THE ESTIMATION OF BACTERIAL DENSITIES FROM DILUTION SERIES

By D. J. FINNEY
Lecturer in the Design and Analysis of Scientific Experiment, University of Oxford

(1) INTRODUCTION

The dilution technique for estimating the density of a bacterial suspension involves taking samples of constant size from each of several dilutions of the suspension, incubating these under standard conditions, and recording presence or absence of bacterial growth for each sample. From the numbers of sterile samples at each dilution an estimate of the number of organisms per unit volume of the original suspension is formed and its precision is assessed. That precision will be much lower than for an estimate based on direct bacterial counts (e.g. by colony counts) in the same samples, but the dilution technique is much simpler and can be used in circumstances that make counts impracticable. Many methods of statistical analysis have been proposed for dilution series; these were excellently reviewed by Eisenhart & Wilson (1943) and by Cochran (1950).

The simplest method is undoubtedly that suggested by Fisher (1922), on which is based Table $VIII_2$ of Fisher & Yates's collection (1948). This method is exceedingly ingenious, and of remarkably high efficiency, but it cannot always be applied and the conditions for its applicability may sometimes be opposed to the best interests of experimental design. Halvorson & Ziegler (1933) suggested the use of the principle of maximum likelihood for estimating the density, and suggestions for the systematic computation of the maximum likelihood estimate have been made by Barkworth & Irwin (1938) and Finney (1947). Mather's (1949) solution to a closely related problem leads to an ideal method of computing the maximum likelihood estimate for a dilution series. This paper is written to show the application of Mather's 'loglog' transformation and to compare the new proposals with Fisher's method.

(2) THE LOGLOG TRANSFORMATION

For convenience of notation the unit of volume may be taken as the size of the sample used in the dilution series. Suppose that, in the original bacterial suspension, the mean density of organisms is μ per unit volume. In a suspension diluted by a factor z (in general $z \leqslant 1$, but the statistical theory remains applicable with $z > 1$) the density will be μz. Two important assumptions must now be made if any statistical estimation procedure is to be valid; these are:

(i) The organisms are distributed entirely at random in the original suspension and in all dilutions used.

(ii) The nature of the culture medium and incubation are such as to ensure visible growth for every sample containing one or more organisms.

https://doi.org/10.1017/S0022172400015333 Published online by Cambridge University Press

By virtue of the first assumption the number of organisms per sample at dilution z follows a Poisson distribution with mean μz. The probability of a sterile sample is, by the second assumption, the probability that a sample contains no organisms; this is the first term of the Poisson series,

$$P = e^{-\mu z}. \tag{1}$$

The general form of experiment will consist of tests of n_i samples at a dilution $z_i\,(i = 1, 2, 3, \ldots, k)$, with records of r_i, the number sterile. The probability of the result observed for dilution z_i is

$$\text{Prob.}\,(r_i) = \frac{n_i!}{r_i!\,(n_i - r_i)!} P_i^{r_i} (1 - P_i)^{n_i - r_i}, \tag{2}$$

where P_i is defined by equation (1). Moreover

$$p_i = r_i/n_i \tag{3}$$

is an estimate of P_i from the data for the corresponding dilution. Write now

$$x = \log_e z, \tag{4}$$

and, following Mather (1949),

$$Y = \log_e(-\log_e P); \tag{5}$$

Y is termed the *loglog* of P. Natural logarithms are more convenient here than logarithms to base 10, and are easily read from Fisher & Yates's Table XXVI (1948). Table 1, an abbreviated form of a table given by Mather, shows the loglog function to sufficient accuracy for the present method. Equation (1) may now be expressed as

$$Y = \log_e \mu + x. \tag{6}$$

Consequently, the estimation of $\log_e \mu$ can be regarded as a form of regression calculation of the loglog of the proportion of sterile plates on x, subject to the restriction that the regression coefficient is unity.

Table 1. *The transformation of proportions to loglogs*

p	0·000	0·001	0·002	0·003	0·004	0·005	0·006	0·007	0·008	0·009
0·00	—	1·93	1·83	1·76	1·71	1·67	1·63	1·60	1·57	1·55
p	0·00	0·01	0·02	0·03	0·04	0·05	0·06	0·07	0·08	0·09
0·0	—	1·53	1·36	1·25	1·17	1·10	1·03	0·98	0·93	0·88
0·1	0·83	0·79	0·75	0·71	0·68	0·64	0·61	0·57	0·54	0·51
0·2	0·48	0·45	0·41	0·39	0·36	0·33	0·30	0·27	0·24	0·21
0·3	0·19	0·16	0·13	0·10	0·08	0·05	0·02	−0·01	−0·03	−0·06
0·4	−0·09	−0·11	−0·14	−0·17	−0·20	−0·23	−0·25	−0·28	−0·31	−0·34
0·5	−0·37	−0·40	−0·42	−0·45	−0·48	−0·51	−0·55	−0·58	−0·61	−0·64
0·6	−0·67	−0·70	−0·74	−0·77	−0·81	−0·84	−0·88	−0·92	−0·95	−0·99
0·7	−1·03	−1·07	−1·11	−1·16	−1·20	−1·25	−1·29	−1·34	−1·39	−1·45
0·8	−1·50	−1·56	−1·62	−1·68	−1·75	−1·82	−1·89	−1·97	−2·06	−2·15
0·9	−2·25	−2·36	−2·48	−2·62	−2·78	−2·97	−3·20	−3·49	−3·90	−4·60
p	0·000	0·001	0·002	0·003	0·004	0·005	0·006	0·007	0·008	0·009
0·97	−3·49	−3·53	−3·56	−3·60	−3·64	−3·68	−3·72	−3·76	−3·81	−3·85
0·98	−3·90	−3·95	−4·01	−4·07	−4·13	−4·19	−4·26	−4·34	−4·42	−4·50
0·99	−4·60	−4·71	−4·82	−4·96	−5·11	−5·30	−5·52	−5·81	−6·21	−6·91

https://doi.org/10.1017/S0022172400015333 Published online by Cambridge University Press

(3) THE METHOD OF CALCULATION

The calculations are not quite those of an ordinary regression, and are closely related to the standard procedure for probit analysis (Finney, 1951), as generalized to tolerance distributions of various types (Finney, 1949)—though the notion of a tolerance distribution is not relevant here. Mather has given most of the theory, and all that need be presented here is the set of rules for computation.

The first step is to tabulate z, n, r and p for each dilution. The *empirical loglog* of each p is then found from Table 1; if $p=0$ or $p=1$, the empirical loglog is infinite and must be omitted. A column of *expected loglogs*, Y, is then added. This is constructed by guessing the density of organisms in the original suspension (or in any of the dilutions), calculating the densities for every dilution to correspond to this guessed value, and writing the Y values as the natural logarithms of these. The differences between successive values of Y will then be equal to the differences between successive values of x (in general, one decimal place in Y suffices). The agreement between values of Y and the empirical loglogs is a check that the guessed density is near the truth. Though a good choice for Y will expedite the calculations, it does not affect the final estimate, and therefore no more objective rule need be given.

Two additional columns, nw and η, are next completed. The *weight*, nw, is the product of n and a *weighting coefficient*, w, defined by

$$w = \frac{e^{2Y}}{e^{e^Y}-1}. \tag{7}$$

The *working deviate*, η, is obtained from any one of the formulae

$$\eta = \eta_0 + p(\eta_1 - \eta_0), \tag{8}$$

$$\eta = \eta_1 - (1-p)(\eta_1 - \eta_0), \tag{9}$$

$$\eta = (1-p)\eta_0 + p\eta_1, \tag{10}$$

where η_0 is defined by

$$\eta_0 = -e^{-Y}, \tag{11}$$

and is known as the *minimum working deviate*; η_1 is defined by

$$\eta_1 = e^{(e^Y-1)} - e^{-Y}, \tag{12}$$

and is known as the *maximum working deviate*, and $(\eta_1 - \eta_0)$ is known as the *range*. Table 2* shows the minimum and maximum working deviates, the range, and the weighting coefficient as functions of Y.

The next step is to find the weighted mean deviate, $\bar{\eta}$, as

$$\bar{\eta} = Snw\eta / Snw. \tag{14}$$

This quantity is subtracted from the logarithm of the guessed density that was used in initiating the calculations, in order to give a revised value for the density at that dilution. A new column of expected loglogs is then based on the revision,

* In his analysis, Mather used a *working loglog*, y, defined by

$$y = Y - \eta, \tag{13}$$

but η is more convenient here. Mather's table corresponding to Table 2 is inaccurate at extremes of Y, perhaps because it was calculated from an inadequate table of logarithms. Table 2 has therefore been based on new and independent calculation from the New York Works Projects Administration tables of the exponential function (1939).

https://doi.org/10.1017/S0022172400015333 Published online by Cambridge University Press

Table 2. *Minimum and maximum working deviates, range, and weighting coefficients*

Y	η_0	η_1	Range	w
2·5	−0·0821	—	16,034·	0·00076
2·4	−0·0907	—	5,559·0	0·00198
2·3	−0·1003	—	2,152·1	0·00463
2·2	−0·1108	—	920·59	0·00980
2·1	−0·1225	—	431·03	0·01895
2·0	−0·1353	—	219·00	0·03376
1·9	−0·1496	—	119·81	0·05587
1·8	−0·1653	69·915	70·080	0·08653
1·7	−0·1827	43·369	43·552	0·12622
1·6	−0·2019	28·387	28·589	0·17448
1·5	−0·2231	19·498	19·721	0·22985
1·4	−0·2466	13·981	14·228	0·29005
1·3	−0·2725	10·417	10·689	0·35223
1·2	−0·3012	8·0309	8·3321	0·41342
1·1	−0·3329	6·3809	6·7138	0·47080
1·0	−0·3679	5·2071	5·5750	0·52204
0·9	−0·4066	4·3097	4·7163	0·57071
0·8	−0·4493	3·7108	4·1601	0·59975
0·7	−0·4966	3·2235	3·7201	0·62471
0·6	−0·5488	2·8456	3·3944	0·64034
0·5	−0·6065	2·5476	3·1541	0·64716
0·4	−0·6703	2·3094	2·9797	0·64598
0·3	−0·7408	2·1164	2·8572	0·63780
0·2	−0·8187	1·9584	2·7771	0·62369
0·1	−0·9048	1·8275	2·7323	0·60473
0·0	−1·0000	1·7183	2·7183	0·58198
−0·1	−1·1052	1·6263	2·7315	0·55638
−0·2	−1·2214	1·5483	2·7697	0·52880
−0·3	−1·3499	1·4817	2·8316	0·49999
−0·4	−1·4918	1·4245	2·9163	0·47057
−0·5	−1·6487	1·3751	3·0238	0·44107
−0·6	−1·8221	1·3323	3·1544	0·41192
−0·7	−2·0138	1·2950	3·3088	0·38345
−0·8	−2·2255	1·2625	3·4880	0·35592
−0·9	−2·4596	1·2339	3·6935	0·32951
−1·0	−2·7183	1·2087	3·9270	0·30435
−1·1	−3·0042	1·1865	4·1907	0·28054
−1·2	−3·3201	1·1669	4·4870	0·25811
−1·3	−3·6693	1·1495	4·8188	0·23708
−1·4	−4·0552	1·1341	5·1893	0·21744
−1·5	−4·4817	1·1203	5·6020	0·19916
−1·6	−4·9530	1·1081	6·0611	0·18220
−1·7	−5·4739	1·0972	6·5711	0·16650
−1·8	−6·0496	1·0874	7·1370	0·15201
−1·9	−6·6859	1·0787	7·7646	0·13866
−2·0	−7·3891	1·0708	8·4599	0·12638
−2·1	−8·1662	1·0638	9·2300	0·11511
−2·2	−9·0250	1·0575	10·083	0·10478
−2·3	−9·9742	1·0518	11·026	0·09532
−2·4	−11·023	1·0468	12·070	0·08667
−2·5	−12·182	1·0422	13·224	0·07876
−2·6	−13·464	1·0381	14·502	0·07155
−2·7	−14·880	1·0344	15·914	0·06497
−2·8	−16·445	1·0310	17·476	0·05898
−2·9	−18·174	1·0280	19·202	0·05352
−3·0	−20·086	1·0253	21·111	0·04856

https://doi.org/10.1017/S0022172400015333 Published online by Cambridge University Press

Table 2 (*continued*)

Y	η_0	η_1	Range	w
−3·1	−22·198	1·0229	23·221	0·04404
−3·2	−24·533	1·0207	25·554	0·03994
−3·3	−27·113	1·0187	28·132	0·03621
−3·4	−29·964	1·0169	30·981	0·03282
−3·5	−33·115	1·0153	34·130	0·02974
−3·6	−36·598	1·0138	37·612	0·02695
−3·7	−40·447	1·0125	41·459	0·02442
−3·8	−44·701	1·0113	45·712	0·02212
−3·9	−49·402	1·0102	50·412	0·02004
−4·0	−54·598	1·0092	55·607	0·01815
−4·1	−60·340	1·0083	61·348	0·01644
−4·2	−66·686	1·0075	67·694	0·01488
−4·3	−73·700	1·0068	74·707	0·01348
−4·4	−81·451	1·0062	82·457	0·01220
−4·5	−90·017	1·0056	91·023	0·01105
−4·6	−99·484	1·0050	100·49	0·01000
−4·7	—	1·0046	110·95	0·00905
−4·8	—	1·0041	122·51	0·00820
−4·9	—	1·0037	135·29	0·00742
−5·0	—	1·0034	149·41	0·00672
−5·1	—	1·0031	165·02	0·00608
−5·2	—	1·0028	182·27	0·00550
−5·3	—	1·0025	201·34	0·00498
−5·4	—	1·0023	222·41	0·00451
−5·5	—	1·0020	245·69	0·00408
−5·6	—	1·0018	271·43	0·00369
−5·7	—	1·0017	299·87	0·00334
−5·8	—	1·0015	331·30	0·00302
−5·9	—	1·0014	366·04	0·00274
−6·0	—	1·0012	404·43	0·00248
−6·1	—	1·0011	446·86	0·00224
−6·2	—	1·0010	493·75	0·00203
−6·3	—	1·0009	545·57	0·00183
−6·4	—	1·0008	602·85	0·00166
−6·5	—	1·0008	666·14	0·00150
−6·6	—	1·0007	736·10	0·00136
−6·7	—	1·0006	813·41	0·00123
−6·8	—	1·0006	898·85	0·00111
−6·9	—	1·0005	993·28	0·00101
−7·0	—	1·0005	1,097·63	0·00091
−7·1	—	1·0004	1,213·0	0·00082
−7·2	—	1·0004	1,340·0	0·00075
−7·3	—	1·0003	1,481·3	0·00068
−7·4	—	1·0003	1,637·0	0·00061
−7·5	—	1·0003	1,809·0	0·00055
−7·6	—	1·0003	1,999·2	0·00050
−7·7	—	1·0002	2,209·3	0·00045
−7·8	—	1·0002	2,441·6	0·00041
−7·9	—	1·0002	2,698·3	0·00037
−8·0	—	1·0002	2,982·0	0·00034
−8·1	—	1·0002	3,295·5	0·00030
−8·2	—	1·0001	3,641·9	0·00027
−8·3	—	1·0001	4,024·9	0·00025
−8·4	—	1·0001	4,448·1	0·00022
−8·5	—	1·0001	4,915·8	0·00020
−8·6	—	1·0001	5,432·7	0·00018
−8·7	—	1·0001	6,003·9	0·00017
−8·8	—	1·0001	6,635·2	0·00015
−8·9	—	1·0001	7,333·0	0·00014
−9·0	—	1·0001	8,104·1	0·00012

https://doi.org/10.1017/S0022172400015333 Published online by Cambridge University Press

and the whole cycle of computation is repeated; the process is iterated until the adjustment $\bar{\eta}$ is negligible.

The final version of the Y column, which may be calculated to additional places of decimals, gives maximum likelihood estimates of the densities at each dilution. In particular, the value of Y corresponding to $z = 1$ (whether or not samples were taken at this level) is, by equation (1), $\log_e m$, where m is the maximum likelihood estimate of μ. Moreover, the precision of the estimate may be expressed by the variance

$$V(\log_e m) = 1/Snw, \tag{15}$$

where the weights are from the last cycle computed. To a first approximation, fiducial limits to μ may be obtained by assuming a normal distribution of errors and the variance as given in equation (15)

(4) A NUMERICAL EXAMPLE

An example will show how very simple Tables 1 and 2 make this apparently complicated process. Fisher & Yates (1948, Example 5·1) reported a series of dilution tests for estimating the numbers of 'rope' spores (*Bacillus mesentericus*) in a potato flour. A suspension of the flour containing 0·04 g./c.c. was diluted by nine successive factors of 2, and five samples of 1 c.c. were withdrawn at each level. The samples were plated and incubated, and the resulting data on numbers of sterile plates are recorded in Table 3. The proportions sterile, p, and the empirical loglogs of p have been entered in the table. The general run of values in the empirical loglog column suggests that an expected value of $-0{\cdot}1$ for $z = 1/32$, corresponding to a density

$$e^{-0{\cdot}1} = 0{\cdot}91$$

at this dilution, would be about correct. The densities at other dilutions may be calculated by considering their ratios to 1/32, and their logarithms are the entries for Y. Since the ratio of successive dilutions is always 2, and $\log_e 2 = 0{\cdot}69$, the Y column is most easily constructed by additions or subtractions of 0·69 to the value $-0{\cdot}1$.

Table 3. *A dilution series for the estimation of the density of rope spores in a potato flour*

Dilution (z) of original suspension of 0·04 g./c.c.	n	r	p	Empirical loglog	First cycle			Second cycle		
					Y	w	η	Y	w	η
1	5	0	0·0	—	3·4	0·000	—	3·4	0·000	—
1/2	5	0	0·0	—	2·7	0·000	—	2·7	0·000	—
1/4	5	0	0·0	—	2·0	0·034	−0·135	2·0	0·034	−0·135
1/8	5	0	0·0	—	1·3	0·352	−0·272	1·3	0·352	−0·272
1/16	5	1	0·2	0·48	0·6	0·640	0·130	0·6	0·640	0·130
1/32	5	2	0·4	−0·09	−0·1	0·556	−0·013	0·0	0·582	0·087
1/64	5	3	0·6	−0·67	−0·8	0·356	−0·133	−0·7	0·383	−0·029
1/128	5	3	0·6	−0·67	−1·5	0·199	−1·120	−1·4	0·217	−0·942
1/256	5	5	1·0	—	−2·2	0·105	1·058	−2·1	0·115	1·064
1/512	5	5	1·0	—	−2·9	0·054	1·028	−2·8	0·059	1·031

https://doi.org/10.1017/S0022172400015333 Published online by Cambridge University Press

When each dilution has the same n, the calculations can be performed with a weight w, instead of nw; the factor n must be introduced at the end, before the variance is evaluated. In Table 3, the w column has been filled by direct entry of weighting coefficients from Table 2. The working deviates were found from equation (8). The first two spaces in the column can be left blank, because the associated weights are negligible. For $z = 1/32$, for example,

$$\eta = -1{\cdot}1052 + 0{\cdot}4 \times 2{\cdot}7315$$
$$= -0{\cdot}013.$$

By summations of w, and of the products $w\eta$,

$$Sw = 2{\cdot}296, \quad Sw\eta = -0{\cdot}1280.$$

Therefore

$$\bar{\eta} = -\frac{0{\cdot}1280}{2{\cdot}296}$$
$$= -0{\cdot}0557.$$

Subtraction of this from the provisional $Y = -0{\cdot}1$ at $z = 1/32$ gave a revised expected loglog of $-0{\cdot}044$ at this dilution. A second cycle of calculations was then performed with a new Y column based on this value but rounded to the nearest 0·1 (interpolation for more exact arithmetical working will seldom be necessary). From the second cycle, details of which are shown in Table 3,

$$Sw = 2{\cdot}382, \quad Sw\eta = 0{\cdot}0012,$$

and therefore

$$\bar{\eta} = 0{\cdot}0005.$$

The further revision of the expected loglog at $z = 1/32$ is $Y = -0{\cdot}0445$. The last adjustment is so small that no more cycles need be computed, as will be apparent from the use to which the calculations are now put.

If μ is the mean number of spores per unit sample (i.e. per c.c.) of the original suspension, the estimate of μ is given by

$$\log_e (m/32) = -0{\cdot}0445,$$

or

$$m = 30{\cdot}6.$$

Also, since

$$Snw = 5 \times Sw$$
$$= 11{\cdot}91$$
$$= 1/(0{\cdot}0290)^2,$$

the standard error to be attached to $\log_e (m/32)$ is 0·290. For 5 % fiducial limits, this standard error should be multiplied by 1·960 and subtracted from or added to $\log_e (m/32)$; the limits are $-0{\cdot}6129$, 0·5239. Hence

$$\text{Lower fiducial limit to } \mu = 32\, e^{-0{\cdot}6129}$$
$$= 17{\cdot}3,$$

and

$$\text{Upper fiducial limit to } \mu = 32\, e^{0{\cdot}5239}$$
$$= 54{\cdot}0.$$

The suspension contained 1 g. of flour per 25 c.c.; the flour is therefore estimated to contain 765 'rope' spores per g., and may be confidently asserted to have not less than 432 nor more than 1350.

https://doi.org/10.1017/S0022172400015333 Published online by Cambridge University Press

(5) COMPARISON WITH FISHER'S METHOD

A common procedure in estimation by a dilution series is to test equal numbers of samples at each of many dilutions arranged in a geometrical progression; dilution factors of 2, 4 and 10 are the favourites. The series is made long enough to ensure that at one extreme the samples are almost always fertile, and at the other they are almost always sterile. Fisher (1922) pointed out that, for such an experiment, a highly efficient estimate is obtainable by equating the total number of sterile plates to its expectation. If z_0 is the first dilution tested, and a the dilution factor, this involves writing

$$n\,(e^{-\mu z_0} + e^{-\mu z_0/a} + e^{-\mu z_0/a^2} + \ldots) = S\,(r), \tag{16}$$

and taking the solution of this equation as the estimate of μ. The method is 87·7 % efficient, in the sense that the maximum likelihood estimate has a variance which is 87·7 % of the variance of the solution of equation (16). Strictly, this statement applies to the limit of the ratio of the two variances as n is made large, and the theory of what happens for small n (or even of what magnitude of n may be considered 'large') has not been studied. Fisher's method is thus equivalent to a discarding of 12·3 % of the data. Its advantage is that, with the aid of special tables, equation (16) can be very easily solved. Fisher & Yates (1948, Table $VIII_2$) have given such tables for $a = 2$, 4, 10, and have described how the precision of the estimate may be assessed. For the data used in Table 3, they obtained an estimate of 760 spores per g., with limits at 407 and 1440 spores per g., a range about 13 % wider than that in §4.

When circumstances require that a long series of dilutions be tested, the decision between Fisher's method and maximum likelihood estimation is primarily one of economics. If the same number of samples is tested at each dilution, use of the Fisher and Yates table with eight samples per dilution will give the same precision as maximum likelihood estimation with seven per dilution. Though Mather's loglog transformation makes the maximum likelihood calculations less laborious than before, the experimental technique might be so simple as to make the use of extra samples the easier way of gaining precision. Nothing in this paper is intended to imply that Fisher's method is now out-moded, for it must always retain certain advantages of speed.

Data from a dilution series experiment, however, are sometimes not suitable for application of Fisher's method. This may happen either by accident or by design. If the total span of dilutions does not extend from the almost certainly fertile to the almost certainly sterile, the theory breaks down. Even in Table 3, the lowest dilutions are scarcely low enough for theoretical justification of the method (as is shown by the fact that w is by no means negligible in the last line of the table), and another two dilutions at least are desirable in order to make use of Fisher & Yates's table valid. Moreover, though Fisher & Yates have explained how their table may be applied when not all the n_i are equal, much of the simplicity of the method is then lost; its use for an experiment with different numbers of samples at each dilution would be difficult to justify theoretically and laborious to compute. Again, the construction of tables for the solution of the generalized form of equation (16)

required when successive dilution factors are not all equal is scarcely likely to be attempted.

Some of these criticisms may seem trivial, but they have important bearings on experimental design. If nothing is known about the value of μ at the time the experiment is planned, a long dilution series of the kind to which Fisher's method may be applied is perhaps ideal. On the other hand, if the experimenter has some idea of the density in his suspension, he will employ his efforts most profitably in tests of samples containing about 1·6 organisms each: at this level, for which the probability of a sterile sample is about 0·20, w is a maximum. For example, if the experimenter had guessed the density of 'rope' spores in the suspension discussed in §4 to be about 40 per c.c., he might have restricted his attention to the dilutions $z = 1/8, 1/16, 1/32, 1/64, 1/128$, in which there would be about 5, 2·5, 1·2, 0·6, 0·3 spores per c.c. With the same amount of labour as in the actual experiment, he could have taken 10 samples from each of these five dilutions. The second cycle in Table 3 indicates that he might then have found

$$\begin{aligned} Snw &= 10 \times (0{\cdot}352 + 0{\cdot}640 + 0{\cdot}582 + 0{\cdot}383 + 0{\cdot}217) \\ &= 21{\cdot}74, \end{aligned}$$

so that his estimate of μ would have been almost twice as precise as that in fact obtained. Had he had still more confidence in his guess, and assigned 5, 10, 20, 10, 5 samples to the five dilutions, Snw would have been further increased to 24·72. In such circumstances adoption of an experimental design that will permit the use of Fisher's method is clearly far from efficient.

As in most fields of research the present state of knowledge about the subject of experimentation determines the best type of experimental design. To choose a design because of the simplicity of the statistical analysis required, rather than because of its efficiency for the job in hand, is seldom wise. Equally undesirable, however, is a choice based upon excessive confidence in unreliable information. For example, if the experimenter investigating the numbers of 'rope' spores had guessed the density to be about 4 spores per c.c., he might have taken 10 samples at each of the dilutions $z = 1, 1/2, 1/4, 1/8, 1/16$; Snw would then have been only 10·26, and his precision would have been rather less than that resulting from use of Fisher's method with the original design. A guess of 1 spore per c.c., or of 1000 spores per c.c. might have had disastrous consequences, unless the experimenter had appreciated its unreliability and adopted a wide span of dilutions.

(6) SUMMARY

A transformation devised by Mather is applied to give a new scheme for computing the maximum likelihood estimate of a bacterial density from the evidence of a dilution series. Tables are provided to expedite the application of the method; with their aid, the calculations take a form similar to, but much simpler than, those for probit analysis of quantal responses.

Maximum likelihood estimation is compared with the method proposed by Fisher. The latter has the advantage of extreme simplicity, at least for series to which existing tables can be applied, and the loss of information involved in its

use may often be compensated by the saving of time in calculation. An experimenter who has fairly reliable prior indications of an approximate value for the density, however, ought to concentrate his attention on dilutions that he believes will contain between 4 and 1/4 organisms per sample; he must not apply Fisher's method to his results, but the maximum likelihood estimate will be so much more precise than any estimate from an experimental design using more extreme dilutions as to repay the small additional labour in computation.

I am indebted to Miss M. Callow and Miss H. M. Wild for the computation of Table 2.

REFERENCES

BARKWORTH, H. & IRWIN, J. O. (1938). Distribution of coliform organisms in milk and the accuracy of the presumptive coliform test. *J. Hyg., Camb.*, **38**, 446–57.

COCHRAN, W. G. (1950). Estimation of bacterial densities by means of the 'most probable number'. *Biometrics*, **6**, 105–16.

EISENHART, C. & WILSON, P. W. (1943). Statistical method and control in bacteriology. *Bact. Rev.* **7**, 57–137.

FINNEY, D. J. (1947). The principles of biological assay. *J.R. statist. Soc.* (Suppl.), **9**, 46–91.

FINNEY, D. J. (1949). The estimation of the parameters of tolerance distributions. *Biometrika*, **36**, 239–56.

FINNEY, D. J. (1951). *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve*, 2nd ed. London: Cambridge University Press.

FISHER, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philos. Trans.* A, **222**, 309–68.

FISHER, R. A. & YATES, F. (1948). *Statistical Tables for Biological, Agricultural and Medical Research*, 3rd ed. Edinburgh: Oliver and Boyd.

HALVORSON, H. O. & ZIEGLER, N. R. (1933). Application of statistics to problems in bacteriology. I. A means of determining bacterial population by the dilution method. *J. Bact.* **25**, 101–21.

MATHER, K. (1949). The analysis of extinction time data in bioassay. *Biometrics*, **5**, 127–43.

NEW YORK WORKS PROJECT ADMINISTRATION (1939). *Tables of the Exponential Function.* New York: Federal Works Agency.

(*MS. received for publication* 6. x. 1950.)