J. Hyg., Camb. (1962), **60**, 379
Printed in Great Britain

On combining estimates of relative potency in bioassay

BY B. M. BENNETT

The Department of Preventive Medicine, University of Washington, Seattle 5, Washington U.S.A.

(*Received* 16 *April* 1962)

INTRODUCTION AND STATEMENT OF PROBLEM

Biological assays based on a quantitative response in particular (e.g. hormone assays) typically involve the estimation of (log) relative potency ($= M$) with respect to some standard preparation on the basis of one or more symmetrical determinations usually at equally spaced (log) dose intervals. It is the purpose of this paper to consider further statistical aspects of estimates of M, and especially confidence limits for M based on a combination of separate assays. These are given as examples of tests of linear hypotheses. Two applications are given on the combining of quantitative bio-assay data.

PARALLEL LINE ASSAYS

We consider first the $(2k)$-point symmetric parallel line assay (Finney, 1952, ch. 4), i.e. a series of quantitative responses $\{y_{is}\}$, $\{y'_{it}\}$ $(i = 1, \ldots, k;\ s, t = 1, \ldots, n)$. The y's are independently and normally distributed with a common unknown variance ($= \sigma^2$) and mean values

$$\left.\begin{aligned} \eta_i &= E(y_{is}) = \alpha_1 + \beta x_i, \\ \eta'_i &= E(y'_{it}) = \alpha'_1 + \beta x'_i, \end{aligned}\right\} \tag{1}$$

for the two groups of biological substances ('Standard' *vs.* 'Test'). Here the x's represent series of equally spaced doses (usually in logs) of the substances involved.

As a test of the hypothesis H_0: $\alpha'_1 = \alpha_1 + \beta M$, if M represents the required (log) relative potency, we use the Student test

$$t = \frac{|(y'_{\cdot\cdot} - y_{\cdot\cdot}) - b(x'_{\cdot} - x_{\cdot} + M)|}{s_f \sqrt{\left[\frac{1}{N} + \frac{1}{N'} + \frac{1}{S_{xx}}\{(x'_{\cdot} - x_{\cdot}) + M\}^2\right]}}, \tag{2}$$

with the following notation, $N = N' = nk$

$$\begin{aligned} ny_{i\cdot} &= \Sigma y_{is}, & n'y'_{i\cdot} &= \Sigma y'_{it}, \\ Ny_{\cdot\cdot} &= \Sigma\Sigma y_{is}, & N'y'_{\cdot\cdot} &= \Sigma\Sigma y'_{it}, \\ kx_{\cdot} &= \Sigma x_i, & kx'_{\cdot} &= \Sigma x'_i, \end{aligned}$$

$$\frac{1}{n} S_{xx} b = \Sigma(x_i - x_{\cdot}) y_{i\cdot} + \Sigma(x'_i - x_{\cdot}) y'_{i\cdot},$$

$$\frac{1}{n} S_{xx} = \Sigma(x_i - x_{\cdot})^2 + \Sigma(x'_i - x'_{\cdot})^2,$$

and s_f^2 is an estimate of the variance based on f degrees of freedom, and usually $f = 2N - 3$. The primed N's will be retained in order to allow for the possibility of differing numbers of observations in the two groups.

The t-test above may be derived easily as a test of the linear hypothesis H_0 (Kolodziejczyk, 1935). This consists of obtaining the relative and absolute minimum Q_r, Q_a, respectively, of the quadratic form

$$Q = \sum_i \sum_s (y_{is} - \alpha_1 - \beta x_i)^2 + \sum_i \sum_t (y'_{it} - \alpha'_1 - \beta x'_i)^2,$$

where Q_r is the restricted minimum of Q if H_0 is true.

It may be verified that the test for the linear hypothesis is

$$F = \frac{(N+N'-3)}{1} \frac{(Q_r - Q_a)}{Q_a} = t^2,$$

which has the F or variance ratio distribution. Here Q_a is defined as

$$Q_a = (N+N'-3)\, s_f^2 = \sum_i \sum_s \{y_{is} - y_{\cdot\cdot} - b(x_i - x_{\cdot})\}^2 + \sum_i \sum_t \{y'_{it} - y'_{\cdot\cdot} - b(x'_i - x'_{\cdot})\}^2.$$

Fieller (1944) showed that the t-test: $|t| \leqslant t_\epsilon$ with f degrees of freedom and level of significance ($=\epsilon$) provided a confidence interval for M as the roots of a certain equation if the combined slope ($=b$) differs significantly from zero. This latter condition was subsequently modified (Fieller *et al.*, 1954) by further distinguishing the cases in which the confidence limits could be (i) inclusive, (ii) exclusive, (iii) non-existent according to the significance level used.

If $M^* = \{x'_{\cdot} - x_{\cdot} + M\}$ we obtain from (2) the following inequality

$$g(M^*) = \lambda_0 M^{*2} - 2\lambda_1 M^* + \lambda_2 < 0, \tag{3}$$

where

$$\lambda_0 = b^2 - \frac{1}{S_{xx}}(s_f^2 t_\epsilon^2)$$
$$\lambda_1 = (y'_{\cdot\cdot} - y_{\cdot\cdot})\, b$$
$$\lambda_2 = (y'_{\cdot\cdot} - y_{\cdot\cdot})^2 - \left(\frac{1}{N} + \frac{1}{N'}\right)(s_f^2 t_\epsilon^2),$$

and the zeros of $g(M^*)$ are the lower and upper confidence limits for M^* provided λ_0 is > 0.

SERIES OF PARALLEL LINE ASSAYS

We consider now a series of c independent $(2k)$-point parallel line assays. The observations will be denoted $\{y_{ris}\}$, $\{y'_{rit}\}$ $(r = 1, \ldots, c;\ i = 1, \ldots, k;\ s, t = 1, \ldots, n_r = n'_r)$ normally and independently distributed with means

$$\left.\begin{aligned} \eta_{ri} &= E(y_{ris}) = \alpha_r + \beta_r x_{ri}, \\ \eta'_{ri} &= E(y'_{rit}) = \alpha'_r + \beta_r x'_{ri}, \end{aligned}\right\} \tag{4}$$

and the hypothesis of a common relative potency ($=M$) for the series of assays is then:

$$\alpha'_r = \alpha_r + \beta_r M \quad (r = 1, \ldots, c).$$

The appropriate test for this linear hypothesis is with appropriate changes of notation from (2)

$$\begin{aligned} F &= \frac{(N+N'-3c)}{c} \frac{Q_r - Q_a}{Q_a} \\ &= \frac{(N+N'-3c)}{cQ_a} \sum_{r=1}^{c} \frac{b_r^2 (M_r - M)^2}{\left[\dfrac{1}{N'_r} + \dfrac{1}{N_r} + \dfrac{1}{S_{rxx}}\{(x'_{r\cdot} - x_{r\cdot}) + M\}^2\right]} \leqslant F_\epsilon, \end{aligned} \tag{5}$$

with $N = \Sigma N_r = N' = \Sigma N'_r$. In (5) F has the variance ratio distribution with c and $(N+N'-3c)$ degrees of freedom, respectively. Also defined in (5) are the sample values

$$M_r + (x'_{r.} - x_{r.}) = (y_{r..} - y_{r..})/b_r,$$

and
$$Q_a = \sum_r \sum_i \sum_s \{y_{ris} - y_{r..} - b_r(x_{ri} - x_{r.})\}^2 + \sum_r \sum_i \sum_t \{y'_{rit} - y'_{r..} - b_r(x'_{ri} - x'_{r.})\}^2, \tag{6}$$

$$S_{rxx} b_r = \sum_{i,s} (x_{ri} - x_{r.}) y_{ri.} + \sum_{i,t} (x'_{ri} - x'_{r.}) y'_{ri.},$$

$$S_{rxx} = \sum_{i,s} (x_{ri} - x_{r.})^2 + \sum_{i,t} (x'_{ri} - x'_{r.})^2.$$

As an equation of degree $2c$ in M, $F(M) = F_\epsilon$ must generally be solved by approximate or iterative methods. When this equation has only two real roots, these form the appropriate confidence interval for M, and are 'shortest' only in the sense that they less frequently cover any 'wrong' value of M than any other set. When there are more than two real roots there are inevitably some difficulties in interpreting the separate intervals formed. It is conjectured that there will usually be two and only two real roots in the case of closely controlled experimental assay data. No formal proof of this is presently available for the zeros of random polynomials from (5).

It is important to note that the value of M which minimizes $F(M)$ is also the maximum likelihood estimate ($= \hat{M}$) since $F(\hat{M})$ is proportional to $(Q_r - Q_a)$ for constant values of Q_a. A similar method was used by Tocher (1952) on the problem of estimating the point of concurrence of regression lines.

If now in each term of the denominator of F in equation (5) we use the approximation $M_r \simeq M$, i.e. replace $(x'_{r.} - x_{r.}) + M$ by $(y'_{r..} - y_{r..})/b_r$, equation (5) becomes

$$F = \frac{N+N'-3c}{cQ_a} \sum_{r=1}^{c} \frac{b_r^2 (M_r - M)^2}{\left[\frac{1}{N_r} + \frac{1}{N'_r} + \frac{1}{S_{rxx}} \frac{(y'_{r..} - y_{r..})^2}{b_r^2}\right]} \leqslant F_\epsilon, \tag{7}$$

and provides approximate $100(1-\epsilon)\%$ confidence limits for M. The use of this approximation gives as a minimizing value for

$$\Sigma w_r (M_r - M)^2 = \Sigma w_r M_r^2 - (\Sigma w_r M_r)^2 / (\Sigma w_r),$$

the weighted mean or estimate $\bar{M} = (\Sigma w_r M_r)/(\Sigma w_r)$ based on the weights

$$w_r = \frac{b_r^2}{\left[\frac{1}{N_r} + \frac{1}{N'_r} + \frac{1}{S_{rxx}} \frac{(y'_{r..} - y_{r..})^2}{b_r^2}\right]} \tag{8}$$

for $r = 1, \ldots, c$.

In the special case where

$$\frac{1}{N^*} = \frac{1}{N_r} + \frac{1}{N'_r} = \text{constant}, \quad S_{xx} = S_{rxx} = \text{constant}, \quad (x'_{r.} - x_{r.}) = \text{constant},$$

equation (5) in $M^* = (x'_. - x_.) + M$ becomes

$$F = \frac{N+N'-3c}{cQ_a} \frac{\sum_{r=1} \{y'_{r..} - y_{r..} - b_r M^*\}^2}{\left(\frac{1}{N^*} + \frac{1}{S_{xx}} M^{*2}\right)} \leqslant F_\epsilon, \tag{9}$$

or regarded as an inequality in M^*

$$G^*(M^*) = \lambda_0^* M^{*2} - 2\lambda_1^* M^* + \lambda_2^* \leqslant 0, \tag{10}$$

where

$$\lambda_1^* = \sum_r b_r^2 - \frac{(cQ_a) F_\epsilon}{(N+N'-3c) S_{xx}},$$

$$\lambda_1^* = \sum_r b_r (y'_{r..} - y_{r..}),$$

$$\lambda_2^* = \sum_r (y'_{r..} - y_{r..})^2 - \frac{(cQ_a) F_\epsilon}{N^*(N+N'-3c)},$$

and the zeros of $G^*(M^*)$ are the lower and upper $100(1-\epsilon)\,\%$ confidence limits for M^* if $\lambda_0^* > 0$. This latter condition is equivalent to a joint test that not all the regression coefficients $\{\beta_r\}$ $(r = 1, \ldots, c)$ are significantly different from zero.

Finally, it should be mentioned that the methods used in the previous sections may be easily extended to the case of combining the results of assays for which the response is quantal or also to the case of slope-ratio assays (Finney, 1952).

COMPARISONS WITH OTHER ESTIMATES

(i) For combining the results of parallel line assays Finney (1952, §14·3) presents a method for determining an average potency from the ratio estimate

$$\bar{M}' = \sum_r w'_r (H_r - b_r G_r) / \Sigma w'_r b_r, \tag{11}$$

$$M_r = \frac{1}{b_r}(H_r - G_r b_r),$$

$$G_r = (x'_{r.} - x_{r.}), \quad H_r = (y'_{r.} - y_{r.}),$$

and the condition that the approximate variance of $\bar{M}' = V(\bar{M}')$ is minimized gives as a choice of weights

$$w'_r = \left(\frac{1}{N_r} + \frac{1}{N'_r}\right)^{-1}.$$

Approximate confidence limits for the ratio $\bar{M}'$ are then available by an application of Fieller's theorem (Finney, p. 373.)

(ii) The limits available from an iterative solution of equation (5) are also to be compared with those provided by the interval estimate: $\bar{M} + t_\epsilon \{V(\bar{M})\}^{\frac{1}{2}}$ based on a weighted average $\bar{M} = \Sigma w_r M_r / \Sigma w_r$, where the reciprocals of the weights are:

$$\frac{1}{w_r} = V(M_r) = \frac{1}{b_r^2}\left[\left(\frac{1}{N_r} + \frac{1}{N'_r}\right) + \frac{(x'_{r.} - x_{r.} + M_r)^2}{S_{rxx}}\right],$$

i.e. proportional to the approximate variance of each sample M_r $(r = 1, \ldots, c)$ and t_ϵ is the normal deviate for level of significance ϵ. Confidence limits obtained by this method are to be considered as approximate only, and are based on an assumed ultimate normality of the distribution of $\bar{M}$. These weights are of course those obtained as in equation (8).

(iii) When different types of assays are to be combined, Irwin (1950) has also proposed a weighted average of the form

$$\bar{M}'' = \pm \Sigma w''_r (H_r - b_r G_r - b_r \log \rho)$$

and determined the weights in order to obtain a minimum variance for this expression.

Example 1

The following example compares the various estimates of M and their respective confidence limits for the case : $c = 2$. The data are those from two assays of vitamin D_3 (Finney, 1952, Table 14.2) using line test scores as responses.

Equation (5) $F(M) = F_\epsilon$ results in a quartic in M, the coefficients of which will not be reproduced here because of limitations of space. The maximum likelihood estimate ($= \hat{M}$) was obtained from a series of computed values of $F(M)$ by interpolation. Equation (5) has only two real roots (0·463, 2·129) and a minimum at $\hat{M} = 1{\cdot}231$. The curve $F(M)$ (of Fig. 1) crosses the horizontal asymptote

$$F(M) = F(\pm\infty) = \frac{1}{cs^2}\sum_r S_{rxx} b_r^2$$

at a negative value ($M = -3{\cdot}75$) and also approaches the horizontal asymptote from below for large positive values of M. The condition

$$F(M) = \frac{1}{cs^2}\sum S_{rxx} b_r^2 > F_\epsilon$$

is equivalent to a test that all regression coefficients are significantly different from zero.

Table 1. *Comparison of estimates of M and confidence limits (2 assays of vitamin* D_3*)*

	Estimate of M	95 % confidence limits
(i). Equation 11	$\overline{M}' = 1{\cdot}213$	(0·694, 1·829)
(ii).	$\overline{M} = 1{\cdot}182$	(0·431, 1·934)
Equation (5)	$\hat{M} = 1{\cdot}231$	(0·463, 2·129)

The exact confidence interval available from equation (5) is seen to be somewhat wider than that available from (i) and (ii) and non-symmetric. The confidence limits are inclusive for significance levels $\epsilon < 0{\cdot}0000\,002$, at which level they then become exclusive $(-3{\cdot}75, +\infty)$. For $\epsilon < 0{\cdot}0000001$ ($F_\epsilon = 33{\cdot}5$) the limits are non-existent.

Example 2

As a further example of the comparison of the various estimates of M and their confidence limits, we consider the data of Smith, Marks, Fieller and Broom on four cross-over assays of a test preparation of insulin (Finney, 1952, table 14.6). The first assay had the additional complication of one missing observation. In computing M the pooled estimate of the variance ($s^2 = 26{\cdot}299$ with 31 degrees of freedom) was used since the variances did not differ significantly from assay to assay. The polynomial function $F(M)$ is of degree eight and is approximately symmetric about $M = 0$ (cf. Fig. 2). It approaches the horizontal asymptotes $F(\pm\infty)$ from below for negative values of M, and from above for positive M with one intersection point at an extremely large positive value ($M = 531{\cdot}3$). In the intervals of M of interest $F(M)$ is parabolic in form with only two intersection points with the horizontal line: $F(M) = F_\epsilon$, so that the confidence limits are generally inclusive for extremely small levels of significance ($\epsilon < 0{\cdot}0000001$) and then exclusive ($-\infty$, 531·3).

The separate estimates of M in both Examples 1 and 2 closely approximate each other and the estimate in Example 2 reflects the essential symmetry of a high-order polynomial about the axis. The criterion of shortest length is not relevant here

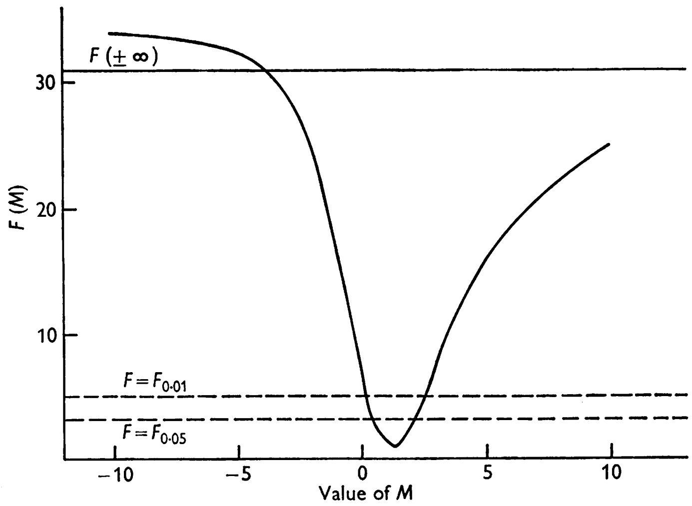

Fig. 1. Example 1 (2 assays of vitamin D_3).

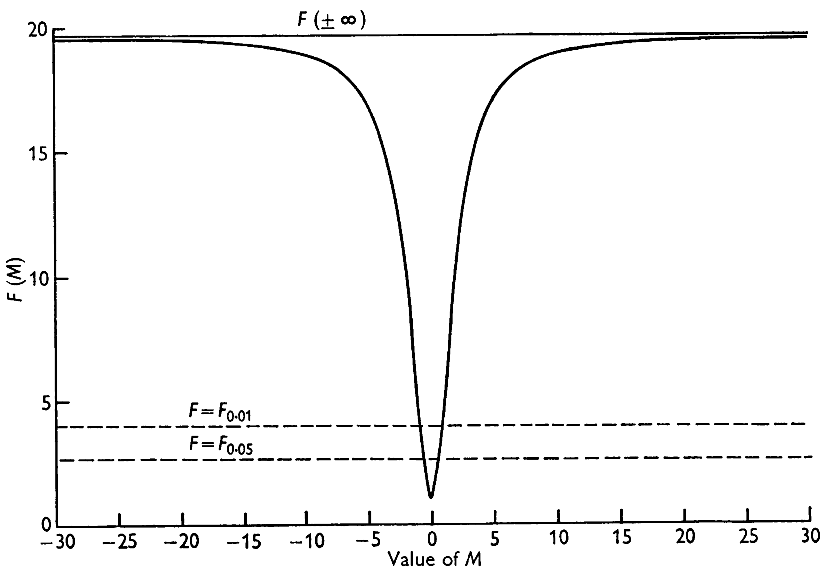

Fig. 2. Example 2 (4 assays of insulin).

in the various comparisons since both methods (i) and (ii) involve the assumption of a ratio of normal variates and approximations to its variance.

It should be noted that Examples 1 and 2 both involve extremely 'good' biological assay data, and this is reflected in the extraordinarily small significance levels required before the corresponding confidence limits available from (5) lose the inclusiveness property. Some attention is also being given to applications to less satisfactory assay data, i.e. where in fact certain of the individual slopes may fail to be significant and more than two real roots occur.

Table 2. *Comparison of estimates of M and confidence limits (4 assays of insulin)*

Method	Estimate of M	95 % confidence limits
(i)	$\overline{M}' = -0{\cdot}084$	$(-0{\cdot}575, 0{\cdot}404)$
(ii)	$\overline{M} = 0{\cdot}035$	$(-0{\cdot}417, 0{\cdot}487)$
Equation 5	$\hat{M} = -0{\cdot}00001$	$(-0{\cdot}591, 0{\cdot}587)$

Example 3

As a further example of the linear hypothesis we consider an appropriate test for equality of relative potency ($M_1 = M_2$) based on two independent parallel line assays assuming $\beta_1 = \beta_2$. This is equivalent to a test of the linear hypothesis: $\alpha_1 - \alpha_2 = \alpha_1' - \alpha_2'$ in the notation of the earlier section. For this hypothesis the appropriate test may be shown to be

$$t = \frac{|(y_{1..}' - y_{1.}) - (y_{2..}' - y_{2..}) - \bar{b}\{(x_{1.}' - x_{1.}) - (x_{2.}' - x_{2.})\}|}{s\sqrt{\left[\left(\frac{1}{N_1} + \frac{1}{N_1'}\right) + \left(\frac{1}{N_2} + \frac{1}{N_2'}\right) + \frac{1}{S_{xx}}\{(x_{1.}' - x_{1.}) - (x_{2.}' - x_{2.})\}^2\right]}}, \tag{12}$$

or the Student t-test with $N + N' - 5$ degrees of freedom based on the estimate $(N + N' - 5)\, s^2 = Q_a$ and combined slope equal to $\bar{b}$ defined as in (6) previously.

Certain portions of this article were prepared while the author was a Fulbright Research Scholar at the University of Aberdeen. The helpful comments of Dr D. J. Finney and computational assistance from the Department of Statistics are gratefully acknowledged.

REFERENCES

FIELLER, E. C. (1944). A fundamental formula in the statistics of biological assay. *Quart. J. Pharm.* **17**, 117.

FIELLER, E. C., CREASY, M. A. & DAVID, S. T. (1954). Symposium on interval estimation. *J. R. statist. Soc.* B **16**, 175.

FINNEY, D. J. (1952). *Statistical Method in Biological Assay*. London: Griffin and Co. Ltd.

IRWIN, J. O. (1950). Biological assays with special reference to biological standards. *J. Hyg., Camb.*, **48**, 216.

KOLODZIEJCZYK, ST. (1935). On an important class of statistical hypotheses. *Biometrika*, **27**, 161.

TOCHER, K. (1952). On the concurrence of a set of regression lines. *Biometrika*, **40**, 297.