BIOMETRIC ANALYSIS OF AN EXPONENTIAL KINETIC MODEL OF DISINFECTION

BY

N. SUNDARARAJ

University of Agricultural Sciences, Bangalore

(Received : November, 1976)

INTRODUCTION

In its broad sense, disinfection includes removal and destruction of micro-organisms both by chemical and physical means. Although many mathematical models are proposed for describing the process of thermal microbial disinfection, the most commonly preferred is the exponential model

$$\eta\ (t_h) = \theta_1\ exp\ (-t_h/\theta_2),\ h = 1, 2, \ldots, n$$

where

$\eta(t_h) =$ the number of viable organisms at any time $t = t_h$

$\theta_1 =$ the number of organisms at time $t = 0$

$\theta_2 =$ a constant proportional to what is known as Decimal Reduction Time (D), that is, the time required to reduce the microbial population by 90%. In fact $D = \theta_2 \ln 10$

In practice a microbiologist obtains experimental observations on the number of organisms kept under constant heat for preselected points of time and desires to estimate the characterising parameter and obtain a measure of its precision. Usually the input (time) variable $t_h\ (h = 1, 2, \ldots, n)$ can be assumed to be measured at each of n points of time with negligible error, but the corresponding numbers of organisms at each of these points, viz., $\eta(t_1),\ \eta(t_2) \ldots \eta(t_n)$ cannot be measured without errors. However, their estimates $y_1, y_2, \ldots y_n$ can be determined. If m_h replicates are made on viable organisms at time $t_h(h = 1, 2, \ldots n)$, then these observations y_{hk} $(k = 1, 2, \ldots m_h;\ h = 1, 2, \ldots n)$ will each contain an unobservable random error $\epsilon_{hk}\ (k = 1, 2, \ldots m_h;\ h = 1, 2, \ldots n)$

so that

$$y_{hk} = \theta_1\ exp(-t_h/\theta_2) + \epsilon_{hk} \qquad \ldots(1.1)$$

$$E\left(\epsilon_{hk}\right)=0,\ \ E\left(\epsilon_{hk}\ \epsilon_{h'k'}\right)=0,\ \ E\left(\epsilon_{hk}^{2}\right)=\sigma^{2}$$
$$h\neq h'=1,\ 2,\ \ldots n$$
$$k\neq k'=1,\ 2,\ \ldots m_{h}$$

The justification for the exponential model structure in (1.1) partly comes from empirical evidence (see for example, Chick 1908, Pflug and Schmidt, 1968, and Meynell & Meynell, 1965). The assumptions about the random components ϵ_{hk} are mainly realized from designing of experiments.

Certain biometric problems then arise : how would the random observations y_{hk} be employed to estimate the unknown parameters θ_1 & θ_2 and how can the statistical uncertainties in the resulting estimators be measured and estimated ?

In this paper a new method will be proposed for setting up a confidence interval on the unknown parameter θ_2 via the least square equation; and therefrom to obtain a measure of an 'effective' standard error for the above point estimate of θ_2. Further, in order to have an appreciation of the proposed method, these results will be compared with those obtained by employing the customary method of analysis *viz.*, the logarithmic transformation of the model (1.1) to a practical data collected by Pflug *et al.* (1969) and reported in Sundararaj (1971) on thermal microbial sterilization of spore— *Bacillus subtilis* var. *niger.*

2. Point Estimators for θ_1 and θ_2

The point estimators considered in this section are the least square estimators. The least square estimators for θ_1 and θ_2 for the intact nonlinear model (1.1) are obtained by minimizing the expression

$$S=\sum_{h}\sum_{k}\left[y_{hk}-\theta_1\,exp\,(-t_h/\theta_2)\right]^2$$

with respect to θ_1 and θ_2. The two least square equations in $\tilde{\theta}_1$ & $\tilde{\theta}_2$ are :

$$\tilde{\theta}_1\sum_{h}m_h\,exp\,(-2t_h/\tilde{\theta}_2)=\sum_{h}\sum_{k}y_{hk}\,exp\,(-t_h/\tilde{\theta}_2)\quad\ldots(2.1)$$

$$\tilde{\theta}_1\sum_{h}m_ht_h\,exp\,(-2t_h/\tilde{\theta}_2)=\sum_{h}\sum_{k}y_{hk}\,t_h\,exp\,(-t_h/\tilde{\theta}_2)\quad\ldots(2.2)$$

whose solutions for $\tilde{\theta}_1$ and $\tilde{\theta}_2$ will provide the desired estimators.

It may be noted that we may eliminate $\tilde{\theta}_1$ between (2.1) & (2.2) to obtain an equation in $\tilde{\theta}_2$ alone, viz.,

$$\frac{\sum\limits_h m_h\, t_h\, exp\ (-2t_h/\tilde{\theta}_2)}{\sum\limits_h m_h\, exp\ (-2t_h/\tilde{\theta}_2)} = \frac{\sum\limits_h \sum\limits_k y_{hk}\, t_h\, exp\ (-t_h/\tilde{\theta}_2)}{\sum\limits_h \sum\limits_k y_{hk}\, exp\ (-t_h/\tilde{\theta}_2)} \quad ...(2.3)$$

Equation (2.3) may be first solved for θ iteratively and then obtain the solution for $\tilde{\theta}_1$ from (2.1).

3. PRECISION OF THE ESTIMATOR AND CONFIDENCE INTERVAL ON θ_2

It may be observed that a straight forward solutions for these problem are not, in general, available in view of the fact that the estimators $\tilde{\theta}_1$ & $\tilde{\theta}_2$ are not linear functions of the observation random variables y_{hk} $(k=1, 2, ...m_h\,;\,h=1, 2, ...n)$. We shall consider here a method of providing, first a confidence interval on θ_2 and therefrom to obtain an estimate of an 'effective' standard error of $\tilde{\theta}_2$.

For this, we recast the least square equation (2.3) in the form

$$0=g(y, \tilde{\theta}_2)=\sum_h \sum_K H(\tilde{\theta}_2)y_{hk} \qquad ...(3.1)$$

where

$$H(\tilde{\theta}_2) = \{[\sum_h m_h\, t_h\, exp\ (-2t_h/\tilde{\theta}_2)]-[\sum_h m_h\, exp\ (-2t_h/\tilde{\theta}_2)]\, t_h\}$$

$$[exp\ (-t_h/\tilde{\theta}_2)] \qquad ...(3.2)$$

Now replacing $\tilde{\theta}_2$ by θ_2 we may construct a new random variable given by

$$g(y, \theta_2) = \sum_h \sum_K H\ (\theta_2)y_{hK} \qquad ...(3.3)$$

It may be easily verified that

$$E\ [g(y, \theta_2)]=0 \qquad ...(3.4)$$
$$V\ [g(y, \theta_2)]=\sigma^2 \sum_h m_h\, H^2(\theta_2) \qquad ...(3.5)$$

The confidence interval on θ_2 may now be constructed as follows. Under the Gaussian assumption, each $y_{hk}(K=1, 2,... m_h\,;\,h=1, 2,...n)$ is normally and independently distributed so that the new random variable $g(y, \theta_2)$ which is a linear function of these

random variables is itself exactly normally distributed with mean zero and variance given by (3.5). Consequentially we may make a precise probability statement of the type

$$P\left[- Z_{\alpha/2} \leqslant \frac{g(y, \theta_2)}{\sqrt{\sigma^2 \underset{h}{\Sigma} m_h H^2(\theta_2)}} \leqslant Z_{(1-\alpha/2)} \right] = (1-\alpha) \quad ...(3.6)$$

In particular therefore the boundary values $\tilde{\theta}_{2,L}$ & $\tilde{\theta}_{2,U}$, which are the solutions of the equations

$$\frac{g(y, \tilde{\theta}_2, L)}{\sqrt{\sigma^2 \underset{h}{\Sigma} m_h H^2(\tilde{\theta}_2, L)}} = -Z_{\alpha/2}$$

and

$$\frac{g(y, \tilde{\theta}_2, U)}{\sqrt{\sigma^2 \underset{h}{\Sigma} m_h H^2(\tilde{\theta}_2, U)}} = Z_{(1-\alpha/2)} \quad ...(3.7)$$

provide a 100 $(1-\alpha)$ per cent confidence interval on θ_2. The search for the numerical values for $\tilde{\theta}_2, L$ & $\tilde{\theta}_2, U$ will involve considerable amount of computational labour, and this should not be a problem when computer facilities are available.

The important implications of equation (3.7) are :

(1) The confidence interval on θ_2, essentially comes from consideration of the distribution of an appropriate linear combinations of the observation random variables y_{hK} $(k=1, 2, ...m_h, h=1, 2, ...n)$ viz.. equation (3.4).

(2) Equation (3.4) assumes the value zero if and only when θ_2 equals $\tilde{\theta}_2$, which is the least square solution for θ_2.

(3) Employing this, we can construct a confidence interval on θ_2 without obtaining an explicit form of the distribution of $\tilde{\theta}_2$, which is generally intractable ; and such a confidence interval will include the least square value of $\tilde{\theta}_2$ for θ_2.

Finally the 'effective' standard error of the estimate may be obtained as

$$SE(\tilde{\theta}_2) = \left| \frac{\tilde{\theta}_{2,U} - \tilde{\theta}_{2,L}}{2Z_{(1-\alpha/2)}} \right| \quad ...(3.8)$$

As is true in practice, the variance σ^2 in (3.5) is generally unknown. Then either it is to be estimated from the observations

themselves or occasionally it may be possible to obtain an independent estimate, say s^2, from previous similar experiments. In the latter case, if it is based on a large number of degrees of freedom, then the central expression in (3.6), with replacing σ^2 by s^2, will have a student t distribution and the confidence limits are obtained from (3.7) by employing appropriate t-limits in place of normal z-limits. In this case the method provides exact confidence limits on θ_2. On the other hand, if such an independent estimate of σ^2 is not available, which is generally the case, it will have to be estimated from the sample observations themselves and the method will involve some approximation since the distribution will not then follow exact t-distribution. However the robustness of the central limit theory should operate to provide a reasonably good approximation to this.

4. THE QUASI-LEAST SQUARE ESTIMATOR FOR θ_2 AND ITS PRECISION

In order to have an appreciation of the foregoing results, the method that is customarily employed for estimating the precision of the point estimator of θ_2 is considered in this section. The method, here called the quasi-least square method, involves the logarithmic transformation of the observation random variable y_{hk} ($k=1, 2,...$ $m_h/h=1, 2,...n$) which renders the model (1.1) to a linear form so that the simple linear regression formulae may be applied. The straight forward application of the least square principle to the logarithmic transformation of the model (1.1) leads to the point estimator $\hat{\theta}$ of θ_2,

$$\hat{\theta}_2 = -\left[\frac{\sum\limits_h m_h (t_h-\bar{t})^2}{\sum\limits_h \sum\limits_K (t_h-\bar{t})l_n y_{hK}}\right] \qquad ...(4.1)$$

whose asymptotic precision by the error propogation method is given by (see Sundararj, 1971)

$$V(\hat{\theta}_2) = \left\{\frac{\sum\limits_h \sum\limits_K (t_h-\bar{t}) \exp (2t_h/\theta_2)}{[\sum\limits_h m_h(t_h-\bar{t})^2]^2}\right\}\left(\frac{\sigma^2 \theta_2^*}{\theta_1^2}\right) \qquad ...(4.2)$$

The estimate of the variance can be obtained by substituting the sample values for the unknown parameters. The above formulae are employed for the data on microbial destruction reported in the following section.

Note : Writing the model (1.1) as $y_{hk}=[\theta_1 exp (-t_h/\theta_2)] [1+\epsilon_{hk}/\theta_1 exp (-t_h/\theta_2)]$ the log transformation leads to $\ln y_{hk}=$ in $\theta_1-th/\theta_2+\epsilon^*_{hk}$ where $\epsilon^*_{hk}=\ln [1+\epsilon_{hk}/\theta_1 exp (-t_h/\theta_2)]$. Thus the least square estimators for θ_1 & θ_2 via minimizing $\sum\sum\epsilon^{*2}_{hk}$ is not a true minimization of $\sum\sum\epsilon^2_{hk}$ and hence termed as quasi-least square estimators.

5. Application

This section concerns with the application of the foregoing formulae to estimate θ_2 and its precision to the experimental data of Pflug & associates (Pflug *et al.* 1969). Briefly, the experiment is as follows.

By means of a twenty μl pipette, 1.0 ml of aqueous spores (*Bacillus subtilis* var. *niger*) suspension (aproximately 1×10^6 spores) was deposited on to each of 3 sterile, stainless strips (planchets), each $\frac{1}{2}''$ square lying flat in a shallow groove running the length of a rectangular copper boat. Seven such copper boats were used, assigned randomly either to no heat or to one of the six heating times listed in Table 1. The samples were conditioned at 23 degree centigrade and 39% relative humidity while treated at 100 degree centigrade and 0.02% relative humidity. The guidelines for the estimation of the number of surviving spores following the heat treatment included the standard dilution method and counting of colonies on all plates. Table 1 summarises the data on the estimated number of spores in the original 50 ml solution.

All the foregoing formulae from section 2 through section 4 have been employed to estimate the parameter θ_2 and its precision, since this parameter is the one which characterizes the effect of the disinfectant, *viz.*, the thermal disinfectant in the present example. Further the numerical results of the analysis are reported in Table 2. However, not in terms of the original units, but in terms of a related concept called the Decimal Reduction Time (*D*), as latter one is usually understood as time required to reduce the microbial population by 90%. The conversion formula from θ_2 to D is simply given by the relationship $D = \theta_2 \ln 10$. The confidence intervals and the standard errors are given in Table 2 for D.

TABLE 2

The estimates of D in Minutes with the SE's, 95% Confidence
Intervals etc., for the data of Table 1

Method	Point Estimates	SE Estimates	Confidence Lower	Limits Upper	Width of the Con. limits
A	103.07	5.69	91.96	114.26	22.30
B	109.44	6.14	97.39	121.48	24.08

Methods :

A : Stands for the proposed new method.

B : Stands for the Quai-Least Square Method.

TABLE 1

The estimated number of spores* in original 50 ml suspension at the times listed in column 2, with two determinations for each Planchet and with 3 Planchets for each boat

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Planchet 1		Planchet 2		Planchet 3		
Boat No.	Times in Min.	Determination No.		Determination No.		Determination No.		Arith. Mean
		(1)	(2)	(1)	(2)	(1)	(2)	
1	0	1275000	1504500	1224000	1504500	1657500	1504500	1445000
2	75	90525	112200	119850	104550	124950	109650	110288
3	75	144075	116025	112200	112200	93075	85425	110500
4	150	18800	22200	20800	21400	16600	15900	19283
5	225	7975	7925	7625	6700	6875	7450	7425
6	300	700	825	975	950	750	475	779
7	300	1050	850	1150	1150	1000	775	996

*Bacillus subtilis var. niger. Samples were conditioned at 23°C and 39% relative humidity while treated at 100°C and 0.02% relative humidity for times listed. Apparatus type—Open, Name—Hot Plate in Globe Box.

The foregoing numerical analysis shows that the new method proposed proves slightly better than the quasi-least square method customarily employed for this model. Further results (not reported here) demonstrated the proposed method compared very well with the maximum likelihood method employed for the intact nonlinear model.

SUMMARY

Parameters characterizing a model intended to describe a biological or a chemical phenomenon often enter the model in a nonlinear fashion; as a result, the statistical problems of point and interval estimation of such parameters are much too complicated in contra-distinction with linear model methods. A novel method has been proposed in this paper for the case of an exponential model-often termed a 'kinetic exponential' model in the context of employing it as a model for thermal disinfection. The merits of this method are compared with the customarily employed method of estimating the parameters after rendering the linear by log-transformation by means of analysis of a practical data on disinfection.

REFERENCES

[1] Chick, H. (1908) : An Investigation of the Laws of Disinfection, Journal of Hygiene, 8, pp 92-158.

[2] Pflug, I.J. & Schmidt, C.F. (1968) : Thermal Destruction of Micro-organisms. Chap. IV in *Disinfection Sterilization & Preservation.* (ed) Carl A. Lawrence & Seymor S. Block, Lea and Febiger, Philadelphia.

[3] Pflug, I. I. *et al.* (1969) : Environmental Biology as related to Planetary Quarantine, *Semiannual Progress Report 3,* NASA Grant NGL 24-005-160, Division of Environmental Health, University of Minnesota, Minneapolis, Minn., USA.

[4] Meynell, G.C. & Meynell, E. (1965) : *Theory and Practice in Experimental Bacteriology,* Cambridge University Press, Cambridge England (Chap. IV).

[5] Sundararaj, N. (1971) : "The Biometric Analysis of Nonlinear Models with Applications to Disinfection" The PH.D. Thesis submitted to the University of Minnesota, USA.