# BIO-ASSAYS WITH NON-ORTHOGONAL DATA

By M. N. Das

*Indian Council of Agricultural Research, New Delhi*

## Introduction

Methods for analysis of non-orthogonal data are, no doubt, available extensively in literature, but those are suitable mainly for agricultural problems. As the purpose of analysis of bio-assay data is somewhat different from that of agricultural and other related problems, it is necessary to extend the technique to cover the special purposes of the analysis of bio-assay data. Excepting for the exact method of analysis illustrated by Finney (1952) for a four-point parallel line assay with five missing observations, no other method seems to be available in the current statistical literature. The object of the present paper is, thus, to present a systematic method of analysis of bio-assay data when they are non-orthogonal in two-way classification.

The fact that the intra block variance in agricultural experiments is dependent on the block size, makes the usefulness of any planned non-orthogonal design with unequal as also large block size doubtful. On the contrary, Finney has said regarding experiments with littermates as experimental units that 'the assumption of an intra-litter variance independent of litter size, is likely to be more nearly correct than would be the corresponding assumption in an agricultural field trial'. This actually points, as Finney has stressed, to the need of evolving designs and the corresponding method of analysis so as to accommodate in the same experiment litters of different sizes and consequently to allow more than one replication within litter, a situation which arises when the litter size is greater than the number of doses. Such designs on the one hand save wastage of animal resources and on the other increase the percision of different estimates over what could have been obtained by adopting the usual designs after rejecting some animals from the different litters.

As the method of analysis suitable for such designs will not be the same as those evolved for the existing designs, it is first necessary to work out a method of computation of the various results required for the interpretation of bio-assay data, when they are, in general,

non-orthogonal, so that the results in the case of such designs can be easily deduced from the general results.

The main hurdle for the analysis of non-orthogonal data by the method of fitting constants, is the solution of a set of normal equations. The details regarding the method of obtaining such equations as also of their solution, are easily available in literature. The aim of the present paper has thus been to evolve techniques of analysis suitable for bio-assays on the assumption that a solution of the normal equations as obtained for the method of fitting constants by the least squares technique, is available.

The additional requirements for the analysis of bio-assay data are (i) to get subdivisions of the adjusted *S.S.* due to doses to supply validity tests for ascertaining the correctness of the assumptions regarding the method of estimation of potency, (ii) to estimate the potency and (iii) to find the fiducial limits of the estimate of potency.

## Validity Tests

If, in a two-way classification with unequal cell frequencies involving treatments (doses) and blocks (litters) $n_{ij}$ denotes the number of observations in the cell defined by the $i$th treatment in the $j$-th block, $T_i$, the sum total of the observations in the $i$-th treatment and $B_j$ that in the $j$-th block, then it can be shown (Das, 1953) that the solution for the treatment effects ($t_i$) from the normal equations can be obtained as:

$$t_i = \sum_k C_{ik} Q_k \qquad (i = 1, 2, \ldots . p)$$

where $Q_k$ is the adjusted total of the $k$-th treatment and is given by

$$T_k - \sum_j \frac{n_{kj} B_j}{n_{.j}}$$

$$n_{.j} = \sum_k n_{kj}, \quad \text{and}$$

$p$ = the number of treatments and $C_{ik}$'s are known constants which are functions of $n_{ij}$.

For validity tests in parallel line assays the *S.S.* due to different contrasts of the dose effects having 1 *d.f.* each are to be found out. Now if $l_i$'s ($i = 1, 2, \ldots . p$) are $p$ numbers such that $\Sigma l_i = 0$, then $\Sigma t_i l_i$ is a contrast of the dose effects. The *S.S.* due to this contrast can be obtained by first squaring it and then dividing by some suitable divisor. The divisor is, of course, the sum of the squares of the coefficients of the observations involved in the contrast which is again

the coefficient of $\sigma^2$, the error variance in the variance of the contrast. It has been shown by the author (Das, 1953) that the variance of any contrast of the treatment effects, say, $\Sigma lt = \sigma^2 \Sigma lq$ where $q_i$ is the solution of the normal equations obtained by replacing $Q$'s by $l$'s, *i.e.*, $q_i = \sum_k C_{ik}\, l_k$.

Hence the *S.S.* due to the contrast $\Sigma lt$ is equal to $(\Sigma lt)^2/\Sigma lq$. If there is another contrast $\Sigma l't$, its *S.S.* will be similarly $(\Sigma l't)^2/\Sigma l'q'$, but it is not necessary that these two components will be orthogonal. As the covariance between the two contrasts $\Sigma lt$ and $\Sigma l't$ is equal to $\sigma^2 \Sigma l'q = \sigma^2 \Sigma lq'$, the two components can be orthogonal only if $\Sigma l'q = 0$.

In slope ratio-assays involving equal number of doses, the components for 'Intersection' and 'Blank' can be obtained similarly by defining the contrasts in terms of the dose effects in place of the dose totals as is done in the orthogonal case. As the present technique can give the *S.S.* only for single *d.f.* the *S.S.* due to regressions which has 2 *d.f.* can not be obtained through this method. It seems the better alternative is to adopt the technique suggested by Finney (1952), though it is hardly a necessity. But through Finney's technique the different contrasts required for such assays can not be obtained directly as they are not mutually orthogonal. The *S.S.* due to regressions can, however, be obtained by modifying some of the contrasts.

## Estimation of Potency

In parallel line assays the log potency less $(\bar{x}_S - \bar{x}_T)$ can be obtained by dividing

$$\frac{t_1 + t_2 + \ldots + t_{k_T}}{k_T} - \frac{t_1' + t_2' + \ldots + T_{k_S}'}{k_S} \text{ by } \frac{L_1}{\Sigma l^2}$$

where $l$'s are the constants defining $L_1$, the regression contrast formed out of the dose effects, $t$'s are the effects of the test preparation and $t'$'s are the effects of the standard preparation. In slope ratio-assays also the regression coefficients can be obtained from $L_S$ and $L_T$ where $L_S$ and $L_T$ are similarly the contrasts for the two regression coefficients obtained out of the dose effects by replacing the dose totals in terms of which the contrasts can be defined in the orthogonal cases, by the dose effects.

The fiducial limits can be obtained with the help of Fieller's theorem which requires the estimation of the variance of each of the contrasts together with their co-variance. It has already been discussed under validity tests as to how to obtain such variances and co-variance.

## A Design Suitable for Bio-Assays

A generalised balanced design has been suggested by the author (1957). It is particularly suitable for Bio-assays and other experiments with litter-mates as experimental units and involving small number of treatments. It has been defined as below:

Taking a block to mean, in general, a group of experimental units like plots, animals, etc., having some common feature, let there be $b$ blocks each accommodating $t$ treatments such that each of $q$ of these treatments occurs within each of these blocks $n \geqslant 0$ times and the remaining $t - q$ treatments, either $s$ or $(s + p)$ times $(s \geqslant 0,\ p \geqslant 0)$ such that the cells, *i.e.*, defined by a block and a treatment, taking the frequency $(s + p)$ form a balanced incomplete block design with parameters $v = t - q$, $b = b$, $r$, $k$ and $\lambda$. To these $b$ blocks add another $b'$ blocks such that the frequency of occurrence of the different treatments is the same within the same block but may differ from block to block. Such a design with $t$ treatments and $b + b'$ blocks has been called a generalised balanced design with two types of replications, as the $q$ treatments in the first set and $t - q$ treatment in the second are differently replicated; and with blocks of different sizes, as the first $b$ blocks have each the same size, say, $K$, while the other blocks may each be different in size.

Besides the randomised block designs and the B.I.B. designs with or without some extra treatments which are present in each block, various types of super-complete designs also come out as particular cases of the general design.

When such designs are applied for bio-assays, the question of choosing the particular design to be adopted as also of fixing the number of doses to be included in each set, can be decided only after the number of litters available as also their sizes are known. By properly choosing these numbers of doses in each set, maximum utilization of the animals is possible in most of the situations. One fact, however, should be considered, *viz.*, in the case of super complete designs, it is always better to spread the extra animals over as many doses as possible; or in other words, spreading the extra animals over larger number of doses gives better precision than concentrating them on fewer doses.

After the number of the doses in the second set has been fixed, care must be taken to select them properly as the precision of estimates of different contrasts among the dose effects required for the purpose of interpretation of the assay, depends on such selection. Though it is better to have the doses in the second set distributed equally over

the two preparations, it need not be at the cost of some animals from each litter which may be required to be thrown away in order to achieve such equity of distribution. If any contrast is based on the dose effects such that all the effects enter the contrast with equal weight, there is not much to select as to the doses to be included in the second set, as is the case in a four-point assay. If, again, the contrast does not utilise all the doses or utilises them with unequal weights as in the contrasts, $L_2$ and $L_2'$ in the eight-point and $L_p$ in other symmetrical assays, the precision of the contrasts depends on the choice of doses to be included in the second set. If the weights for certain doses are greater in a contrast, then the greater the number of observations available for these doses, the more will be the precision of the contrast. Thus, in a six-point assay if the number of doses in the second set be four, the precision of the regression, $b$, will be greater if the four doses are taken as $S_1$, $T_1$, $S_3$ and $T_3$ rather than $S_2$, $T_2$, $S_3$ and $T_3$, provided the frequency $s$ is greater than zero and $n \leqslant s$. When $s$ is zero, the situation will be reversed.

The following illustration shows the usefulness of such a design in avoiding wastage of animal resources.

In a six-point assay if litters of sizes 8 and 6 are available, then instead of throwing away two animals from each of the litters of size 8, the following design which is a particular case of the general design, can be adopted.

*A Super-complete Design*

(Numbers in the table indicate the frequencies of occurrence of the doses in the different litters)

| Litters | Doses | | | | | |
|---|---|---|---|---|---|---|
| | $S_1$ | $T_1$ | $S_2$ | $T_2$ | $S_3$ | $T_3$ |
| $L_1$ .. | 1 | 1 | 2 | 2 | 1 | 1 |
| $L_2$ .. | 1 | 1 | 2 | 1 | 2 | 1 |
| $L_3$ .. | 1 | 1 | 2 | 1 | 1 | 2 |
| $L_4$ .. | 1 | 1 | 1 | 2 | 2 | 1 |
| $L_5$ .. | 1 | 1 | 1 | 2 | 1 | 2 |
| $L_6$ .. | 1 | 1 | 1 | 1 | 2 | 2 |
| $L_7$ .. | 1 | 1 | 1 | 1 | 1 | 1 |
| $L_8$ .. | 1 | 1 | 1 | 1 | 1 | 1 |

It will be seen that the two extra animals have been distributed over only four of the doses, two from each preparation, which has helped in keeping the number of litters required for balance, small, namely, 6. Otherwise, had the two extra animals been distributed over all the six doses, the number of litters required for balance would have been 15.

The solution of the normal equations for the general design can be obtained from

$$t_l = \frac{1}{R_1}\left[Q_l + \frac{n\,(R - bn)\,\Sigma Q_m}{K(A - vB)}\right]$$

$$t_m = \frac{1}{A}\left[Q_m + \frac{B\Sigma Q_m}{A - vB}\right]$$

where $t_l$ denotes treatments in the first set, and $t_m$, those in the second set,

$$A = R_2 - \frac{p^2(r - \lambda)}{K},$$

$$B = \frac{R\,(s - n) + p\,(\lambda p + sr)}{K},$$

$R_1$ and $R_2$ are respectively the replications of the first and second set of treatments and $R = bs + rp$.

For this design the variance of $\Sigma\, lt$ where the $l$'s are 4 times the constants defining the regression contrast and $t$'s the dose effects, is $\cdot 393\,\sigma^2$ when the doses in the second set are $S_2$, $T_2$, $S_3$ and $T_3$. Had the doses in the second set been $S_1$, $T_1$, $S_3$ and $T_3$ its variance would have been $\cdot 340\,\sigma^2$ which is less than the previous variance and hence the precision of $b$ becomes more for the second choice of doses in the second set. If, again, all the extra animals were allotted to two doses only, *viz.*, either $S_1$ and $T_3$ or $T_1$ and $S_3$, the variance of the contrast becomes $\cdot 345\,\sigma^2$. If the frequency $s$ is 0 in the above design, the variances of the contrasts corresponding to the above two choices become $\cdot 345\,\sigma^2$ and $\cdot 381\,\sigma^2$ respectively. Again, variance of $L_p$ which has equal weights for all the doses is the same, *viz.*, $\cdot 562\,\sigma^2$ for both the choices of the doses.

Similarly

$$\text{Var}\,(L_2) = \frac{1592}{1833}\,\sigma^2 \text{ for the choice } S_2,\ T_2,\ S_3 \text{ and } T_3$$

$$= \frac{144}{117}\,\sigma^2 \text{ for the choice } S_1,\ T_1,\ S_3 \text{ and } T_3.$$

and

$$\text{Var}\,(L_2') = 1{\cdot}073\,\sigma^2 \text{ for the first choice}$$
$$= 1{\cdot}229\,\sigma^2 \text{ for the second choice.}$$

Thus, though the first choice is not good for regression and parallelism, it increased the precision of the other validity tests and the second choice produced just the opposite results.

## An Example

The different steps in the analysis of non-orthogonal data in bio-assays have been illustrated by means of the following example.

The data analysed were obtained from a six-point parallel line bio-assay for the estimation of the potency of an unknown preparation of vitamin D and have been reported by Bliss (1952). In the original assay there were twelve litters of six rats each such that the different rats in each litter were treated by six doses of vitamin, *viz.*, three from each of a standard and the test preparations. The response measured

*Table of Observations*

(Scores of degrees of healing)

| Litters | Doses in $\mu$g. | | | | | |
|---|---|---|---|---|---|---|
| | (1) 2·5 (*S*) | (2) 2·5 (*T*) | (3) 5 (*S*) | (4) 5 (*T*) | (5) 10 (*S*) | (6) 10 (*T*) |
| $L_1$ | 2 | 3 | 8 (7) | 9 (9) | 8 | 7 |
| $L_2$ | 6 | 3 | 4 (8) | 5 | 9 (11) | 8 |
| $L_3$ | 4 | 4 | 6 (7) | 6 | 12 | 9 (10) |
| $L_4$ | 6 | 9 | 11 | 14 (10) | 10 (13) | 13 |
| $L_5$ | 10 | 8 | 15 | 8 (8) | 17 | 10 (12) |
| $L_6$ | 4 | 5 | 10 | 11 | 13 (5) | 13 (9) |
| $L_7$ | 11 | 3 | 4 | 6 | 9 | 15 |
| $L_8$ | 2 | 5 | 9 | 8 | 14 | 6 |
| $L_9$ | 12 | 15 | 10 | 18 | 9 | 15 |

was the degree of healing in the split tibia of each of the rats, scored on an arbitrary scale.

The design adopted was a randomised block with six doses (treatments) and twelve litters. What has been done to get a particular case of the design is that data from nine of the litters have been taken as they are. Then the observations in each of the six of these nine litters have been increased by two more taken from four of the doses in each of three other litters and allotted to the corresponding doses of the former six litters such that the twelve observations thus transferred formed a balanced incomplete block design with parameters $v = 4$, $b = 6$, $r = 3$, $k = 2$ and $\lambda = 1$. The table above gives the data as well as the design which now consists of three litters of six rats each, together with six more litters of eight rats each, four of which were treated by four of the doses while the remaining four were treated by the other two doses in equal numbers. The figures in brackets are the transferred observations from the corresponding doses in other litters.

The parameters of the design are shown below:—

$t = 6,\ q = 2,\ n = s = p = ni = 1,\ b' = 3,\ v = 4,\ b = 6,\ r = 3,$
$k = 2,\ \lambda = 1.$

$R_1 = 9,\ R = 9,\ R_2 = 12,\ K = 8,\ i = 1, 2, 3.$

The purpose of analysis is first to obtain the error *S.S.* and the adjusted *S.S.* due to the doses and then to get components of the latter sum of squares for the purpose of validity tests to ascertain if the different assumptions underlying the method of estimation of potency through a parallel line assay based on quantitative response, are tenable. Though the estimation of the effect of individual dose and the variance of their difference is not directly required in bio-assay analysis, these have to be obtained while finding the necessary components of the adjusted *S.S.* due to doses.

The various steps in the analysis are shown below.

$$A = R_2 - \frac{p^2 (r - \lambda)}{K} = 12 - \frac{2}{8} = \frac{47}{4}.$$

$$B = \frac{R (s - n) + p (\lambda p + sr)}{K} = \frac{4}{8} = \tfrac{1}{2}.$$

$$A - vB = \frac{39}{4}.$$

| | Dose totals with No. of obs. ($T_i$) | Litter totals with No. of obs. ($B_j$) | Adjusted dose totals ($Q_i$) |
|---|---|---|---|
| 1 | 60 (9) | 53 (8) | −19·625 |
| 2 | 52 (9) | 54 (8) | −27·625 |
| 3 | 99 (12) | 58 (8) | −1·250 |
| 4 | 112 (12) | 86 (8) | 4·000 |
| 5 | 130 (12) | 88 (8) | 24·125 |
| 6 | 127 (12) | 70 (8) | 20·375 |
| 7 | | 48 (6) | |
| 8 | | 44 (6) | |
| 9 | | 79 (6) | |

$$\Sigma Q_m = -1{\cdot}250 + 4{\cdot}000 + 24{\cdot}125 + 20{\cdot}375 = 47{\cdot}250$$

Hence,

$$t_m = \frac{1}{A}\left(Q_m + \frac{B\Sigma Q_m}{A - vB}\right)$$

$$= \frac{4}{47}(Q_m + 2{\cdot}423)$$

where

$m = 3, 4, 5$ and $6$ and $t$ stands for dose effects.

Thus,

$t_3 = {\cdot}100, t_4 = {\cdot}547$

$t_5 = 2{\cdot}259, \; t_6 = 1{\cdot}940.$

Again,

$$t_l = \frac{1}{R_1}\left[Q_l + \frac{n(R - nb)\,\Sigma Q_m}{K(A - vB)}\right]$$

$$= \frac{1}{9}(Q_l + 1{\cdot}798) \text{ where } l = 1, 2.$$

Hence $t_1 = -1{\cdot}979, \; t_2 = -2{\cdot}866.$

The check that $\Sigma t = 0$, is satisfied.

Now the adjusted *S.S.* due to the doses can be obtained from $\Sigma tQ$ and is equal to 214·100.

For the purpose of validity tests the *S.S.* due to the following contrasts among the dose effects are to be obtained.

(1) $L_p = (-t_1 + t_2 - t_3 + t_4 - t_5 + t_6)$ due to preparation.

(2) $L_1 = (-t_1 - t_2 + t_5 + t_6)$ due to regression of response on log dose.

(3) $L'_1 = (t_1 - t_2 - t_5 + t_6)$ due to parallelism of the two regression lines corresponding to the two preparations.

(4) $L_2 = (t_1 + t_2 - 2t_3 - 2t_4 + t_5 + t_6)$ due to quadratic component of the regression.

(5) $L'_2 = (t_1 - t_2 - 2t_3 + 2t_4 + t_5 - t_6)$ due to difference between the two quadratic components of regression—one from each preparation.

The sum of the squares due to each of the contrasts can be obtained from $(\Sigma lt)^2/\Sigma lq$ where $\Sigma l = 0$ and $q_i$ is a solution of the normal equations obtained by replacing $Q$'s by $l$'s in the normal equations. The different components of the adjusted *S.S.* due to doses, thus obtained, have been presented in the analysis of variance table given below.

It will be seen that the total of the five components is not equal to the adjusted *S.S.* due to the doses obtained from $\Sigma tQ$. This is so, because even though the dose contrasts are mutually orthogonal, the corresponding components of the *S.S.* need not be mutually orthogonal in the case of non-orthogonal data.

Total *S.S.* (cr) = 6016.

Litter *S.S.* = 5387·959 − 5096·970 = 290·989.

Within cell *S.S.* = 66·000.

The interaction *S.S.* = 6016 − 5387·959 − 66·000 − 214·100.
= 347·941.

Regarding choice and isolation of proper error for testing the different contrasts, Finney (1952, §14·4) has shown that for testing $L_p$, the proper error is $L_p \times$litters interaction component and for $L_1$ it is $L_1 \times$litters component. For testing the other components as also for finding the fiducial limits of $M$, either the intra litter error or this pooled with $L_1' \times$ litter, $L_2 \times$ litter and $L_2' \times$ litter components of interaction is the best error in situations, where the sensitivity of the animals does not differ from litter to litter.

*Analysis of Variance Table*

| Due to | *d.f.* | *S.S.* | *M.S.* |
|---|---|---|---|
| Litters (unadjusted) .. | 8 | 290·989 | .. |
| Preparations .. .. | 1 | 1·024 | 1·024 |
| Regressions .. .. | 1 | 208·234 | 208·234 |
| Parallelism .. .. | 1 | ·822 | ·822 |
| Quadratic Component .. | 1 | 3·505 | 3·505 |
| Deviation from Quadratic .. | 1 | 4·109 | 4·109 |
| Doses .. .. .. | 5 | 214·100 | 42·820 |
| Interaction .. .. | 40 | 347·941 | 8·698 |
| Error .. .. .. | 12 | 66·000 | 5·500 |

As the interaction *M.S.* and error *M.S.* are of the same order these may be pooled and used as error. The results show that the response is linearly related with log dose and that the variation due to deviation from parallelism is not significant. Hence the regression coefficient estimated from the above analysis can be used for estimating the potency $R$ from the relation

$$\text{Log}_{10}R = \frac{(\bar{y}_T - \bar{y}_S)}{b} \log_{10}{}^2 \text{ as } \bar{x}_S = \bar{x}_T,$$

where $\bar{y}_T - \bar{y}_S$ is the average difference between the preparations and $b$, the regression coefficient. An estimate of $\bar{y}_T - \bar{y}_S$ can be obtained from $L_p/3$ and $b$ can be estimated from $L_1/4$. Thus, $b$ has been obtained as 9·044/4 and $L_p$ as − ·759.

Hence,

$$\text{Log}_{10}R = -\frac{\cdot 759 \times \cdot 301 \times 4}{3 \times 9 \cdot 044} = \bar{1} \cdot 966,$$

*i.e.*,

$$R = \cdot 925.$$

Though for complete data $\bar{y}_T - \bar{y}_S$ is positive, it has turned out to be negative in the present case due to the redistribution of the observations over the litters and hence $R$ has come out less than unity.

$$\text{Variance } L_p = \frac{238}{423}\sigma^2 = \frac{238}{423} \times 7{\cdot}960$$

$$\text{Variance } L_1 = \frac{720}{1833} \times 7{\cdot}960$$

Co-variance $(L_p, L_1) = 0$.

Hence assuming $g$ negligible fiducial limits of

$$\frac{L_p}{L_1} = \frac{-{\cdot}759}{9{\cdot}044} = -{\cdot}0839,$$

are

$$-{\cdot}0839 \pm \frac{1{\cdot}96 \times 2{\cdot}82}{9{\cdot}044}\left\{\frac{238}{423} + {\cdot}0839^2 \times \frac{720}{1833}\right\}^{\frac{1}{2}}$$

*i.e.*, $-{\cdot}5422$, ${\cdot}3746$.

So the fiducial limits of $\log_{10}R$ are $4/3 \times {\cdot}301$ times the limits of $L_p/L_1$, *i.e.*, $\bar{1}{\cdot}7824$ and ${\cdot}1502$.

Hence the fiducial limits of $R$ are ${\cdot}6059$ and $1{\cdot}4132$.

While discussing a four-point Oestrone assay with seven litters of rats having five missing observations, Finney (1952, § 4.13 and § 4.19) suggested an alternative method for the exact analysis of the assay. Instead of estimating the dose contrasts through estimation of individual dose effects, he first defined three independent variates such that the three partial regressions of the response on these variates estimated from within block variances and covariances, give the estimates of three contrasts among the dose effects.

The same data can be analysed also through the method of fitting constants easily without much involvement.

The normal equations for the data after eliminating the litter effects can be written by following Das (1953) as

$$\frac{17}{3}t_1 + \frac{t_2}{3} = Q_1$$

$$\frac{14}{3}t_2 = Q_2$$

$$\frac{17}{3}t_3 + \frac{t_2}{3} = Q_3$$

and $t_4 = -(t_1 + t_2 + t_3)$ where the $t$'s are the dose effects in the order in which they have been presented. The solution offers no difficulty and is obtained as

$$t_1 = \frac{3}{17}\left\{Q_1 - \frac{Q_2}{14}\right\}$$

$$t_2 = \frac{3}{14} Q_2$$

$$t_3 = \frac{3}{17}\left\{Q_3 - \frac{Q_2}{14}\right\}$$

and

$$t_4 = \frac{3}{17}\left\{Q_4 - \frac{Q_2}{14}\right\}.$$

The adjusted dose totals have been obtained as

$Q_1 = -91\frac{2}{3}$, $Q_2 = 95\frac{1}{3}$, $Q_3 = -76$ and $Q_4 = 72\frac{1}{3}$

whence the dose effects have been estimated as

$t_1 = -17{\cdot}3782$, $t_2 = 20{\cdot}4285$, $t_3 = -14{\cdot}6134$ and $t_4 = 11{\cdot}5630$.

Hence $\bar{y}_T - \bar{y}_S = \frac{1}{2}(17{\cdot}3782 - 20{\cdot}4285 - 14{\cdot}6134 + 11{\cdot}5630)$

$= -3{\cdot}0503$

and $b = \frac{1}{4}(17{\cdot}3782 + 20{\cdot}4285 + 14{\cdot}6134 + 11{\cdot}5630) = 15{\cdot}9957$.

and these are exactly the corresponding figures obtained by Finney for getting $M - \bar{x}_S + \bar{x}_T = -3{\cdot}0503/15{\cdot}9957 = -{\cdot}1907$.

The adjusted dose $S.S.$, $\Sigma tQ$ comes out to $5487{\cdot}53$.

The component of $S.S.$ due to parallelism (say) is obtainable from

$$\frac{(-17{\cdot}3782 - 20{\cdot}4285 + 14{\cdot}6134 + 11{\cdot}5630)^2}{90/119} = 178{\cdot}85.$$

as $\Sigma lq = 90/119$.

Finney did not actually obtain this component but suggested it can be obtained ' by finding out how much of 5487 is left when a regression on $x_1$ and $x_2$ is formed after omitting $x_3$'. The component as obtained through Finney's method comes to $178{\cdot}95$, the difference being due to approximations.

The following table shows the values of the $q$'s required to get the variance of the contrasts corresponding to the different contrasts.

| | $l_1$ | $l_2$ | $l_3$ | $l_4$ | Solution of normal equations obtained by replacing $Q$'s by $l$'s: $q_1$ | $q_2$ | $q_3$ | $q_4$ |
|---|---|---|---|---|---|---|---|---|
| $L_p$ | −1 | −1 | 1 | 1 | $\frac{-39}{17\times14}$ | $\frac{-3}{14}$ | $\frac{45}{17\times14}$ | $\frac{45}{17\times14}$ |
| $L_1$ | −1 | 1 | −1 | 1 | $\frac{-45}{17\times14}$ | $\frac{3}{14}$ | $\frac{-45}{14\times17}$ | $\frac{39}{17\times14}$ |
| $L_1'$ | 1 | −1 | −1 | 1 | $\frac{45}{17\times14}$ | $\frac{-3}{14}$ | $\frac{-39}{14\times17}$ | $\frac{45}{17\times14}$ |

Now variance $L_p$ $= \sigma^2 \Sigma lq = \frac{90}{119}\sigma^2$

Variance $L_1$ $= \sigma^2 \Sigma l'q' = \frac{90}{119}\sigma^2$

Co-variance $(L_p, L_1)$ $= \sigma^2 \Sigma lq' = \frac{-6}{119}\sigma^2$

Hence variance $\frac{L_p}{2}$ $= \cdot 18908\,\sigma^2$

Variance $(b)$ $= \text{Var}\left(\frac{L_1}{4}\right) = \frac{90}{16\times119}\sigma^2 = \cdot 04727\,\sigma^2$

and Co-variance $\left(b, \frac{L_p}{2}\right)$ $= \frac{-6}{8\times119}\sigma^2 = -\cdot 00630\,\sigma^2$

and these are exactly the figures obtained by Finney through the other method.

In general, the results from both the approaches are identical, as they are basically the same, *viz.*, that of fitting constants. In the method described in the paper the constants in the regression equation are the dose effects with $x_i$'s as pseudo variates where $x_i$ takes the value unity for the observations against the $i$th dose and zero elsewhere. The regression equation under this system may be written as

$$Y = \Sigma d_i x_i + \text{litter affects and error.}$$

Now, if $\Sigma d_i l_{ij} / \Sigma l_{ij}^2 = L_j$ be $(K-1)$ mutually orthogonal contrasts of dose effects where $k$ is the number of doses and $\Sigma di = 0$, the regression equation becomes identical with

$$Y = \Sigma L_j Z_j + \text{litter effects and error}$$

where $$Z_j = \sum_i l_{ij} x_i$$

As Finney estimates the constants $L_j$'s with $Z_j$'s as pseudo variates through the same method of fitting constants after eliminating the litter effects as has been done here to estimate $di$'s which are linear functions of $L_j$'s, the two approaches cannot but give the same results, as it follows from the theory of linear estimation that the linear functions of best estimates are also the best estimates of the corresponding linear functions of the parameters (Bose, 1943).

While the present technique can exploit the advantages of various types of designs with balancing or near balancing of doses over litters for easy algebraic solution of the normal equations, perhaps Finney's method cannot always take advantage of such situations to get solution of normal equations. Moreover, for the different validity tests it becomes necessary to solve as many more sets of equations.

## SUMMARY

Methods of analysing non-orthogonal bio-assay data have been presented systematically. A non-orthogonal but balanced design accommodating litters of different sizes has been suggested for such assays. The different steps in the analysis of non-orthogonal bio-assay data have been illustrated by means of examples.

## REFERENCES

1. Bliss, C. I. .. *Statistics of Bio-assay*, Academic Press, I.N.C. Publishers, New York, 1952, pp. 497–98.
2. Bose, R. C. .. *Class Notes in the M.Sc. Class of the Calcutta University*, 1943.
3. Das, M. N. .. *J. Ind. Soc. Agric. Statist.*, 1953, **5**, 161–78.
4. ———— .. *Ibid.*, 1957, **9**, 18–30.
5. Finney, D. J. .. *Statistical Methods in Biological Assay*, Charles Griffin & Co., Ltd., 1952.