SOME IMPROVEMENTS IN WEIGHING AND OTHER EXPERIMENTAL TECHNIQUES[1]

By Harold Hotelling

Columbia University

When several quantities are to be ascertained there is frequently an opportunity to increase the accuracy and reduce the cost by combining suitably in one experiment what might ordinarily be considered separate operations. The theory of design of experiments developed as a branch of modern mathematical statistics, and of which fundamental considerations are set forth in R. A. Fisher's book [1], provides many improvements of this kind. Since the main interests of Fisher and other originators of this theory have been in biology, the applications so far made have been chiefly biological in character, excepting for certain economic and social investigations involving stratified sampling. The possibilities of improvement of physical and chemical investigations through designed experiments based on the theory of statistical inference have scarcely begun to be explored.

The following example is due to F. Yates [2]. A chemist has seven light objects to weigh, and the scale also requires a zero correction, so that eight weighings are necessary. The standard error of each weighing is denoted by σ, the variance therefore by σ^2. Since the weight assigned to each object by customary techniques is the difference between the reading of the scale when carrying that object and when empty, the variance of the assigned weight is $2\sigma^2$, and its standard error is $\sigma\sqrt{2}$.

The improved technique suggested by Yates consists of weighing all seven objects together, and also weighing them in groups of three so chosen that each object is weighed four times altogether, twice with any other object and twice without it. Calling the readings from the scale $y_1, \cdots, y_8$ we then have as equations for determining the unknown weights $a, b, \cdots, g$,

$$\begin{array}{l}
a + b + c + d + e + f + g = y_1 \\
a + b + c \qquad\qquad\qquad\quad\;\; = y_2 \\
a \qquad\quad\;\; + d + e \qquad\quad\;\; = y_3 \\
a \qquad\qquad\qquad\;\; + f + g = y_4 \\
\quad\;\; b \quad\;\; + d \quad\;\; + f \quad\;\; = y_5 \\
\quad\;\; b \qquad\quad\;\; + e \quad\;\; + g = y_6 \\
\qquad\;\; c + d \qquad\qquad\; + g = y_7 \\
\qquad\;\; c \quad\;\; + e + f \quad\;\; = y_8 .
\end{array}$$

---

[1] Presented at the Wellesley meeting of the Institute of Mathematical Statistics, Aug. 13, 1944.

Any particular weight is found by adding together the four equations containing it, subtracting the other four, and dividing by 4. Thus

$$a = \frac{y_1 + y_2 + y_3 + y_4 - y_5 - y_6 - y_7 - y_8}{4}.$$

The variance of a sum of independent observations is the sum of their variances, as is well known, and the variance of c times an observation is c^2 times the variance of that observation. Taking $c = \frac{1}{4}$ for the first four terms in the expression for a and $c = -\frac{1}{4}$ for the others gives for the variance of a by this method $\sigma^2/2$, which is only one-fourth that for the direct method. The standard error, or probable error, has been halved. If a degree of accuracy is required calling for repetition a certain number of times of the weighings by the direct method, then only one-fourth as many weighings are needed by Yates' method to procure the same accuracy in the average.

A further improvement, which does not seem to have been mentioned in the literature, will be obtained if Yates' procedure is modified by placing in the other pan of the scale those of the objects not included in one of his weighings. Calling the readings in this case $z_1, \cdots, z_8$, we have

$$\begin{aligned}
a + b + c + d + e + f + g &= z_1\\
a + b + c - d - e - f - g &= z_2\\
a - b - c + d + e - f - g &= z_3\\
a - b - c - d - e + f + g &= z_4\\
-a + b - c + d - e + f - g &= z_5\\
-a + b - c - d + e - f + g &= z_6\\
-a - b + c + d - e - f + g &= z_7\\
-a - b + c - d + e + f - g &= z_8.
\end{aligned}$$

From these equations,

$$a = \frac{z_1 + z_2 + z_3 + z_4 - z_5 - z_6 - z_7 - z_8}{8},$$

with a like expression for each of the other unknowns. The variance of each unknown by this method is $\sigma^2/8$. The standard error is half that by Yates' method, or a quarter of its value by the direct method of weighing each object separately. The number of repetitions required to procure a particular standard error in the mean is one-sixteenth that by the direct method.

A simpler example illustrating the same point is that of two objects to be weighed, with a scale already corrected for bias. Again let σ^2 be the variance of an individual weighing. If we weigh the two objects together in one pan of the scale, and then in opposite pans, we have as equations for the unknown weights a and b,

$$a + b = z_1, \qquad a - b = z_2,$$

whence

$$a = (z_1 + z_2)/2, \qquad b = (z_1 - z_2)/2.$$

The variances of a and b by this method are both equal to $\sigma^2/2$, half the value when the two objects are weighed separately. The means found from a number of pairs of weighings of sums and differences have the same precision as those found from twice as many pairs of weighings of the objects separately.

Further economies of effort, or gains in accuracy, are possible with larger numbers of weighings and of objects to be weighed. These improvements can to some extent be applied also to other types of measurement, as of distances, since it is sometimes possible to measure the sum of a number of such quantities, or the difference between two such sums, with approximately the same accuracy as a single one of them. The outstanding case, however, seems to be that of weighing on a balance objects light enough so that their aggregate weight is below the maximum for which the balance was designed, since in this case it is quite reasonable to assume that the several recorded results all have the same standard error σ and that they are independent.

In what follows, some principles underlying the design of efficient schemes of this kind will be developed and applied to obtain some additional plans. However no comprehensive general solution has been reached; this appears to be a matter for further mathematical research. Also, we leave aside in this paper the problem of estimating the error variance. All this discussion is based on the minimization of the actual variance. In order to utilize the results it is necessary that this variance be either known a priori or estimated from the residuals from the least-square solution. The latter type of estimate is in some ways more satisfactory, since it refers to the actual experiment rather than to some previous experiments which may not have been made under exactly the same conditions. But in order to have such an estimate it is necessary that the number of observations exceed the number of unknowns, and desirable that the excess shall have a large enough value to insure a stable estimate of the error variance σ^2. The appropriate test for significance, or determination of confidence limits for the unknowns, must then utilize the Student distribution or its generalization, the variance ratio distribution, which take full account of the instability caused by an inadequate number of degrees of freedom for estimating σ.

It is only when σ is known exactly apart from the experiment being designed that the criteria we here consider are exactly applicable. In other cases there may need to be a balancing, in the design of the experiment, between the desiderata of *minimum* variance and of *accurately known* variance, with the accuracy of this knowledge depending on the number of available degrees of freedom. A theory of design taking full account of this consideration would require a use of the power functions of the Student distribution and the variance ratio distribution, discovered respectively by R. A. Fisher [3] and P. C. Tang [4].

We shall denote by N the number of weighings to be made, and by p the number of objects to be weighed. In order that it be possible to determine the unknown weights from the observations it is necessary that $p \leqq N$, and if a possible bias in the scale must be eliminated by means of the same data it is necessary that $p \leqq N - 1$. Supposing these conditions to be satisfied, we shall show, among other things, that the minimum possible variance for one of the unknowns is σ^2/N; that the experiment may be arranged so that a selected one of the unknowns has exactly this minimum variance excepting when N is odd and a bias must be allowed for also; and that for some, but not all, combinations of p and N, this minimum variance is attained for all the unknowns simultaneously. This minimum value σ^2/N is of course equal to the variance of the mean of N weighings of one object alone, disregarding the rest; but it will be seen below that by complex experiments of the kind indicated, determinations from the same number of weighings of the other weights also can at the same time be made with some finite variance, which may or may not have the minimum value.

The following notation will be used in the proof. Let $x_{i\alpha} = 1$ or -1 if the ith object is included in the αth weighing by being placed respectively in the left- or right-hand pan, and let $x_{i\alpha} = 0$ if the ith object is not included in the αth weighing. Here $i = 1, 2, \cdots, p$ and $\alpha = 1, \cdots, N$. Let y_α be the result recorded for the αth weighing, let Δ_α be the error in this result, and let b_i be the true weight of the ith object, so that we have the N equations

$$x_{1\alpha}b_1 + x_{2\alpha}b_2 + \cdots + x_{p\alpha}b_p = y_\alpha + \Delta_\alpha, \tag{1}$$

provided there is no bias, or if by y_α we mean the observed weight corrected for a bias known a priori. Under these conditions the estimate of each of the b_i's having the properties of zero bias and minimum variance is that provided by the method of least squares. This statement, which does not depend on any assumption of a normal or other particular form of distribution of the errors, has been known long but not widely, since there is an easier derivation of the method by the application to the normal distribution to the method of maximum likelihood. Its proof, due originally to Laplace, has appeared in many forms in the work of Gauss and later authors [5]; the latest version is by the present writer [6].

Letting S stand for summation over all the N weighings we put

$$a_{ij} = Sx_{i\alpha}x_{j\alpha}, \qquad g_i = Sx_{i\alpha}y_\alpha, \tag{2}$$

and write the normal equations in the form

$$\Sigma a_{ij}b_j = g_i,$$

where Σ stands for a sum with respect to j from 1 to p. From the usual theory of least squares (cf. for example the reference last cited) it is known that the standard error of the determination of b_1 from these equations—which is the minimum possible standard error of b_1 for any way of combining the observations—is σ times the square root of A_{11}/A, where

$$A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \cdot & \cdot & \cdot & \cdot \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{vmatrix},$$

and A_{11} is the minor of A obtained by deleting the first row and column.

The matrices of A and of A_{11} are known to be positive definite or semi-definite. The semi-definite case is excluded by the consideration that the normal equations shall actually determine the unknowns. Hence the inverse of the latter matrix exists and is positive definite. But this inverse, which we may write

$$d = \begin{bmatrix} d_{22} & \cdots & d_{2p} \\ \cdot & \cdot & \cdot \\ d_{p2} & \cdots & d_{pp} \end{bmatrix},$$

consists of the coefficients in the identity

$$A/A_{11} = a_{11} - \sum_{i,j=2}^{p} d_{ij} a_{i1} a_{j1} .$$

which is obtained by expanding A with reference to its first row and first column. The positive definite character of d therefore leads to the following

LEMMA: *If $a_{12}, \cdots, a_{1p}$ $(= a_{21}, \cdots, a_{p1}$ respectively) are free to vary while the other elements of A remain fixed, the maximum value of A/A_{11} is a_{11}, and is attained when and only when $a_{12} = a_{13} = \cdots = a_{1p} = 0$.*

From this it is evident that the variance of b_1, namely $\sigma^2 A_{11}/A$, cannot be less than σ^2/a_{11}, and will reach this value only if the experiment is so arranged that the elements after the first in the first row and column of A are all zero. That such an arrangement is possible may be seen by a consideration of the matrix

$$X = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{p1} \\ x_{12} & \cdot & \cdot & x_{p2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{1N} & \cdot & \cdot & x_{pN} \end{bmatrix}$$

whose elements are restricted to be 1's, -1's and 0's. The condition $a_{i1} = 0$, by (2), means simply that $Sx_{i\alpha}x_{1\alpha} = 0$, a condition which may be expressed by saying that the first column of X is orthogonal to the ith column. The condition that the variance of b_1 have its minimum value σ^2/a_{11} is thus, according to the lemma, that the first column of X shall be orthogonal to all the others. The *minimum minimorum* of this variance will be reached if the first row of X is not only orthogonal to all the others, but consists entirely of 1's and -1's, so that $a_{11} = N$. The value of this minimum minimorum is σ^2/N.

If there is a possible bias b_0 this procedure needs to be modified by the addition of b_0 to the left member of (1) and subsequent treatment of this term like the others, putting $x_{0\alpha} = 1$ in (2), and modifying X by adjoining a column of 1's. The necessary and sufficient condition that the variance of b_1 shall equal σ^2/N is then that the column

$$\begin{matrix} x_{11} \\ x_{12} \\ \cdots \\ x_{1N} \end{matrix}$$

shall consist entirely of 1's and -1's and shall be orthogonal to a column consisting of 1's, and to all the other columns of X.

If no bias needs to be eliminated the experiment can be arranged so that the variance of b_1 is σ^2/N merely by filling up the first column of X with 1's and -1's in any arbitrary manner, and then choosing the later columns so as to be orthogonal to this first one, and so that all are linearly independent. This can be accomplished, for example by choosing the first element in all the columns to be the same as that in the first column; choosing the ith element in the ith column ($i = 2, 3, \cdots, p$) to be the negative of the ith element in the first column; and making all the other elements of X zero.

When a bias is to be eliminated, so that there is a column of 1's in X corresponding to b_0, it is necessary that N be even in order that the column of X corresponding to b_1 may consist of 1's and -1's in equal numbers, without any 0's, a condition essential for the orthogonality between these two columns with the maximum value N for a_{11}. Supposing N even, let us assign the value 1 to each of the first $N/2$ elements of the column corresponding to b_1 and the value -1 to the last $N/2$ elements of this row. The remaining rows of X may then be filled up by the same method as that indicated above for the case in which there is no bias. The variance of b_1 will then take its theoretical minimum value σ^2/N.

If N is odd and there is a possible bias, the column of X corresponding to b_1 can be filled up with 1's and -1's in equal numbers, with a single zero, and the remaining columns can be made orthogonal to it. The variance of b_1 in this case will be $\sigma^2/(N - 1)$.

The method suggested above for filling up the later columns of X is convenient for the proof, but is not usually to be recommended in practice, since other methods will in all but the simplest cases give smaller standard errors for the unknowns other than the first. For some values of N and p it is possible to determine all the unknowns with equal and minimum variance. These are the cases in which all the columns of X can be made mutually orthogonal and without zeros, excepting that the column corresponding to b_0 may contain some zeros. Thus for $N = 4$ the scheme of weighing represented by the matrix

$$X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix},$$

whose columns are all mutually orthogonal, may be applied to weigh three objects when there is a possible bias, or four where there is not, with variance $\sigma^2/4$ for each of the unknowns in either case. The matrix $X'X$ of the normal equations has the form

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

Calling the results of the weighings y_0, y_1, y_2, y_3 in the case of possible bias we have for the unknowns the expressions

$$b_1 = (y_0 + y_1 - y_2 - y_3)/4$$

$$b_2 = (y_0 - y_1 + y_2 - y_3)/4$$

$$b_3 = (y_0 - y_1 - y_2 + y_3)/4.$$

The complete orthogonality exemplified by this design has several advantages besides the fact that the variance of each of the unknowns has the same minimum value as if all the weighings were to be devoted to it alone (or half the value of the variance of this unknown if half the weighings were devoted to it plus bias and half to determining the bias). The diagonal form of the matrix $X'X$ means that the labor of solving normal equations, which is sometimes formidable, is reduced to the trivial task of dividing by N. Also, the diagonal form of this matrix implies that its inverse is also of diagonal form, from which it follows that the estimates of the different unknowns are statistically independent. Consequently the variances, or standard errors, of linear functions of the unknowns are easy to find. Thus the variance of the difference between the estimates of two of the weights is simply the sum of their variances. But of course if the main object of the experiment is to determine a particular difference of this kind, or any other linear function of the weights, a different design should be sought to minimize the particular variance which is of interest.

In contrast to the satisfactory design possible with four weighings, no complete orthogonality is possible with six weighings, or with any odd number, if the number of objects to be weighed is the maximum possible for the number of weighings and if each object is actually to enter into each weighing in one pan or the other. For $N = 3$ and bias known to be zero consider the scheme

$$X = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix},$$

which corresponds to weighing two objects, first together in one pan, then in opposite pans, and then weighing one alone. Calling b_1 the weight of the object that has been on the scale through all three weighings and b_2 the other we have the estimates

$$b_1 = (y_1 + y_2 + y_3)/3$$

$$b_2 = (y_1 - y_2)/2,$$

with respective variances

$$\sigma_1^2 = \sigma^2/3, \qquad \sigma_2^2 = \sigma^2/2.$$

Thus the first weight is determined with the minimum possible variance but the second is not.

An alternative method of weighing under these same conditions is to weigh both objects in one pan together twice and to weigh them in opposite pans once. This gives

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix},$$

with the normal equations

$$3b_1 + b_2 = y_1 + y_2 + y_3$$

$$b_1 + 3b_2 = y_1 + y_2 - y_3,$$

whose solution is

$$b_1 = (y_1 + y_2 + 2y_3)/4$$

$$b_2 = (y_1 + y_2 - 2y_3)/4,$$

and variances

$$\sigma_1^2 = \sigma_2^2 = \tfrac{3}{8}\sigma^2.$$

Thus the weights are by this method determined with equal accuracy, which is better than by the preceding method for one of the objects but worse for the other. To choose between the two methods it is therefore appropriate to take into consideration the relative accuracy desired in the weights of the two objects. Either method is better than weighing the objects separately.

Either of these two X matrices can also be made the basis for weighing a single object when the scale is suspected of having a bias. The weight of this object will be estimated as b_2, and will have the variance $\frac{1}{2}\sigma^2$ by the first method, or $\frac{3}{8}\sigma^2$ by the second. Thus the second method is distinctly superior in this case.

Orthogonality between columns obviously requires both negative and positive signs, corresponding to weighings in both pans of the balance. Thus the experimental designs of maximum efficiency for weighing on a balance are not available with a spring scale, or in making measurements of any kind in which it is not possible to arrange that the quantities read off are differences. In such cases the elements of X are restricted to be 1 or 0. Let us now consider some of the simplest cases of this kind, assuming for simplicity that $\sigma = 1$. We shall deal only with cases in which there is no bias.

For $N = 3$, $p = 2$ the simple experiment of weighing one object twice and the other once yields variances $\frac{1}{2}$ and 1 respectively. All other designs are in this case less satisfactory, with the possible exception of that specified by

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$

with $b_1 = (y_1 + 2y_2 - y_3)/3$ and $b_2 = (y_1 - y_2 + 2y_3)/3$ having each the variance of $\frac{2}{3}$.

For $N = 3$, $p = 3$ the most efficient design is given by

$$X = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix},$$

with $b_1 = (y_1 + y_2 - y_3)/2$, and b_2 and b_3 given by cyclic permutation in this formula. The variance of each unknown is $\frac{3}{4}$.

For $N = 4$, $p = 3$ a design having an advantage in some situations is that given by

$$X = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

(together of course with those obtained by permutations of rows and of columns, as is to be understood throughout). The normal equations are

$$\begin{aligned} 3b_1 + 2b_2 + 2b_3 &= y_1 + y_2 + y_3 \\ 2b_1 + 3b_2 + 2b_3 &= y_2 + y_3 + y_4 \\ 2b_1 + 2b_2 + 3b_3 &= y_1 + y_3 + y_4 . \end{aligned}$$

An expeditious method of solution in this as in many similar cases is to add them all together and then subtract an appropriate multiple of the sum from each of the normal equations in turn. The variance of each unknown found by this experiment is $5/7 = .714$. The simple experiment consisting of weighing one of the objects twice and the others once each yields variances in one case larger and in two cases smaller than this.

For $N = p = 4$ the cyclic arrangement

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

leads to variances all equal to 7/9.

For $N = 5$, $p = 2$ the most efficient design appears to be

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

The variance of each unknown is in this case 1/3.

REFERENCES

[1] R. A. FISHER, *The Design of Experiments*, Oliver and Boyd, London and Edinburgh, 1935 and 1937.

[2] F. YATES, "Complex experiments," *Jour. Roy. Stat. Soc. Supp.*, Vol. 2 (1935), pp. 181–247, including discussion. The reference is to p. 211.

[3] R. A. FISHER, "The sampling error of estimated deviates. . . ," *Mathematical Tables*, Vol. 1 (1931), pp. xxvi–xxxv. London, British Association for the Advancement of science.

[4] P. C. TANG, "The power function of the analysis of variance tests with tables and illustrations of their use," *Stat. Res. Mem.*, Vol. 2 (1939), pp. 126–149.

[5] Cf. WHITTAKER and ROBINSON, *Calculus of Observations*, Section 115.

[6] HAROLD HOTELLING, "Problems of prediction," *Am. Jour. of Sociology*, Vol. 48 (1942), pp. 61–76.