Biometrika (1964), **51**, 3 *and* 4, *p.* 313
Printed in Great Britain

A generalized multivariate analysis of variance model useful especially for growth curve problems*

By RICHARD F. POTTHOFF AND S. N. ROY†
University of North Carolina

SUMMARY

The usual MANOVA (multivariate analysis of variance) model (see equation (1)) may be generalized (equation (3)) by allowing for the appending of a post-matrix in the expectation equation. As explained in § 1, this generalized model (3) is applicable particularly to many kinds of growth curve problems, as well as to other problems. Section 2 is theoretical, and develops techniques of analysis under the generalized model. A numerical example involving growth curves is worked out in § 3.

1. Nature of the problem

1·1. *Motivation of the problem: growth curve analysis*

As a simple example of the general type of growth curve problem to which the techniques of this paper may be particularly suited, consider the data of Table 1. A certain measurement in a dental study was made on each of 11 girls and 16 boys at ages 8, 10, 12 and 14. We will assume that the (4×4) variance matrix of the 4 correlated observations is the same for all 27 individuals. Such matters as the following might be of interest:

(*a*) Should the growth curves be represented by second degree equations in time (t), or are linear equations adequate?

(*b*) Should two separate curves be used for boys and girls, or do both have the same growth curve?

(*c*) Can we obtain confidence band(s) for the expected growth curve(s)?

We will return to these questions in § 3, where the data of Table 1 will be used in a numerical example illustrating the techniques to be presented in § 2.

Growth curve situations similar to the one just described have been considered from various angles by some previous writers, including Wishart (1938), Box (1950), Rao (1958), Leech & Healy (1959), Rao (1959), Healy (1961), Elston & Grizzle (1962) and Bock (1963). All of these growth curve situations involve successive, and therefore in general correlated, measurements on the same individuals, usually animals; except in the simpler situations, the individuals are divided into two or more groups, where the different groups may represent, e.g. different treatments which are to be compared. The different previous approaches to the growth curve problem deal mainly with rather specialized aspects of it, and so a major purpose of the present paper is to furnish new tools for growth curve analysis which are of sufficiently general applicability. These tools, which are derived from the theory of multivariate normal analysis of variance and which are based on polynomial models for the growth curves, provide both confidence bands for the growth curves and over-all tests of significance for various kinds of compound hypotheses, even under the more complex types of experimental designs.

* This research was supported by the Mathematics Division of the Air Force Office of Scientific Research.

† The Editors have heard with deep regret of Professor Roy's sudden death on 23 July 1964.

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

Table 1. *Measurements* on 11 girls and 16 boys, at 4 different ages*

	Girls					Boys			
	Age in years					Age in years			
Individual	8	10	12	14	Individual	8	10	12	14
1	21	20	21·5	23	1	26	25	29	31
2	21	21·5	24	25·5	2	21·5	22·5	23	26·5
3	20·5	24	24·5	26	3	23	22·5	24	27·5
4	23·5	24·5	25	26·5	4	25·5	27·5	26·5	27
5	21·5	23	22·5	23·5	5	20	23·5	22·5	26
6	20	21	21	22·5	6	24·5	25·5	27	28·5
7	21·5	22·5	23	25	7	22	22	24·5	26·5
8	23	23	23·5	24	8	24	21·5	24·5	25·5
9	20	21	22	21·5	9	23	20·5	31	26
10	16·5	19	19	19·5	10	27·5	28	31	31·5
11	24·5	25	28	28	11	23	23	23·5	25
					12	21·5	23·5	24	28
					13	17	24·5	26	29·5
					14	22·5	25·5	25·5	26
					15	23	24·5	26	30
					16	22	21·5	23·5	25
Mean	21·18	22·23	23·09	24·09	Mean	22·87	23·81	25·72	27·47

* These data were collected by investigators at the University of North Carolina Dental School. Each measurement is the distance, in millimeters, from the centre of the pituitary to the pteryo-maxillary fissure. The reason why there is an occasional instance where this distance decreases with age is that the distance represents the relative position of two points.

1·2. *Theoretical statement of the problem*

The basic theoretical problem which we will consider is actually more general than that of growth curve analysis. At present, however, it appears that the principal application of our theoretical results is to growth curves.

The usual MANOVA model (see, e.g. Roy, 1957, Chapter 12) is

$$E[\mathbf{X}(n \times p)] = \mathbf{A}(n \times m)\,\boldsymbol{\xi}(m \times p), \tag{1}$$

where the different rows of $\mathbf{X}$ are distributed mutually independently and the p elements in any row follow a multivariate normal distribution with unknown variance matrix $\boldsymbol{\Sigma}(p \times p$, and positive definite) and mean vector as specified by (1), $\mathbf{A}$ being a matrix of known constants and $\boldsymbol{\xi}$ being a matrix of unknown parameters. A hypothesis of the form

$$\mathbf{C}(s \times m)\,\boldsymbol{\xi}(m \times p)\,\mathbf{V}(p \times u) = \mathbf{0}(s \times u) \tag{2}$$

(where $\mathbf{C}$ and $\mathbf{V}$ are matrices of known constants and $\mathbf{0}$ is the null matrix) can be tested under the model (1), and confidence bounds (as well as estimates) are also available (see § 2·1 below).

The model (1) can be generalized by appending to it a post-matrix: we consider a more general model

$$E[\mathbf{X}_0(n \times q)] = \mathbf{A}(n \times m)\,\boldsymbol{\xi}(m \times p)\,\mathbf{P}(p \times q), \tag{3}$$

in which the different rows of $\mathbf{X}_0$ are distributed mutually independently and the q elements in any row follow a multivariate normal distribution with unknown variance matrix $\boldsymbol{\Sigma}_0(q \times q$, and positive definite) and mean vector as specified by (3), $\mathbf{A}$ and $\mathbf{P}$ being matrices of known constants and $\boldsymbol{\xi}$ being a matrix of unknown parameters. The theoretical problem

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

of this paper is to find a technique for testing a hypothesis of the form (2) under the generalized expectation model (3), and to obtain related confidence bounds; estimators also will be obtained. This problem is dealt with in §2·2.

1·3. *Examples of growth curve applications*

The generalized MANOVA model (3) is especially applicable to growth curve situations. We now present some examples showing how the model (3) and hypothesis of the form (2) fit in with growth curve applications.

(i) We start with the simplest situation. We have a group of n animals, all subject to the same conditions, which are each observed at q points in time, $t_1, t_2, \ldots, t_q$. The q observations on a given animal are not independent, but rather are assumed to be multivariate normal with unknown variance matrix $\mathbf{\Sigma}_0$. The growth curve is assumed to be a polynomial in time of degree $p-1$, so that the expected value of the measurement of any animal at time t will be $\xi_0 + \xi_1 t + \xi_2 t^2 + \ldots + \xi_{p-1} t^{p-1}$. The matrix $\mathbf{A}$ is $n \times 1$ and contains all 1's; the matrix $\boldsymbol{\xi}$ is $1 \times p$ and consists of $\xi_0, \xi_1, \xi_2, \ldots, \xi_{p-1}$; the element in the kth row and lth column of the matrix $\mathbf{P}$ is t_l^{k-1}. We may wish to test hypotheses concerning the ξ's, or obtain a confidence band for the growth curve.

(ii) Now let us generalize the model (i) and suppose that, instead of one group of animals, we have m groups of animals, with n_j animals in the jth group, and with each group being subjected to a different treatment. Animals in all groups are measured at the same q points in time and are assumed to have the same variance matrix $\mathbf{\Sigma}_0$. The growth curve associated with the jth group is $\xi_{j0} + \xi_{j1} t + \xi_{j2} t^2 + \ldots + \xi_{j,p-1} t^{p-1}$. The matrix $\mathbf{A}$ this time will contain m columns, and will consist of n_1 rows $(1, 0, \ldots, 0)$, n_2 rows $(0, 1, 0, \ldots, 0), \ldots$, and n_m rows $(0, \ldots, 0, 1)$. The (j, k) element of $\boldsymbol{\xi}$ will be $\xi_{j,k-1}$; $\mathbf{P}$ will be the same matrix as in (i) above.

If (2) is to be the hypothesis that all m growth curves are equal, we set $u = p$, $\mathbf{V} = \mathbf{I}$, $s = m-1$, and we take $\mathbf{C}$ to be a matrix whose last column contains all minus ones and whose first $(m-1)$ columns constitute the identity matrix. If (2) is to be the hypothesis that all m growth curves are equal except possibly for the additive constant ξ_{j0}, then we take $\mathbf{C}$ to be the same $(m-1) \times m$ matrix as before, and $\mathbf{V}$ to be a $p \times (p-1)$ matrix whose first row contains all 0's and whose last $(p-1)$ rows constitute the identity matrix. If (2) is to be the hypothesis that all m growth curves are actually of degree $(p-2)$ or less, we set $s = m$, $\mathbf{C} = \mathbf{I}$, $u = 1$, and take $\mathbf{V}$ to be a $p \times 1$ vector with all 0's except for a 1 as the last element. Many other possible hypotheses besides these three can also be tested.

(iii) We consider next a generalization of the model (ii) to a situation where there are two sets of treatment effects instead of just one. As an example, suppose we have a set of 3 diets and a set of 2 temperatures which we wish to test simultaneously. Let us assume that there is no interaction between diet and temperature, and that the growth curve for an animal subjected to the αth diet and the βth temperature can be represented by an equation of the form

$$(\xi_{\alpha0} + \xi_{\alpha1} t + \xi_{\alpha2} t^2 + \ldots + \xi_{\alpha,p-1} t^{p-1}) + (\mu_{\beta0} + \mu_{\beta1} t + \mu_{\beta2} t^2 + \ldots + \mu_{\beta,p-1} t^{p-1}). \quad (4)$$

Let there be $n_{\alpha\beta}$ animals subjected to the combination of the αth diet and the βth temperature ($\alpha = 1, 2, 3$; $\beta = 1, 2$). We assume that every animal, no matter which of the 6 groups he is in, is measured at the same q points in time and has the same unknown variance matrix $\mathbf{\Sigma}_0$. $\boldsymbol{\xi}$ will be a $5 \times p$ matrix whose (j, k) element is $\xi_{j,k-1}$ if $j = 1, 2$, or 3, and is $\mu_{j-3,k-1}$ if $j = 4$ or 5. $\mathbf{A}$ will contain $n_{\alpha\beta}$ rows consisting of 1's in the αth and $(\beta+3)$th positions and 0's in the other three positions, for $\alpha = 1, 2, 3$ and $\beta = 1, 2$. $\mathbf{P}$ will be the same matrix as in (i) and (ii).

Hypotheses analogous to those indicated under (ii) can be tested. For example, if (3) is

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

to be the hypothesis that there is no difference in the effects of the 2 temperatures, then we set $\mathbf{V} = \mathbf{I}$ and take $\mathbf{C}$ to be the row vector $(0, 0, 0, 1, -1)$.

Actually, the set-up which has just been described can be generalized still further. Not only can we make the obvious generalization which would enable us to use general numbers of diets and temperatures (rather than the specific numbers 3 and 2, respectively), but we can consider entirely different types of designs, including, in particular, various kinds of incomplete block designs and factorial designs. Just about any of the usual designs can be accommodated if a growth curve model based on the sum of several different polynomials of like degree, each representing a particular effect as in (4), is appropriate.

(iv) We consider finally a second generalization of the model (ii) which is in a different direction from the generalization (iii). Suppose, as in (ii), that we have m groups of animals with every animal being measured at (say) q' points in time, but suppose that, instead of our measuring only a single characteristic associated with growth, we measure more than one such characteristic. In other words, we now have a multi-response instead of a single-response situation. As a simple illustration, suppose that we are measuring both height and weight of every animal at every point in time rather than just measuring weight. Let us assume that the height growth curve for the jth group of animals is

$$\xi_{j0} + \xi_{j1} t + \xi_{j2} t^2 + \ldots + \xi_{j,\,p_1-1} t^{p_1-1},$$

and that the weight growth curve for the jth group is $\gamma_{j0} + \gamma_{j1} t + \gamma_{j2} t^2 + \ldots + \gamma_{j,\,p_2-1} t^{p_2-1}$. We will have $q = 2q'$ observations on every animal—q' observations on height and q' observations on weight. These $q = 2q'$ observations have variance matrix $\mathbf{\Sigma}_0$; not only do we expect the q' height observations to be correlated among themselves and the q' weight observations to be correlated among themselves, but we also expect the height observations to be correlated with the weight observations.

The $\mathbf{A}$ matrix will be the same as in (ii). $\boldsymbol{\xi}$ will be $m \times p$, where $p = p_1 + p_2$, and the (j, k) element of $\boldsymbol{\xi}$ will be $\xi_{j,\,k-1}$ if $k \leqslant p_1$ and $\gamma_{j,\,k-p_1-1}$ if $k > p_1$. We assume that $\mathbf{X}_0$ is arranged so that the first q' columns contain the height measurements and the last q' columns the weight measurements. Then $\mathbf{P}[(p_1 + p_2) \times 2q']$ will have a submatrix $\mathbf{P}_1(p_1 \times q')$ in its upper left-hand corner, a submatrix $\mathbf{P}_2(p_2 \times q')$ in its lower right-hand corner, and 0's elsewhere, where $\mathbf{P}_1$ and $\mathbf{P}_2$ are matrices like the $\mathbf{P}$ matrix that was associated with models (i), (ii) and (iii).

If (2) is to be the hypothesis that the height growth curves are the same for all m groups, we set $s = m-1$ and $u = p_1$, take $\mathbf{C}$ to be a matrix whose last column contains all minus ones and whose first $(m-1)$ columns constitute the identity matrix, and take $\mathbf{V}$ to be a matrix whose first p_1 rows constitute the identity matrix and whose last p_2 rows contain all 0's. If (2) is to be the hypothesis that both the height growth curves and the weight growth curves are the same for all m groups, we take $\mathbf{C}$ as before and set $u = p$, $\mathbf{V} = \mathbf{I}$. Other hypotheses can also be tested.

The set-up we have presented can easily be generalized to handle a situation where height measurements are made on each animal at points $t_{11}, t_{12}, \ldots, t_{1q_1}$, and weight measurements are made at points $t_{21}, t_{22}, \ldots, t_{2q_2}$, the two sets of points not necessarily being the same. More importantly, the model can also be generalized in an obvious manner to handle measurements on an arbitrary number of different characteristics instead of on just two as considered in our illustration.

If it is desired to generalize in directions (iii) and (iv) at the same time, this is also possible.

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

1·4. *Other applications*

The model (3) may have uses outside of growth curves. For example, if each row of $\mathbf{X}_0$ represents a group of measurements taken at neighbouring points in space rather than in time, then the theory developed for the model (3) might be put to use in certain practical situations, e.g. measurements of the intensity of light or sound at various distances from the source. One could assume a response equation of a certain degree in the co-ordinate(s) of space. The space in question could be one-, two-, or three-dimensional (unlike time, which can only be one-dimensional).

It may be helpful to point out that perhaps the simplest possible application of the generalized MANOVA model (3) is for the case where $m = p = 1$, $q = 2$, and both $\mathbf{A}$ and $\mathbf{P}$ consist of all 1's. Such a model means that we have n pairs of observations from a bivariate normal population with unknown variance matrix $\mathbf{\Sigma}_0(2\times 2)$ and with the two variates having a common but unknown mean ξ.

2. Solution of the problem

This section attacks the problem of analysis under the generalized MANOVA model (3). Our method of approach will be to reduce this problem to a simpler one—that of analysis under the model (1)—which has already been treated in the literature. Since we will be making use of the standard results for this model (1), we start off by briefly summarizing them (for more details see, e.g. Roy (1957), Chapter 12).

2·1. *Analysis under the usual MANOVA model*

We may test the hypothesis (2) under the model (1). (Incidentally, hypotheses in other forms may also be tested; see Appendix A.) It is assumed in (2) that $\mathbf{C}$ has rank $s\,(\leqslant m)$ and $\mathbf{V}$ has rank $u\,(\leqslant p)$. Let r (where $r \leqslant m$, $r < n$, $s \leqslant r$) be the rank of $\mathbf{A}$ in (1), and let the matrix $\mathbf{A}_1(n\times r)$ be the first r columns of $\mathbf{A}$; we assume that the columns of $\mathbf{A}$ are arranged in such a way that $\mathbf{A}_1$ is of rank r. Define $\mathbf{A}_2(n\times \overline{m-r})$ to be the last $(m-r)$ columns of $\mathbf{A}$. Also define the matrix $\mathbf{C}_1(s\times r)$ to be the first r columns of $\mathbf{C}$, and define $\mathbf{C}_2(s\times \overline{m-r})$ to be the last $(m-r)$ columns of $\mathbf{C}$. For the hypothesis (2) to be testable in the strong sense (Roy & Roy, 1958–59), we must have $\mathbf{C}_2 = \mathbf{C}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{A}_2$; hence rank of $\mathbf{C}_1$ = rank of $\mathbf{C} = s$. We define the matrices $\mathbf{S}_h(u\times u)$ and $\mathbf{S}_e(u\times u)$ by

$$\mathbf{S}_h = \mathbf{V}'\mathbf{X}'\mathbf{A}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{C}_1'[\mathbf{C}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{C}_1']^{-1}\mathbf{C}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{X}\mathbf{V} \tag{5a}$$

and

$$\mathbf{S}_e = \mathbf{V}'\mathbf{X}'[\mathbf{I}-\mathbf{A}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1']\mathbf{X}\mathbf{V}. \tag{5b}$$

Three tests of the hypothesis (2) under the model (1) have been proposed: Roy's test uses the largest characteristic root of $(\mathbf{S}_h\mathbf{S}_e^{-1})$ as its test statistic; a concept developed by Hotelling (1951) can be extended to yield a test (called the sum-of-the-roots test) based on the trace of $(\mathbf{S}_h\mathbf{S}_e^{-1})$; and a concept (the λ-criterion) developed by Wilks (1932) can be generalized to obtain a test which uses essentially the statistic $|\mathbf{S}_e|/|\mathbf{S}_h+\mathbf{S}_e| = 1/|\mathbf{S}_h\mathbf{S}_e^{-1}+\mathbf{I}|$. Very little is known as to the relative power of these three tests, except that no one of the three tests is uniformly better than either one of the others.† However, Roy's test has the

† Empirical verification of this statement has been obtained in a computer investigation, carried out by R. Gnanadesikan and colleagues at Bell Telephone Laboratories, for the case where s^* of equation (7a) is 2. Theoretical verifications of the admissibility of the largest-root test and the sum-of-the-roots test were obtained respectively by Mikhail (1960) and Ghosh (1963).

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

advantage that the distribution of the test statistic under the null hypothesis is known exactly, and has been tabulated (see Heck, 1960); also, as of now, associated confidence bounds are available only for Roy's test.

Simultaneous $100(1-\alpha)$ % confidence bounds on the functions $\mathbf{b}'\mathbf{C}\boldsymbol{\xi}\mathbf{V}\mathbf{f}$ for all $\mathbf{b}(s\times 1)$ and all $\mathbf{f}(u\times 1)$ are given (see Roy, 1957, p. 101, formula (14.6.3)) by

$$\mathbf{b}'\mathbf{C}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{X}\mathbf{V}\mathbf{f}\pm\{[h_\alpha/(1-h_\alpha)]\,[\mathbf{b}'\mathbf{C}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{C}_1'\mathbf{b}]\,[\mathbf{f}'\mathbf{S}_e\mathbf{f}]\}^{\frac{1}{2}}, \tag{6}$$

where h_α stands for the $(1-\alpha)$ fractile of the distribution tabulated by Heck (1960) with the three parameters (denoted by s, m, and n in Heck's notation, but to be denoted respectively by s^*, m^* and n^* in our notation) equal to

$$s^* = \min(s,u), \tag{7a}$$

$$m^* = \tfrac{1}{2}(|s-u|-1), \tag{7b}$$

and

$$n^* = \tfrac{1}{2}(n-r-u-1). \tag{7c}$$

In the special case where s^* (7a) is 1, the expression $[h_\alpha/(1-h_\alpha)]$ in (6) may be replaced by $[(2m^*+2)F_\alpha/(2n^*+2)]$, where F_α stands for the $(1-\alpha)$ fractile of the F distribution with $(2m^*+2)$ and $(2n^*+2)$ degrees of freedom.

The estimator of $\mathbf{b}'\mathbf{C}\boldsymbol{\xi}\mathbf{V}\mathbf{f}$ is of course the first term of (6).

2·2. *Analysis under the generalized MANOVA model*

2·2·1. *Reduction of the problem to the previous case*

We consider now the problem of the generalized MANOVA model (3). With respect to the matrix $\mathbf{P}$ in (3), we shall assume that $p\leqslant q$, and that $\mathbf{P}$ is of the full rank p. If $\mathbf{P}$ were not a matrix of the full rank, it would always be possible to re-write the model by re-defining $\boldsymbol{\xi}$ and $\mathbf{P}$ in such a way that the new $\mathbf{P}$ *would* be of the full rank.

The original observations $\mathbf{X}_0$ obtained under the model (3) may be subjected to a transformation of the form

$$\mathbf{X} = \mathbf{X}_0\mathbf{G}^{-1}\mathbf{P}'(\mathbf{P}\mathbf{G}^{-1}\mathbf{P}')^{-1}, \tag{8}$$

where we allow $\mathbf{G}(q\times q)$ to be any symmetric positive definite matrix, or any other non-singular matrix such that $\mathbf{P}\mathbf{G}^{-1}\mathbf{P}'$ is of the full rank. The matrix $\mathbf{X}(n\times p)$ defined by (8) will then be such that the different rows of $\mathbf{X}$ will be distributed mutually independently and the p elements in any row will follow a multivariate normal distribution with (unknown) positive definite variance matrix

$$\boldsymbol{\Sigma}(p\times p) = [\mathbf{P}(\mathbf{G}')^{-1}\mathbf{P}']^{-1}\mathbf{P}(\mathbf{G}')^{-1}\boldsymbol{\Sigma}_0\mathbf{G}^{-1}\mathbf{P}'(\mathbf{P}\mathbf{G}^{-1}\mathbf{P}')^{-1}$$

and with mean vector as specified according to the equation $E(\mathbf{X}) = \mathbf{A}\boldsymbol{\xi}$.

Thus we see that $\mathbf{X}$ (8) observes all the conditions of the model (1), no matter what choice is made for $\mathbf{G}$, subject to the limitations noted. Hence, a valid test of the hypothesis (2) under the model (3) may be obtained by substituting $\mathbf{X}$ (8) into equations (5) and then using the resulting matrices $\mathbf{S}_h$ and $\mathbf{S}_e$ to calculate whichever of the three test statistics (see § 2·1) is desired. Also, simultaneous confidence bounds under the model (3) can clearly be obtained via the transformation (8) and the formula (6). Similarly, estimators are available.

The approach just described provides the only presently existing technique of analysis under the model (3) for the general case. However, Rao (1959) did develop a method of analysis just for the special case of (3) where $m = 1$ and $\mathbf{A}(n\times 1)$ consists of all 1's (see Appendix C for further discussion).

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

2·2·2. *Choice of the matrix* $\mathbf{G}$

A curious feature of the technique of analysis just described is that it is valid for any choice of the matrix $\mathbf{G}$, so long as $\mathbf{G}$ is non-singular and so long as $\mathbf{P}\mathbf{G}^{-1}\mathbf{P}'$ is non-singular. We now consider the problem of how to choose $\mathbf{G}$.

In the first place, let us note that, for the case $p = q$, (8) becomes $\mathbf{X} = \mathbf{X}_0\mathbf{P}^{-1}$; hence for this particular case there is no need even to choose $\mathbf{G}$. Thus, when $p = q$, analysis under the model (3) amounts to virtually the same thing as analysis under the usual model (1).

When $p < q$, however, the choice of $\mathbf{G}$ affects the power of tests, the width of confidence intervals, and the variance of estimators. There are then several possible approaches to the problem of choosing $\mathbf{G}$.

(i) A very simple way of choosing $\mathbf{G}$ is to set $\mathbf{G} = \mathbf{I}(q \times q)$. Such a choice of $\mathbf{G}$ has the advantage that it will simplify the calculations to an extent; in fact, for growth curve models, the computational benefits of using orthogonal polynomials in the $\mathbf{P}$ matrix will exist only if we take $\mathbf{G} = \mathbf{I}$. We shall see shortly, however, that if some information about $\mathbf{\Sigma}_0$ is available, then it may not be best to choose $\mathbf{G} = \mathbf{I}$.

On the other hand, though, the choice $\mathbf{G} = \mathbf{I}$ has also a second advantage, on top of the computational advantage. All alternative approaches to the problem of choosing $\mathbf{G}$ (see below) involve some element of arbitrariness, to greater or lesser degree; but no such arbitrary factor enters in if we just take $\mathbf{G} = \mathbf{I}$. Some experimenters would regard the arbitrariness as something seriously to be avoided even though one has to pay a high price for avoiding it (as one very well may), whereas other experimenters would not object to it at all. Experimenters in the former category can just disregard (ii)–(v) below, and can fully utilize the results of this paper simply by setting $\mathbf{G} = \mathbf{I}$ always.

Wishart (1938), among others, essentially starts out by employing the transformation (8) with $\mathbf{G} = \mathbf{I}$ and then treating the resulting n rows of $\mathbf{X}$ as estimates of polynomial growth curve coefficients based respectively on the n animals; but from here he proceeds in a direction different from ours.

(ii) Among all estimators of the form $\mathbf{d}'(1 \times n)\,\mathbf{X}_0(n \times q)\,\mathbf{w}(q \times 1)$,

$$\mathbf{b}'\mathbf{C}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{X}_0\mathbf{\Sigma}_0^{-1}\mathbf{P}'(\mathbf{P}\mathbf{\Sigma}_0^{-1}\mathbf{P}')^{-1}\mathbf{V}\mathbf{f} \tag{9}$$

is the minimum variance unbiased estimator of $\mathbf{b}'\mathbf{C}\boldsymbol{\xi}\mathbf{V}\mathbf{f}$ under the model (3) (see Appendix B for proof). In (6) as well as in (5 a), we are in effect using

$$\mathbf{C}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{X}\mathbf{V} = \mathbf{C}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{X}_0\mathbf{G}^{-1}\mathbf{P}'(\mathbf{P}\mathbf{G}^{-1}\mathbf{P}')^{-1}\mathbf{V} \tag{10}$$

to estimate $\mathbf{C}\boldsymbol{\xi}\mathbf{V}$ under the model (3). In comparing (10) with (9), one would immediately be tempted to conclude that $\mathbf{G} = \mathbf{\Sigma}_0$ represents the optimal choice of $\mathbf{G}$ (based on a criterion of minimum variance unbiased estimation).

Unfortunately, however, we of course cannot make the choice $\mathbf{G} = \mathbf{\Sigma}_0$ due to the fact that $\mathbf{\Sigma}_0$ is unknown. But if we have available, before the experiment is run, a fairly good guess of $\mathbf{\Sigma}_0$, then we might normally prefer to set $\mathbf{G}$ equal to this guess of $\mathbf{\Sigma}_0$ rather than to choose $\mathbf{G} = \mathbf{I}$.

Observe that the analysis we make will still be valid even though $\mathbf{G}$ differs radically from $\mathbf{\Sigma}_0$. However, sensitivity would be affected: we would tend to suspect that, speaking very roughly, the more $\mathbf{G}$ differs from $\mathbf{\Sigma}_0$, the worse the power of the test will be and the wider the confidence intervals will be; estimators remain unbiased but increase in variance as $\mathbf{G}$ gets 'farther' from $\mathbf{\Sigma}_0$.

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

Note also, though, that the value of $\mathbf{X}$ (8) is not affected if $\mathbf{G}$ is multiplied by a scalar constant; in other words, using the matrix $k\mathbf{G}$ in (8) will produce exactly the same result as using $\mathbf{G}$.

(iii) The $\mathbf{G}$ we choose does not necessarily have to be non-stochastic. We may use a stochastic $\mathbf{G}$ so long as it is stochastically independent of $\mathbf{X}_0$ for, since the pertinent distributions relating to tests and confidence bounds are mathematically independent of $\mathbf{G}$, it follows that stochastically independent variation of $\mathbf{G}$ is irrelevant.

So far as we know, however, it is generally *not* legitimate to use a stochastic $\mathbf{G}$ which is not stochastically independent of $\mathbf{X}_0$. Thus, for example, one could not use for $\mathbf{G}$ an estimate of $\mathbf{\Sigma}_0$ which is derived from $\mathbf{X}_0$.

(iv) If, however, experimental information is available outside of the one experiment which produced $\mathbf{X}_0$, then a stochastic $\mathbf{G}$ obtained from such information would be independent of $\mathbf{X}_0$. For instance, let E denote the experiment which yields the observations $\mathbf{X}_0$ under the model (3), with the variance matrix $\mathbf{\Sigma}_0$, and suppose there is available a different, and independent, experiment E' whose results can somehow be utilized to get an estimate of $\mathbf{\Sigma}_0$. Then this estimate of $\mathbf{\Sigma}_0$ can be used for $\mathbf{G}$. E' does not have to have the same model (3) as E, nor in fact does E' even have to be based on a model of the general form (3).

(v) In certain situations something else can be done. Let $\mathbf{X}_0(E_1)$ and $\mathbf{X}_0(E_2)$ be the respective observations from two similar (independent) experiments E_1 and E_2. For simplicity, we consider the case where E_1 and E_2 are performed under exactly the same model (3), with the same unknown variance matrix $\mathbf{\Sigma}_0$. Let $\mathbf{G}(E_j)$ $(j = 1, 2)$ denote an estimate of $\mathbf{\Sigma}_0$ derived from E_j; one could, for example, use the obvious estimate

$$\mathbf{G}(E_j) = \frac{1}{n-r}\mathbf{X}_0'(E_j)\,[\mathbf{I}-\mathbf{A}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1']\,\mathbf{X}_0(E_j),$$

which is unbiased. Now if $\psi[\mathbf{X}_0, \mathbf{G}]$ denotes the test statistic to be employed for testing the hypothesis (2) under the model (3), and if R_α denotes the critical region, so chosen that $P\{\psi \in R_\alpha\} = \alpha$ if (2) is true, then we clearly could utilize either of the critical regions

$$\psi[\mathbf{X}_0(E_2),\, \mathbf{G}(E_1)] \in R_\alpha \tag{11a}$$

or

$$\psi[\mathbf{X}_0(E_1),\, \mathbf{G}(E_2)] \in R_\alpha \tag{11b}$$

for testing (2). However, consider as an alternative the test (likewise of level α) which rejects the hypothesis (2) if

$$\psi[\mathbf{X}_0(E_2),\, \mathbf{G}(E_1)] \in R_{\frac{1}{2}\alpha} \quad \text{and/or} \quad \psi[\mathbf{X}_0(E_1),\, \mathbf{G}(E_2)] \in R_{\frac{1}{2}\alpha} \tag{12}$$

and accepts otherwise. It is conjectured that this rather peculiar test (12) will generally be more powerful than either of the tests (11a) or (11b).

Confidence bounds related to (12) are also available. We compute an interval (6) using $\mathbf{X}_0(E_2)$ and $\mathbf{G}(E_1)$, and then again using $\mathbf{X}_0(E_1)$ and $\mathbf{G}(E_2)$, both with $h_{\frac{1}{2}\alpha}$ substituted for h_α. Then a final interval is obtained by taking the intersection of (i.e. the points common to) these two initial intervals. The set of such final intervals has simultaneous confidence coefficient $\geqslant 1-\alpha$.

3. NUMERICAL EXAMPLE

We return finally to the problem of analysing the data of Table 1. In effect, Table 1 is the matrix $\mathbf{X}_0(n \times q$, or $27 \times 4)$.

Before we can subject $\mathbf{X}_0$ to the transformation (8), we must first choose $\mathbf{G}(4 \times 4)$. This $\mathbf{G}$ we will base on an informed guess of $\mathbf{\Sigma}_0$ (see (ii), (iii) and (iv) under the subdivision 2·2·2).

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

We actually know very little about $\boldsymbol{\Sigma}_0$, but it can be presumed that the $q = 4$ observations on an individual are serially correlated. Perhaps the simplest serial correlation model is the one (see, e.g. Koopmans, 1942, § 2) in which the correlation coefficient between any two observations d periods of time apart is equal to ρ^d, and in which the variance is constant with respect to time; under such a model, $\boldsymbol{\Sigma}_0$ would be this constant times $(1-\rho^2)$, times

$$\frac{1}{1-\rho^2}\times\begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}. \tag{13}$$

Thus it seems appropriate to use (13) for $\mathbf{G}$. It remains only to select ρ. Unfortunately, the results of the experiment which produced $\mathbf{X}_0$ may not legitimately be used to estimate this ρ. However, we may look at a different experiment. In another dental study, the height of the ramus bone of 20 boys was measured at ages 8, $8\frac{1}{2}$, 9 and $9\frac{1}{2}$ (see Elston & Grizzle, 1962, p. 155); assuming that observations d periods of time apart have correlation coefficient ρ_1^d, and assuming an arbitrary growth curve (i.e. we avoid here the restrictive assumption of a first- or second-degree polynomial), we can obtain from these data the maximum-likelihood estimate of ρ_1 by solving a certain cubic equation analogous to equation (11) of Koopmans (1942), the coefficients of this cubic equation being calculated from the diagonal and next-to-diagonal elements of the 4×4 residual variance matrix. This estimate turns out to be $\hat{\rho}_1 = 0{\cdot}953$. Since the unit of time associated with ρ is four times as long as that associated with ρ_1 (2 years versus $\frac{1}{2}$ year), we expect roughly that $\rho = \rho_1^4$, whence our estimate of ρ is $\hat{\rho} = \hat{\rho}_1^4 = (0{\cdot}953)^4 = 0{\cdot}824$. Admittedly, this means of getting $\hat{\rho}$ might be questioned, but it appears that a $\mathbf{G}$ (13) based on some such $\hat{\rho}$ would be more sensible than a $\mathbf{G}$ equal to $\mathbf{I}$ (since the latter is the same thing as a $\mathbf{G}$ (13) with ρ estimated to be 0). We have to remember that all our procedures will still be valid even if $\mathbf{G}$ is not 'close' to $\boldsymbol{\Sigma}_0$. If we use (13), with $\rho = 0{\cdot}824$, for $\mathbf{G}$, then

$$\mathbf{G}^{-1} = \begin{bmatrix} 1 & -\rho & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & -\rho & 1 \end{bmatrix} \quad \text{with } \rho = 0{\cdot}824. \tag{14}$$

We consider now the questions (*a*)–(*c*) raised in § 1·1:

(*a*) In the model we assume quadratic equations in t for the growth curves, and then we test whether the coefficients of the second-order terms are 0. Thus in (3) we take $m = 2$, $p = 3$, $\mathbf{A}(27\times 2)$ to be a matrix composed of 11 (1, 0) rows followed by 16 (0, 1) rows,

$$\boldsymbol{\xi} = \begin{bmatrix} \xi_{10} & \xi_{11} & \xi_{12} \\ \xi_{20} & \xi_{21} & \xi_{22} \end{bmatrix}, \quad \text{and} \quad \mathbf{P} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \\ 9 & 1 & 1 & 9 \end{bmatrix}.$$

This is the set-up described in (ii) of § 1·3; in (2) we take $s = 2$, $u = 1$, $\mathbf{C} = \mathbf{I}(2\times 2)$, and $\mathbf{V} = (0, 0, 1)'$, so that (2) specifies that $\xi_{12} = \xi_{22} = 0$. Now

$$r = m = 2, \quad \mathbf{A}_1 = \mathbf{A} \quad \text{and} \quad \mathbf{C}_1 = \mathbf{C} = \mathbf{I}.$$

We must compute $\mathbf{S}_h$ and $\mathbf{S}_e$ (5) with $\mathbf{X}$ as given by the transformation (8). Incidentally, we might remark that equation (8) is formally identical with the formula for the weighted estimates of multiple regression coefficients (see, e.g. Scheffé, 1959, § 1·5), so that any

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

computer program which has already been devised for the latter could clearly be used for computing (8). In our calculations here, we note that

$$\mathbf{A}_1'\mathbf{X}_0 = \begin{bmatrix} 233\cdot0 & 244\cdot5 & 254\cdot0 & 265\cdot0 \\ 366\cdot0 & 381\cdot0 & 411\cdot5 & 439\cdot5 \end{bmatrix} \tag{15}$$

and

$$\mathbf{X}_0'\mathbf{X}_0 = \begin{bmatrix} 13{,}443\cdot00 & 13{,}962\cdot25 & 14{,}891\cdot00 & 15{,}734\cdot50 \\ & 14{,}611\cdot75 & 15{,}517\cdot75 & 16{,}438\cdot75 \\ & & 16{,}609\cdot75 & 17{,}525\cdot75 \\ & & & 18{,}581\cdot25 \end{bmatrix}, \tag{16}$$

and from (14) we get

$$\mathbf{G}^{-1}\mathbf{P}'(\mathbf{P}\mathbf{G}^{-1}\mathbf{P}')^{-1}\mathbf{V} = (0\cdot0625,\ -0\cdot0625,\ -0\cdot0625, 0\cdot0625)'. \tag{17}$$

Hence, combining (5), (8), (15), (16) and (17), we have

$$\mathbf{S}_h(1\times1) = 0\cdot04135 \quad \text{and} \quad \mathbf{S}_e(1\times1) = 0\cdot4069.$$

Also, (7) gives us $s^* = 1$, $m^* = 0$ and $n^* = 11\frac{1}{2}$. In general, Roy's test rejects the hypothesis (2) if $c_{\max}(\mathbf{S}_h\mathbf{S}_e^{-1})$ exceeds $h_\alpha/(1-h_\alpha)$, where h_α is found from Heck's (1960) tables with parameters (7) and $c_{\max}$ denotes the maximum characteristic root. However, when s^* (7*a*) is 1, Roy's test is identical with the other two tests based respectively on Hotelling's and Wilks' techniques: all three then reduce to an F-test, which rejects (2) if

$$[(2n^*+2)/(2m^*+2)]\,c_{\max}(\mathbf{S}_h\mathbf{S}_e^{-1})$$

exceeds the $(1-\alpha)$ fractile of the F distribution [D.F. = $(2m^*+2)$, $(2n^*+2)$]. Thus we calculate $25S_h/(2S_e) = 1\cdot27$, which is far below the $F_{0\cdot05}(2, 25)$ value of 3·39.

(*b*) Since we failed to reject the hypothesis $\xi_{12} = \xi_{22} = 0$ in (*a*), we will use a model based on linear, rather than quadratic, growth curves in attacking questions (*b*) and (*c*). Thus in (3) we take $\mathbf{A}$ exactly as before, but now we have $p = 2$ rather than 3,

$$\boldsymbol{\xi} = \begin{bmatrix} \xi_{10} & \xi_{11} \\ \xi_{20} & \xi_{21} \end{bmatrix}, \quad \text{and} \quad \mathbf{P} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \end{bmatrix}.$$

To answer question (*b*) as to whether the girls' curve and boys' curve are the same, we set $s = 1$, $u = 2$, $\mathbf{C} = (-1, 1)$, and $\mathbf{V} = \mathbf{I}(2\times2)$ in (2), so that we are testing $\xi_{10} = \xi_{20}$, $\xi_{11} = \xi_{21}$. Since $\mathbf{G}^{-1}$, $\mathbf{A}_1'\mathbf{X}_0$ and $\mathbf{X}_0'\mathbf{X}_0$ are naturally the same as before ((14), (15), (16)) we calculate

$$\mathbf{G}^{-1}\mathbf{P}'(\mathbf{P}\mathbf{G}^{-1}\mathbf{P}')^{-1} = \begin{bmatrix} 0\cdot425170 & 0\cdot074830 & 0\cdot074830 & 0\cdot425170 \\ -0\cdot165880 & -0\cdot002360 & 0\cdot002360 & 0\cdot165880 \end{bmatrix}' \tag{18}$$

and then use (5*b*), (8), (15), (16), and (18) to obtain

$$\mathbf{S}_e = \begin{bmatrix} 96\cdot2009 & -0\cdot4426 \\ -0\cdot4426 & 3\cdot4195 \end{bmatrix}, \quad S_e^{-1} = \begin{bmatrix} 0\cdot010401 & 0\cdot001346 \\ 0\cdot001346 & 0\cdot292614 \end{bmatrix},$$

and

$$\mathbf{C}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{X}\mathbf{V} = [-1,\quad 1]\begin{bmatrix} 22\cdot640 & 0\cdot4846 \\ 25\cdot111 & 0\cdot7665 \end{bmatrix} = [2\cdot471,\quad 0\cdot2819].$$

Now the non-zero characteristic roots of $(\mathbf{S}_h\mathbf{S}_e^{-1})$ (a $u\times u$ matrix) are always equal to the non-zero roots of the $s\times s$ matrix

$$\mathbf{C}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{X}\mathbf{V}\mathbf{S}_e^{-1}\mathbf{V}'\mathbf{X}'\mathbf{A}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{C}_1'[\mathbf{C}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{C}_1']^{-1}. \tag{19}$$

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

In the present case, (19) is a scalar, whose value we find to be 0·5779. From (7) we get $s^* = 1$, $m^* = 0$ and $n^* = 11$. Hence

$$[(2n^*+2)/(2m^*+2)]\, c_{\max}(\mathbf{S}_h \mathbf{S}_e^{-1}) = [24/2] \times 0{\cdot}5779 = 6{\cdot}93$$

is referred to the F tables (D.F. = 2, 24); since $F_{0\cdot005} = 6{\cdot}66$, we reject the null hypothesis and conclude that two separate curves are required.

(*c*) To get confidence bounds, including bands for the expected curves, we use (6) with $s = 2$, $u = 2$, $\mathbf{C} = \mathbf{C}_1 = \mathbf{I}(2 \times 2)$, and $\mathbf{V} = \mathbf{I}(2 \times 2)$, and with $(\mathbf{A}_1'\mathbf{A}_1)^{-1}$, $(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{X}$ and $\mathbf{S}_e$ the same as in (*b*) above. From (7) we have $s^* = 2$, $m^* = -\frac{1}{2}$ and $n^* = 11$; thus we refer to Chart III of Heck (1960, p. 630) to obtain $h_{0\cdot05} = 0{\cdot}297+$, using $\alpha = 0{\cdot}05$. For the bands, we take $\mathbf{f} = (1, t)'$, remembering that t = age minus 11. We set $\mathbf{b}' = (1, 0)$, $(0, 1)$ and $(-1, 1)$ for getting bands around the girls' curve $[\xi_{10} + \xi_{11} t]$, the boys' curve $[\xi_{20} + \xi_{21} t]$, and the difference between the two curves $[(\xi_{20} - \xi_{10}) + (\xi_{21} - \xi_{11}) t]$ respectively. The confidence band for the girls' curve is then given by $(22{\cdot}640 + 0{\cdot}4846t) \pm \{[0{\cdot}423]\,[1/11]\,[\mathbf{f}'\mathbf{S}_e\mathbf{f}]\}^{\frac{1}{2}}$, i.e.

$$(22{\cdot}640 + 0{\cdot}4846t) \pm 0{\cdot}1961(96{\cdot}201 - 0{\cdot}885t + 3{\cdot}420t^2)^{\frac{1}{2}}; \tag{20a}$$

for the boys' curve it is

$$(25{\cdot}111 + 0{\cdot}7665t) \pm 0{\cdot}1626(96{\cdot}201 - 0{\cdot}885t + 3{\cdot}420t^2)^{\frac{1}{2}}; \tag{20b}$$

and for the difference between the two curves it is

$$(2{\cdot}471 + 0{\cdot}2819t) \pm 0{\cdot}2547(96{\cdot}201 - 0{\cdot}885t + 3{\cdot}420t^2)^{\frac{1}{2}}. \tag{20c}$$

If, in addition, we want a confidence interval, simultaneously with (20*a*)–(20*c*), for, say, the difference in growth rate between girls and boys $(\xi_{21} - \xi_{11})$, this can be obtained by using (6) with $\mathbf{b}' = (-1, 1)$, $\mathbf{f} = (0, 1)'$, and everything else the same as before, so that the resulting interval is $0{\cdot}2819 \pm \{[0{\cdot}423]\,[(1/11) + (1/16)]\,[3{\cdot}4195]\}^{\frac{1}{2}}$, i.e.

$$0{\cdot}2819 \pm 0{\cdot}4711. \tag{20d}$$

It should be remembered that, since this interval (20*d*) is a member of a set of simultaneous intervals, it is wider than a simple confidence interval on $(\xi_{21} - \xi_{11})$ would be. The 95 % confidence coefficient applies simultaneously for all four of the relations (20) and simultaneously for all t.

If the bounds represented by (20*c*) and (20*d*) are not needed, then bands for the girls' and boys' curves which are better than (20*a*) and (20*b*) can be obtained via a combination of Roy's formula (6) with a general technique described by Dunn (1961). First, we compute (6) just as before, except with $s = 1$, $\mathbf{b}'(1 \times 1) = 1$, $\mathbf{C} = \mathbf{C}_1 = (1, 0)$ and $\alpha = 0{\cdot}025$. This gives us the band $(22{\cdot}640 + 0{\cdot}4846t) \pm \{[0{\cdot}360]\,[1/11]\,[\mathbf{f}'\mathbf{S}_e\mathbf{f}]\}^{\frac{1}{2}}$, i.e.

$$(22{\cdot}640 + 0{\cdot}4846t) \pm 0{\cdot}1809(96{\cdot}201 - 0{\cdot}885t + 3{\cdot}420t^2)^{\frac{1}{2}}, \tag{21a}$$

for the girls' curve. Note that now $s^* = 1$, $m^* = 0$ and $n^* = 11$, so that $h_\alpha/(1 - h_\alpha)$ in (6) is taken to be $(2/24)\, F_{0\cdot025}(2, 24) = (2/24) \times 4{\cdot}32 = 0{\cdot}360$. Next we compute (6) exactly as we did for (21*a*), except with $\mathbf{C} = \mathbf{C}_1 = (0, 1)$ instead of $(1, 0)$. This gives us the band

$$(25{\cdot}111 + 0{\cdot}7665t) \pm 0{\cdot}1500(96{\cdot}201 - 0{\cdot}885t + 3{\cdot}420t^2)^{\frac{1}{2}} \tag{21b}$$

for the boys' curve. Since (21*a*) is a 97·5 % confidence band for the girls' curve and (21*b*) is a 97·5 % confidence band for the boys' curve, it follows (see, e.g. Dunn, 1961) that the two bands (21) hold with confidence coefficient $\geqslant$ 95 % simultaneously for both curves and all t.

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

Thus, by comparing (21 *a*) with (20 *a*) and (21 *b*) with (20 *b*) or, more simply, by just comparing 0·360 with 0·423, we see that (21 *a*) and (21 *b*) are preferable to (20 *a*) and (20 *b*) if no bounds like (20 *c*) and (20 *d*) are desired.

Discussion

One might (as the referee has done) question our particular choice of $\mathbf{G}$ and ρ, and ask what the effect of using some other ρ, instead of 0·824, in (14) would have been.* Therefore, we have re-calculated most of the results in (*a*)–(*c*) with $\rho = 0$ (i.e. $\mathbf{G} = \mathbf{I}$), and with $\rho = 0{\cdot}615$. This latter value is the root of the cubic maximum-likelihood equation whose coefficients are determined from the residual variance matrix $\mathbf{X}_0'[\mathbf{I}-\mathbf{A}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{A}_1']\mathbf{X}_0$. Strictly speaking, it is not legitimate to allow the data (i.e. $\mathbf{X}_0$) to influence the choice of $\mathbf{G}$ (or ρ); however, it is nonetheless instructive, for comparison purposes, to find out what happens when $\rho = 0{\cdot}615$. Incidentally, we suspect that many experimenters would be inclined to estimate ρ from the data, even though this is not strictly proper; such a practice would presumably be defended by propounding that any resulting inaccuracies would, for practical purposes, be negligible.

In (*a*), it turns out that the expression (17) has the same value whatever ρ is, so that the test statistic in (*a*) is invariant under choice of ρ. In (*b*), we obtain F-values of 6·31 and 6·72 with $\rho = 0$ and 0·615 respectively (compared with 6·93 for $\rho = 0{\cdot}824$). In (*c*), the confidence bands for the girls' curve and boys' curve which compare respectively with (21 *a*) and (21 *b*) are

$$(22{\cdot}648+0{\cdot}4795t)\pm 0{\cdot}1809(94{\cdot}479+3{\cdot}407t+2{\cdot}958t^2)^{\frac{1}{2}}$$

and

$$(24{\cdot}969+0{\cdot}7844t)\pm 0{\cdot}1500(94{\cdot}479+3{\cdot}407t+2{\cdot}958t^2)^{\frac{1}{2}},$$

respectively, with $\rho = 0$, and are

$$(22{\cdot}643+0{\cdot}4838t)\pm 0{\cdot}1809(94{\cdot}828+0{\cdot}193t+3{\cdot}301t^2)^{\frac{1}{2}}$$

and

$$(25{\cdot}059+0{\cdot}7694t)\pm 0{\cdot}1500(94{\cdot}828+0{\cdot}193t+3{\cdot}301t^2)^{\frac{1}{2}},$$

respectively, with $\rho = 0{\cdot}615$; over the 6-year period from age 8 ($t = -3$) to age 14 ($t = 3$), the average width of the confidence band, exclusive of the factor $2\times 0{\cdot}1809$ or $2\times 0{\cdot}1500$, is 10·15 for $\rho = 0$, 10·22 for $\rho = 0{\cdot}615$, and 10·31 for $\rho = 0{\cdot}824$.

Now we must recognize that no firm conclusions about how to choose $\mathbf{G}$ or ρ can possibly be drawn just from the results of a single experiment. Bearing this in mind, we may note that, in (*b*), $\rho = 0{\cdot}824$ gives the best result and $\rho = 0$ the poorest, whereas in (*c*) the situation is exactly reversed. Thus, in both (*b*) and (*c*), $\rho = 0{\cdot}615$ (supposedly an optimal value of ρ) gives a result intermediate between the other two results.

Far more striking than the differences among the results obtained via these three disparate ρ-values, however, are their similarities. The results are so close together as to suggest that, for this particular example, it may make little practical difference what value of ρ is used.

Arbitrary choice of $\mathbf{G}$ or ρ cannot really be avoided, since, in a certain sense, even the simple choice $\mathbf{G} = \mathbf{I}$ is itself arbitrary. However, when the choice of ρ has as little effect on the results as in our example here, then the arbitrariness can hardly cause extreme concern.

The authors are indebted to Dr James E. Grizzle of the Department of Biostatistics, University of North Carolina, for some helpful comments.

* Incidentally, the referee has pointed out that, on the basis of a certain significance test, neither of the values $\rho = 0{\cdot}824$ and 0 happen to be compatible with the data.

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

The authors also wish to thank Dr Robert Mellinger Nelson of the Department of Orthodontics of the University of North Carolina Dental School for permission to use the data which appear in Table 1.

References

BOCK, R. DARRELL (1963). Multivariate analysis of variance of repeated measurements. *Problems in Measuring Change* (C. W. Harris, Editor), pp. 85–103. Madison, Wisconsin: University of Wisconsin Press.

BOX, G. E. P. (1950). Problems in the analysis of growth and wear curves. *Biometrics*, **6**, 362–89.

DUNN, OLIVE JEAN (1961). Multiple comparisons among means. *J. Amer. Statist. Ass.* **56**, 52–64.

ELSTON, R. C. & GRIZZLE, J. E. (1962). Estimation of time-response curves and their confidence bands. *Biometrics*, **18**, 148–59.

GHOSH, M. N. (1963). Admissibility of some tests of MANOVA. Abstract No. 4. *Ann. Math. Statist.* **34**, 1131.

HEALY, M. J. R. (1961). Experiments for comparing growth curves. Abstract No. 752. *Biometrics*, **17**, 333.

HECK, D. L. (1960). Charts of some upper percentage points of the distribution of the largest characteristic root. *Ann. Math. Statist.* **31**, 625–42.

HOTELLING, HAROLD (1951). A generalized T test and measure of multivariate dispersion. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 23–41. Berkeley and Los Angeles: University of California Press.

KOOPMANS, TJALLING (1942). Serial correlation and quadratic forms in normal variables. *Ann. Math. Statist.* **13**, 14–33.

LEECH, F. B. & HEALY, M. J. R. (1959). The analysis of experiments on growth rate. *Biometrics*, **15**, 98–106.

MIKHAIL, WADIE F. (1960). On the monotonicity and admissibility of some tests in multivariate analysis. Unpublished Ph.D. thesis, University of North Carolina.

RAO, C. RADHAKRISHNA (1958). Some statistical methods for comparison of growth curves. *Biometrics*, **14**, 1–17.

RAO, C. RADHAKRISHNA (1959). Some problems involving linear hypotheses in multivariate analysis. *Biometrika*, **46**, 49–58.

ROY, S. N. (1957). *Some Aspects of Multivariate Analysis*. New York: John Wiley and Sons.

ROY, S. N. & ROY, J. (1959). A note on a class of problems in 'normal' multivariate analysis of variance. *Ann. Math. Statist.* **30**, 577–81.

ROY, S. N. & ROY, J. (1958–59). A note on testability in normal Model I. ANOVA and MANOVA. *Calcutta Mathematical Society Golden Jubilee Commemoration Volume*, Part II, pp. 333–9.

SCHEFFÉ, HENRY (1959). *The Analysis of Variance*. New York: John Wiley and Sons.

WILKS, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika*, **24**, 471–94.

WISHART, JOHN (1938). Growth rate determinations in nutrition studies with the bacon pig, and their analysis. *Biometrika*, **30**, 16–28.

Appendix A

Hypotheses in a form other than (2) can be tested under the model (1), and hence also under the model (3). Roy & Roy (1959) show how to test the hypothesis

$$\boldsymbol{\xi}(m \times p) = \mathbf{B}(m \times l)\,\boldsymbol{\eta}(l \times p),$$

where $\mathbf{B}$ is a matrix of specified constants and $\boldsymbol{\eta}$ is a matrix of unknown parameters. A hypothesis more general than this, of the form

$$\boldsymbol{\xi}(m \times p)\,\mathbf{V}(p \times u) = \mathbf{B}(m \times l)\,\boldsymbol{\eta}(l \times u)$$

(where $\mathbf{V}$ and $\mathbf{B}$ contain constants and $\boldsymbol{\eta}$ contains parameters), is considered by Roy & Roy (1958–59).

Appendix B

We prove that the estimator (9) has the property claimed for it.

Necessary and sufficient conditions for $\mathbf{d}'\mathbf{X}_0\,\mathbf{w}$ to be an unbiased estimator of $\mathbf{b}'\mathbf{C}\boldsymbol{\xi}\mathbf{V}\mathbf{f}$ are

$$k\mathbf{d}'\mathbf{A} = \mathbf{b}'\mathbf{C} \quad \text{and} \quad (1/k)\,\mathbf{P}\mathbf{w} = \mathbf{V}\mathbf{f},$$

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011

where k is any scalar; there is no loss of generality if we assume $k = 1$. We have also

$$\text{var}(\mathbf{d}'\mathbf{X}_0\mathbf{w}) = \mathbf{d}'[\text{var}(\mathbf{X}_0\mathbf{w})]\mathbf{d} = \mathbf{d}'[(\mathbf{w}'\boldsymbol{\Sigma}_0\mathbf{w})\mathbf{I}]\mathbf{d} = (\mathbf{d}'\mathbf{d})(\mathbf{w}'\boldsymbol{\Sigma}_0\mathbf{w}).$$

Thus, in order for $\mathbf{d}'\mathbf{X}_0\mathbf{w}$ to be the unbiased estimator of $\mathbf{b}'\mathbf{C}\boldsymbol{\xi}\mathbf{V}\mathbf{f}$ with the smallest variance, we must (a) choose $\mathbf{d}$ so as to minimize $\mathbf{d}'\mathbf{d}$ subject to $\mathbf{d}'\mathbf{A} = \mathbf{b}'\mathbf{C}$, and ($b$) choose $\mathbf{w}$ so as to minimize $\mathbf{w}'\boldsymbol{\Sigma}_0\mathbf{w}$ subject to $\mathbf{P}\mathbf{w} = \mathbf{V}\mathbf{f}$. The problem ($a$) is encountered under the usual model (1), and the solution is of course $\mathbf{d} = \mathbf{A}_1(\mathbf{A}_1'\mathbf{A}_1)^{-1}\mathbf{C}_1'\mathbf{b}$, which verifies the terms in front of $\mathbf{X}_0$ in (9). To solve (b), we let $\boldsymbol{\lambda}(p \times 1)$ be a vector of Lagrangian multipliers, and differentiate $\mathbf{w}'\boldsymbol{\Sigma}_0\mathbf{w} - \boldsymbol{\lambda}'(\mathbf{P}\mathbf{w} - \mathbf{V}\mathbf{f})$ with respect to $\mathbf{w}$, obtaining the equation system $2\mathbf{w}'\boldsymbol{\Sigma}_0 - \boldsymbol{\lambda}'\mathbf{P} = \mathbf{0}'$; we end up finally with the solution

$$\boldsymbol{\lambda} = 2(\mathbf{P}\boldsymbol{\Sigma}_0^{-1}\mathbf{P}')^{-1}\mathbf{V}\mathbf{f} \quad \text{and} \quad \mathbf{w} = \boldsymbol{\Sigma}_0^{-1}\mathbf{P}'(\mathbf{P}\boldsymbol{\Sigma}_0^{-1}\mathbf{P}')^{-1}\mathbf{V}\mathbf{f}.$$

This verifies the remaining terms in (9).

Incidentally, for the simple bivariate normal common-mean model described in § 1·4, it can easily be proved that the substitution $\mathbf{G} = \boldsymbol{\Sigma}_0$, in addition to resulting in minimum variance unbiased estimates, also leads to a test of the hypothesis $\boldsymbol{\xi} = 0$ whose power is maximal with respect to $\mathbf{G}$.

APPENDIX C

For a special case of the model (3), Rao (1959) developed techniques of analysis, and also mentioned growth curve applications; he also mentioned the simple set-up we described in (i) of § 1·3. Our techniques (see § 2) for this special case of (3) are different from Rao's; hence our approach in this paper is not a generalization of Rao's (1959) approach. However, a generalization of this approach, if it were ever to be done, would no doubt be an important result.*

Rao's techniques can be used in the special case of (3) where $m = 1$ and the n elements of $\mathbf{A}$ are all equal; also we must have $n > u+q-p$. For this case, our test of the hypothesis (2) is based on an F with u and $(n-u)$ D.F., whereas Rao's test is based on an F with u and $(n-u-q+p)$ D.F. In the degenerate situation $p = q$, our $F_{u,\,n-u}$ statistic will be identical with Rao's $F_{u,\,n-u}$ statistic. But when $p < q$, the two tests are of course different. To compare the power of the two tests will require further investigation, but at present we have reason to believe that one test will be better in some situations and the other test in other situations.

Observe that, if $u < n \leqslant u+q-p$, then Rao's test cannot be used at all, whereas our test is still available.

* After this had been written, such a generalization was worked out at the University of North Carolina by C. G. Khatri.

Downloaded from biomet.oxfordjournals.org at University of Massachusetts on February 3, 2011