THE ESTIMATION OF THE PARAMETERS OF AN EXPONENTIALLY DECLINING POPULATION

By C. C. SPICER

Central Public Health Laboratory, Colindale Avenue, London, N.W. 9

In certain kinds of bacteriological work it is necessary to study the rate of killing of organisms by disinfectants which act according to the exponential law, the number surviving after a time t being given by the equation

$$N(t) = N(0)\, e^{-kt}.$$

The estimate of the constant k given by the ordinary least-squares fit of log N against t is not a good one, since it takes no account of the accuracy of the observations and in general gives far too much weight to the smaller values of N. In many applications this is not an important objection since the experiment can be arranged to correct for it; however, in other cases this may not be so. For example, in the testing of air disinfectants the usual technique is such that later samples often contain small numbers of organisms with occasional zero values.

In the air disinfection problem which gave rise to this note the observed bacterial counts are samples from a very much larger population, and it can be assumed that the inevitable correlated errors between successive samples are negligible compared with the errors of sampling, which are of the Poisson form. Similar conditions hold in some other bacteriological applications, and in any case the estimate of k suggested here is rationally weighted and seems preferable to the least-squares fit.

With the assumption of independent errors each count can be regarded as a Poisson variable with expectation

$$\bar{n}_t = \nu_0\, e^{-kt},$$

where ν_0 is the baseline count. The likelihood of the observed counts is then the product of a series of Poisson terms as follows

$$e^L \alpha \frac{e^{-\nu_0}(\nu_0)^{n_0}}{n_0!} \times \ldots \times e^{-\nu_0 e^{-kt}} \frac{(\nu_0 e^{-kt})^{n_t}}{n_t!} \times \ldots \times \text{ etc.}$$

and the log likelihood is

$$L = \text{const.} - \nu_0 \Sigma e^{-kt} + \log \nu_0 \Sigma n_t - k \Sigma t n_t, \tag{1}$$

from which estimates of k and ν_0 can be obtained in the usual way by differentiating and equating to zero. This gives

$$\frac{\Sigma t e^{-kt}}{\Sigma e^{-kt}} = \frac{\Sigma t n_t}{\Sigma n_t}, \tag{2}$$

and ν_0 is found from the equation

$$\nu_0 = \frac{\Sigma n_t}{\Sigma e^{-kt}}. \tag{3}$$

https://doi.org/10.1017/S0022172400044533 Published online by Cambridge University Press

These equations can be solved by trial and error using a table of e^x, but, for the purpose of the present application, a table of the function on the left-hand side of (2) was computed and is given in Table 1. It covers the range from 0·01 to 1·0 and enables a sufficiently accurate value to be obtained by interpolation when the values of t are spaced at equal intervals. The constant k has dimensions t^{-1}, and for comparability with other such values must be converted to its proper value if necessary. For example, if equally spaced counts are made at intervals of 3 min. and a value of k (say k_1) is found using the table, then its value in (hr.)$^{-1}$

Table 1. *Table of values of the function* $\sum_{t=0}^{r} te^{-kt} \Big/ \sum_{0}^{r} e^{-kt}$

r \ k	0·01	0·02	0·03	0·04	0·05	0·06	0·07	0·08	0·09
1	0·4975	0·4961	0·4922	0·4898	0·4876	0·4849	0·4826	0·4800	0·4775
2	0·9965	0·9879	0·9794	0·9732	0·9667	0·9599	0·9534	0·9467	0·9401
3	1·4897	1·4758	1·4620	1·4499	1·4376	1·4250	1·4127	1·4002	1·3878
4	1·9815	1·9609	1·9395	1·9199	1·9001	1·8801	1·8603	1·8405	1·8207
5	2·4732	2·4426	2·4120	2·3833	2·3544	2·3253	2·2965	2·2676	2·2388
6	2·9614	2·9209	2·8796	2·8401	2·8005	2·7606	2·7212	2·6817	2·6424
7	3·4491	3·3959	3·3421	3·2902	3·2383	3·1861	3·1345	3·0829	3·0316
8	3·9347	3·8675	3·7997	3·7338	3·6678	3·6018	3·5365	3·4713	3·4066
9	4·4192	4·3360	4·2524	4·1708	4·0893	4·0079	3·9272	3·8470	3·7675
10	4·9016	4·8010	4·7000	4·6012	4·5026	4·4042	4·3069	4·2102	4·1145
11	5·3826	5·2627	5·1427	5·0250	4·9078	4·7910	4·6756	4·5611	4·4480
12	5·8614	5·7211	5·5806	5·4424	5·3049	5·1683	5·0334	4·8998	4·7680
13	6·3393	6·1763	6·0135	5·8533	5·6941	5·5362	5·3804	5·2265	5·0750
14	6·8149	6·6281	6·4414	6·2577	6·0754	5·8948	5·7169	5·5415	5·3692
15	7·2893	7·0766	6·8644	6·6556	6·4487	6·2441	6·0429	5·8449	5·6508
16	7·7618	7·5218	7·2826	7·0472	6·8143	6·5843	6·3585	6·1369	5·9201
17	8·2325	7·9636	7·6959	7·4324	7·1721	6·9155	6·6641	6·4178	6·1776
18	8·7020	8·4023	8·1043	7·8113	7·5222	7·2378	6·9596	6·6879	6·4234
19	9·1695	8·8376	8·5080	8·1839	7·8647	7·5513	7·2454	6·9473	6·6580
20	9·6352	9·2697	8·9067	8·5502	8·1996	7·8561	7·5215	7·1963	6·8372

r \ k	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
1	0·4750	0·4502	0·4256	0·4013	0·3775	0·3543	0·3318	0·3100	0·2891	0·2689
2	0·9334	0·8675	0·8029	0·7402	0·6798	0·6223	0·5678	0·5167	0·4690	0·4248
3	1·3754	1·2528	1·1342	1·0214	0·9154	0·8173	0·7274	0·6460	0·5728	0·5073
4	1·8009	1·6068	1·4222	1·2507	1·0944	0·9544	0·8307	0·7227	0·6289	0·5481
5	2·2101	1·9306	1·6701	1·4346	1·2271	1·0478	0·8951	0·7662	0·6579	0·5671
6	2·6033	2·2255	1·8815	1·5800	1·3235	1·1098	0·9339	0·7900	0·6722	0·5756
7	2·9806	2·4929	2·0601	1·6933	1·3922	1·1500	0·9567	0·8027	0·6791	0·5793
8	3·3423	2·7343	2·2099	1·7804	1·4404	1·1755	0·9699	0·8092	0·6824	0·5809
9	3·6886	2·9515	2·3343	1·8467	1·4737	1·1915	0·9773	0·8126	0·6839	0·5815
10	4·0198	3·1459	2·4370	1·8965	1·4964	1·2014	0·9815	0·8143	0·6846	0·5818
11	4·3362	3·3194	2·5212	1·9337	1·5117	1·2074	0·9837	0·8152	0·6849	0·5819
12	4·6381	3·4736	2·5897	1·9611	1·5219	1·2110	0·9850	0·8156	0·6850	0·5819
13	4·9260	3·6102	2·6452	1·9813	1·5287	1·2132	0·9866	0·8158	0·6851	0·5820
14	5·2001	3·7307	2·6898	1·9960	1·5332	1·2145	0·9860	0·8159	0·6851	0·5820
15	5·4608	3·8367	2·7255	2·0066	1·5361	1·2153	0·9862	0·8159	0·6851	0·5820
16	5·7085	3·9297	2·7540	2·0143	1·5380	1·2157	0·9863	0·8159	0·6851	0·5820
17	5·9437	4·0110	2·7766	2·0198	1·5393	1·2160	0·9864	0·8160	0·6851	0·5820
18	6·1667	4·0819	2·7945	2·0237	1·5401	1·2162	0·9864	0·8160	0·6851	0·5820
19	6·3780	4·1435	2·8086	2·0265	1·5406	1·2162	0·9864	0·8160	0·6851	0·5820
20	6·5779	4·1970	2·8197	2·0285	1·5409	1·2163	0·9864	0·8160	0·6851	0·5820

https://doi.org/10.1017/S0022172400044533 Published online by Cambridge University Press

will be $20k_1$, and in (min.)$^{-1}$ it will be $\frac{1}{3}k_1$. If the results justify it, more accurate estimates of k and ν_0 can be obtained by solving the following simultaneous equations:

$$\left.\begin{aligned} \delta k \frac{\partial^2 L}{\partial k^2} + \delta \nu_0 \frac{\partial^2 L}{\partial k\, \partial \nu_0} = -\frac{\partial L}{\partial k}, \\ \delta k \frac{\partial^2 L}{\partial k\, \partial \nu_0} + \delta \nu_0 \frac{\partial^2 L}{\partial {\nu_0}^2} = -\frac{\partial L}{\partial \nu_0}, \end{aligned}\right\} \qquad (4)$$

and adding the corrections δk and $\delta \nu_0$ to the first approximations. The values of the differential coefficients are:

$$\frac{\partial^2 L}{\partial k^2} = -\nu_0 \Sigma t^2 e^{-kt},$$

$$\frac{\partial^2 L}{\partial k\, \partial \nu_0} = \Sigma t e^{-kt},$$

$$\frac{\partial^2 L}{\partial {\nu_0}^2} = -\frac{\Sigma n_t}{{\nu_0}^2}.$$

Table 2. *Successive counts of a micrococcus exposed to hexanediol*

t (min.)	Count
0	58
$\frac{1}{18}$	34
$\frac{1}{9}$	10
$\frac{1}{6}$	10
$\frac{2}{9}$	7
$\frac{5}{18}$	3
$\frac{1}{3}$	5
$\frac{7}{18}$	2

The variances of the estimates are obtained using the determinant of the information matrix

$$I = \begin{vmatrix} \dfrac{\partial^2 L}{\partial k^2} & \dfrac{\partial^2 L}{\partial k\, \partial \nu_0} \\ \dfrac{\partial^2 L}{\partial k\, \partial \nu_0} & \dfrac{\partial^2 L}{\partial {\nu_0}^2} \end{vmatrix}, \qquad (5)$$

from which are computed

$$\left.\begin{aligned} \operatorname{var}(k) = -\frac{\partial^2 L}{\partial {\nu_0}^2} \Big/ I, \\ \operatorname{var}(\nu_0) = -\frac{\partial^2 L}{\partial k^2} \Big/ I. \end{aligned}\right\} \qquad (6)$$

The goodness of fit of the observations to the hypothesis of logarithmic decay can be decided by a χ^2 test comparing the observed and calculated counts. If there are *r counts, χ^2 will have $(r-2)$ degrees of freedom.*

The calculations may be illustrated by some results, kindly given by Mr T. Nash, which show the killing effect of hexanediol on a micrococcus. Counts were made at intervals of $\frac{1}{18}$ min., with the results shown in Table 2. Giving t the values 0, 1, 2, ..., 7 the sums $\Sigma(tn_t)$ and $\Sigma(n_t)$ are computed, and then the ratio

$$\frac{\Sigma(tn_t)}{\Sigma(n_t)} = \frac{171}{129} = 1{\cdot}325.$$

https://doi.org/10.1017/S0022172400044533 Published online by Cambridge University Press

Entering Table 1 with $r=7$ this value of $\Sigma te^{-kt}/\Sigma e^{-kt}$ is found to be given by a k between 0·5 and 0·6, since

$$\frac{\Sigma te^{-kt}}{\Sigma e^{-kt}}=1\cdot3922 \quad \text{if} \quad k=0\cdot5$$

and
$$\frac{\Sigma te^{-kt}}{\Sigma e^{-kt}}=1\cdot1500 \quad \text{if} \quad k=0\cdot6.$$

The first estimate of k is therefore

$$k_1=0\cdot6-\frac{1\cdot325-1\cdot150}{1\cdot392-1\cdot150}\times0\cdot1$$
$$=0\cdot528.$$

Using this value of k, ν_0 is found to be 53·50.

To find the standard errors of k and ν_0 it is necessary to calculate

$$\Sigma te^{-kt}=3\cdot176,$$
$$\nu_0\Sigma t^2e^{-kt}=550\cdot24,$$
$$\frac{\Sigma nt}{\nu_0^2}=0\cdot04506,$$

from which is derived the determinant of the information matrix

$$I=\begin{vmatrix} -550\cdot24 & 3\cdot176 \\ 3\cdot176 & -0\cdot04506 \end{vmatrix}$$
$$=550\cdot24\times0\cdot04506-(3\cdot176)^2$$
$$=14\cdot729.$$

The variances of k and ν_0 are then

$$\text{var }(k)=\frac{0\cdot04506}{14\cdot729}=0\cdot00306,$$
$$\text{var }(\nu_0)=\frac{550\cdot24}{14\cdot729}=37\cdot4.$$

Finally, then,
$$k=0\cdot528\pm0\cdot055,$$
$$\nu_0=53\cdot5\pm6\cdot12.$$

As the observations were made at intervals of $\frac{1}{18}$ min. the value of k in (min.)$^{-1}$ is $18\times0\cdot528=9\cdot51$, of which the standard error will be $(18\times0\cdot055)=0\cdot99$. The corrections introduced by a second approximation are negligible compared with these standard errors.

It sometimes happens that a series of baseline counts are available before the disinfectant is applied. In this case the method of estimation is unaltered but the equations may be written

$$\nu_0=\frac{\sum_1^r n_t+\Sigma n_0}{s+\sum_1^r e^{-kt}} \tag{7}$$

https://doi.org/10.1017/S0022172400044533 Published online by Cambridge University Press

and
$$\frac{\sum_1^r te^{-kt}}{s+\sum_1^r e^{-kt}} = \frac{\sum_1^r tn_t}{\Sigma n_0 + \sum_1^r n_t}, \tag{8}$$

where Σn_0 is the sum of all the counts made before the application of the killing agent, and s is the number of such counts. The information matrix is

$$I = \begin{vmatrix} -\nu_0 \sum_1^r t^2 e^{-kt} & \sum_1^r te^{-kt} \\ \sum_1^r e^{-kt} & -\frac{\Sigma n_0 + \sum_1^r n_t}{\nu_0^2} \end{vmatrix}, \tag{9}$$

and the variances of k and ν_0 are

$$\left.\begin{aligned} \operatorname{var}(k) &= \frac{\Sigma n_0 + \sum_1^r n_t}{\nu_0^2} \Big/ I, \\ \operatorname{var}(\nu_0) &= \nu_0 \sum_1^r t^2 e^{-kt}/I. \end{aligned}\right\} \tag{10}$$

These equations are identical in form with those given for a single baseline count, the estimate of ν_0 now being a weighted mean of those provided by the baseline counts and by the killing curve respectively.

The estimates of k and ν_0 given above seem to be generally preferable to those found by fitting a straight line, as they are more reasonably weighted and also, given the table, more easily calculated. On the other hand, the variances may be disturbed by the occurrence of correlated errors among the counts, and if such errors cannot be discounted the variance formulae cannot be said to give more than an order of magnitude.

The case of an exponential decay curve whose total number is known at the outset has not been encountered in bacteriology though a haematological example has been pointed out to me by Dr Armitage. The problem seems to merit some discussion as it provides a further possible estimate of k which is very easy to compute. An analogous but more complex situation has been discussed by Kendall (1949, 1950) and by Moran (1951, 1953). The exponential decay process depends essentially on the fact that the deaths of individual organisms are quite independent, the probability that an organism is still alive at time t being

$$p(t) = e^{-kt}.$$

Now, if there are n_t organisms still alive at time t then the probability that there will be $n_{t+\tau}$ alive at time $t+\tau$ is

$$\binom{n_t}{n_{t+\tau}} (e^{-k\tau})^{n_{t+\tau}} (1-e^{-k\tau})^{n_t - n_{t+\tau}}.$$

https://doi.org/10.1017/S0022172400044533 Published online by Cambridge University Press

If a series of observations is taken at equal intervals of time the overall likelihood of the observed results will be proportional to

$$\begin{aligned}&(e^{-k\tau})^{n_1}\,(1-e^{-k\tau})^{n_0-n_1}\\ \times\,&(e^{-k\tau})^{n_2}\,(1-e^{-k\tau})^{n_1-n_2}\\ &\vdots\\ \times\,&(e^{-k\tau})^{n_r}\,(1-e^{-k\tau})^{n_{r-1}-n_r},\end{aligned}$$

and maximizing the log likelihood gives for the estimate of k

$$\frac{\partial L}{\partial k} = -\tau\sum_1^r n_i + \frac{(n_0-n_r)\,\tau e^{k\tau}}{1-e^{-k\tau}}, \tag{11}$$

whence

$$e^{-k\tau} = \frac{n_1+n_2+\ldots+n_r}{n_0+n_1+\ldots+n_{r-1}}$$

$$k = \frac{1}{\tau}\log\left\{\frac{n_0+n_1+\ldots+n_{r-1}}{n_1+n_2+\ldots+n_r}\right\}. \tag{12}$$

The estimate of k given by this expression is based on a weighted mean of the individual values of n_k/n_{k-1}, the weights being equal to the denominator in each case.

The variance of the estimate is given by differentiating the likelihood expression again

$$\left.\begin{aligned}\frac{\partial^2 L}{\partial k^2} &= -\frac{\tau^2(n_0-n_r)\,e^{-k\tau}}{(1-e^{-k\tau})^2} = -\frac{1}{\text{var }(k)},\\ \text{var }(k) &= \frac{4\sinh^2(\frac{1}{2}k\tau)}{\tau^2(n_0-n_r)}.\end{aligned}\right\} \tag{13}$$

It has been pointed out to me by Dr Armitage that this variance is appropriate, in the present case, to the results of single experiments provided that the number of organisms is reasonably large.

The methods discussed in this note are put forward as being preferable to an unweighted estimate of the decay constant. Probably both the suggested models are too simple to be completely applicable to any practical situation, but the estimators are simple and logical enough to deserve consideration in their own right.

SUMMARY

Methods are proposed for measuring the constants of an exponentially declining population under conditions in which the least-squares method is unsuitable. A table is given to facilitate the calculations.

My thanks are due to Miss Shirley Johnson for helping to compute Table 1, to Mr T. Nash who brought this problem to my notice and to Dr P. Armitage for his comments on the MS.

https://doi.org/10.1017/S0022172400044533 Published online by Cambridge University Press

REFERENCES

KENDALL, D. G. (1949). Stochastic processes and population growth. *J.R. statist. Soc.* B, **11**, 230–64.

KENDALL, D. G. (1950). An artificial realisation of a simple birth and death process. *J.R. statist. Soc.* B, **12**, 116–19.

MORAN, P. A. P. (1951). Estimation methods for evolutive processes. *J.R. statist. Soc.* B, **13**, 141–6.

MORAN, P. A. P. (1953). The estimation of the parameters of a birth and death process. *J.R. statist. Soc.* B, **15**, 241–5.

(*MS. received for publication* 8. II. 56)

https://doi.org/10.1017/S0022172400044533 Published online by Cambridge University Press