ON THE CALCULATION OF THE ERROR OF BIOLOGICAL ASSAYS

By J. O. IRWIN, *Of the Statistical Staff, Medical Research Council*

(With 3 Figures in the Text)

1. Introduction

It is now realized by a number of workers that the formula which it has been customary to use hitherto for calculating the limits of error of the dose corresponding to a given response (e.g. the median effective dose) or of the result of a biological assay is only approximate. The usual method gives an approximation to the standard error of the logarithm of the dose in question, or of the result, and then uses it in conjunction with the table of the normal probability integral to find the limits of error. The approximation is obtained by the usual statistical method of treating, as Karl Pearson used to say, statistical differentials as mathematical differentials. Bliss (1935) was, I think, the first to realize that exact fiducial limits could be obtained and actually calculated them in a numerical example—but without discussing their theoretical and practical implications. Their mathematical derivation was first published for particular cases by C. Eisenhart (1939) and by E. C. Fieller (1940) in the appendix to his important paper on the biological standardization of insulin. Fieller also gave the general result with a promise (if circumstances allow) to discuss it more fully elsewhere. Later a method for calculating exact fiducial limits was given in the 'British Standard Method for the biological assay of vitamin D_3 by the chick test' (1940).

Work done since 1939 shows that the approximate formula becomes grossly inadequate when the estimate of the slope of the log. dose-response line is not well determined. If the slope is not significant at a given level of probability, then the fiducial limits for the result are, at that level, 0–∞. This is only common sense, since nothing can be learnt from an assay which does not show any differentiation in response between different levels of dosage. The approximate formula, however, still gives finite limits of error in this case. It may, then, be extremely misleading.

I have in my possession a memorandum which I sent to Mr Fieller in 1939 dealing with the matter at some length and containing a number of points hitherto unpublished. The memorandum is the basis of the following discussion.

2. Fiducial limits

The interpretation to be given to limits of error (as usually calculated) of say 67–150 % at ($P=0{\cdot}99$) for the result of an assay is that in the long run in a proportion P of assays similar to the one actually performed, with samples of the same material, the result will lie between 67 and 150 % of its true value. We may also invert this statement, and say that in the long run, in a proportion P of experiments the true value will lie between 67 and 150 % of the observed result. We then have limits which vary from experiment to experiment because the observed result varies, but which are calculable by a definite rule and which are such that, in the long run, in a proportion P of experiments the true value will lie between them. Such limits are called 'fiducial limits'. If the standard error of the logarithm of the result were known exactly, then the fiducial limits at a given probability

https://doi.org/10.1017/S0022172400012705 Published online by Cambridge University Press

level would, on repetition of the experiment, always be the same percentage of the observed result. The customary statement would be exact instead of only approximate. Actually this is not true. Exact fiducial limits can be calculated by the following method, and they can be expressed as percentages of the observed result. It should be realized, however, that these limits, whether expressed in absolute units or as percentages, would vary on repetition of the experiment. We know that, when they are expressed in absolute units, in a proportion P of experiments, in the long run, the true value would lie between them. We have therefore, in the practical situation in which only one experiment has been performed, a degree of confidence represented by $P=0{\cdot}99$—say—that the true value does lie between them.

3. Fiducial limits for the dose corresponding to a given response

Let k doses be given to groups of animals. If the response is measurable, let y be the average response of a group of animals to a particular dose, x the logarithm to the base 10 of that dose. If the response is quantal, let y be the normal equivalent deviation or probit corresponding to the percentage ($100p$) of animals in the group which react. We are dealing with the case when the relation between y and x is linear.

Let the observed regression line be

$$\eta=\bar{y}+b\,(x-\bar{x}),$$

and let the true value of y (or η) for given x be Y.

When the response is measurable let s^2 be the estimate of variance of the responses of animals on the same dose, based on ν degrees of freedom. Let t be the value of the variate which for ν degrees of freedom corresponds to $\tfrac{1}{2}p_1+\tfrac{1}{2}p_1=p_1$ (0·05 or 0·01 say). For a quantal response let t be the corresponding normal deviate, and let $s^2=1$.

Let w be the weight of the observation y. When the response is measurable, w will be equal to n, the number of animals on the dose in question. When it is quantal, $w=nz^2/pq$, where

$$p=\int_{-\infty}^{\eta}\frac{1}{\sqrt{(2\pi)}}\,e^{-\frac{1}{2}x^2}\,dx,\quad q=1-p,\quad z=\frac{1}{\sqrt{(2\pi)}}\,e^{-\frac{1}{2}\eta^2}.$$

Now the variance of η is

$$V(\eta)=V(\bar{y})+(x-\bar{x})^2\,V(b)$$
$$=A+B\,(x-\bar{x})^2,$$

where

$$A=V(\bar{y})=\frac{s^2}{S\,(w)},\quad B=V(b)=\frac{s^2}{S\,\{w\,(x-\bar{x})^2\}},$$

the summation S being over the number k of doses used.

Since $(\eta-Y)/(A+B\,(x-\bar{x})^2)^{\frac{1}{2}}$ is distributed in the 't' distribution with ν degrees of freedom (or normally in the quantal case) in a proportion $P=(1-p_1)$ of experiments, we shall have

$$\bar{y}+b\,(x-\bar{x})-t\,(A+B\,(x-\bar{x})^2)^{\frac{1}{2}}<Y<\bar{y}+b\,(x-\bar{x})+t\,(A+B\,(x-\bar{x})^2)^{\frac{1}{2}}. \tag{1}$$

If (1) is true

$$\{Y-\bar{y}-b\,(x-\bar{x})\}^2<t^2\,\{A+B\,(x-\bar{x})^2\},$$

or

$$(x-\bar{x})^2\,(b^2-t^2B)-2b\,(x-\bar{x})\,(Y-\bar{y})+(Y-\bar{y})^2-t^2A<0,$$

or

$$(b^2-t^2B)\left\{(x-\bar{x})-\frac{b\,(Y-\bar{y})}{b^2-t^2B}\right\}^2<t^2\left\{A+\frac{B\,(Y-\bar{y})^2}{b^2-t^2B}\right\}.$$

(i) If $b^2>t^2B$, that is if the slope is significant, we find

$$\left|(x-\bar{x})-\frac{b\,(Y-\bar{y})}{(b^2-t^2B)}\right|<\frac{t}{(b^2-t^2B)}\,\{A\,(b^2-t^2B)+B\,(Y-\bar{y})^2\}^{\frac{1}{2}},$$

https://doi.org/10.1017/S0022172400012705 Published online by Cambridge University Press

or $$\left[\bar{x}+\frac{b\,(Y-\bar{y})}{b^2-t^2B}-\frac{t}{(b^2-t^2B)}\{A\,(b^2-t^2B)+B\,(Y-\bar{y})^2\}^{\frac{1}{2}}\right]$$
$$<x<\left[\bar{x}+\frac{b\,(Y-\bar{y})}{b^2-t^2B}+\frac{t}{(b^2-t^2B)}\{A\,(b^2-t^2B)+B\,(Y-\bar{y})^2\}^{\frac{1}{2}}\right]. \qquad (2)$$

This gives the fiducial limits for x which are always real, since $A\,(b^2-t^2B)+B\,(Y-\bar{y})^2$ is positive.

If the response is quantal and fiducial limits for the median effective dose are required we put $Y=0$ (normal equivalent deviations) or 5 (probits).

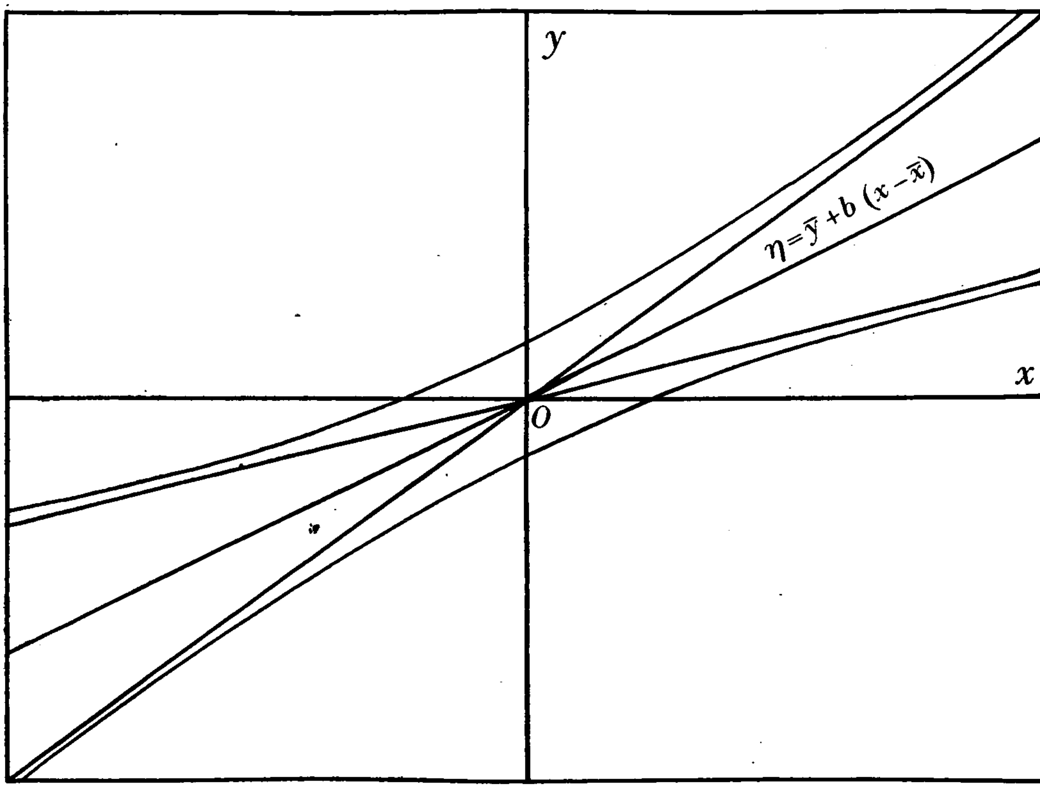

Fig. 1. Here the slope is significantly greater than zero. O is the point $\bar{x}$, $\bar{y}$, where x = log. dose, y = response. The fiducial limits for x corresponding to any given Y and a given value of P (0·95 or 0·99 say) are the points in which the corresponding abscissa cuts the hyperbola.

The fiducial limits for the values x corresponding to different values of Y lie on the hyperbola
$$(x-\bar{x})^2\,(b^2-t^2B)-2b\,(x-\bar{x})\,(Y-\bar{y})+(Y-\bar{y})^2=t^2A. \qquad (3)$$

The asymptotes are obtained by equating the left-hand side of (3) to zero and are
$$\left.\begin{aligned}Y&=\bar{y}+(x-\bar{x})\,(b+t\sqrt{(B)})\\ Y&=\bar{y}+(x-\bar{x})\,(b-t\sqrt{(B)})\end{aligned}\right\}. \qquad (4)$$

The slopes of *both* asymptotes are of the same sign, positive if $b>0$ or negative if $b<0$. Fig. 1 illustrates the position.

(ii) If $b^2<t^2B$, we find
$$\left|(x-\bar{x})+\frac{b\,(Y-\bar{y})}{t^2B-b^2}\right|>\frac{t}{(t^2B-b^2)}\{B\,(Y-\bar{y})^2-A\,(t^2B-b^2)\}^{\frac{1}{2}},$$

https://doi.org/10.1017/S0022172400012705 Published online by Cambridge University Press

whence $$x > \left[\bar{x} - \frac{b\,(Y-\bar{y})}{t^2B-b^2} + \frac{t}{(t^2B-b^2)}\{B\,(Y-\bar{y})^2 - A\,(t^2B-b^2)\}^{\frac{1}{2}}\right],$$

or $$x < \left[\bar{x} - \frac{b\,(Y-\bar{y})}{t^2B-b^2} - \frac{t}{(t^2B-b^2)}\{B\,(Y-\bar{y})^2 - A\,(t^2B-b^2)\}^{\frac{1}{2}}\right]. \quad (5)$$

Thus in this case in a proportion $P=(1-p_1)$ of experiments x will be *outside* the limits given by (5). The fiducial limits corresponding to probability $(1-p_1)$ ($=0\cdot95$ or $0\cdot99$ say)

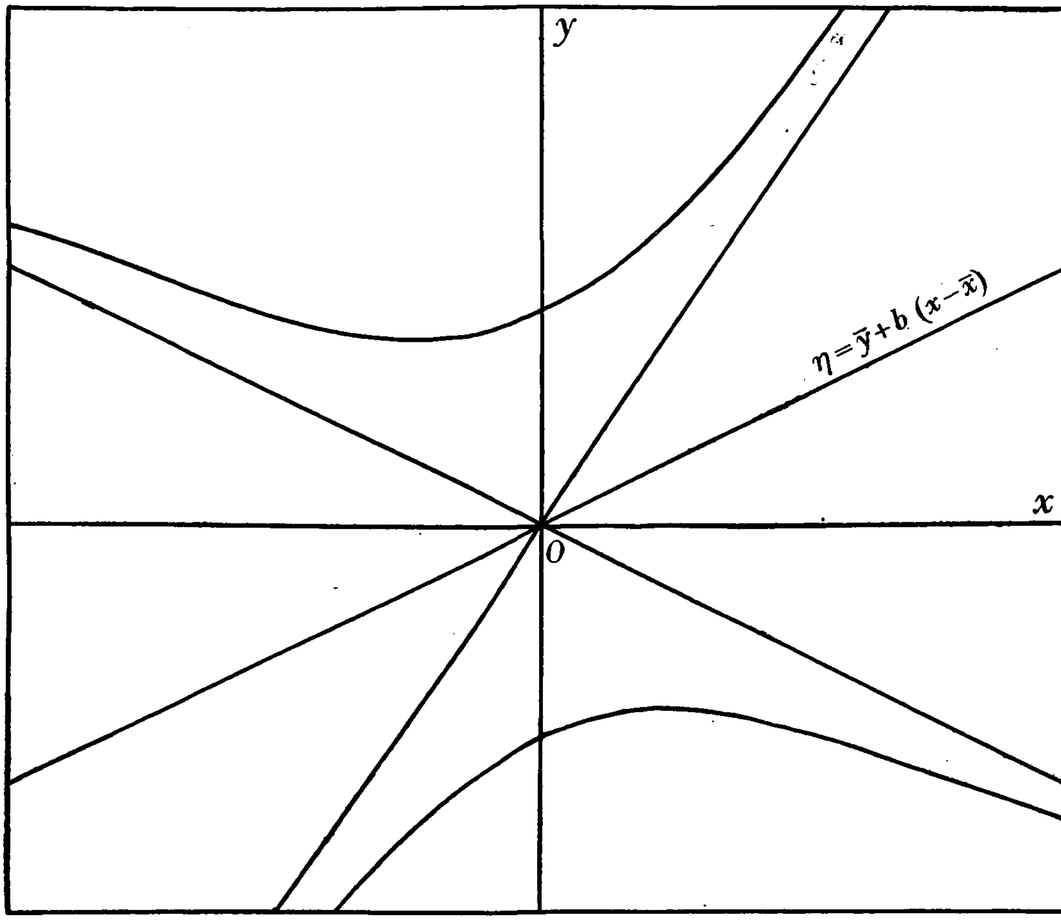

Fig. 2. Here the slope is positive but not significantly different from zero. O is the point $\bar{x}$, $\bar{y}$ where $x=\log$. dose, $y=$response. The fiducial limits for x corresponding to any given Y and a given value of P (0·95 or 0·99 say) are infinite. Fiducial limits corresponding to 1–P (0·05 or 0·01 say) would be given by the points in which the appropriate abscissa cuts the hyperbola, but these are of no practical interest.

are *infinite*. The interval between the limits (5) gives us the fiducial range corresponding to probability p_1; this tells us that in a proportion p_1 of experiments (say once in twenty or once in a hundred times) the $\bar{x}$ corresponding to given Y will lie between these limits.

The limits are real only if $$(Y-\bar{y})^2 \geqslant \frac{A}{B}(t^2B-b^2). \quad (6)$$

The lower asymptote (4) has a negative slope, if b is positive and if (6) is satisfied the horizontal line through Y cuts the same branch of the hyperbola twice. Fig. 2 illustrates the position.

https://doi.org/10.1017/S0022172400012705 Published online by Cambridge University Press

If $(Y-\bar{y})^2=\frac{A}{B}(t^2B-b^2)$, the horizontal line through Y is a tangent to the hyperbola and the fiducial limits corresponding to p_1 (0·05 or 0·01 say) happen to coincide in the particular experiment. Of course they would not coincide in all repetitions of the experiment. The important conclusion is that the fiducial limits corresponding to $(1-p_1)$ (0·95 or 0·99) are infinite. The limits corresponding to p_1 are only of theoretical interest.

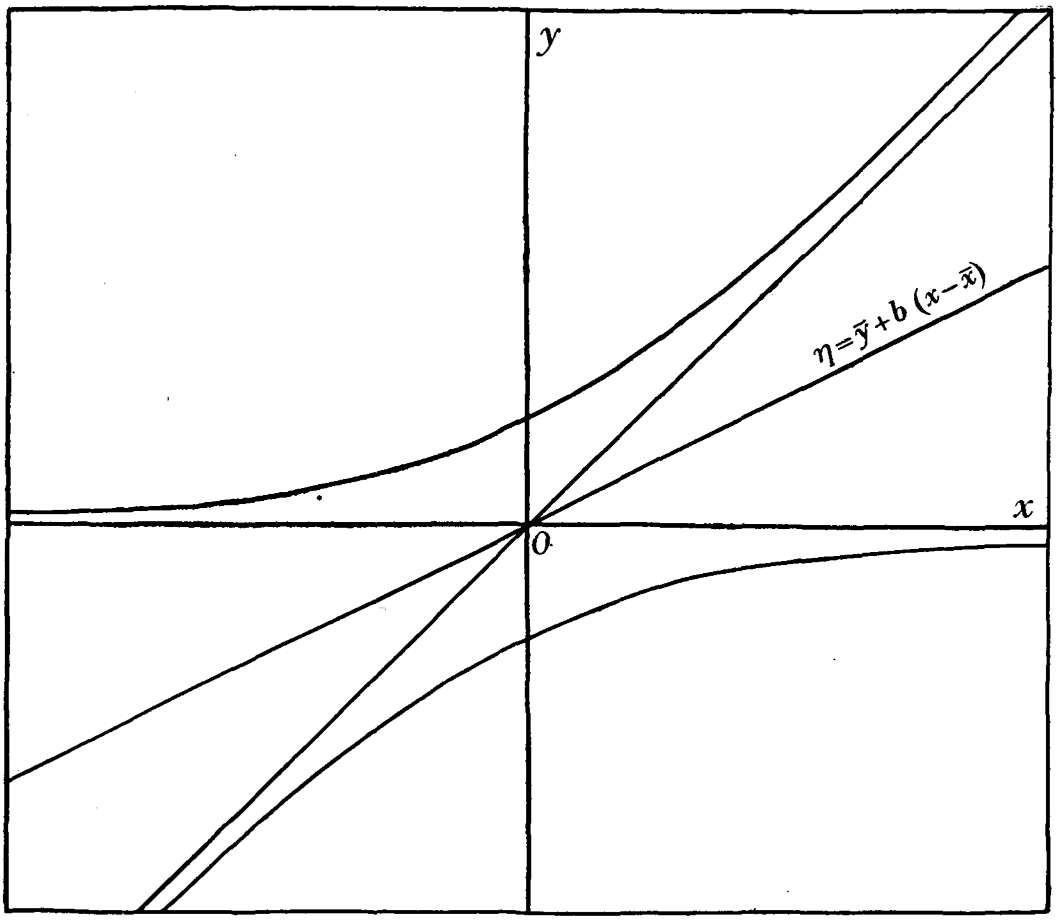

Fig. 3. Here the slope just reaches the significance level. O is the point $\bar{x}$, $\bar{y}$ where $x=$log.dose, $y=$response. The fiducial limits for x corresponding to any given Y and a given value of P (0·95 or 0·99 say) are the points in which the corresponding abscissa cuts the hyperbola. Here one fiducial limit is finite and the other infinite unless $Y=\bar{y}$ when both are infinite.

(iii) If $b^2=t^2B$, one asymptote is horizontal and we find

$$x>\left(\bar{x}+\frac{Y-\bar{y}}{2b}-\frac{t^2A}{2b\,(Y-\bar{y})}\right)\quad \text{if}\quad bY>b\bar{y},$$

$$x<\left(\bar{x}+\frac{Y-\bar{y}}{2b}-\frac{t^2A}{2b\,(Y-\bar{y})}\right)\quad \text{if}\quad bY<b\bar{y}.$$

If $bY>b\bar{y}$, there is a finite lower fiducial limit and the other limit is infinite, if $bY<b\bar{y}$ there is a finite upper fiducial limit and the other is infinite. If $Y=\bar{y}$ both limits are infinite. Fig. 3 illustrates the position, for $b>0$.

https://doi.org/10.1017/S0022172400012705 Published online by Cambridge University Press

The answer to our original problem is that there are always finite real fiducial limits if the slope is significant, otherwise they are infinite except in the critical case where the slope is just significant, in which case at least one limit is infinite.

4. Fiducial limits for the result of an assay

When we are fitting two parallel straight lines as in comparing a test and standard preparation, suppose we have for the observed regression lines

Standard preparation $$\eta_1 = \bar{y}_1 + b\,(x - \bar{x}_1),$$

Test preparation $$\eta_2 = \bar{y}_2 + b\,(x - \bar{x}_2),$$

Then $$\eta_1 - \eta_2 = (\bar{y}_1 - \bar{y}_2) + b\,\{(x_1 - \bar{x}_1) - (x_2 - \bar{x}_2)\}. \tag{8}$$

When $\eta_1 - \eta_2 = 0$, $$x_1 - x_2 = \bar{x}_1 - \bar{x}_2 - \frac{(\bar{y}_1 - \bar{y}_2)}{b}, \tag{9}$$

which is our estimate of the relative potency of doses of test and standard bearing an assigned ratio to one another. Fiducial limits for $x_1 - x_2$ are required.

Write $\eta_1 - \eta_2 = \eta$, $x_1 - x_2 = x$, $\bar{y}_1 - \bar{y}_2 = \bar{y}$, and let Y be the true value of η corresponding to given x. Fiducial limits for x corresponding to given Y, as a rule to $Y = 0$, are required. Since $$\eta = \bar{y} + b\,(x - \bar{x}), \tag{10}$$

the sampling variance of η is given by $$V(\eta) = A + B\,(x - \bar{x})^2, \tag{11}$$

where $$A = s^2 \left\{\frac{1}{S_1\,(w)} + \frac{1}{S_2\,(w)}\right\}, \quad B = s^2/[S_1\,\{w\,(x - \bar{x})^2\} + S_2\,\{w\,(x - \bar{x})^2\}].$$

Here s^2 and w have the same meaning as before and the summation S_1 is over the doses of the standard and S_2 over the doses of the test preparation.

Now $\dfrac{\eta - Y}{\sqrt{(V(\eta))}}$ is distributed as 't' with ν degrees of freedom (or normally in the quantal case).

Hence in a proportion $(1 - p_1)$ of assays, if the slope is significant $$\bar{y} + b\,(x - \bar{x}) - t\,\{A + B\,(x - \bar{x})^2\}^{\frac{1}{2}} < Y < \bar{y} + b\,(x - \bar{x}) + t\,\{A + B\,(x - \bar{x})^2\}^{\frac{1}{2}}, \tag{12}$$

whence $$\left[\bar{x} + \frac{b\,(Y - \bar{y})}{b^2 - t^2 B} - \frac{t}{(b^2 - t^2 B)}\,\{A\,(b^2 - t^2 B) + B\,(Y - \bar{y})^2\}^{\frac{1}{2}}\right]$$
$$< x < \left[\bar{x} + \frac{b\,(Y - \bar{y})}{b^2 - t^2 B} + \frac{t}{(b^2 - t^2 B)}\,\{A\,(b^2 - t^2 B) + B\,(Y - \bar{y})^2\}^{\frac{1}{2}}\right]. \tag{13}$$

When $Y = 0$ these limits become

$$\bar{x}_1 - \bar{x}_2 - \frac{b\,(\bar{y}_1 - \bar{y}_2)}{b^2 - t^2 B} - \frac{t}{(b^2 - t^2 B)}\,\{A\,(b^2 - t^2 B) + B\,(\bar{y}_2 - \bar{y}_1)^2\}^{\frac{1}{2}},$$

and $$\bar{x}_1 - \bar{x}_2 - \frac{b\,(\bar{y}_1 - \bar{y}_2)}{b^2 - t^2 B} + \frac{t}{(b^2 - t^2 B)}\,\{A\,(b^2 - t^2 B) + B\,(\bar{y}_2 - \bar{y}_1)^2\}^{\frac{1}{2}} \tag{14}$$

The various cases which arise can be distinguished as before, and we reach the conclusion that provided the slope is significant and consequently $b^2 > t^2 B$, there are finite real fiducial limits given by (14).

5. Relation between the exact fiducial limits and the approximate formula

The limits (14) can be written

$$\bar{x}_1-\bar{x}_2+\frac{\bar{y}_2-\bar{y}_1}{b}+\frac{t^2B\,(\bar{y}_2-\bar{y}_1)}{b\,(b^2-t^2B)}$$
$$\pm t\left[\frac{A}{b^2}+\frac{B\,(\bar{y}_2-\bar{y}_1)^2}{b^4}+\frac{At^2B}{b^2(b^2-t^2B)}+B\,(\bar{y}_2-\bar{y}_1)^2\left\{\frac{1}{(b^2-t^2B)^2}-\frac{1}{b^4}\right\}\right]^{\frac{1}{2}}, \quad (14')$$

and the limits (2) can be put in a similar form.

Here the relation between the exact and approximate formulae is clearly shown, for the first two members in each term give the limits hitherto used. It may be noted that if $\bar{y}_2>\bar{y}_1$ and $b>0$ the approximate formula places the middle of the 'fiducial range' too low, if $\bar{y}_2<\bar{y}_1$, too high. The same applies to formula (2) according as $Y\gtrless\bar{y}$. Thus the approximate formula, in addition to underestimating the width of the fiducial range, may bias its position.

Fieller (1940) writes $$C=\frac{b^2}{b^2-t^2B},$$

and obtains (14′) in the form

$$\bar{x}_1-\bar{x}_2-\frac{C\,(\bar{y}_1-\bar{y}_2)}{b}\pm t\sqrt{(C)}\left[\frac{A}{b^2}+\frac{CB\,(\bar{y}_1-\bar{y}_2)^2}{b^4}\right]^{\frac{1}{2}} \quad (14'')$$

This is more elegant than (14′) and shows that the approximate formula replaces C by unity. In B.S.I. specification No. 911 it is suggested that at the 5 % level the exact formula should always be used unless $(b^2/B)>60$ or the slope is greater than 7·75 times its standard error. This makes $C<1{\cdot}07$, if $t=2{\cdot}0$. If this criterion is to be applied it is little further trouble to calculate C and use the exact formula, and this seems the best thing to do.

Nevertheless the approximate sampling variance of the logarithm of the result

$$\left[\frac{A}{b^2}+\frac{B\,(\bar{y}_2-\bar{y}_1)^2}{b^4}\right]$$

still has one use of some importance. When we wish to combine the results of several assays of the same substance carried out by the same method, the best procedure available is to weight the logarithms of the results with the reciprocals of their sampling variances to obtain a weighted mean. The approximate sampling variance of the weighted mean will be the reciprocal of the sum of these weights.

6. A numerical comparison

Table 1 gives a comparison of the exact and approximate fiducial limits for eleven determinations in different laboratories of the median-fertility dose of vitamin E (dl-α-tocopheryl acetate). Four doses were used and the response was the percentage of positively mated female rats which produced a litter. The results ranged from 0·55 to 1·71 mg. the mean being 1·00 mg. (E. M. Hume *et al.* 1941). In laboratories 3, 4 and 5 there is a very material difference between the approximate and exact results. In laboratories 4 and 5 the slope (b) is only $2\frac{3}{4}$ times its standard error and the approximate formula consequently underestimates the fiducial range particularly at the ($P=0{\cdot}99$) level. In both cases $\bar{y}>5$

consequently the lower limit is too high in the approximate formula. This is also the case in laboratory (3) where, although the slope is moderately well determined, there is considerable bias in $\bar{y}$. It is interesting to note the contrast between laboratories (3) and (7) where the bias in $\bar{y}$ is the same, but the slope in the latter case is so well determined that the bias has practically no effect. It is clearly desirable to use sufficient animals to get a well determined slope if accurate results are to be obtained.

I have many similar results for biological assays of vitamin A by the rat-growth method, but those will shortly be published in a separate report.

Table 1. *Comparison of exact and approximate fiducial limits for the median fertility dose of vitamin E*

			Fiducial limits %			
			Approx.		Exact	
Laboratory	Mean probit $\bar{y}$	b/S.E. of b	$P=0{\cdot}95$	$P=0{\cdot}99$	$P=0{\cdot}95$	$P=0{\cdot}99$
1	5·32	4·74	86–117	82–122	83–116	76–122
2a	4·91	3·76	82–123	77–131	80–128	71–148
2b	5·03	4·05	85–118	80–125	83–121	75–131
3	5·75	3·64	72–139	65–155	53–128	29–136 !
4	5·71	2·74	58–172	49–204	19–143	0–155 !
5	5·35	2·76	85–117	81–123	72–117	24–127 !
6a	5·13	5·17	88–114	85–118	88–116	83–122
6b	5·27	3·90	82–123	77–131	76–122	63–131
7	5·77	5·36	87–116	83–121	85–116	80–123
8	4·89	4·03	84–119	80–125	83–123	77–137
9	5·06	3·98	85–117	81–123	83–119	75–129

REFERENCES

BLISS, C. I. (1935). *Ann. Appl. Biol.* **22**, 325.

EISENHART, C. (1939). *Ann. Math. Statist.* **10**, 162.

FIELLER, E. C. (1940). *J. Roy. Statist. Soc.* Suppl. **7**, 49.

British Standard Method for the biological assay of vitamin D_3 by the chick method (1940). B.S.I. Specification 911.

HUME, E. M. *et al.* (1941). *Nature, Lond.*, 18 Oct.

(*MS. received for publication* 30. x. 42.—Ed.)

https://doi.org/10.1017/S0022172400012705 Published online by Cambridge University Press