

A Multivariate Analysis of Covariance

Prepared Under Contract No. DA-36-034-ORD-1517 (RD)

(Experimental Designs for Industrial Research)

by

H. F. Smith

Institute of Statistics

Mimeo Series No. 157

September, 1956

## A Multivariate Analysis of Covariance\*

by

H. Fairfield Smith  
North Carolina State College

DeLury (1948) presented an example of covariance analysis with several characters which became somewhat complicated, both because the magnitude of treatment effects increased with their duration, and because one of the variables which originally had been intended for use as a concomitant control was affected by treatments. No satisfactory form of analysis for dealing with these complications was reached, and the problems raised seem worth re-examining.

The weights of hind-leg muscles of rats, one intact (v), one denervated (u), were observed at 4, 8, and 12 days after operation, followed by treatment twice per day with one of four drugs (of which D was merely control saline). There were four randomly chosen rats in each of the twelve groups. Initial (x) and final body weights (y) were observed.

The objective of the experiment was to determine whether the drugs delay atrophy of denervated muscles. Apparently it had been hoped that drugs would not affect intact muscles, so that they could be used as concomitant control. But inspection of the data revealed that drugs seriously affected both intact muscles (v) and final body weights (y). The question was then asked whether a drug might delay the rate of decay of denervated muscle relative to the shrinkage to be expected if there is no drug x denervation interaction. The common empirical definition of null interaction for additive effects is however inappropriate and irrelevant to the question: shrinkages caused by drug and denervation cannot possibly be additive; and there is no hypothesis to define how the two effects should superimpose in absence of interaction.

---

\*Part of a general review of analysis of covariance methods sponsored by the Office of Ordnance Research, U. S. Army, under Contract DA-36-034-ORD-1517.

Having observed that  $v$  and  $y$  are affected by treatments, DeLury duly noted that initial body weight ( $x$ ) is the only variable available for unambiguous use as concomitant control in the usual way. He nevertheless continued with an endeavor to use error regression on  $v$  or  $y$  in similar manner as a means of allowing for drug effect on healthy muscle in an attempt to answer the above question. Much of his paper is a discussion of the difficulties into which one may be led by such procedure. In effect he reached the conclusion that it is not a very meaningful form of attack.

He then endeavoured to develop a more elaborate model which might better represent a plausible relation among the variables, and to deduce from it a relation among the observations which could be put to statistical test. The idea is good, but the derived regression fails as a method of testing it for a number of reasons. The chief is that  $v$  is brought into the hypothesis as an empirical substitute for unknown functions of time representing the rates of growth or shrinkage of intact muscle with each drug, a function which can only operate between time treatments. Contrariwise the test ultimately proposed involved evaluating the derived regression from variation within treatments where  $v$  can only represent varying sizes of rats and has nothing whatever to do with variation over time as required by the hypothesis. The consequence is that the coefficients of the regression proposed for testing are not even remotely estimates of the respective symbols of the hypothesis. In fact, the proposed test differs from the endeavours of sections I and II to use  $x$  and  $v$  as concomitant variables in the usual way only in that it is incorrectly weighted by times of observation.

The usual analysis of covariance procedure has been frequently used to "adjust" the experimental variate for two characters such as  $x$  and  $v$  (or  $x$  and  $y$ ), one of which has been affected by treatments. But no paper which I have seen seems

to recognize that the adjustments are incompatible; that estimates of the adjusted means are estimates of an artifact which cannot exist. If drugs differentially affect  $v$  it is impossible that animals on different drugs can have simultaneously both equal  $x$  and equal  $v$ : a value of  $u$  adjusted for such a postulate seems meaningless, being analogous to estimating a function of  $x$  and  $x^2$  for  $x = a$  and  $x^2 = b^2$  when  $a \neq b$ . What then should be done with data of this sort?

I think we must look at the intact muscle as a dependent experimentally varied quantity with the same status as the denervated muscle. We could make a split-plot type of analysis to compare the effect of drugs on the average of both muscle conditions (main-plots) and on the differences between the two. But knowing a priori that differences will increase with time, it is almost certain that their rate of increase must be affected by treatment if only because the potentiality for shrinkage is limited. One may therefore forecast that denervation x drug interactions, as defined by the usual additive model for a split-plot analysis, would be complicated. When interactions are complex, the only thing to do is to study the individual treatment combinations, summarizing where a systematic pattern can be found. Here we have one simple quantitative treatment, time, whose effect should be systematically progressive. We may therefore reduce the number of descriptive statistics required by considering regressions on that factor

Shrinkage being inevitably limited it cannot progress linearly. Proportionate differences may however be longer sustained, and by using logarithms we may be able to flatten the curves enough for linear regressions to be at least first approximations. One procedure (considered in part by DeLury at p. 166) would be to combine the three times for each drug into estimates of mean, linear and curvature effects, and then compare these between drugs, hoping that curvatures may be small and negligible. But since at zero time (which was not observed) all muscles must

start from similar sizes for equal sized rats, the means and linear terms would be telling the same story. Can we concentrate both pieces of information in one statistic? Let  $Z_{it}$  denote the observed mean for any character ( $X = \log_{10} x$ ,  $Y = \log y$ ,  $V = \log v$ ,  $U = \log u$ ) with drug  $i$  at time  $t = 1, 2, 3$ . Assuming a linear regression on time the least squares estimate of the mean size at zero time for each drug is  $Z_{i0}$  as defined in table 1. This linear function of  $Z_{it}$  is orthogonal to the usual estimate of curvature,  $Q_i$ , and the third linear function orthogonal to these two is that defined as  $Z_{i\tau}$ . If the linear regression be adequate all  $Q_i$  should be zero, all  $Z_{i0}$  should be equal within sampling error, and most of the drug effects should appear in  $Z_{i\tau}$ . It is obviously similar to the linear function suggested by Fisher (1935, sec. 50) to evaluate effects of fertilizer qualities compared at several levels. If the rate of shrinkage with each drug were estimated as usual by  $b_i'' = \frac{1}{2}(Z_{i3} - Z_{i1})$ , then  $Z_{i\tau} = Z_{i0} + (7/3)b_i''$ , or in other words  $Z_{i\tau}$  is, assuming linear regression, an estimate of the 'yield' at  $\tau = 7/3$ , or at 9 1/3 days since 4 days have been taken as the unit of time.

Table 1. Definition of time effects

	$Z_{i1}$	$Z_{i2}$	$Z_{i3}$
$3Z_{i0}$	4	1	-2
$6Z_{i\tau}$	1	2	3
$Q_i$	1	-2	1

However, given the extra information that all  $Z_{i0}$  should be equal the  $b_i''$  as defined above are not efficient estimates of the slopes. If we assume the regression model

$$Z_{it} = \alpha + \beta_i t \quad (1)$$

with the constant  $\alpha$  common to all drugs, it may easily be shown that the least square

estimate of  $\alpha$  is  $Z_{.0}$  and of  $\beta_1$  is  $(3/7)(Z_{1\tau} - Z_{.0})$ , where  $Z_{.0} = \frac{1}{4} \sum_1^4 Z_{i0} / 4 =$

$\sum_1^3 (1/3)(4Z_{.1} + Z_{.2} - 2Z_{.3})$ ,  $Z_{.t}$  being the means over all drugs at each time. This estimate of  $\beta_1$  (from observations at three times) has  $\text{var}(\hat{\beta}_1) = (5/28) \text{var}(Z_{1t})$ , whereas the variance of  $b_1''$ , as defined above, would be  $\text{var}(b_1'') = (\frac{1}{2}) \text{var}(Z_{1t})$ .

If the linear regression model does not in fact fit,  $Z_{1\tau}$  and contrasts between them are still valid for the specified weighted means of observations at the given times, although interpretation may not be quite so simple and they will no longer contain all the information about drug effects. For example if regressions are curved with true means given by

$$Z_{it} = \alpha + \beta_1 t + \gamma_1 t^2$$

(which may be adequate for interpolation between observed times even if not the true curves); then (provided the approximating parabola does not have a maximum or minimum within the observed range)  $Z_{1\tau}$ , as defined for  $t = 1, 2, \text{ and } 3$ , estimates the yield at approximately

$$\tau = \frac{7}{3} \left( 1 + \frac{\gamma}{4.2\beta + 19.6} \right)$$

This will not usually differ much from  $7/3$  specified by the linear hypothesis. The estimate  $Z_{i0}$  will be biased from the true yield ( $\alpha$ ) at zero time by  $-(10/3)\gamma_1$ .

Sums of squares and products of effects as defined in table 1 (with appropriate divisors) lead to the analysis of variance and covariance exhibited in table 2. Since  $Z_{.0}$  and  $Z_{.t}$  are estimates of yields, as distinguished from contrasts, they are not independent of origin and their squares are of no interest. For convenience of computation table 2 is based on X and Y reduced by 2.25, V and U increased by 0.25; and all multiplied by  $10^3$ .

Average curvature,  $Q_{\cdot}$ , with only one degree of freedom, is conveniently evaluated by a t test (which of course gives the same test of significance as would the "reduced square"). By the usual formula the estimate  $Q'_{\cdot}$  for a character Z adjusted for X is

$$Q'_{\cdot}(Z) = Q_{\cdot}(Z) - b_{ZX} Q_{\cdot}(X) \quad (2)$$

where  $b_{ZX}$  indicates the regression of Z on X from the error row of table 2.

Correspondingly its variance is

$$\text{var}(Q_{\cdot}(Z)) = \left( \frac{6}{16} + \frac{(6.3)^2}{143891} \right) s_Z^2$$

where  $s_Z^2$  = variance of a single observation about the error regression on X, table 2

$6 = 1^2 + 2^2 + 1^2 =$  s. sq. of coefficients of  $Z_{\cdot t}$  in the linear function  $Q_{\cdot}$ ,

$16 =$  number of rats in each  $Z_{\cdot t}$ ,

$6.3 = Q_{\cdot}(X)$ ,

$143891 =$  S.Sq.(X) from which the regression was evaluated (table 2) henceforward denoted " $\text{dev}^2 X$ ".

This leads for each character to (for logarithms  $\times 10^3$ )

$$Q'_{\cdot}(Y) = 9.41 \pm 16.75$$

$$Q'_{\cdot}(V) = -23.37 \pm 27.88$$

$$Q'_{\cdot}(U) = -9.00 \pm 39.99$$

$$Q'_{\cdot}(V-U) = -14.38 \pm 31.88$$

Clearly average curvatures are undetectable relative to experimental error.

Table 2. Analysis of variance and covariance: sums of squares and products.

	d.f.	X <sup>2</sup>	Y <sup>2</sup>	V <sup>2</sup>	U <sup>2</sup>	(V-U) <sup>2</sup>	XY	XV	XU
<u>Means</u>									
Z <sub>.o</sub>	1	46577	26235	927886	1078933	5693	34956	207890	224173
Z <sub>.τ</sub>	1	373843	185265	4021876	872377	1148006	263173	1226193	571080
Q <sub>.</sub>	1	106	495	1021	32	693	229	-329	-58
<u>Between drugs</u>									
Z <sub>i<sub>o</sub></sub>	3	16653	4512	18576	19476	1236	7881	14828	12276
Z <sub>i<sub>τ</sub></sub>	3	15436	180337	560061	329064	48863	49549	77819	50430
Q <sub>i</sub>	3	11605	367	1649	5593	2602	867	1413	2590
Total treatments	12	464220	397212	5531068	2305474	1207093	356655	1527814	860491
Within treats.	36	143891	96071	125008	260693	105782	100298	86923	126703

Table 3. Reduced mean squares

	d.f.	Y <sup>2</sup>	V <sup>2</sup>	U <sup>2</sup>	(V-U) <sup>2</sup>	(V-Y) <sup>2</sup> *
<u>Between drugs</u>						
Z <sub>o</sub>	3	510.0	2198.7	3577.5	1200.3	2308.6
Z <sub>τ</sub>	3	36439.5	147409.6	81234.6	19632.3	40960.2
Q <sub>i</sub>	3	1487.4	1272.3	3218.3	937.1	1046.0
<u>Within treatments</u>						
Regression	1	69911.6	52509.3	111567.9	10997.4	
Error (s <sub>Z</sub> <sup>2</sup> )	35	747.4	2071.4	4260.7	2708.1	1133.4
<u>Regression coeff. on X (b<sub>ZX</sub>)</u>						
		.667668	.604090	.880548	-.276458	-.09295
		±.07207	±.11998	±.17208	±.13719	±.0886

\* Unreduced since regression is negligible; error mean square has 36 d.f. (The reduced error mean square would be 1130.)

Table 4. Estimates of initial weights and of rates of shrinkage (or growth) for each drug at average size of rat ( $X_{..} = 2.3435$ ).

	$(X_{.0} - X_{..})$ or $(X_{i1} - X_{..})$	Y'	V'	U'	(V-U)'	(V-Y) #
Z <sub>.0</sub>	-.0111	2.3192	.1245	(.1564)*	(.0319)*	-2.1940
s.e.		.01048	.01744	.02502	.01994	.01286
Regression per 9 1/3 days ( $B_{Zt}$ )						
D	.0201	.0572	.0696	-.1398	.2413	.0104
B	.0216	.0391	-.0108	-.2154	.2364	-.0520
C	-.0205	-.0326	-.0902	-.2111	.1528	-.0570
A	-.0143	-.0613	-.2205	-.3558	.1671	-.1432
s.e.		.03158	.02261	.02702	.01644	.01660
s.e. of a diff.		.01227	.02042	.02929	.02335	.01485

\*Regression for U evaluated as  $U'_{i\tau} - V'_{.0}$ .

\*Regression for V-U evaluated as  $V'_{i\tau} - U'_{i\tau}$ ; that is, assuming the theoretical value  $(V - U)_{.0} = 0$ .

Regression per 9 1/3 days:  $B_{Zt} = Z_{i\tau} - Z_{.0}$ , the latter symbols being defined in table 1.

Note that s.e. of a difference being sometimes less than s.e. of a single  $B_{Zt}$  is not an error but is due to correlation between the estimates. It is the s.e. of a difference between two  $Z_{i\tau}$ .

Treatments are, -- D: saline (control); B: moderate quinidine; C: moderate atropine; A: heavy atropine; injected twice daily during period of observation.

X =  $\log_{10} x$ : x = live body weight at time of operation

Y =  $\log y$ : y = body weight (at time of killing) in grams.

V =  $\log v$ : v = intact muscle weight (gm.)

U =  $\log u$ : u = denervated muscle weight.

Primes indicate values adjusted for X.

# Not adjusted for X -- regression on X is negligible and its theoretical value may be zero.

Table 3 presents the reduced sums of squares, derived in the usual way, to test significance of variation between drugs after adjusting to common X. The results are clear cut without need for reference to the F table. No differential effects due to drugs are detectable for estimates at zero time or for curvatures. Only two of these ten mean squares exceed their respective error variances, and they have individually  $P > .1$ . Therefore, taken along with the earlier finding that average curvatures are also approximately zero, the linear hypothesis is adequate to describe the data to the degree of approximation permitted by experimental error. Differential effects on rates of shrinkage are very highly significant for all characters.

Estimates: Let Z stand for any one of the three characters, Y, U, V or of a difference between two of them, as required;  $Z_{it}$  and  $X_{it}$  for the means of four observations with drug i at time t; replacing i by a dot indicates averaging over drugs. ( $Z_{.0}$  is of course the same function of  $Z_{.t}$  as is  $Z_{i0}$  of  $Z_{it}$ , etc.) A prime will indicate values adjusted by the error regression,  $b_{ZX}$ , to estimate the corresponding mean at  $X = X_{..}$ , the overall mean of X. Adjusted treatment means are

$$Z'_{it} = Z_{it} - b_{ZX}(X_{it} - X_{..})$$

and adjusted effects are the respective functions of  $Z_{it}^!$ . As usual,  $b_{ZX}$  is uncorrelated with  $Z_{it}$  and we regard functions of  $X$  as given constants. The estimates of interest are shown in table 4. They and their error variances are obtained as follows:

Variation between  $Z_{10}^!$  being ascribable to chance we can accept their mean,  $Z_{.0}^!$ , as the estimate of size of each character at zero time. It is

$$Z_{.0}^! = Z_{.0} - b_{ZX}(X_{.0} - X_{..})$$

$X_{..}$  does not drop out, as for example in equation (2) it commonly does, because  $Z_{.0}$  is not a "contrast", i.e. a linear function whose coefficients sum to zero, but is an estimate of a quantity of similar magnitude to  $Z$  itself.  $Z_{.0}^!$  would be more accurately estimated at  $X = X_{.0}$ , where the regression correction would vanish. But  $X_{.0}$  would not be a central value for other statistics and it seems more convenient to consider everything adjusted as usual to the overall mean  $X_{..}$ . Assuming as we do a common  $b_{ZX}$  for all treatments overlapping adjustments anyway cancel from contrasts which are of principal interest. Following the usual rules (as described above for  $Q$ ) the variance is

$$\text{var}(Z_{.0}^!) = \left[ \frac{21}{9.16} + \frac{(X_{.0} - X_{..})^2}{\text{dev}^2 X} \right] s_z^2 = .146884 s_z^2$$

The estimates and their standard errors are shown at the top of table 4.

If the linear hypothesis is adequate we expect  $V_{.0}' = U_{.0}'$ ; and this is not disproven ( $t = 1.6$ ).  $Y_{.0}^!$  is less than  $X_{..}$  by  $.02424 \pm .01048$  and this is significant ( $t = 2.31$ , 35 d.f.,  $P$  ca. .03). If both  $x$  and  $y$  were similarly observed live weights these should be equal. But the variability of  $Y$  within treatments is so strikingly less than that of  $X$  that these body weights seem somehow to have been differently observed:  $y$  may have been dead weights after cleaning out excreta.

Since  $s_V^2 < s_U^2$ ,  $V'_{.0}$  will be preferred as the estimate of initial muscle weight. Nothing is gained by averaging  $V'_{.0}$  and  $U'_{.0}$ . Owing to their close correlation the variance of their unweighted mean would be greater than that of  $V'_{.0}$ ; and optimum weighting, accepting observed variances and covariances as if exact, would give an apparent gain too trivial to compensate for uncertainty introduced by estimated weights.

$V'_{.0} - Y'_{.0}$  estimates  $\log(v_0/y_0)$  where  $v_0$  are muscle and body weights of an average rat before treatment. As should be expected, if the model adequately fits the data, it is practically the same as the logarithm of the mean ratio for the control treatments D at all times, namely -- 2.1870.

Drug effects are summarized by  $Z'_{i\tau}$ . However rates of shrinkage are more interesting figures. Table 4 therefore presents

$$B'_{i\tau} = Z'_{i\tau} - X'_{.0}$$

where  $B'_{i\tau}$  is the regression of Z on time in logarithm per 9 1/3 days = the estimate of  $(7/3)\beta_i$  of (1). Capital B is used to distinguish from regressions within treatments for which b has been used.  $Z'_{i\tau}$ , if wanted can be obtained by adding  $Z'_{.0}$ , which is listed at the top of each column, except for the following modification for U.

Having decided to prefer  $V'_{.0}$  to  $U'_{.0}$  as the estimate of initial muscle size, the estimates given in table 4 for rates of shrinkage of U are

$$B'_{i\tau}(U) = U'_{i\tau} - V'_{.0}$$

These have the further advantage of leading to

$$B'_{i\tau}(V-U) = (V-U)'_{i\tau} = V'_{i\tau} - U'_{i\tau}$$

which are the preferable estimates for comparisons between the two types of muscle.

These regression coefficients are (individually) less accurately evaluated than the  $Z'_{i\tau}$ , but this does not affect comparisons from which the common term drops out. A difference between two is identically the same as between the corresponding  $Z_{i\tau}$ , namely, for  $i \neq j$ ,

$$(B'_{i\tau} - B'_{j\tau}) = (Z'_{i\tau} - Z'_{j\tau}) = Z_{i\tau} - Z_{j\tau} - b(X_{i\tau} - X_{j\tau}).$$

Since the three parts are uncorrelated the variance is

$$\left[ \frac{2.14}{4.36} + \frac{(X_{i\tau} - X_{j\tau})^2}{\text{dev}^2 X} \right] s_Z^2$$

Evaluation of a separate error for every comparison being a nuisance, the average variance for all six possible comparisons may be used as an approximate error variance for all. It is given by (Finney, 1946)

$$\frac{2.7}{72} \left( 1 + \frac{M.Sq.(X_{i\tau})}{\text{dev}^2 X} \right) s_Z^2 = .20140 s_Z^2$$

where  $M.Sq.(X_{i\tau})$  is  $15436/3$  from table 2.

Error variances of the  $B'_{i\tau}$  individually would only be of interest for comparison to other experiments, or to set individual confidence intervals. They may however be worth quoting as further illustration of the general principle for estimating variances of such quantities.

Reverting to observed uncorrelated quantities

$$B'_{i\tau} = Z_{i\tau} - Z_{.0} - b_{ZX}(X_{i\tau} - X_{.0})$$

whence, again averaging the variances of the regression adjustments,

$$\text{ave. var.}(B'_{i\tau}) = s_Z^2 \left[ \frac{14}{36.4} + \frac{21}{9.16} + \frac{\sum_1 (X_{i\tau} - X_{.0})^2}{4 \text{dev}^2 X} \right] = .24682 s_Z^2 \quad (3)$$

Following the above definitions there are two special cases.

$$(1) B'_{i\tau}(U) = U_{i\tau} - V_{.0} - b_{UX}(X_{i\tau} - X_{..}) + b_{VX}(X_{.0} - X_{..})$$

and the last two terms are correlated, leading to

$$\begin{aligned} \text{ave. var } (B'_{i\tau}(U)) &= \left[ \frac{7}{72} + \frac{\sum (X_{i\tau} - X_{..})^2}{4 \text{ dev}^2 X} \right] s_U^2 + \left[ \frac{7}{48} + \frac{(X_{.0} - X_{..})^2}{\text{dev}^2 X} \right] s_V^2 \\ &\quad - \frac{2(X_{.0} - X_{..})(X_{.0} - X_{..})}{\text{dev}^2 X} s_{UV} = 730.22 \times 10^{-6} \end{aligned}$$

where  $s_{UV}$  is the error covariance of U and V after adjusting for X, which can be derived from the error row of table 2 as

$$\left[ \sum_{UV} - \frac{\sum_{UX} \cdot \sum_{VX}}{\sum X^2} \right] \frac{1}{35}$$

or from table 3 as

$$\frac{1}{2}(s_U^2 + s_V^2 - s_{V-U}^2) = 1812.0 \quad (4)$$

(Using  $U_{.0}$  and (3) the variance would be  $1051.6 \times 10^{-6}$ .)

(2) Since we assume theoretically  $(V-U)_{.0} = 0$ ,

$$B'_{i\tau}(V-U) = Z'_{i\tau} = Z_{i\tau} - b_{ZX}(X_{i\tau} - X_{..})$$

Hence the general formula

$$\text{ave. var } (Z'_{i\tau}) = \left[ \frac{114}{36.4} + \frac{\sum (X_{i\tau} - X_{..})^2}{4 \text{ dev}^2 X} \right] s_Z^2 = .099853 s_Z^2 \quad (5)$$

applies also for  $B'_{i\tau}$  in the special case  $Z = (V-U)$ .

The individual (adjusted) treatment means and the time regressions are shown in figure 1. Some supplementary data for 16 and 20 days after operation, later obtained from another paper (Solandt et al, 1943), are also shown in the figure. The graph suggests that the same linear model could be satisfactorily fitted for

the extended period. Indeed, considering the well known risks of extrapolation, the regressions fitted to the earlier data alone continue to do remarkably well. The only serious discrepancies are the muscle weights for treatment A at 16 days, which show deviations  $.177 \pm .034$  for V and  $.185 \pm .048$  for U; but these observations were for the only 2 rats surviving to 16 days out of 25 (excluding the 12 killed at days 4 to 12.) originally put on heavy atropine; hence a possible interpretation is that they may have been selected for resistance to the drug. Apart from that one treatment on U and V the mean square deviations from the extrapolated regressions, weighted by numbers of observations, were

Character	No. of devs.	Mean square ( $\times 10^6$ )	Ratio to $s_z^2$
Y	7*	1782	2.38
V	6+	5582	2.69
U	5+§	1712	.40

\*Including  $A_4$ .

+Excluding  $A_4$ .

§ Excluding also  $C_4$  for which published data are evidently a misprint copy of the v observations.

The mean square ratios for Y and V would be just about the 5 per cent significance level if they were honest F values. However after taking into account error of the extrapolated regression estimates and correlations (both ignored in these estimates) the discrepancies would not be significant. Exact testing would require re-fitting the model to the entire data.

The contrasts (V-U) and (V-Y) are of course easily derived from the differences among means of the respective characters; but their variances depend on correlations between the characters. Instead of worrying with covariances we can carry through computations as above with  $Z = (V-U)$ , etc., as a single character; hence the reason for including squares of these differences in tables 2 and 3. To save space their sums of products with X are not shown since they can be easily derived from others.

The comparison  $(V-Y) = \log (v/y)$  is incidental to the purpose of the experiment but to consider how the drugs affect muscle as compared to body size may be of some interest. As we should expect the ratio remains constant (within experimental error) for all times on "drug" D (control). Drugs shrink muscle proportionately more than body weight, as is again reasonable since presumably the weight of skeleton must remain unaffected.

Comparison of V and U was a prime purpose of the experiment. Before doing the experiment its authors seem to have hoped that one or more of the drugs might delay muscle atrophy. In fact they cause denervated muscle to shrink still faster than with no drug, but they have this effect also on intact muscle and the question has been asked whether or not their shrinking effect on the denervated muscle is as great as on intact muscle. This raises the supplementary question of what may be meant by "as great" since there must be a limit to the potential rate of shrinkage. By drug x denervation interaction is here implied that comparisons between rates of shrinkage on intact and denervated muscles under different drugs do not vary by more than could be ascribed to the lower potentiality for further shrinkage in atrophying muscle. If we had a clear cut hypothesis as to how drug and denervation effects should superimpose if the drug has no effect on atrophy, testing would be straightforward. Since we have no such hypothesis we can only seek to map out from the observations what in fact happens. Interaction as defined would be indicated by a significant reversal of effects; that is, if for two drugs, i and i', the contrasts  $U_{it} - U_{i't}$  and  $V_{it} - V_{i't}$  were both significant and with opposite signs. Less stringently we may postulate that in absence of interaction the difference  $V_{it} - U_{i't}$ , or equivalently  $U_{it}$ , should change smoothly and monotonically with intensity of drug effect. The latter may be measured by  $V_{i't}$ ,  $(V_{it} + U_{i't})$ , or by

body weight ( $Y_{i\tau}$ ). There being no reason to suppose that the relation should be linear a curve may have to be considered before significant deviations can be claimed to indicate interaction.

Whether an a priori hypothesis may be available for test, or the 'null' interaction relation of U to V has to be inferred from the data, U and V must both be regarded as joint experimental variates. Neither formulation can treat V as a concomitant variable as in ordinary analysis of covariance. One might be tempted to test for a monotonic relation by studying deviations of U or of (V-U) from its regression on V or (V+U). But the following circumstances indicate that such procedure could be misleading. For simplicity of discussion suppose that the relation may be linear. Firstly the structural relation specified by hypothesis should conform simultaneously to any of the forms.

$$U = \alpha + \beta V \quad (6)$$

$$U - V = \alpha + (\beta - 1)V \quad (7)$$

$$U = \left( \alpha + \beta(U + V) \right) / (1 + \beta)$$

$$U - V = \left( 2\alpha - (1 - \beta)(U + V) \right) / (1 + \beta)$$

Using regressions only (6) and (7) would be the same and would not conform with the others. Secondly deviations from regressions (6) and (7) would be identical, whereas the error variances to which one would be led to compare them might be very different, as they are in the present example. The tests for deviations to which each would lead cannot both be correct; the implication is that both would be wrong.

No satisfactory general method for fitting a structural curve seems available. When groups can be demarcated independently of possible random errors the Nair-Shrivastava (1942) method of averages might serve. In the present example a curve

will be no better than a straight line. Two or three forms of curve were tried but only increased the mean square deviation. Since reasonably reliable estimates of the error variances and covariance are available the Kummell line may be fitted.

With correlated errors, the simplest procedure seems to be to make a linear transformation to variates with uncorrelated errors and equal error variance. The estimates of error variances after adjusting for X are, from table 3 and equation 4

$$s_V^2 = 2071.4, s_U^2 = 4260.7, s_{UV} = 1812.0.$$

The corresponding values for  $Z_{i\tau}^!$  are proportional to these, the average values being given by multiplying by .099853 (equation 5). Choosing V as one variate the other is

$$W = \pm \frac{\sigma_V^2 U - \sigma_{UV} V}{(\sigma_U^2 \sigma_V^2 - \sigma_{UV}^2)^{1/2}}$$
$$= .76968V - .87987U$$

when the above estimates are substituted for the  $\sigma$ 's, and the negative sign is chosen. For brevity let deviations from the means of  $V_{i\tau}^!$  and  $W_{i\tau}^!$  (or equivalently of the corresponding  $B_{i\tau}^!$ , table 4) be denoted

$$y_{2i} = V_{i\tau}^! - V_{\cdot\tau}^!$$

$$y_{1i} = W_{i\tau}^! - W_{\cdot\tau}^!$$

which we assume to be uncorrelated with common variance .09985  $s_V^2 = 206.8$  for logarithms multiplied by  $10^3$ . The values of y are:

Treatment	$y_2$	$y_1$
D	132.58	22.24
B	52.16	26.81
C	-27.19	-38.00
A	-157.55	-11.05

The angle  $\beta$  of the structural line with the  $y_2$  axis is now estimated from

$$\tan 2\hat{\beta} = \frac{2\sum y_{1i}y_{2i}}{\sum y_{2i}^2 - \sum y_{1i}^2} = .33065; \quad \beta = 9^{\circ}8.4'$$

The estimated line is

$$y_1 = .1609 y_2$$

$$U'_T = -.1486 + .6919 V'_T \quad (8)$$

Let deviations from the line be denoted

$$x_{1i} = y_{1i} \cos\hat{\beta} - y_{2i} \sin\hat{\beta}$$

and deviations from the origin along the line be

$$x_{2i} = y_{1i} \sin\hat{\beta} + y_{2i} \cos\hat{\beta} = \xi_i + \epsilon$$

where  $\epsilon$  is a random variable with variance  $\sigma^2$ . On the null hypothesis that deviations from the line are only random error the expectation of  $\sum x_{1i}^2$  is (Smith, 1956)

$$\sigma^2(n-2)(1 - \omega - O(\omega^2))$$

where here  $n = 4$  and  $\omega = \sigma^2 / \sum \xi_i^2$ . Estimating  $\sum \xi_i^2 = \sum x_{2i}^2 - n\sigma^2 = 46176.1$ ,  $\omega = .00448$ , the estimate of  $\sigma^2$  from deviations from the line is

$$s^2(x_1) = \sum x_{1i}^2 / 1.9910 = 820.3.$$

The variance ratio for deviations from the line/internal error variance is therefore  $820.3/206.8 = 3.97$ . Assuming that the sampling distribution of the ratio may be approximately that of F with 2 and 35 degrees of freedom, significance slightly beyond the 5 per cent point is suggested. Since there seems to have been virtually no a priori knowledge to indicate what sort of result to expect from this experiment the practical conclusion at this level of significance for a test statistic whose exact distribution is not known might be to reserve judgment. As already noted, after these computations were done some supplementary data were obtained. It can be seen from figure 1 that if all data were used the estimates of  $U_{\tau}$  for drug B and  $V_{\tau}$  for drug C would each be increased. Both alterations would bring the points closer to the fitted structural line and deviations re-computed for all the data probably would not be significant. So far as present evidence goes the relation (8) seems to describe very well the relation between U and V after 9 days on variable doses of atropine or quinidine. Differential effect of drugs on denervated versus intact muscles is not yet demonstrated.

The foregoing analysis is defective in using estimates of error  $\sigma_{pq}$  for scaling as if they were exactly known. However with over 30 d.f. for error such defect may not be serious and the procedure seems substantially better than some alternatives which might be suggested.

If a regression analysis were used the mean square deviation of  $U'_{i\tau}$  or of  $(U - V)_{i\tau}$  from their regressions on  $V'_{i\tau}$  is 1081.3, leading to variance ratios relative to the respective error variances of 2.54 or 4.00. The latter ratio happens to be near that formerly obtained, but this depends on coincidental circumstances which could not be forecast. The reason is that the correlation of errors in U and V is such that the structural analysis minimizes a sum of squares of

deviations which are not far from vertical on the scales of U and V and hence are similar to those minimized by the regression analysis, that the regression of  $(V-U)_\tau$  on  $V_\tau$  is not far from zero, and that the random errors of these two variates are almost uncorrelated. An unusual feature is that the estimate of the slope of the structural line for  $U'_\tau$  on  $V'_\tau$ , viz. .6919 is less than that of the regression which is .6983. This has happened because the random errors are highly correlated with even steeper regression, viz. 8748.

An alternative approach to the structural analysis would be to note that the Kummell line corresponds to the first canonical variate and hence that deviations from it correspond to the second canonical variate. Presumably it would not be too difficult to modify a canonical analysis for covariance on a concomitant variable (cf. Cochran and Bliss, 1948) and to obtain an approximate  $\chi^2$  test for the second canonical variate. The  $\chi^2$  approximations for these multivariate tests however suppose the error variances to be exactly known (Rao, 1952, sec. 9d), whereas the variance ratio here derived allows that the estimate of error is also a random variable. The above analysis furthermore shows in elementary manner exactly what is being done in contrast to the mystification which some workers find in canonical variate terminology.

No matter what conclusions were reached from the given experiment, after discovery that drugs affected intact muscles a further experiment would be needed to give a convincing answer to the main question if there seemed any chance that a drug might be beneficial. DeLury (p. 167) notes that atrophy is irreversible whereas weight loss due to drugs is recoverable. It might happen however that drug shrinkage is not recoverable by denervated muscle; and since any reasonable definition of "beneficial" would seem to require that at some stage a drugged denervated muscle

should be larger than an undrugged one, conclusive evidence that a drug delays atrophy could only be given after withdrawing the drug and observing  $u$  after  $v$  has recovered. If  $u$  failed to recover the practical conclusion could only be that the drug had intensified atrophy even if its immediate effect was in some sense less than its effect on intact muscle.

An analysis of covariance with several characters, already discussed some years ago in *Biometrics*, is re-examined. The point is first made that to apply standard routine analysis of covariance reduction when an "independent" variable is affected by treatments may be equivalent to estimating an artifact with little or no meaning.

The example studied has one concomitant variable,  $x$  = body size, representing variability of the experimental material before treatments begin, and several experimental variates. The treatments are combinations of time and drugs. A method of analysis is proposed which first studies regressions on time for each drug and picks out a convenient function of time observations in which is concentrated all available information about contrasts between drugs. Relations of the dependent variates between drugs is then studied by means of a structural relation, that is one relative to which the random errors of both of two correlated variates are given due consideration.

The important distinction is made between a regression associated with random variation within treatments, and regressions or structural relations associated with variation between treatments. The former is that used in ordinary analysis of covariance to remove some of the extraneous random variation. The latter derive from treatment effects. Variation between quantitative treatments (in this example, time) can be summarized by regression of the experimental variates on the independent

variable defined by the treatments. Joint variation of two experimental variates between qualitative treatments must be described by a structural relation. In the example there arose the question whether such variation of two characters could be supposed to be functionally related, the contrary condition being that a given change in one variate may be associated with variable amounts of change in the other, depending on the quality of causative treatment. It is demonstrated that to evaluate this in terms of a regression analysis, instead of by a structural analysis, may be seriously misleading.

Literature Cited

- Cochran, W. G. and Bliss, C. I. (1948) Discriminant functions with covariance. Ann. Math. Stat. 19:151.
- DeLury, D. B. (1948) The analysis of covariance. Biometrics 4:153-170.
- Finney, D. J. (1946) Standard error of yields adjusted for regression on an independent measurement. Biometrics Bull. 2:53-55.
- Fisher, R. A. (1935) Design of experiments. Edinburgh.
- Nair, K. R. and Shrivastava, M. P. (1942) On a simple method of curve fitting. Sankhya 6:121.
- Rao, C. R. (1952) Advanced statistical methods in biometric research. Wiley.
- Smith, H. F. (1956) Estimating a linear functional relation. Insti. Stat. Mimeo Series No. 145.
- Solandt, D. Y., DeLury, D. B. and Hunter, J. (1943) The effect of atropine and quinidine sulphate on atrophy and fibrillation in denervated skeletal muscle. Am. J. Physiol. 140:247-255.

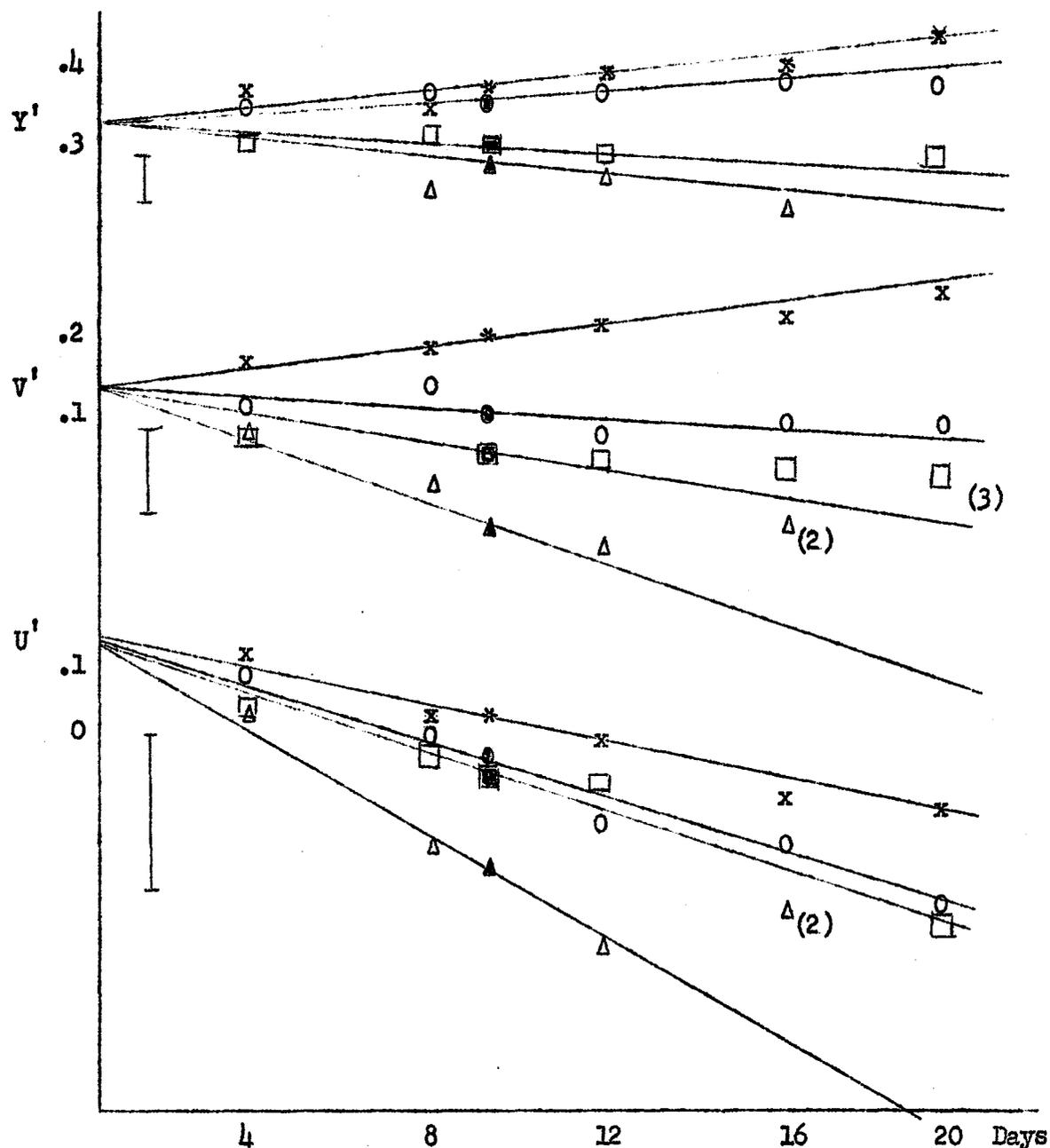


Figure 1. Treatment means (adjusted to  $X_{..}$ ) and time regressions fitted as described in the text.

Drugs	D	B	C	A
Treatment Means	x	O	□	△
$Z_{1T}$	*	●	■	▲

S.E. of a Tr. Mean x 2  $\left[ \right]$   
 (n) = No. of obs. when < 4

Rough diagram only.  
 Not exactly to scale.