambda_1} E_{\underset{\sim}{\omega}}\psi \leqq \inf_{\lambda=\lambda_1} E_{\underset{\sim}{\omega}}\psi \leqq \beta_{\lambda_1}(\lambda_1) = \inf_{\lambda\geqq\lambda_1} \beta_{\lambda_1}(\lambda) .$$

The second inequality is due to the fact that φ_{λ_1} is most powerful invariant test against the alternatives $\lambda=\lambda_1$, and the last inequality follows from the monotonicity of β_{λ_1} .
We have proved

Theorem (7.6). The test φ_{λ_1} of the theorem (7.5) is most powerful among invariant level ϵ test for the hypothesis $\lambda \leqq \lambda_0$ against the alternatives $\lambda=\lambda_1$. It is a maximin test against the alternatives $\lambda \geqq \lambda_1$.

We finally express the test statistic by the original variables. For an arbitrary matrix $\underset{\sim}{M}^{H\times H} = (M_{hk})$ we have

$$\begin{aligned}\sum_{i=2}^{I} \underset{\sim}{Z}_i'\underset{\sim}{M}\underset{\sim}{Z}_i &= \sum_{i=2}^{I} \sum_{h=1}^{H} \sum_{k=1}^{H} M_{hk}Z_{ih}Z_{ik} \\ &= \sum_{h=1}^{H} \sum_{k=1}^{H} M_{hk} \sum_{i=2}^{I} Z_{ih}Z_{ik} .\end{aligned} \tag{7.7}$$

By (4.18), with J=1 and $\hat{\underset{\sim}{B}}_{i\bullet} = \hat{\underset{\sim}{B}}_i$,we have $\sum_{i=2}^{I} Z_{ih}Z_{ik}$ $= \sum_{i=1}^{I} (\hat{B}_{ih}-\hat{B}_{\bullet h})(\hat{B}_{ik}-\hat{B}_{\bullet k})$, and by substitution in (7.7) we get

$$\sum_{i=2}^{I} \underset{\sim}{Z}_i'\underset{\sim}{M}\underset{\sim}{Z}_i = \sum_{i=1}^{I} (\hat{\underset{\sim}{B}}_i-\hat{\underset{\sim}{B}}_\bullet)'\underset{\sim}{M}(\hat{\underset{\sim}{B}}_i-\hat{\underset{\sim}{B}}_\bullet) . \tag{7.8}$$

By (4.13) we have

$$W = \sum_{i=1}^{I} S_i^2 , \tag{7.9}$$

where S_i^2 is the sum of squared residuals in regression no. i . We now substitute (7.8) and (7.9) in (7.2) and get

$$F_{\lambda_1} = \frac{\sum_{i=1}^{I} (\hat{\underset{\sim}{B}}_i - \hat{\underset{\sim}{B}}_\cdot)' (\lambda_1 \underset{\sim}{M}_B + \underset{\sim}{M}_W)^{-1} (\hat{\underset{\sim}{B}}_i - \hat{\underset{\sim}{B}}_\cdot) + \sum_{i=1}^{I} S_i^2}{\sum_{i=1}^{I} (\hat{\underset{\sim}{B}}_i - \hat{\underset{\sim}{B}}_\cdot)' (\lambda_o \underset{\sim}{M}_B + \underset{\sim}{M}_W)^{-1} (\hat{\underset{\sim}{B}}_i - \hat{\underset{\sim}{B}}_\cdot) + \sum_{i=1}^{I} S_i^2} .$$

The problem of testing $\lambda \leqq \lambda_o$ is essentially the same as that studied by Spjøtvoll (1967) in connection with unbalanced classical model II . He found results analogous to our theorems (7.5) and (7.6).

Now consider the case $J > 1$, nested classification, when $\underset{\sim}{\Sigma}_C = \sigma_C^2 \underset{\sim}{M}_C$ and $\underset{\sim}{\Sigma}_D = \sigma_D^2 \underset{\sim}{M}_D$, where $\underset{\sim}{M}_C$ and $\underset{\sim}{M}_D$ are known matrices. The above technique may be applied without essential changes if the variance ratio σ_C^2/σ_D^2 is known a priori and we want to test the hypothesis $\sigma_D^2 \leqq \lambda_o \sigma^2$. When σ_C^2/σ_D^2 is unknown, we can find no optimal test procedures, but of course we can apply the tests derived in subsection 7 B .

8. A sketch of some further extensions of the theory.

8 A . In the present paper we have mainly been concerned with inference problems which possess an (in some sense) optimal solution. Numerous estimation and test problems which are not treated here may be attacked by well known multivariate techniques.

We mention in particular the problem of comparing several samples, (e.g. corresponding to different preparations in example 1 C). Like the classical model II , the present model may be extended to higher levels of nesting, and the results in this paper may be generalized correspondingly.

References.

Anderson, T.W. (1958): Introduction to Multivariate Statistical Analysis, Wiley, New York.

Basu, D. (1955): On statistics independent of a complete sufficient statistic, Sankhya **15**, 377-380.

Frislid, K., Haram, E.M., Norberg, R. and Øie, S. (1973): Inter- and intraindividual variation in absorption from tablets and solutions in man, Pharmaceutica Acta Helvetiae, **48**, 610-621.

Herbach, L. (1959): Properties of model II-type analysis of variance tests, A:Optimum nature of the F-test for model II in balanced case. Ann.Math.Statist., **30**, 939-959.

Lehmann, E. (1959): Testing Statistical Hypotheses, Wiley, New York.

Miller, R.G. (1966): Simultaneous Statistical Inference, McGraw-Hill, New York.

Rao, C.R. (1965): The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves, Biometrica, **52**, 447-458.

Scheffé, H. (1959): The Analysis of Variance, Wiley, New York.

Spjøtvoll, E. (1967): Optimum invariant tests in unbalanced variance component models, Ann.Math.Statist., **38**, 422-428.

Sverdrup, E. (1953): Similarity, unbiasedness, minimaxibility and admissibility of statistical test procedures. Skand. Aktuarietidskr., **36**, 64-86.

Sverdrup, E. (1965): Estimates and test procedures in connection with stochastic models for deaths, recoveries and transfers between different states of health, Skand. Aktuarietidskr., **48**, 184-211.

Swamy, P.A.V.B. (1971): Statistical inference in random coefficient regression models, Lecture Notes in Operations Research and Mathematical Systems, **55**, Springer-Verlag, Berlin.

Wagner, J.W. (1961): Biopharmaceutics: Absorption aspects. Journal of Pharmaceutical Sciences, **50**, 359-387.

Wald, A. (1947): A note on regression analysis, Ann.Math.Statist. **18**, 586-589.