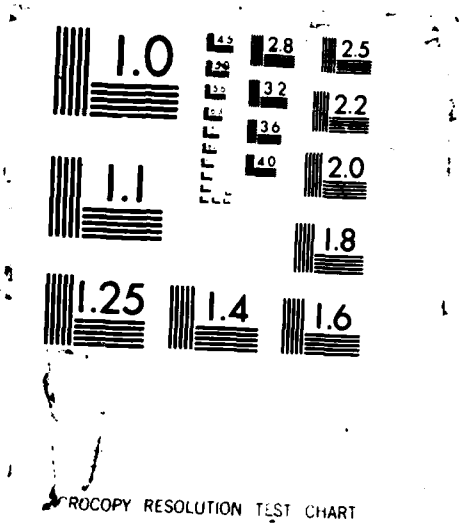
AD-A186 842 LINEAR BAYES ESTIMATORS OF THE POTENCY CURVE IN BIOASSAY(U) COLUMBIA UNIV NEW YORK L KUO 1984 AFOSR-RR-87-1083 $AFOSR-87-0072 1/1

UNCLASSIFIED F/G 6/1 NL


1.0
1.1
1.25
1.4
1.6
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.8
ROCOPY RESOLUTION TEST CHART

DTIC FILE COPY

SECURITY

AD-A186 042

OCUMENTATION PAGE

1a. REPOR UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 87-1083
6a. NAME OF PERFORMING ORGANIZATION Columbia University 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code) Columbia University New York, NY 10027	7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448 BLDG 410
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR 8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR 87-0072
8c. ADDRESS (City, State and ZIP Code) BLDG 410 Bolling AFB DC 20332-6448	10. SOURCE OF FUNDING NOS PROGRAM ELEMENT NO. 61102F; PROJECT NO. 2304; TASK NO. A5; WORK UNIT NO.

11. TITLE (Include Security Classification)
Linear Bayes Estimators of the Potency Curve in Bioassay

12. PERSONAL AUTHOR(S)
Lynn Kuo

13a. TYPE OF REPORT Journal	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB GR	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The prior distribution on the class of potency curves in quantal bioassay is assumed to be Ferguson's Dirichlet distribution. Given the integrated squared error loss and the quantal response data, we derive the Bayes estimator in a linear space generated by the data. Some numerical examples and asymptotic results are also given.

DTIC ELECTE SEP 29 1987 S D

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL B. Woodruff	22b. TELEPHONE NUMBER (Include Area Code) (202) 767- 5027 22c. OFFICE SYMBOL NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE. UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

Linear Bayes Estimators of the Potency Curve in Bioassay

by

LYNN KUO

Division of Statistics, University of California, Davis
Davis, California 95616 U.S.A.

SUMMARY

Tne prior distribution on the class of potency curves in quantal bioassay is assumed to be Ferguson's Dirichlet distribution. Given the integrated squared error loss and the quantal response data, we derive the Bayes estimator in a linear space generated by the data. Some numerical examples and asymptotic results are also given.

Some Key words: Bayesian nonparametric estimation; Linear estimators; Mixtures of Dirichlet processes; Quantal Bioassay.

1. INTRODUCTION

The Bayesian nonparametric approach to estimating the tolerance distribution in quantal bioassay has received some attention. The computational difficulty in evaluating these Bayes estimators has hindered their applications. This paper explores the linear Bayes approach to the bioassay problem. These linear Bayes estimators can be computed easily by using statistical software which has the capability of inverting a matrix.

Let us state the quantal bioassay problem as follows: The experimenter intends to test the potency of a stimulus by giving subjects injections of the stimulus at different levels; namely, he chooses L dosage levels, $t_1,\ldots,t_L$, and treats $n_1,\ldots,n_L$ subjects at these levels respectively. Each subject possesses a fixed tolerance level. If a stimulus exceeds a subject's tolerance level, the subject responds positively. If not, there is no response. Therefore we observe the number of positive responses at each level. These numbers are denoted by $k_1,\ldots,k_L$. The potency curve F is the distribution of tolerance levels; i.e. F is defined by the probability F(t) of getting a positive response to a dosage at level t for all t. The objective of this article is to make inferences about the potency curve F.

The maximum likelihood estimator of F subject to the monotonicity restriction may be found, as in Barlow, Bartholomew, Bremner and Brunk (1972), to be a weighted isotonic regression. If we have some prior information of F, we would expect to do better by a Bayesian nonparametric method. Kraft and Van Eeden (1964) derive an estimate of F by using a tailfree process prior. The Bayes estimator of F using Ferguson's Dirichlet process prior (1973) has been studied by Ramsey (1972), Antoniak (1974), Bhattacharya (1981), Disch (1981), and Ammann (1984). Ramsey discusses the posterior distribution and

its mode. He also presents interesting examples to illustrate the desirable features of the Bayes estimators. Bhattacharya obtains the asymptotic posterior distribution by means of a Markov Chain formulation. Ammann generalizes the Dirichlet prior to a class of priors which uses processes with independent increments.

Antoniak (1974) has treated this problem with integrated loss function $L(F,\hat{F}) = \int (F-\hat{F})^2 dW$ where W is a known weight function. He has shown that the posterior distribution of F given the data is a mixture of Dirichlet process distributions, and he has derived the Bayes estimator of F for two dosage levels. Unfortunately, this mixture becomes increasingly intractable when the number of stimulus levels increases. It is shown by Antoniak, when L = 2, that a double summation (two indices) can be used to express this mixture; when L = 3, six indices of summation are needed. In general, L(L-1) indices are needed to represent this mixture. It already takes a substantial amount of bookkeeping for the mixture even in the case L = 3. Therefore, it is important to search for methods of approximating the Bayes rule.

Disch (1981) proposes two methods for approximating the posterior distribution. One method uses a single Dirichlet to approximate the mixture of Dirichlets. The parameters of the single Dirichlet are defined to be a function of $\sum_{i<\ell} k_i$ and $\sum_{i>\ell} (n_i-k_i)$. The other method is a two-step method. The two steps involve deleting a subset of observational doses and using the above method of single Dirichlet approximation again. As pointed out by Disch, the drawback of the two step method is that the approximation depends on which subset of doses is being deleted.

A linear Bayes estimator is proposed here to approximate the Bayes estimator. Given the loss function $L(F,d) = \int(F(t)-d(t))^2\, dW(t)$, we restrict the decision rules $d(t)$ in the linear space generated by $k_1,\ldots,k_L$ and 1. A linear Bayes rule is the Bayes rule in the linear space. The solution is given by pointwise minimization. For each t, we find constants $\lambda_1,\ldots,\lambda_L,\lambda_0$, depending on t, to minimize $E(F(t) - \lambda_0 - \lambda_1 k_1 - \ldots - \lambda_L k_L)^2$. It is much easier to evaluate the linear Bayes rule than to evaluate the Bayes rule $E(F(t)|k_1,\ldots,k_L)$. The former evaluation requires only partial specification of F.

This approach of approximating Bayes rules in some nonparametric problems by Bayes estimators within the class of linear combinations of a given set of functions on the sample space has been proposed by Goldstein (1975). He applies his result to derive linear polynomial estimators of the mean and moments of the unknown distribution function. However, Goldstein does not specifically treat the bioassay problem. The objective and loss function studied here are also different from Goldstein's. Integrated loss function is used here, since the objective is to estimate the whole distribution function.

To use Ferguson's process indexed by α, the statistician needs to specify α. The measure $\alpha(-\infty,t]$ can be rewritten as $MF_0(t)$, where $M = \alpha(-\infty,\infty)$ and $F_0(t) = \alpha(-\infty,t]/M$. It is known that $EF(t) = F_0(t)$ and $\mathrm{var}(F(t)) = F_0(t)(1-F_0(t))/(M+1)$ where F is the unknown random distribution chosen by the Ferguson's process. The function F_0 represents the statistician's prior belief on the shape of F, and M represents the degree of concentration of F around F_0. Large M indicates that F concentrates more around F_0.

The linear Bayes rule $\hat{d}$ derived in this paper can be summarized as follows. For t to be one of the stimulus levels, say $t = t_j$, the linear Bayes rule $\hat{d}(t_j)$ is a linear combination of $F_0(t_j)$, and $k_i/n_i - F_0(t_i)$ for i=1,...,L. The coefficients are ratios of two determinants of LxL matrices where the entries are functions of M, F_0, and $n_1,\ldots n_L$. For values of t not at the drug levels, say for example $t_j < t < t_{j+1}$, $\hat{d}(t)$ can be interpolated from $\hat{d}(t_j)$ and $\hat{d}(t_{j+1})$ using F_0. Perhaps the most interesting result of this paper is this interpolating formula. It reduced the problem of evaluating $\hat{d}(t)$ for each t to the problem of just evaluating L of them, $\hat{d}(t_1),\ldots,\hat{d}(t_L)$.

One of the drawbacks of the linear Bayes estimator is that it may not be nondecreasing. Fortunately, nonmonotonicity does not occur very often. If monotonic restriction is a great concern to the user, the pool-adjacent-violators algorithm (Barlow, Bartholomew, Bremner and Brunk, 1972, pp. 13-18) can be used for obtaining the isotonic regression on $\hat{d}(t_1),\ldots,\hat{d}(t_L)$. Nevertheless, this rule is not Bayes to the original problem. It is also not clear what to do for t not at the drug levels. An alternative formulation of linear Bayes estimation under order restriction might be of some interest and could be explored later. The present paper adopts the one without an order restriction for computational simplicity. This formulation enables us to derive the interpolating formula given in Theorem 2 for the linear Bayes rule.

The general discussion on the Dirichlet process and the mathematical model for the quantal bioassay is given in Section 2. The derivation of the linear Bayes estimator is given in Section 3. Some asymptotic results and limiting behavior for $M \to 0$ or $M \to \infty$ for the linear Bayes rules are given in Sections 4 and 5. A numerical example comparing the Bayes estimators and

linear Bayes estimators for L = 2 is given in Section 6. A SAS program is given in the appendix to illustrate the ease of computing linear Bayes estimators in practice. The program computes the linear Bayes estimators for all possible observations for $n_1 = n_2 = 3$. The PROC MATRIX procedure which has the capability of inverting a matrix is used. Another numerical example given in Section 7 using the data in Cox (1970, p.9) illustrates the use of linear Bayes estimators in practice.

2. PRELIMINARY WORK

A brief description of the Dirichlet process and the quantal bioassay problem is given here.

Ferguson (1973) constructs a prior on the space of distribution functions. The prior indexed by α chooses a random distribution function P. This is denoted by $P \sim D(\alpha)$. The prior can be defined as follows. Given any measurable partition $B_1,\ldots,B_k$ of R^1, the joint distribution of the random probabilities $P(B_1),\ldots,P(B_k)$ is a Dirichlet distribution with parameters $\alpha(B_1),\ldots,\alpha(B_k)$.

Given the loss function

$$L(F,\hat{F}) = \int (F(t) - \hat{F}(t))^2 \, dW(t),$$

it is shown by Ferguson that the Bayes estimator of F for the no sample problem is

$$\hat{F}(t) = EF(t) = \alpha(-\infty,t]/\alpha(R) = MF_0(t)/M = F_0(t).$$

The Bayes estimator of F, when we observe a sample $x_1,\ldots,x_n$ from F, is given by

$$\hat{F}_n(t) = E(F(t)|x_1,\ldots,x_n) = \{\alpha(-\infty,t] + \sum_{i=1}^{n} I(x_i \le t)\}/(M+n)$$

$$= \frac{M}{M+n} F_0(t) + \frac{n}{M+n} F_n(t) ,$$

where F_n is the sample empirical distribution function.

Recall the bioassay problem, k_i is the number of subjects react positively at the stimulus level t_i. Therefore, the likelihood function is given by

$$L(F) = \prod_{i=1}^{L} \binom{n_i}{k_i} \{F(t_i)\}^{k_i} \{1-F(t_i)\}^{n_i-k_i} . \qquad (2.1)$$

Given the Ferguson's prior $F \sim D(\alpha)$ and $L = 2$, it is shown by Antoniak that the posterior distribution of F is a mixture of Dirichlet processes. Define $\beta_1 = MF_0(t_1)$, $\beta_2 = M(F_0(t_2)-F_0(t_1))$, and $\beta_3 = M(1-F_0(t_2))$. It is shown by Antoniak that the posterior distribution of $F(t_1)$, $F(t_2)-F(t_1)$, is given by

$$\sum_{i=0}^{k_2} \sum_{j=0}^{n_1-k_1} a_{ij} D(\beta_1+k_1+i,\ \beta_2+n_1-k_1+k_2-i-j,\ \beta_3+n_2-k_2+j), \qquad (2.2)$$

where the mixing indices are

$$a_{ij} = b_{ij} / \sum_{i=0}^{k_2} \sum_{j=0}^{n_1-k_1} b_{ij} \qquad \text{and}$$

$$b_{ij} = \binom{n_1-k_1}{j}\binom{k_2}{i} \frac{\Gamma(\beta_1+k_1+i)\ \Gamma(\beta_2+n_1-k_1+k_2-i-j)\ \Gamma(\beta_3+n_2-k_2+j)}{\Gamma(\beta_1)\ \Gamma(\beta_2)\ \Gamma(\beta_3)} .$$

Antoniak has also given a good heuristic explanation for the mixture. While it is known that k_2 observations fall in $(0,t_2]$ and n_1-k_1 observations fall in $(t_1,1]$, we don't know how many of the k_2 fall in $(0,t_1]$ and how many of the n_1-k_1 fall in $(t_2,1]$. Let i denote the number of observations in $(0,t_1]$ arising from k_2, and let j denote the number of observations in $(t_2,1)$ arising

from n_1-k_1. Since i and j are unknown, the posterior distribution is a mixture over all possible values of i and j given by the double summation in the expression (2.2).

The Bayes estimators can be obtained as

$$\hat{F}(t_1) = \sum_{i=0}^{k_2} \sum_{j=0}^{n_1-k_1} a_{ij} \frac{\beta_1 + k_1 + i}{M + n_1 + n_2},$$

$$\hat{F}(t_2) = \sum_{i=0}^{k_2} \sum_{j=0}^{n_1-k_1} a_{ij} \frac{\beta_1 + \beta_2 + n_1 + k_2 - j}{M + n_1 + n_2}.$$

When L = 3 , it takes six summations to describe this mixture. In general it takes L(L-1) summations to describe the mixture. To circumvent the difficulty in computing the Bayes estimators, the linear Bayes estimator is proposed and derived in the next section.

3. DERIVATION OF THE LINEAR BAYES ESTIMATOR

Given the loss function $L(F,d) = \int(F(t)-d(t))^2 dW(t)$, the linear Bayes estimator is obtained by minimizing the function $E(F(t)-d(t))^2$ for each t. Let us first restrict our attention to estimating F(t) at the stimulus levels, namely when $t = t_j$, j=1,...,L. The following Theorem 1 says that $\hat{d}(t_j)$ is actually a sum of $F_0(t_j)$ and a linear combination of $k_i/n_i - F_0(t_i)$, i=1,...,L, where the coefficients depending on $t_1,\ldots,t_L$ can be computed from the determinants of certain covariance matrices.

We need the following two notations for Theorem 1. Let cov(k) denote the covariance matrix of $k_1,\ldots k_L$ and let A(i,t) denote this matrix with the i^{th} column replaced by $[\text{cov}(k_1,F(t)), \text{cov}(k_2,F(t)),\ldots,\text{cov}(k_L,F(t))]^T$, namely

$$A(i,t) = \begin{bmatrix} & & & \underline{i^{th}\ \text{column}} & & \\ \text{var } k_1 & \text{cov}(k_1,k_2) & \cdots & \text{cov}(k_1,F(t)) & \cdots & \text{cov}(k_1,k_L) \\ \text{cov}(k_1,k_2) & \text{var } k_2 & \cdots & \text{cov}(k_2,F(t)) & & \\ \vdots & & & & & \vdots \\ \text{cov}(k_1,k_L) & \cdots & & \text{cov}(k_L,F(t)) & \cdots & \text{var } k_L \end{bmatrix}.$$

Theorem 1: In the bioassay problem, let us assume

(1) the potency curve $F(t)$ is a random distribution chosen from the Dirichlet process with parameter $MF_0(t)$ $(-\infty<t<\infty)$;

(2) $n_1,\ldots,n_L$ subjects are treated at stimulus levels $t_1<t_2<\ldots<t_L$ respectively, and $k_1,\ldots,k_L$ are the observed numbers of subjects that react positively at each level;

(3) the loss function is $L(F,d) = \int(F(t)-d(t))^2\, dW(t)$.

If we restrict our decision rules to the class of linear combinations of $k_1,\ldots,k_L$ and 1 for each t, then the Bayes rule in this class, evaluated at the stimulus levels $t_j, j=1,\ldots,L$ is given as follows:

$$\hat{d}(t_j) = F_0(t_j) + \sum_{i=1}^{L} n_i \hat{\lambda}_i(j)\,(k_i/n_i - F_0(t_i)) \qquad (3.1)$$

where

$$\hat{\lambda}_i(j) = |A(i,t_j)|/|\text{cov } k| \ . \qquad (3.2)$$

$|*|$ denotes the determinant of matrix *.

PROOF. Given $t = t_j$, we seek $\lambda_0(j),\ldots,\lambda_L(j)$ to minimize

$$\phi(\lambda) = E(F(t_j) - \lambda_0(j) - \sum_{i=1}^{L} \lambda_i(j)\, k_i)^2,$$

where E denotes the expectation over the joint distribution of the observations and the parameter $F(t_j)$. Differentiating $\phi(\lambda)$ with respect to $\lambda_0(j),\ldots,\lambda_L(j)$ respectively, we obtain L+1 normal equations:

$$E(F(t_j) - \hat{\lambda}_0(j) - \sum_{i=1}^{L} \lambda_i(j)k_i)\, k_\ell = 0,\ \ell=0,\ldots,L,$$

where $k_0 = 1$.

Therefore, from the equation for $\ell = 0$, we obtain

$$\hat{\lambda}_0(j) = EF(t_j) - \sum_{i=1}^{L} \lambda_i(j)\, E(k_i). \tag{3.3}$$

Substituting this expression for λ_0 into each of the L equations with $\ell=1,\ldots,L$ and solving the system of L linear equations, we obtain

$$\hat{\lambda}_i(j) = |A(i,t_j)|/|\text{cov } k| \quad i=1,\ldots,L.$$

Moreover, from equation (3.3)

$$\hat{\lambda}_0(j) = F_0(t_j) - \sum_{i=1}^{L} \hat{\lambda}_i(j) n_i F_0(t_i),$$

$$\hat{d}(t_j) = \hat{\lambda}_0(j) + \sum_{i=1}^{L} \hat{\lambda}_i(j)k_i$$

$$= F_0(t_j) + \sum_{i=1}^{L} n_i\hat{\lambda}_i(j)\, (k_i/n_i - F_0(t_i)).$$

Remark: The coefficients $\hat{\lambda}_1(j),\ldots,\hat{\lambda}_L(j)$ could also be derived by matrix multiplication, i.e.

$$\begin{bmatrix} \hat{\lambda}_1(j) \\ \vdots \\ \hat{\lambda}_L(j) \end{bmatrix} = (\text{cov } k)^{-1} \begin{bmatrix} \text{cov}(k_1, F(t_j)) \\ \vdots \\ \text{cov}(k_L, F(t_j)) \end{bmatrix} . \tag{3.4}$$

Let us compute the entries of cov k, and $A(i,t_j)$ which are needed to evaluate $\hat{\lambda}_i(j)$.

For $i=1,\ldots,L$,

$$\begin{aligned}
\text{var } k_i &= E[\text{var}(k_i|F)] + \text{var}[E(k_i|F)] \\
&= E[n_i F(t_i)(1 - F(t_i))] + \text{var}[n_i F(t_i)] \\
&= n_i F_0(t_i) - n_i F_0(t_i)(MF_0(t_i) + 1)/(M+1) + n_i^2 F_0(t_i)(1-F_0(t_i))/(M+1) \\
&= n_i(M+n_i) F_0(t_i)(1-F_0(t_i))/(M+1) ,
\end{aligned}$$

$$\begin{aligned}
\text{cov}(k_i, k_j) &= Ek_i k_j - Ek_i Ek_j \\
&= n_i n_j EF(t_i)F(t_j) - n_i n_j F_0(t_i)F_0(t_j) \\
&= \begin{cases} n_i n_j F_0(t_i)(1-F_0(t_j))/(M+1) & \text{if } i < j, \\ n_i n_j F_0(t_j)(1-F_0(t_i))/(M+1) & \text{if } i > j; \text{ and} \end{cases}
\end{aligned}$$

$$\begin{aligned}
\text{cov}(k_i, F(t_j)) &= Ek_i F(t_j) - Ek_i EF(t_j) \\
&= n_i EF(t_i)F(t_j) - n_i EF(t_i)\, EF(t_j) \\
&= \begin{cases} n_i F_0(t_i)(1-F_0(t_j))/(M+1) & \text{if } i < j, \\ n_i F_0(t_j)(1-F_0(t_i))/(M+1) & \text{if } i > j. \end{cases}
\end{aligned} \tag{3.5}$$

Next, we want to estimate $F(t)$ for t other than the drug levels, for example $t_j<t<t_{j+1}$. We can find $\hat{d}(t)$ for each t by the same method as in Theorem 1, where t_j is replaced by t in equations (3.1) and (3.2). If $\hat{d}(t)$ can easily be derived from F_0, $\hat{d}(t_j)$, and $\hat{d}(t_{j+1})$, then time and effort can be saved in estimating $F(t)$, since we no longer need to compute the determinant of $A(i,t)$ for each t. The following theorem enables us to linearly interpolating $\hat{d}(t)$ by using $\hat{d}(t_j)$ and $\hat{d}(t_{j+1})$.

Theorem 2: With the assumptions of Theorem 1, and restriction of decision rules to linear combinations of $k_1,\ldots,k_L$, and 1, the Bayes rule for $F(t)$ is given by:

$$\hat{d}(t) = \frac{F_0(t_{j+1})-F_0(t)}{F_0(t_{j+1})-F_0(t_j)}\,\hat{d}(t_j) + \frac{F_0(t) - F_0(t_j)}{F_0(t_{j+1}) - F_0(t_j)}\,\hat{d}(t_{j+1})$$

$$\text{for } t_j < t < t_{j+1},\ j=0,\ldots,L \quad \text{with} \quad F_0(t_0) = 0 = \hat{d}(t_0),$$

$$F_0(t_{L+1}) = 1 = \hat{d}(t_{L+1}).$$

PROOF. By the same method as Theorem 1, we obtain

$$\hat{d}(t) = F_0(t) + \sum_{i=1}^{L} n_i\,\hat{\lambda}_i(t)\,(k_i/n_i - F_0(t_i))$$

$$= \hat{\lambda}_0(t) + \sum_{i=1}^{L} \hat{\lambda}_i(t) k_i, \qquad (3.6)$$

where $\hat{\lambda}_0(t) = F_0(t) - \sum_{i=1}^{L} \hat{\lambda}_i(t)\, n_i\, F_0(t_i)$, and $\hat{\lambda}_i(t), i=1,\ldots,L$ is obtained from the right side of the equation (3.4) with $F(t_j)$ replaced by $F(t)$. Note that the entries $\mathrm{cov}(k_i,F(t))$ can be computed similarly as in equation (3.5).

$$\operatorname{cov}(k_i,F(t)) = \begin{cases} n_i\, F_0(t_i)(1-F(t))/(M+1), & \text{if } t_i < t, \\ n_i\, F_0(t)(1-F_0(t_i))/(M+1), & \text{if } t_i > t. \end{cases}$$

For $t_j < t < t_{j+1}$, it can be verified that for $i=1,\ldots,L$

$$\operatorname{cov}(k_i,F(t)) = \operatorname{cov}(k_i,F(t_j))\frac{F_0(t_{j+1})-F_0(t)}{F_0(t_{j+1})-F_0(t_j)} + \operatorname{cov}(k_i,F(t_{j+1}))\frac{F_0(t)-F_0(t_j)}{F_0(t_{j+1})-F_0(t_j)}. \quad (3.7)$$

Then by a straightforward matrix multiplication using equations (3.4), (3.6) and (3.7), we can show

$$\hat{d}(t) = \frac{F_0(t_{j-1})-F_0(t)}{F_0(t_{j-1})-F_0(t_j)}\hat{d}(t_j) + \frac{F_0(t)-F_0(t_j)}{F_0(t_{j+1})-F_0(t_j)}\hat{d}(t_{j+1}).$$

4. ASYMPTOTIC RESULTS FOR $\hat{d}(t)$.

By evaluating $\hat{\lambda}_i(j)$, we are able to show $\hat{d}(t_j)$ is an asymptotically unbiased and consistent estimator of $F(t_j)$.

Theorem 3: Given the assumptions of Theorem 1, let $n_i \to \infty$ for all $i=1,\ldots,L$. Then $E[\hat{d}(t_j)] \to F(t_j)$, and $\hat{d}(t_j)$ is a weakly consistent estimator of $F(t_j)$ for $j=1,\ldots,L$.

PROOF. From equation (3.1), we have:

$$\hat{d}(t_j) = F_0(t_j) + \sum_{i=1}^{L} n_i\hat{\lambda}_i(j)(k_i/n_i - F_0(t_i)). \quad (4.1)$$

It can be verified, by evaluating equation (3.2), that

$$n_i\hat{\lambda}_i(j) = \begin{cases} O(1/n_j) & \text{if } i \neq j, \\ 1 + \sum_{i=1}^{L} O(1/n_i) & \text{if } i = j. \end{cases}$$

Therefore,

$$E(\hat{d}(t_j)|F) = F_0(t_j) + \sum_{i=1}^{L} n_i\hat{\lambda}_i(j)(F(t_i) - F_0(t_i))$$

$$= F_0(t_j) + \Big(1 + \sum_{i=1}^{L} O(1/n_i)\Big)(F(t_j) - F_0(t_j))$$

$$+ \sum_{i\neq j}^{L} O(1/n_j)(F(t_i) - F_0(t_i))$$

$$\to F(t_j) \text{ as } n_i \to \infty \quad \text{for } i=1,\ldots,L. \tag{4.2}$$

Then, from equations (4.1) and (4.2),

$$\text{var}(\hat{d}(t_j)|F) = E\left[\sum_{i=1}^{L} n_i\hat{\lambda}_i(j)((k_i/n_i) - F(t_i))\right]^2$$

$$= \sum_{i=1}^{L} n_i^2\hat{\lambda}_i(j)^2 \,\text{var}(k_i/n_i)$$

$$+ \sum\sum_{i\neq\ell} n_i\hat{\lambda}_i(j)\, n_\ell\hat{\lambda}_\ell(j)\, \text{cov}(k_i/n_i, k_\ell/n_\ell)$$

$$= (1 + \sum_{i=1}^{L} O(1/n_i))^2 \cdot F(t_j)(1-F(t_j))/n_j$$

$$+ \sum_{i \neq j} O(1/n_j^2) \cdot F(t_i)(1 - F(t_i))/n_i$$

$$\to 0 \text{ as } n_i \to \infty \text{ for } i=1,\ldots,L.$$

(Note that given F, k_i is independent of k_ℓ.)

Therefore,

$\hat{d}(t_j)$ is a consistent estimator of $F(t_j)$.

5. LIMITING BEHAVIOUR

Given the case $M \to 0$, we will show $\hat{d}(t_j) \to k_j/n_j$. As $M \to \infty$, a straightforward evaluation of $\hat{\lambda}_i(j)$ will show that $\hat{\lambda}_i(j) \to 0$. Also, $\hat{d}(t_j) \to F_0(t_j)$. These results are expected. The evaluation for $M \to \infty$ is omitted.

Corollary 1: From the result of Theorem 1,

$$\lim_{M \to 0} \hat{d}(t_j) = k_j/n_j \quad \text{for} \quad j=1,\ldots,L. \tag{5.1}$$

PROOF. First, let us show, for a fixed j:

$$\lim_{M \to 0} \hat{\lambda}_i(j) = \begin{cases} 0 & \text{if } i \neq j \\ 1/n_j & \text{if } i=j \end{cases} \quad \text{for } i=1,\ldots,L, \tag{5.2}$$

where $\hat{\lambda}_i(j)$ are defined by equation (3.2).

We need to consider the following three cases:

(1) If $i < j$, then $\hat{\lambda}_i(j) = |A(i,t_j)|/|\text{cov } k|$.

When $M \to 0$, the j^{th} column is proportional to the i^{th} column in the numerator.

Therefore, $\lim_{M \to 0} \hat{\lambda}_i(j) = 0$.

(2) If $i > j$, then $\lim_{M \to 0} \hat{\lambda}_i(j) = 0$.

This is verified in a similar manner to (1).

(3) If $i=j$, then $\hat{\lambda}_j(j) = |A(j,t_j)|/|\text{cov } k|$.

When $M \to 0$, we can factor n_j from the j^{th} column of the denominator, and show the ratio of the two remaining determinants $\to 1$. Therefore, $\lim_{M\to 0} \hat{\lambda}_j(j) = 1/n_j$.

From equation (3.1), we have

$$\lim_{M\to 0} \hat{d}(t_j) = \lim_{M\to 0} \left[F_0(t_j) + \sum_{i=1}^{L} n_i \hat{\lambda}_i(j)(k_i/n_i - F_0(t_i))\right]$$

$$= F_0(t_j) + (k_j/n_j - F_0(t_j)) = k_j/n_j.$$

REMARK:

Antoniak has given an example for $M \to 0$ with $n_1 = n_2 = 100$, $k_1 = 1, k_2 = 99$, $t_1 = 1/3$, $t_2 = 2/3$, and $F_0(t) = t \in [0,1]$. He obtains the Bayes estimate $\hat{F}(1/3) = \hat{F}(2/3) = 1/2$ and $\hat{F}(t)$ is piecewise linear on $[0,1]$ ($\hat{F}(0) = 0$, $\hat{F}(1) = 1$). The linear Bayes estimate in this case is given by $\hat{d}(1/3) = 1/100$, $\hat{d}(2/3) = 99/100$ and $\hat{d}(t)$ is piecewise linear ($\hat{d}(0) = 0$, $\hat{d}(1) = 1$). Paradoxically, the linear Bayes estimate $\hat{d}$ in this case is intuitively more reasonable than the usual Bayes rule $\hat{F}$.

6. COMPARISON OF LINEAR BAYES TO BAYES ESTIMATES

Comparisons between Bayes and linear Bayes estimates are made in this section. For L=2, the Bayes estimator of F can be evaluated as in Section 2. Let us assume $n_1 = n_2 = 3$, i.e. three subjects are tested at each stimulus level. The linear Bayes estimates and the Bayes estimates evaluated at the two drug levels with different M are given in Table 1 and Table 2 for all the possible values of k_1 and k_2. The prior probabilities for each of the events are also

given in the last column. The prior probabilities are evaluated by E(L(F)) with L(F) given by equation (2.1) with appropriate n_1, n_2, k_1 and k_2. Both tables use the same F_0, standard normal distribution, M=1 and M=10 are used respectively in Table 1 and Table 2. The stimulus levels are administered at -1 and 1. It can be seen that the linear Bayes estimates approximate the Bayes estimates fairly well for M=10. For M=1, the approximation is satisfactory for the cases $k_1 \leq k_2$. The worst case occurs for the case $k_1 = 3$ and $k_2 = 0$. Fortunately, the prior probability for this event to occur is only 0.0002.

7. NUMERICAL EXAMPLE

The data for this example are taken from Cox (1970 p. 9). In all, 150 subjects are tested at 5 different concentrations of the stimulus with 30 subjects at each of the dosages. The numbers of deaths are recorded respectively at the dosage levels. Three different priors, N(2,0.5), N(2,1) and N(2,4) with various M are used for computing the linear Bayes estimates evaluated at the five dosages.

It can be observed that the linear Bayes estimators are not monotonic in M. As M increases, the linear Bayes estimates approach the prior guess F_0 as expected. When M=1, the linear Bayes estimates are relatively insensitive to the choice of the prior guess. They are all close to the isotonic maximum likelihood estimates.

8. CONCLUSIONS

The nonparametric Bayesian method is applied to estimate the potency curve in the quantal bioassay problem. To circumvent the computational difficulty in evaluating the Bayes rule, the linear Bayes estimator is proposed. The linear Bayes estimate can be obtained easily by linearly interpolating the estimates evaluated at the dosage levels. To obtain the estimates at the dosage levels, it is required to invert a LxL matrix which can be done by many standard statistical packages. Some numerial examples are given to demonstrate the ease of obtaining the linear Bayes estimate.

Table 1. Comparison between linear Bayes and Bayes estimates

M=1 $F_0(t_1) = 0.1587$ $F_0(t_2) = 0.8413$

Observations		Linear Bayes		Bayes		Prior Probability
k_1	k_2	$\hat{d}(t_1)$	$\hat{d}(t_2)$	$\hat{F}(t_1)$	$\hat{F}(t_2)$	pr
0	0	0.0101	0.2089	0.0227	0.1964	0.0599
0	1	0.0221	0.4572	0.0372	0.4422	0.0632
0	2	0.0342	0.7055	0.0419	0.6978	0.0925
0	3	0.0462	0.9538	0.0415	0.9585	0.5180
1	0	0.2584	0.2209	0.1655	0.3138	0.0049
1	1	0.2704	0.4692	0.2368	0.5029	0.0092
1	2	0.2825	0.7175	0.2855	0.7145	0.0163
1	3	0.2945	0.9658	0.3022	0.9581	0.0925
2	0	0.5067	0.2330	0.3084	0.4313	0.0012
2	1	0.5187	0.4813	0.4085	0.5916	0.0037
2	2	0.5308	0.7296	0.4971	0.7632	0.0092
2	3	0.5428	0.9779	0.5578	0.9628	0.0632
3	0	0.7550	0.2450	0.4512	0.5488	0.0002
3	1	0.7670	0.4933	0.5687	0.6916	0.0012
3	2	0.7791	0.7416	0.6862	0.8345	0.0049
3	3	0.7911	0.9899	0.8036	0.9773	0.0599

Table 2. Comparison between linear Bayes and Bayes estimates

M=10 $F_0(t_1) = 0.1587$ $F_0(t_2) = 0.8413$

Observations		Linear Bayes		Bayes		Prior Probability
k_1	k_2	$\hat{d}(t_1)$	$\hat{d}(t_2)$	$\hat{F}(t_1)$	$\hat{F}(t_2)$	pr
0	0	0.0941	0.6431	0.0992	0.6380	0.0078
0	1	0.1052	0.7195	0.1087	0.7161	0.0523
0	2	0.1164	0.7960	0.1178	0.7946	0.1824
0	3	0.1276	0.8724	0.1264	0.8736	0.3823
1	0	0.1705	0.6543	0.1617	0.6631	0.0028
1	1	0.1817	0.7307	0.1767	0.7357	0.0207
1	2	0.1929	0.8071	0.1913	0.8087	0.0797
1	3	0.2040	0.8836	0.2054	0.8822	0.1824
2	0	0.2469	0.6655	0.2242	0.6882	0.0006
2	1	0.2581	0.7419	0.2444	0.7556	0.0048
2	2	0.2693	0.8183	0.2643	0.8233	0.0207
2	3	0.2805	0.8948	0.2839	0.8913	0.0523
3	0	0.3234	0.6766	0.2867	0.7133	0.0001
3	1	0.3345	0.7531	0.3118	0.7758	0.0006
3	2	0.3457	0.8295	0.3369	0.8383	0.0028
3	3	0.3569	0.9059	0.3620	0.9008	0.0078

Table 3. Linear Bayes estimates for Cox's data

t_i Log_2(Concentration)	0	1	2	3	4
No. of death	2	8	15	23	27
k_i/n_i	0.067	0.267	0.500	0.767	0.900
(a) $F_0 = N(2,0.5)$					
$F_0(t_i)$	0.0000	0.0228	0.5	0.9773	1.000
M	$\hat{d}(0)$	$\hat{d}(1)$	$\hat{d}(2)$	$\hat{d}(3)$	$\hat{d}(4)$
1	0.0645	0.2606	0.5005	0.7706	0.9032
30	0.0334	0.1594	0.5022	0.8491	0.9499
100	0.0155	0.0894	0.5007	0.9125	0.9769
300	0.0061	0.0497	0.5002	0.9506	0.9909
500	0.0038	0.0397	0.5001	0.9605	0.9943
(b) $F_0 = N(2,1)$					
$F_0(t_i)$	0.0228	0.1587	0.5	0.8413	0.9772
1	0.0656	0.2633	0.5007	0.7676	0.9024
30	0.0476	0.2149	0.5026	0.7939	0.9375
100	0.0349	0.1854	0.5011	0.8172	0.9584
300	0.0277	0.1694	0.5003	0.8313	0.9697
500	0.0259	0.1654	0.5002	0.8350	0.9725
(c) $F_0 = N(2,4)$					
$F_0(t_i)$	0.3085	0.4013	0.5	0.5987	0.6915
1	0.0794	0.2692	0.5019	0.7576	0.8922
30	0.2115	0.3380	0.5020	0.6667	0.7829
100	0.2675	0.3741	0.5008	0.6276	0.7307
300	0.2931	0.3910	0.5003	0.6097	0.7064
500	0.2990	0.3949	0.5002	0.6055	0.7007

ACKNOWLEDGEMENT

This work was partially supported by the U.S. Air Force Office of Scientific Research. The author would like to thank Professor Thomas Ferguson and Dr. Michael P. Cohen for discussions and helpful comments on this work.

REFERENCES

Ammann, L.P. (1984). Bayesian nonparametric inference for quantal response data. Ann. Statist., 12, 636-645.

Antoniak, C.E. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems, Ann. Statist. 2, 1152-1174.

Barlow, R.E., Bartholomew, D.J., Bremner, J.M, Brunk, H.D. (1972). Statistical Inference Under Order Restrictions, New York: John Wiley and Sons.

Bhattacharya, P.K. (1981). Posterior distribution of a Dirichlet process from quantal response data. Ann. Statist. 9, 803-811.

Cox, D.R. (1970). Analysis of Binary Data, London: Chapman and Hall.

Disch, D. (1981). Bayesian nonparametric inference for effective doses in quantal-response experiment. Biometrics 37, 713-722.

Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. Ann. Statist. 1, 209-230.

Goldstein, M. (1975). Approximate Bayes solutions to some nonparametric problems. Ann. Statist. 3, 512-517.

Goldstein, M. (1975). A note on some Bayesian nonparametric estimates. Ann. Statist. 3, 736-740.

Kraft, C. and van Eeden, C. (1964). Bayesian bio-assay. Ann. Statist. 35, 886-890.

Ramsey, F.L. (1972). A Bayesian approach to bio-assay. Biometrics 28, 841-858.

APPENDIX: SAS source program for computing the linear Bayes estimates.

```
DATA OBS;
INPUT O1-O2;
CARDS;
0 0
0 1
0 2
.
.
.
3 2
3 3
;
DATA SIZE;
INPUT S1-S2;
CARDS;
3 3
.
.
.
3 3
;
DATA PARAM;
P1=PROBNORM(-1);
P2=PROBNORM(1);
M=1;
PROC MATRIX;
FETCH O DATA=OBS;
FETCH N DATA=SIZE;
FETCH B DATA=PARAM;
M=B(1,3);
L=2;
DO K=1 TO 16;
   X=J(L,L);
   C=L(L,L);
   A=J(L,L);
   LAMDA=J(1,L);
   EST=J(1,L);
 DO I=1 TO L;
   X(I,I)=N(K,I)*(M+N(K,I))*B(1,I)*(1-B(1,I))#/(M+1);
   END;
```

APPENDIX: Contined....

```
DO I=2 TO L;
  DO J=1 TO I-1;
  X(I,J)=N(K,I)*N(K,J)*B(1,J)*(1-B(1,I))#/(M+1);
  X(J,I)=X(I,J);
  END;
  END;
DO J=1 TO L;
  DO I=1 TO L;
  IF I<=J THEN DO;
  C(I,J)=N(K,I)*B(1,I)*(1-B(1,J))#/(M+1);
  END;
  ELSE DO;
  C(I,J)=N(K,I)*B(1,J)*(1-B(1,I))#/(M+1);
  END;
  END;
END;
DO J=1 TO L;
A(,J)=INV(X)*C(,J);
BB=B(1,1 2);
LAMDA(1,J)=B(1,J)-(A(,J)#N(K,)')'*BB';
EST(1,J)=A(,J)'*O(K,)'+LAMDA(1,J);
END;
OK=O(K,);
PRINT OK A LAMDA EST;
FREE X C A LAMDA EST;
END;
```

END

DATE
FILMED

DEC.

1987