A GENERALIZED MODIFIED-χ^2 ANALYSIS OF CATEGORICAL DATA FROM A COMPLEX DILUTION EXPERIMENT

By

Gary G. Koch and H. Dennis Tolley

Department of Biostatistics
University of North Carolina, Chapel Hill, N. C.

Institute of Statistics Mimeo Series No. 864

MARCH 1973

A Generalized Modified-χ^2 Analysis of Categorical Data From a Complex Dilution Experiment

by

Gary G. Koch and H. Dennis Tolley
Department of Biostatistics
School of Public Health
University of North Carolina
Chapel Hill, N. C.

ABSTRACT

A method in the current literature used to estimate the density of bacteria in a sample solution is the most probable number (MPN) procedure.

This paper considers the change in density of bacteria in the sample through time. The MPN with its variance for a fixed time point is just one module of the experiment and is treated as a generalization of the unrestricted maximum likelihood estimates used for cell probabilities in linear categorical data analysis. The MPN variance estimates are used to fit an exponential decay model over time by weighted least squares. When the models fit, comparisons of decay rates between population can be undertaken.

1. INTRODUCTION

Many situations in public health studies involve the estimation of bacteria density in a solution. Moreover, in certain applications, such statistical problems are further complicated because the estimation of the density at a given time point is but one module of a larger experiment concerned with decay rates or extinction times of bacteria. These experiments require one to consider models involving density estimates at several points in time.

As indicated by Finney [1964], the two major procedures for bacteria enumeration are the colony count method and the quantal response method. The colony count method assumes that the progeny of each bacteria grow in discernable colonies which are counted after an incubation time. From these counts, estimates of the density are formed. Because all bacteria are not suitable for colony count methods quantal response data have been used to form a variety of estimates. Although the procedure illustrated in this paper is potentially applicable to colony count methods, we will consider only the quantal response method here. Data of this second type are generated by the inoculation of several sterile tubes (or plates) for each aliquot taken from a sequence of serial dilutions of the original solution. From the number of fertile tubes (i.e., tubes showing growth after incubation), density estimates are derived. Cornell and Speckman [1967] review this statistical problem in detail; and their conclusions indicate that the maximum likelihood estimate has satisfactory properties for both large and small sample sizes in such experiments.

Mather [1949] used simple density estimates at different points in time to analyze extinction rates of bacteria. Samples of bacteria were

exposed to bactericide for x=12(2)26 minutes. At the end of x minutes the proportion π_x of sterile samples was estimated. Mather then applied the exponential decay model $\log(-\log \pi_x) = \mu + \beta x$ to the observed results. Epstein [1967] gave a theoretical justification for this heuristic analysis by considering extinction time as an extreme value problem.

Mather's approach is concerned with one dilution per time point and thus direct use of the observed proportions is apparent. However, with serial dilution procedures, there are several proportions for density estimation at each time point. Consequently a reliable direct use of these quantities to fit the decay model is not immediate. One can apply a weighted least squares approach to all aspects of this problem. However, as Cornell and Speckman [1967] have suggested, this method can give biased results when there are small sample sizes at each time point, particularly in situations where either all (or none) of the tubes at certain dilutions are positive (or negative). Thus, an alternative methodology is needed.

In this paper, we illustrate a unified approach for estimating the parameters of an exponential decay model. This is accomplished by first determining the maximum likelihood estimates of the bacteria density at each time point and then applying the methodology of Grizzle, Starmer, and Koch [1969] (subsequently abbreviated GSK) for categorical data to fit exponential decay functions to these determinations across time. In this way, the principles of weighted least squares can be used to test the goodness of fit of such models and to make comparisons among the decay rates for experiments conducted on different populations.

2. THE MODEL

Let us first consider the estimation of the density for a single population at a single time. Enumeration of bacteria by the maximum likelihood method (or what is essentially the same as the 'most probable number' (MPN) procedure as named by McCrady [1915]) is based on two assumptions:

1. The distribution of the individual bacteria cells is random without aggregation of any kind, and hence the number of bacteria in a small unit of the solution follows a Poisson distribution.
2. Growth will ensue in a sterile tube with the introduction of one or more bacteria.

Under these assumptions the probability that an inoculated tube is sterile (shows no growth after incubation) is $\exp(-\lambda z)$ where λ is the density per inoculator unit in the original solution and z is the concentration of the inoculator used. Hence the probability of a fertile tube is $\{1-\exp(-\lambda z)\}$.

At the time point for which the density estimate is desired, we create k dilutions $z_1, \ldots, z_k$ from a portion of original solution. At the i-th dilution we inoculate n_i tubes. The dilution and inoculation of aliquot processes are undertaken so that individual tube responses are independent. From these assumptions, it follows that the number x_i of fertile tubes for the i-th dilution has the binomial distribution in (2.1). Thus, the

$$\Pr\{x_i = r\} = \binom{n_i}{r} [1-\exp(-\lambda z_i)]^r [\exp(-\lambda z_i)]^{n_i-r} \tag{2.1}$$

k-tuple $\underset{\sim}{x}' = (x_1, \ldots, x_k)$ has the likelihood function in (2.2).

$$L(\underset{\sim}{x}, \lambda) = \prod_{i=1}^{k} \binom{n_i}{x_i} [1-\exp(-\lambda z_i)]^{x_i} [\exp(-\lambda z_i)]^{n_i-x_i} \tag{2.2}$$

The 'most probable number' of bacteria is the value of λ, say $\hat{\lambda}$, which maximizes (2.2) or equivalently the value $\hat{\lambda}$ which solves the equation (2.3).

$$\sum_{i=1}^{k} \frac{\{x_i\, z_i \exp(-\hat{\lambda}\, z_i)\}}{\{1-\exp(-\hat{\lambda}\, z_i)\}} = \sum_{i=1}^{k} (n_i - x_i)\, z_i \qquad (2.3)$$

If $k = 1$, then $\hat{\lambda} = -(1/z_1)\{\log_e[(n - x_1)/n]\}$. However, for $k \geq 2$, the solution $\hat{\lambda}$ to (2.3) must be reached by iterative methods and algorithms for doing this have been given by Finney [1964] and Peto [1953]. Since the left hand side of (2.3) is a monotonic decreasing function of λ, this equation can also be easily solved by computer methods based on successive approximation techniques. This latter approach is used in the example in Section 3. In accordance with traditional maximum likelihood theory, the asymptotic variance of $\hat{\lambda}$ is given by expression (2.4) which is the reciprocal of the

$$\text{Var}(\hat{\lambda}) = \left[E\left\{\frac{d^2}{d\lambda^2} \log L(\underset{\sim}{x}, \lambda)\right\}\right]^{-1} = \left[\sum_{i=1}^{k} \frac{n_i\, z_i^2 \exp(-\lambda\, z_i)}{\{1-\exp(-\lambda\, z_i)\}}\right]^{-1} \qquad (2.4)$$

Fisher Information Statistic. Thus, a consistent estimate w for the variance of $\hat{\lambda}$ can be obtained by substituting $\hat{\lambda}$ for λ as indicated in (2.5). This result

$$w = \left[\sum_{i=1}^{k} \frac{(n_i\, z_i^2 \exp(-\hat{\lambda}\, z_i))}{\{1-\exp(-\hat{\lambda}\, z_i)\}}\right]^{-1} \qquad (2.5)$$

is analogous to that used by Peto [1953].

Next let us suppose that there are s populations of bacteria and for the h-th population we take u_h sets of observations at the time points $x_1, x_2, \ldots, x_{u_h}$. For the j-th time point of the h-th population we form the estimates $\hat{\lambda}_{hj}$ and w_{hj} on the basis of (2.3) and (2.5). Following Mather,

we assume an exponential decay model (2.6) characterizes the variation

$$\log_e (\lambda_{hj}) = (\mu_h + \beta_h x_{hj}) \tag{2.6}$$

of the λ_{hj} over time. If $y_{hj} = \log_e (\hat{\lambda}_{hj})$, then a consistent estimate v_{hj} for the variance of y_{hj} is given by (2.7) which is based on a first order

$$v_{hj} = \frac{w_{hj}}{(\hat{\lambda}_{hj})^2} \tag{2.7}$$

Taylor Series approximation.

The y_{hj} formed here may be analyzed in the same way as the generalized log-linear functions of cell proportions considered by GSK since both represent unrestricted maximum likelihood estimators of their corresponding asymptotic expectations. However, this application assumes the validity of (2.2) at each time point, and hence involves a somewhat more complex underlying model than the standard product multinomial distribution. In addition, the y_{hj} and their estimated variance v_{hj} must be initially determined by a suitably specialized computer sub-routine for successive approximation as opposed to direct matrix multiplication. Otherwise, the use of weighted least squares to fit linear models which characterize the variation of the y_{hj} over time in the various populations can be applied in exactly the same manner as outlined in GSK. For this purpose, we define in (2.8) the vector $\underset{\sim}{y}_h$ of the

$$\underset{\sim}{y}_h = \begin{bmatrix} y_{h1} \\ y_{h2} \\ \cdots \\ y_{hu_h} \end{bmatrix}, \quad \underset{\sim}{V}_{y,h} = \begin{bmatrix} v_{h1} & 0 & \cdots & 0 \\ 0 & v_{h2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & v_{hu_h} \end{bmatrix} \tag{2.8}$$

y_{hj} for the h-th population and its corresponding estimated covariance matrix $V_{y,h}$. Similarly, with respect to the model (2.6), we define in (2.9) the design matrix $\underset{\sim}{X}_h$

$$\underset{\sim}{X}_h = \begin{bmatrix} 1 & x_{h1} \\ 1 & x_{h2} \\ \cdots & \cdots \\ 1 & x_{hu_h} \end{bmatrix}, \quad \underset{\sim}{\gamma}_h = \begin{bmatrix} \mu_h \\ \beta_h \end{bmatrix} \tag{2.9}$$

and the parameter vector $\underset{\sim}{\gamma}_h$. Hence, in matrix notation, the model (2.6) can be written as indicated in (2.10) where "E_A" means "asymptotic expectation."

$$E_A\{\underset{\sim}{y}_h\} = \underset{\sim}{X}_h \underset{\sim}{\gamma}_h \tag{2.10}$$

The weighted least squares estimates $\underset{\sim}{g}_h$ are obtained from (2.11). A consistent

$$\underset{\sim}{g}_h = (\underset{\sim}{X}_h' \underset{\sim}{V}_{y,h}^{-1} \underset{\sim}{X}_h)^{-1} \underset{\sim}{X}_h' \underset{\sim}{V}_{y,h}^{-1} \underset{\sim}{y}_h \ . \tag{2.11}$$

estimate for the variance of $\underset{\sim}{g}_h$ is given in (2.12). A goodness of fit

$$\underset{\sim}{V}_{g,h} = (\underset{\sim}{X}_h' \underset{\sim}{V}_{y,h}^{-1} \underset{\sim}{X}_h)^{-1} \tag{2.12}$$

statistic for assessing the extent to which the exponential decay model characterizes the respective determinations from the h-th population is the residual sum of squares in (2.13) which has approximately a chi-square

$$X_h^2 = SS\ (E_A\{\underset{\sim}{y}_h\} = \underset{\sim}{X}_h \underset{\sim}{\gamma}_h) = \underset{\sim}{y}_h' \underset{\sim}{V}_{y,h}^{-1} \underset{\sim}{y}_h - \underset{\sim}{g}_h' (\underset{\sim}{X}_h' \underset{\sim}{V}_{y,h}^{-1} \underset{\sim}{X}_h) \underset{\sim}{g}_h \tag{2.13}$$

distribution with D.F. = $(u_h - 2)$ in large samples under the hypothesis that the model (2.10) fits for the h-th population where h = 1, 2, ..., s.

If such models are satisfactory for each of the s populations, then we can compare their corresponding parameters by testing appropriate hypotheses in terms of a general model for the combined set of experiments. Thus, we define in (2.14) the composite vector $\underset{\sim}{y}$ of the $\underset{\sim}{y}_h$ and the corresponding

$$\underset{u\times u}{\underset{\sim}{y}} = \begin{bmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \\ \cdots \\ \underset{\sim}{y}_s \end{bmatrix}, \qquad \underset{u\times u}{\underset{\sim}{V}_y} = \begin{bmatrix} \underset{\sim}{V}_{y,1} & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{V}_{y,2} & \cdots & \underset{\sim}{0} \\ \cdots & \cdots & \cdots & \cdots \\ \underset{\sim}{0} & \underset{\sim}{0} & \cdots & \underset{\sim}{V}_{y,s} \end{bmatrix} \tag{2.14}$$

estimated covariance matrix $\underset{\sim}{V}_y$ whre $u = \sum_{h=1}^{s} u_h$. If we then form the overall (uxt) design matrix $\underset{\sim}{X}$ and composite parameter vector $\underset{\sim}{\gamma}$ as shown in (2.15)

$$\underset{\sim}{X} = \begin{bmatrix} \underset{\sim}{X}_1 & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{X}_2 & \cdots & \underset{\sim}{0} \\ \cdots & \cdots & \cdots & \cdots \\ \underset{\sim}{0} & \underset{\sim}{0} & \cdots & \underset{\sim}{X}_s \end{bmatrix}, \quad \underset{\sim}{\gamma} = \begin{bmatrix} \underset{\sim}{\gamma}_1 \\ \underset{\sim}{\gamma}_2 \\ \cdots \\ \underset{\sim}{\gamma}_s \end{bmatrix} \tag{2.15}$$

where $t = 2s$ is the total number of parameters, then the exponential decay models in (2.10) for the s populations can be simultaneously expressed by the overall model in (2.16). In this context, the weighted least squares estimator $\underset{\sim}{g}$ for $\underset{\sim}{\gamma}$

$$E_A\{\underset{\sim}{y}\} = \underset{\sim}{X}\,\underset{\sim}{\gamma} \tag{2.16}$$

is given by (2.17) and the corresponding estimate for the covariance matrix $\underset{\sim}{V}_g$ of

$$\underset{\sim}{g} = (\underset{\sim}{X}'\,\underset{\sim}{V}_y^{-1}\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}'\,\underset{\sim}{V}_y^{-1}\,\underset{\sim}{y} \tag{2.17}$$

$\underset{\sim}{g}$ is given by (2.18). Similarly, a goodness of fit statistic for the model

$$\underset{\sim}{V}_g = (\underset{\sim}{X}'\,\underset{\sim}{V}_y^{-1}\,\underset{\sim}{X})^{-1} \tag{2.18}$$

(2.16) is the overall residual sum of squares in (2.19) which has approximately

$$X^2 = SS(E_A\{\underset{\sim}{y}\} = \underset{\sim}{X}\,\underset{\sim}{\gamma}) = \underset{\sim}{y}'\,\underset{y}{V}^{-1}\,\underset{\sim}{y} - \underset{\sim}{g}'(\underset{\sim}{X}'\,\underset{\sim}{V}_y^{-1}\,\underset{\sim}{X})\,\underset{\sim}{g} \tag{2.19}$$

a chi-square distribution with D.F. = (u-t) in large samples under the hypothesis that the model fits. At this point, one can note that for the $\underset{\sim}{X}$ defined by (2.15), the goodness of fit statistic in (2.19) is the sum of the s statistics obtained from (2.13) and hence D.F. = (u-2s).

Under the assumption that the model (2.16) does suitably account for the variation in the vector $\underset{\sim}{y}$, tests of linear hypotheses with respect to the parameters comprising $\underset{\sim}{\gamma}$ can be undertaken. In particular, for a general hypothesis of the form H_0: $\underset{\sim}{C}\underset{\sim}{\gamma} = \underset{\sim}{0}$ where $\underset{\sim}{C}$ is a known (dxt) matrix of full rank $d \leq t$, a suitable test statistic is given by (2.20) which has

$$X^2 = SS(\underset{\sim}{C}\ \underset{\sim}{\gamma} = 0) = \underset{\sim}{g}'\ \underset{\sim}{C}'[\underset{\sim}{C}(X'\ V_y^{-1}\ X)^{-1}\ C']\underset{\sim}{C}\ \underset{\sim}{g} \qquad (2.20)$$

approximately a chi-square distribution with D.F. = d in large samples under H_0. In this manner, the intercept parameters μ_h and slope parameters β_h for the respective populations can be readily compared with one another. The results of these tests can then be used as the basis for the formulation of a final model which reflects the manner in which the exponential decay of bacteria density is influenced by factors pertaining to differences among the populations. This is accomplished by constructing appropriate $\underset{\sim}{X}$ matrices with full rank t (i.e., $\underset{\sim}{\gamma}$ is a vector of t parameters) and fitting (2.16) in this more general context. The results given in (2.17), (2.18), (2.19), and (2.20) may then be directly applied to this revised model. Finally, predicted values $\hat{y}_{hj}$ for the bacteria density at each time point in each population can be determined from this analysis by use of the expression in (2.21) where $\underset{\sim}{\hat{y}}$ is the composite vector of all the $\hat{y}_{hj}$. These values represent

$$\underset{\sim}{\hat{y}} = \underset{\sim}{X}\ \underset{\sim}{g} = \underset{\sim}{X}[X'\ \underset{\sim}{V}_y^{-1}\ \underset{\sim}{X}]^{-1}\ \underset{\sim}{X}'\ \underset{\sim}{V}_y^{-1}\ \underset{\sim}{y} \qquad (2.21)$$

improved estimates of the log bacteria densities which are obtained by applying a fitted model to the results of the entire experiment. Also, the corresponding residuals $(y_{hj} - \hat{y}_{hj})$ indicate possible trouble spots with respect to conclusions about comprehensive relationships.

3. Example

To illustrate the use of this methodology, we shall consider data from Schiemann [1972] who was concerned with decay rates for Leptospira autumnalis. In particular, two intersecting sets of experiments were conducted.

I. Experiments P1 and P2 were undertaken to compare pH = 7.4 with pH = 8.0 at temperature T = 20°C.

II. Experiments T1, T2, and T3 were undertaken to compare T = 20°C with T = 25°C and T = 30°C at pH = 7.4.

Each experiment consisted of a number of serial dilution modules at different time points with each module involving the inoculation of n = 10 tubes at k = 3 different dilutions. The resulting data for the number of positive tubes are shown in Table 1.

The maximum likelihood estimates $\hat{\lambda}$ for each of the respective modules are readily determined from (2.3) by using a simple successive approximation computer program. The corresponding estimates of variance w are obtained from (2.5) by substitution. These results can then be converted to the appropriate log forms y and v as previously noted. Both $\hat{\lambda}$ and y as well as their estimated standard errors ($\sqrt{w}$ and $\sqrt{v}$ respectively) are also shown in Table 1.

TABLE 1

Survival of Leptospira Autumnalis

	Elapsed Time (Hrs.)	No. of Positive Tubes For Dilution Factor 0.1	0.01	0.001	0.0001	Estimated Density $\hat{\lambda}$	Estimated s.e. for $\hat{\lambda}$ $\sqrt{v}$	Log($\hat{\lambda}$) y	Estimated s.e. for y $\sqrt{v}$	Final Model Predicted $\hat{y}$
	30.17		10	8	2	1723.8	608.3	7.45	.35	7.40
	35.92		10	6	2	1086.4	408.8	6.99	.38	7.22
	42.08		10	7	1	1180.8	437.0	7.07	.37	7.02
Experiment P1	48.08		10	4	0	493.2	200.3	6.20	.41	6.84
pH = 7.4	54.33		10	5	0	621.7	256.8	6.43	.41	6.64
T = 20° C	59.92		10	6	0	792.4	318.7	6.68	.40	6.47
	66.00		10	4	1	589.7	243.6	6.38	.41	6.28
	74.83		10	3	0	399.1	154.9	5.99	.39	6.00
	84.17		10	3	0	399.1	154.9	5.99	.39	5.71
	30.00		10	6	2	1086.5	408.8	6.99	.38	6.77
	35.75		10	4	2	699.6	286.7	6.55	.41	6.47
	41.92		10	3	0	399.1	154.9	5.99	.39	6.15
	47.92		10	3	0	399.1	154.9	5.99	.39	5.85
Experiment P2	54.00		10	2	0	329.1	129.9	5.80	.37	5.52
pH = 8.0	59.75		8	4	0	216.1	75.3	5.38	.35	5.22
T = 20° C	65.83		6	1	0	92.2	35.9	4.52	.39	4.90
	74.72		7	1	0	116.2	42.8	4.76	.37	4.43
	84.00		1	0	1	18.9	13.6	2.94	.72	3.95
	94.00		1	0	0	9.4	9.4	2.25	1.00	3.42
	24.08		10	6	1	935.9	363.7	6.84	.39	7.10
	36.08		10	5	0	621.7	256.8	6.43	.41	6.73
	42.25		10	4	0	493.2	200.3	6.20	.41	6.53
	48.08		10	2	1	386.9	149.0	5.96	.39	6.35
	54.17		10	6	0	792.4	318.7	6.68	.40	6.16
	60.17		10	5	0	621.7	256.8	6.43	.41	5.97
Experiment T1	66.25		10	4	1	589.7	243.6	6.38	.41	5.78
pH = 7.4	72.17		9	2	0	272.1	77.5	5.40	.35	5.66
T = 20° C	78.25		7	0	0	99.6	38.0	4.60	.38	5.09
	96.08		7	0	0	99.6	38.0	4.60	.38	4.95
	102.00		8	0	0	127.6	46.2	4.85	.36	4.66
	107.50		5	0	0	59.9	26.7	4.09	.45	4.49
	114.00		5	0	0	59.9	26.7	4.09	.45	4.29
	138.00	10	3	0		39.9	15.5	3.69	.39	3.53
	150.00	10	2	0		32.9	12.2	3.49	.37	3.16
	4.08		10	10	0	2397.9	854.7	7.78	.36	7.66
	12.00		10	9	0	1704.8	601.9	7.44	.35	7.26
	23.92		10	7	0	1012.2	386.7	6.92	.38	6.67
	29.83		10	3	1	473.6	191.0	6.16	.40	6.38
	35.92		10	3	0	399.1	154.9	5.99	.39	6.08
Experiment T2	42.08		10	2	1	386.9	149.0	5.96	.39	5.77
pH = 7.4	47.92		10	0	0	231.2	80.9	5.44	.35	5.48
T = 25° C	60.00		6	2	0	107.1	40.2	4.67	.38	4.89
	66.00		6	0	0	77.8	31.8	4.35	.41	4.59
	71.92		4	1	0	56.4	25.7	4.03	.46	4.29
	80.17		4	1	0	56.4	25.7	4.03	.46	3.89
	88.33		0	1	0	9.1	9.2	2.20	1.02	3.48
	95.92		1	0	0	9.4	9.4	2.25	1.00	3.11
	107.33	9	1	0		19.3	6.7	2.96	.35	2.54
	2.00		10	8	1	1504.6	536.7	7.32	.36	7.69
	3.83		10	10	1	2872.5	1054.1	7.96	.37	7.54
	15.75		10	4	0	493.2	200.3	6.20	.41	6.55
	19.83		10	6	0	792.4	318.7	6.68	.40	6.22
Experiment T3	35.75		9	1	0	192.9	67.1	5.26	.35	4.90
pH = 7.4	41.92		7	0	0	99.6	38.0	4.60	.38	4.39
T = 30° C	47.75		3	0	0	31.5	18.1	3.45	.58	3.91
	53.75		2	1	0	30.0	17.6	3.40	.59	3.41
	65.83		1	0	0	9.4	9.4	2.25	1.00	2.41
	71.75		1	0	0	9.4	9.4	2.25	1.00	1.92
	80.00	2	0	0		2.0	1.4	.69	.70	1.24
	88.17	1	0	0		.9	.9	-.06	1.00	.56

Exponential decay models of the form (2.10) were fitted to the estimated values y for log λ which were associated with each experiment. The resulting goodness of fit X^2-statistics determined from (2.13) were all non-significant (α = .25) and are given with corresponding D.F. in the last two columns of Table 2. Thus, such models adequately describe the relationship between bacteria density and time. The estimates of the intercept and slope parameters for this model together with their estimated standard errors are also shown in Table 2. Finally, tests of significance pertaining to

Table 2

Estimated Parameters and Goodness of Fit Statistics
for Within Experiment Exponential Decay Models

Experiment	Estimate of intercept	Estimated s.e.	Estimate of slope	Estimated s.e.	Residual X^2-statistic	D.F.
P1	7.99	0.42	-0.0254	0.0074	3.26	7
P2	8.79	0.46	-0.0601	0.0082	5.54	8
T1	7.65	0.25	-0.0291	0.0028	11.44	13
T2	7.90	0.21	-0.0502	0.0036	6.19	12
T3	7.97	0.22	-0.0852	0.0059	7.38	10

comparisons among the respective experiments can be undertaken by applying (2.20) to appropriate linear contrasts of the intercept and slope parameters. The corresponding X^2-statistics are shown in Table 3 and suggest the

Table 3

Tests of Hypotheses for Comparisons of Intercept and Slope Parameters

Hypothesis	D.F.	X^2
$\mu_{P1} = \mu_{P2}$	1	1.68
$\mu_{T1} = \mu_{T2} = \mu_{T3}$	2	1.05
$\beta_{P1} = \beta_{T1}$	1	0.22
$\beta_{P1} = \beta_{P2}$	1	9.85
$\beta_{T1} = \beta_{T2} = \beta_{T3}$	2	79.07

following conclusions

1. The intercept parameters of experiments P1 and P2 are the same.
2. The intercept parameters of experiments T1, T2, and T3 are the same.
3. The slope parameters of experiments P1 and T1 are the same. Hence, the decay rates are the same.
4. The slope parameters of experiments P1 and P2 are significantly different ($\alpha = .01$). Hence, it follows that pH has a definite effect on the decay rate of bacteria density.
5. The slope parameters of experiments T_1, T_2, and T_3 are significantly different ($\alpha = .01$). Hence, it follows that temperature has a definite effect on the decay rate of bacteria density.

On the basis of these results, a final model was fitted to the estimated y values for the combined experiments with the respective parameters being

i. a common intercept parameter for experiments P1 and P2
ii. a common intercept parameter for experiments T1, T2, and T3
iii. a common slope parameter β_1 for experiments P1 and T1
iv. a slope parameter β_{P2} for experiment P2
v. a slope parameter β_{T2} for experiment T2
vi. a slope parameter β_{T3} for experiment T3

The goodness of fit statistic from (2.19) for this model was $X^2 = 36.55$ with D.F. = 54 which is non-significant ($\alpha = .25$). Thus, this model provides a satisfactory framework which accounts for the variation of the y's in the combined experiments in the sense of clarifying the roles of pH and temperature on the exponential decay of bacteria density. The estimated parameters and corresponding estimated standard errors are given in Table 4 while tests of significance for various comparisons are given in Table 5.

Table 4

Estimated Parameters and Standard Errors for Final Model

Parameter	Estimate	Estimated s.e.
μ_P	8.34	0.15
μ_T	7.86	0.12
β	-0.0313	0.0017
β_{P2}	-0.0524	0.0035
β_{T2}	-0.0495	0.0027
β_{T3}	-0.0828	0.0046

Table 5

Tests of Hypotheses for Final Model

Hypothesis	D.F.	X^2
$\mu_P = \mu_T$	1	10.40
$\beta_1 = \beta_{P2}$	1	43.91
$\beta_1 = \beta_{T2} = \beta_{T3}$	2	181.45
$\mu_P = \mu_T, \beta_{P2} = \beta_{T2} = \beta_{T3} = \beta_1$	4	251.30
$\mu_P = \mu_T, \beta_{P2} = \beta_{T2} = \beta_{T3} = \beta_1 = 0$	5	640.68

Thus, the conclusions which follow from the analysis of this final model are similar to those derived from the preliminary one. However, it should be noted that the estimated parameters in Table 4 have smaller estimated standard errors than the corresponding ones in Table 2. This gain in precision results from the efficient combination of certain similar features in the variation of the data from the separate experiments. Finally, predicted values $\hat{y}$ based on the final model are shown in the last column of Table 1. These are determined from the matrix equation (2.21) and provide a useful descriptive indication of the extent to which the fitted model approximates the actual y values.

ACKNOWLEDGMENTS

This research was supported by the National Institutes of Health, Institute of General Medical Sciences Grants GM-70004-02 and GM-0038-19. The authors would like to thank Dr. Don Schiemann for providing us with the data used in Section 3.

REFERENCES

Cornell, R. G. and Speckman, J. A. [1967]. Estimation for a simple exponential model. Biometrics 23, 717-737.

Epstein, B. [1967]. Bacterial extinction time as an extreme value phenomenon. Biometrics 23, 835-839.

Finney, D. J. [1964]. Statistical Methods in Biological Assay. 2nd Edn., Section 21.5. Hafner Publishing Company, New York.

Grizzle, J. E., Starmer, C. F., and Koch, G. G. [1969]. Analysis of categorical data by linear models. Biometrics 25, 489-504.

Mather, K. [1949]. The analysis of extinction time data in bioassay. Biometrics 5, 127-143.

McCrady, M. H. [1915]. The numerical interpretation of fermentation-tube results. J. Infectious Diseases 17, 183-212.

Peto, S. [1953]. A dose response equation for the invasion of micro-organisms. Biometrics 9, 320-335.

Schiemann, D. A. [1972]. Viability maintenance in Leptospira autumnalis Akiyami A. Ph.D. dissertation, Department of Environmental Sciences and Engineering, University of North Carolina.