NONPARAMETRIC BAYESIAN BIOASSAY INCLUDING ORDERED POLYTOMOUS RESPONSE

by

Alan E. Gelfand

Lynn Kuo

TECHNICAL REPORT No. 458

SEPTEMBER 14, 1992

Prepared Under Contract

N00014-92-J-1254 ((NR-042-267)

FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4065

NONPARAMETRIC BAYESIAN BIOASSAY INCLUDING ORDERED POLYTOMOUS RESPONSE

by

Alan E. Gelfand

Lynn Kuo

TECHNICAL REPORT No. 458

SEPTEMBER 14, 1992

Prepared Under Contract

N00014-92-J-1254 ((NR-042-267)

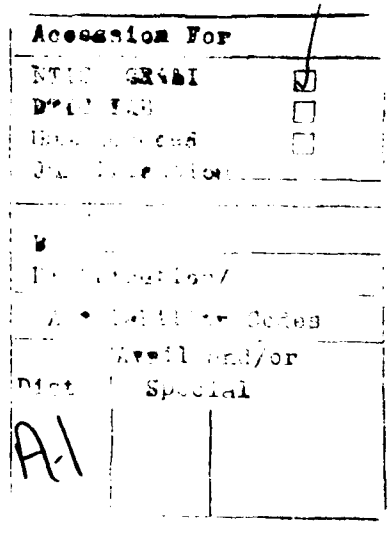
FOR THE OFFICE OF NAVAL RESEARCH

Professor Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA 94305-4065



NONPARAMETRIC BAYESIAN BIOASSAY INCLUDING ORDERED POLYTOMOUS RESPONSE

by

Alan E. Gelfand Lynn Kuo

Department of Statistics, U–120
University of Connecticut
Storrs, Connecticut 06269–3120, U.S.A.

SUMMARY

Previous attempts at implementing fully Bayesian nonparametric bioassay have enjoyed limited success due to computational difficulties. We show here how this problem may be generally handled using a sampling based approach to develop desired marginal posterior distributions and their features. A useful extension is presented which treats the case of ordered polytomous response. Illustrative examples are provided.

Key Words: bioassay, Dirichlet process prior, Gibbs sampler, potency curve, product Beta prior.

1. Introduction

We first formalize the quantal bioassay problem as follows. An experimenter wishes to investigate the potency of a stimulus by administering it at k dosage levels, $t_1, t_2,...,t_k$. A total of n_i subjects are treated at level t_i with the number of positive responses obtained denoted by X_i, $i=1,2,...,k$. The potency curve F is the distribution of tolerance levels, that is $F(t)$ is the probability of achieving a positive response at dosage level t. We seek to make inference about F. Parametric models as, for example, discussed in Finney (1978) typically specify F as a family of logit or probit forms. We take a nonparametric setting only assuming F to be an arbitrary right–continuous non–decreasing function whose range is $[0,1]$.

Taking a Bayesian inferential framework, we note that the fully Bayesian nonparametric approach to estimating the tolerance distribution in a quantal bioassay has not previously been implemented due to computational difficulties. Antoniak (1974) has shown that, under Ferguson's Dirichlet process prior (1973), the posterior distribution of the potency curve is a mixture of Dirichlet process distributions. Unfortunately this mixture becomes increasingly intractable as the number of observed dosage levels increases. In her unpublished dissertation, of M.N. Wesley abandons calculation of the posterior expectation of the potency curve at the observed dosage levels because of such difficulties, estimating, instead, the posterior mode. Kuo (1988), again in deference to these difficulties, obtains linear Bayes estimates under squared error loss. Related previous Bayesian work includes Kraft and Van Eeden (1964), Ramsey (1972), Bhattacharya (1981), Disch (1981) and Ammann (1984).

This paper has two objectives. First, we show how fully Bayesian analysis to obtain for any t, the posterior distribution of $F(t)$ given the data $X = (X_1, X_2,...,X_k)$ can be straightforwardly implemented under two rich classes of prior specification. In addition, features of these distributions such as mode, expectations and quantiles can be readily obtained. Second, we provide a useful extension of the quantal bioassay model which allows ordered polytomous response arising from stochastically ordered potency curves. Missing data can be readily handled. To our knowledge no literature on Bayesian approaches to this problem exists.

The format is the following. In Section 2 we develop the details to implement the first objective. Section 3 provides an illustrative example using the Dirichlet process prior with a data set from Cox and Snell (1989). In Section 4 we develop the aforementioned extension while in Section 5 we provide an illustrative example adapted from Maxwell (1961) for the case of two ordered response levels utilizing a product–Beta class of priors. Section 6 offers brief summary and discussion.

2. Models, Distribution Theory and Inference For the Basic Bioassay Problem

We assume the responses, X_i, to the dosage levels t_i are independently distributed as Binomial $Bi(n_i,p_i)$ with $p_i = F(t_i)$ where $F(t)$ is an unknown underlying potency curve yielding the probability of a response at dosage level t. Hence the likelihood at $X = x$ is

$$L(p;x) = \prod_{i=1}^{k} \binom{n_i}{x_i} p_i^{x_i} (1-p_i)^{n_i-x_i} \qquad (1)$$

where $p = (p_1,\ldots,p_k)$.

Interest focuses on inference regarding the p_i's and the function F. Often F is given a parametric form $F(t;\theta)$ such as a scale and location logistic or probit curve. This literature is large. See, for example, Finney (1978) for discussion and references. Here we assume F belongs to the nonparametric class of right–continuous non–decreasing functions taking values in $[0,1]$.

The Bayesian framework requires specification of a prior which represents our beliefs regarding $F(t)$. Practically this necessitates the specification of a joint measure for p with density denoted by $h(p)$ which by assumptions on F is over the set $S^k = \{p: 0 \le p_1 \le \ldots \le p_k \le 1\}$. One family of priors which has been discussed in this context in the literature is the Dirichlet process prior introduced by Ferguson (1973). Kuo (1988) provides a recent summary of this discussion. This specification assumes that F is close to some given F_0 with closeness quantified by a given precision M. More precisely, for any t, the induced prior on $F(t)$ is a Beta distribution, $Be[MF_0(t), M\{1-F_0(t)\}]$. Since $E\{F(t)\} = F_0(t)$ and $Var\{F(t)\} = F_0(t)\{1-F_0(t)\}/(M+1)$, F_0 and M have interpretations which facilitate their specification from prior information. In particular F_0 is usually taken to be a standard distribution whose median agrees with our prior guess for the LD50 and whose spread provides rough agreement with our prior expectations at other dosage levels. M is chosen to reflect our confidence in F_0 in accord with the extent of our prior experience. In practice we might experiment with several choices of M to see the effect on posterior features of interest.

The induced prior on p is an ordered Dirichlet

$$\pi_D(p)=\frac{\Gamma\left(\sum_{i=1}^{k+1}\gamma_i\right)}{\prod_{i=1}^{k+1}\Gamma(\gamma_i)}\cdot p_1^{\gamma_1-1}(p_2-p_1)^{\gamma_2-1}\cdots(p_k-p_{k-1})^{\gamma_k-1}(1-p_k)^{\gamma_{k+1}-1} \qquad (2)$$

where $\gamma_i = M\{F_o(t_i) - F_o(t_{i-1})\}$, i=1,...,k+1. Here we assume $t_o = -\infty$, $F_o(t_o) = 0$, $t_{k+1} = \infty$, $F_o(t_{k+1}) = 1$, so that $\Sigma\gamma_i = M$. Marginal and conditional distributions and expectations under (2) are routine.

A second class of priors for p is the product–Beta. This family has been discussed in unpublished work of L. Sharples and takes the form

$$\pi_B(p)=c_k(\alpha,\beta)\prod_{i=1}^{k} p_i^{\alpha_i-1}(1-p_i)^{\beta_i-1}, \quad \alpha_i>0,\ \beta_i>0 \qquad (3)$$

where $\alpha = (\alpha_1,...,\alpha_k)$, $\beta = (\beta_1,...,\beta_k)$ and c_k is the normalizing constant under restriction to S^k. Sharples shows that c_k can be expressed as a finite multidimensional summation.

Note that π_B offers mathematical convenience in that it is conjugate with respect to (1) while π_D is not. The family (3) is a flexible class of priors but it is not clear how to select α and β in accord with prior information. That is, unlike π_D, π_B is not induced in any obvious way from a prior on F since it is specified within dosage levels rather than across them. However, π_B can be chosen to reflect to prior information about F in terms of a given F_o and given precision. Letting $e^{(i)}$ denote a row vector having value 1 at the i^{th} coordinate, 0's elsewhere, expectations involving the p_i may be formally given, e.g., $E_B(p_i) = c_k(\alpha, \beta)/c_k(\alpha + e^{(i)}, \beta)$. This suggests equating $E_B(p_i) = F_o(t_i)$, i = 1,...,k. Moreover, $M_i \equiv \alpha_i + \beta_i$ can be viewed as a precision

parameter analogous to M above. The magnitude of M_i reflects our confidence in the value $F_0(t_i)$. With specification of M_i we obtain k equations in k unknowns. Unfortunately, explicit calculation of the $E_B(p_i)$ will be infeasible except in very special cases making solution of this system of equations virtually impossible. However, the conditional distribution of p_i given p_j, $j \neq i$ is obviously a Beta distribution, $Be(\alpha_i, \beta_i)$ restricted to $[p_{i-1}, p_{i+1}]$. Though, again, the mean of this conditional distribution will not be available explicitly the mode is readily obtained It is $\rho_i \equiv (\alpha_i - 1)/(M_i - 2)$ provided $\alpha_i > 1$, $\beta_i > 1$ and $p_{i-1} \leq \rho_i \leq p_{i+1}$. Taking ρ_i as an approximation to the marginal mode for p_i we equate it to $F_0(t_i)$ whence $\alpha_i = (M_i - 2)\, F_0(t_i) + 1$ and $\beta_i = M_i - \alpha_i$.

Objects of primary interest for Bayesian inference are the marginal posterior distributions of $p_i|X$ and, at a specified t, of $F(t)|X$. For such distributions the mean or the mode provide point estimates while appropriate quantiles provide interval estimates. As noted in the introduction, computation of such distributions and estimates has proved very difficult. However the sampling based approach discussed in the context of hierarchical Bayes models in recent papers by Gelfand and Smith (1990), and Gelfand et. al. (1990) is ideally suited to this problem. This approach, known as the Gibbs sampler, is an iterative Markovian updating scheme which dates at least to Metropolis et al (1953). We do not review details here merely remarking that implementation requires sampling from so called complete conditional distributions. For the remainder of this section we develop these distributions under both π_B and π_D and indicate how the desired posterior density estimates and features are obtained.

We note that in our illustrations sampling is conducted with v

independent parallel replications each taken to r iterations. Choice of v determines how close our density estimate is to the exact density at the r^{th} iteration the order of convergence being $0(v^{-1})$. Choice of r determines how close the latter density is to the actual marginal posterior density with convergence at an exponential rate (Geman and Geman, 1984; Tanner and Wong, 1987). Settings for r and v to achieve smooth converged estimates vary with the application and require diagnostic assessment as in Gelfand et. al., (1990). For the examples in sections 3 and 5 r = 20 and v = 1000.

Each of the priors (2) and (3) requires a minor twist to facilitate implementation of the sampling approach. For π_D, although (1) and (2) are not conjugate with respect to the p_i, introduction of a set of unobserved multinomial variables simplifies the required sampling. Let $Z_i = (Z_{i1},...,Z_{i,k+1})$ ~ Mult(n_i, λ), i=1,...,k, where $\lambda=(\lambda_1,...,\lambda_{k+1})$ with $\lambda_j = p_j - p_{j-1}$, $p_0 \equiv 0$, $p_{k+1} \equiv 1$. The variable Z_{ij} denotes, amongst the n_i individuals receiving dosage level t_i, the unobserved number who would have responded to dosage level t_j but not to dosage level t_{j-1}.

The Gibbs sampler may be implemented using random draws from the complete conditional distribution of $p|X, Z_1,...,Z_k$ and from the complete conditional distributions of $Z_i|X, p, Z_j$, $j\neq i$. The former is an ordered Dirichlet updating (2)

$$\frac{\Gamma\left(\sum_{j=1}^{k+1} \overline{\gamma}_j\right)}{\prod_{j=1}^{k+1} \Gamma(\overline{\gamma}_j)} \prod_{j=1}^{k+1} \lambda_j^{\overline{\gamma}_j - 1}$$

where $\bar{\gamma}_j = \gamma_j + \Sigma Z_{ij}$. Thus, the complete conditional density for p_i, over the set $[p_{i-1}, p_{i+1}]$, denoted by $g_D(p_i|X, Z_1,...,Z_k, p_j, j\neq i)$, is that of $\Delta_i + p_{i-1}$ where $\Delta_i \sim (p_{i+1}-p_{i-1})\text{Beta}(\bar{\gamma}_i, \bar{\gamma}_{i+1})$.

The complete conditional distribution for Z_i is a product of two multinomials. That is, writing $Z_i = (Z_i(1), Z_i(2))$ where $Z_i(1) = (Z_{i1},...,Z_{ii})$, $Z_i(2) = (Z_{i,i+1},...,Z_{i,k+1})$, $Z_i(1)$ and $Z_i(2)$ are conditionally independent given X_i with $Z_i(1) \sim \text{Mult}(X_i, p_i^{-1}\lambda(1))$, $Z_i(2) \sim \text{Mult}(n_i-X_i, (1-p_i)^{-1}\lambda(2))$, $\lambda(1) = (\lambda_1,...,\lambda_i)$, $\lambda(2) = (\lambda_{i+1},...,\lambda_{k+1})$.

After v independent replications each to the r^{th} iteration we obtain $(p_s^{(r)}, Z_{1s}^{(r)},...,Z_{ks}^{(r)})$, $s=1,...,v$. The resulting estimate for the marginal posterior density of p_i is

$$\hat{f}_D(p_i|X) = v^{-1} \sum_{s=1}^{v} g_D(p_i|X, Z_{1s}^{(r)},...,Z_{ks}^{(r)}, p_{js}^{(r)}, j\neq i) \qquad (4)$$

Similarly the posterior mean of p_i is estimated using the mean of g_D leading to

$$E_D(p_i|X) = v^{-1} \sum_{s=1}^{v} [p_{i-1,s}^{(r)} + (p_{i+1,s}^{(r)} - p_{i-1,s}^{(r)})\{\bar{\gamma}_{is}/(\bar{\gamma}_{is}+\bar{\gamma}_{i+1,s})\}] \qquad (5)$$

where $\bar{\gamma}_{js} = \gamma_j + \Sigma Z_{ijs}^{(r)}$, $j=1,...,k+1$.

A posterior density estimate for $F(t)$ at say $t=t^*$ may be obtained by including $F(t^*)$ as an additional model parameter. More precisely we revise the prior to include $F(t^*)$ and to take the form $\pi_D(p) \cdot h_D(F(t^*)|p)$. Paralleling (2) if $t^* \in [t_i, t_{i+1}]$, h is naturally the density of $p_i + \Delta^*$ where $\Delta^* \sim (p_{i+1} -$

$p_i)$ Beta$(\gamma^*, \gamma_{i+1} - \gamma^*)$ with $\gamma^* = M\{F_0(t^*) - F_0(t_i)\}$. Since there is no data at dosage level t^*, the complete conditional distribution for $F(t^*)$ is $h_D(F(t^*)|p)$ as well. Therefore the posterior density estimate for $F(t^*)$ is $v^{-1} \Sigma\, h_D(F(t^*)|p_s^{(r)})$. A convenient by–product of this specification is that the posterior mean of $F(t^*)$ is

$$E_D\{F(t^*)|X\} = E_D\{p_i + (p_{i+1} - p_i)\,\gamma^*/\gamma_{i+1}|X\}$$

$$= \frac{\gamma_{i+1} - \gamma^*}{\gamma_{i+1}} E_D(p_i|X) + \frac{\gamma^*}{\gamma_{i+1}} E_D(p_{i+1}|X) \qquad (6)$$

Expression (6) shows that once the posterior means for the p_i's have been computed using (5), an elementary interpolation formula enables $E_D(F(t)|X)$ for any t. This formula was first noted by Antoniak (1974). The unpublished Ph.D. theses of M.N. Wesley also discusses (6) as well as computation of $E_D(p_i|X)$. In fact, she obtains exact formulas for these expectations for up to four dosage levels but evaluation of these formulas even for three levels is a forbidding task. By comparison, the estimate (5) is routine to evaluate regardless of the number of dosage levels. Moreover, using the above estimates of the posterior densities for the p_i and $F(t)$ we may straightforwardly obtain other features of these densities such as modes and quantiles.

For π_B, examination of (1) and (3) reveals that the complete conditional distribution of $p_i|X, p_j, j\neq i$, is that of a Beta$(\alpha_i + X_i, \beta_i + n_i - X_i)$ restricted to $[p_{i-1}, p_{i+1}]$ which we denote by $g_B(p_i|X_i, p_{i-1}, p_{i+1})$. Sampling from g_B by retaining appropriately restricted draws from the unrestricted Beta

distribution will be very inefficient when $p_{i+1} - p_{i-1}$ is small. One–for–one sampling from g_B may be achieved through the use of the incomplete Beta function which is incorporated into many scientific subroutine libraries. If $G(y) \equiv P\{Y \leq y \mid Y\text{-Beta}(a,b)\}$ then Z is distributed as Beta(a,b) restricted to [c,d] if $Z = G^{-1}[G(c) + U\{G(d) - G(c)\}]$ where U is a $U(0,1)$ variate. After v replications, each to the r^{th} iteration we obtain $p_s^{(r)}$, $s=1,\ldots,v$. To obtain a posterior density estimate for p_i we would compute

$$\hat{f}_B(p_i \mid X) = v^{-1} \sum_{s=1}^{v} g_B(p_i \mid x_i, p_{i-1,s}^{(r)}, p_{i+1,s}^{(r)}) \qquad (7)$$

Use of (7) is preferred to a kernel density estimate based upon the $p_{is}^{(r)}$ since, by using the conditional structure of the model, substantially smaller v is needed (see Gelfand and Smith, 1990). Though standardization is required to obtain each g_B in (7) this requires only a univariate numerical integration and can be done quite rapidly using simple trapezoidal or Simpson's rule integration. Since moments of g_B are not available explicitly, an expression such as (5) is unavailable for $E_B(p_i \mid X)$; posterior moments of p_i are most easily calculated using the $p_{is}^{(r)}$.

A posterior density estimate for $F(t)$ at say $t=t^*$ may be obtained by including $F(t^*)$ as an additional model parameter. More precisely, as with π_D, we revise the prior to include $F(t^*)$ and to take the form $\pi_B(p) \cdot h_B(F(t^*) \mid p)$. Paralleling the development below (3) if $t^* \in [t_i, t_{i+1}]$, h_B is naturally a Beta(α^*, β^*) restricted to $[p_i, p_{i+1})$ where with $M^* \equiv \alpha^* + \beta^*$ specified, $\alpha^* = (M^*-2)F_0(t^*) + 1$ and $\beta^* = M^* - \alpha^*$. Again, in the absence of data at dosage level t^*, the complete conditional distribution for $F(t^*)$ is $h_B(F(t^*) \mid p)$

so that the posterior density estimate for $F(t^*)$ becomes $v^{-1} \Sigma h_B(F(t^*)|p_s^{(r)})$. Again one dimensional numerical integration to standardize h_B is required. Unfortunately no interpolation formula analogous to (6) is available for $E_B(F(t)|X)$.

3. A Numerical Example

The data for this example are taken from an illustrative table in Cox and Snell (1989 p.7). In all, there are 150 subjects at 5 different stimulus levels, 30 subjects at each level. The data are given in Table 1. A probit model N(2,.25) was chosen as the prior shape F_0 with four choices for strength of prior belief, M = 1, 30, 100, 300. Table 1 supplies the maximum likelihood estimate and prior guess, $F_0(t_i)$, for each stimulus level. Using the π_D prior, (2), Table 1 also shows the nonparametric Bayes estimates, (5), of $F(t_i)$ for each value of M using the Gibbs sampler with a total of r = 20 iterations and v = 1000 replications. The supplied standard deviation (S.D.) measures the variation amongst these replications and also is an estimate of the marginal posterior standard deviation of $F(t_i)$. Interval estimates for the $F(t_i)$ were obtained using the empirical distribution of the 1000 estimates from the 1000 replications. In particular, equal tail 95 percent confidence intervals are given here.

Under π_D the Bayes estimate of F(t) can be obtained using the interpolation formula (6). Figure 1 graphs the Bayes estimate of the entire potency curve using this formula for M=1 and M=100 as well as the curve F_0. With increasing M, the Bayes estimate approaches the prior F_0. Figure 2 plots, for M = 1, the Bayes estimate with 95 percent posterior confidence band

developed from pointwise interval estimates using five points between each observed stimulus level. With increasing M the bands would become narrower.

4. Ordered Potency Curves

We extend the model of Section 2 to allow ordered polytomous response arising from stochastically ordered potency curves. Suppose for a given stimulus we observe whether or not each of two events occurred such that one is contained within the other. For example, event A might be "patient died" while event C might be "patient's condition worsened" whence $A \subset C$. Other illustrations include: in response to a dosage of medication, A is "improvement in a particular metabolic index", C is "improvement or no change in this index"; in response to hours of tutoring, A is a "high pass" on a standardized exam, C is a "pass" on this exam. We show here how the approach of Section 2 can be extended to handle such situations with further extension to more than two nested events being obvious. The case of non–nested events as for example if A is a decline in blood pressure while C is a decline in cholesterol level is not handled here.

Formalizing notation, suppose at level t_i, $i=1,2,...,k$ we observe n_i subjects with event A occurring X_i times, event C occurring Y_i times with $A \subset C$ so that $X_i \leq Y_i$. We model this situation with two underlying potency curves $F_A(t)$, $F_C(t)$ which are stochastically ordered, i.e., $F_A(t) \leq F_C(t)$. Letting $p_i \equiv F_A(t_i) \leq q_i \equiv F_C(t_i)$ we assume that the joint distribution of X_i and Y_i is specified through $(X_i, Y_i - X_i) \sim \text{Mult}(n_i; p_i, q_i - p_i, 1 - q_i)$. Hence with $p = (p_1,...,p_k)$ and $q = (q_1,...,q_k)$ the likelihood at $X = x$, $Y = y$ is

$$\prod_{i=1}^{k} \frac{n_i!}{x_i!(y_i-x_i)!(n_i-y_i)!}\, p_i^{x_i}(q_i-p_i)^{y_i-x_i}(1-q_i)^{n_i-y_i} \tag{8}$$

Interest focuses on inference regarding the p_i's and q_i's as well as F_A and F_C. Prior specification is restricted to the set $T^k = \{(p,q) : 0 \le p_1 \le \ldots \le p_k \le 1, 0 \le q_1 \le q_2 \le \ldots \le q_k \le 1, p_i \le q_i \text{ for all } i\}$. We consider two families of priors extending π_D and π_B respectively.

In extending π_D the Dirichlet process prior is replaced by a product Dirichlet process prior with stochastic order. We do not attempt formal definition or investigation of such measures here. Rather, if F_A is close to some given $F_{A,o}$ with precision M_A and F_C is close to some given $F_{C,o}$ with precision M_C then for any t we assume that the joint prior on $(F_A(t), F_C(t))$ is of the form

$$d(t)\{F_A(t)\}^{M_A F_{A,o}(t)-1}\{1-F_A(t)\}^{M_A\{1-F_{A,o}(t)\}-1}\cdot$$

$$\{F_C(t)\}^{M_C F_{C,o}(t)-1}\{1-F_C(t)\}^{M_C\{1-F_{C,o}(t)\}-1} \tag{9}$$

over $0 \le F_A(t) \le F_C(t) \le 1$. Expression (9) is a product of Beta forms standardized by $d(t)$ under the order restriction. Similarly the induced prior on (p, q) over T^k takes the form of a product of ordered Dirichlets,

$$\Pi_D(\underline{p},\underline{q}) = c(\gamma, \eta) \prod_{j=1}^{k+1} \Delta_j^{\gamma_j-1}\, \epsilon_j^{\eta_j-1} \tag{10}$$

where $\Delta_j = p_j - p_{j-1}$, $\epsilon_j = q_j - q_{j-1}$, $\gamma_j = M_A\{F_{A,o}(t_j) - F_{A,o}(t_{j-1})\}$, $\eta_j =$

$M_C\{F_{C,o}(t_j) - F_{C,o}(t_{j-1})\}$ and $c(\gamma,\eta)$ is the standardizing constant.

π_B is extended to produce a form over T^k which is conjugate with (8), that is,

$$\pi_B(p,q) = c(\alpha, \beta, \delta) \prod_{j=1}^{k} p_j^{\alpha_j-1} (q_j-p_j)^{\beta_j-1} (1-q_j)^{\delta_j-1} \qquad (11)$$

where again c is the standardizing constant. Linking $\pi_B(p,q)$ to prior information can be done similar to that for $\pi_B(p)$ using the conditional modes of p_i and of q_i-p_i. Assuming $\alpha_i + \beta_i + \delta_i = M_i$, M_i specified, with $\alpha_i, \beta_i, \delta_i > 1$ we obtain $\alpha_i = (M_i-3)\, F_{A,o}(t_i)$, $\beta_i = (M_i-3)\{F_{C,o}(t_i) - F_{A,o}(t_i)\}$ and $\delta_i = M_i - \alpha_i - \beta_i$.

Implementation of our sampling approach for the prior (10) is simplified by the introduction of unobserved multinomial variables as in Section 2. Associating Z_i^A with event A, Z_i^C with event C, i=1,...,k the distribution theory proceeds similarly to that in Section 2 leading to marginal posterior density estimates for p_i and q_i analogous to (4). Unfortunately the restriction $p_i \le q_i$ precludes explicit calculation of complete conditional moments so that there is no analogue to (5); we use the $p_{i,s}^{(r)}$, s=1,...,v, to obtain posterior expectations. Posterior density estimates for $F_A(t^*)$ and $F_C(t^*)$ along with interpolation formulas for posterior expectations can be developed paralleling the discussion below (5) leading to (6). For the prior (11) distribution theory and sampling parallel that above (7). Posterior density estimates analogous to (7) arise along with posterior density estimates for $F_A(t^*)$ and for $F_C(t^*)$. The use of (11) is illustrated with an example in Section 5.

It is worth noting that, regardless of prior, missing data can be readily

handled. For instance, should X_i be unavailable we need only include it as an additional parameter in the model. Complete conditional distributions for each of the other parameters when X_i is given as well will be exactly as before while the complete conditional distribution of X_i will be $Bi(Y_i, p_i/q_i)$.

5. A Second Numerical Example

The data in Table 2 is taken from Maxwell (1961, p. 70). In his table each of 223 boys is classified into one of five different age groups and is assigned a rating on a four point scale $(-1, 0, 1, 2)$ for the symptom of "disturbed dreams". A rating of 2 denotes the most severe suffering from bad dreams while a rating of -1 denotes no suffering at all. To simplify the illustration we combine the two intermediate categories, ratings 0 and 1, into one to obtain Table 2.

In the context of Section 4 we envision two underlying "potency" curves. One is associated with the event A, "rank -1 is assigned", that is, no suffering from bad dreams. The other is associated with the event C, "rank -1, 0 or 1 is assigned", that is, no suffering or mild suffering from bad dreams. With these definitions, curves for the incidence of A and of C which are monotonically increasing with age seem reasonable. Obviously $A \subset C$ thus ordering the curves. Labeling the five age categories from youngest to oldest as 0, 1, 2, 3, and 4, for category i, X_i counts the incidence of A, Y_i counts the incidence of C. The top part of Table 3 converts Table 2 to this notation.

The product Beta prior, (11), is taken here. For illustration we assume here that all prior precisions M_i are equal to 10, that $F_{A,o}$ is $N(2,4)$ and that $F_{C,o}$ is $N(1,4)$. For each age category prior incidence probabilities associated

with these two curves appear in Table 3 along with α_i, β_i and δ_i.

Table 3 next lists the Bayes estimates (posterior means) based upon r=20 iterations and v=1000 replications. As in Table 1 standard deviations and 95% equal tail interval estimates for each p_i and q_i are also given. Lastly this table gives the maximum likelihood estimate for (p, q). The maximum likelihood estimate for p_2 badly violates the monotonicity assumption but a standard two sample test reveals that it is not significantly smaller than that for p_1 even at the 0.1 level. As analternative estimate we include the isotonic regression of the maximum likelihood estimate using the pooled–adjacent–values–algorithm (Robertson, Wright and Dykstra, 1988). This estimate is still unsatisfying since strict monotonicity might be anticipated.

6. Summary and Discussion

Nonparametric bioassay is inherently attractive in freeing us from parametric specification of a form for the presumed underlying potency curve. Similarly, Bayesian inference is attractive for the bioassay problem since we often have prior knowledge regarding the potency curve, e.g., the LD50, the steepness, etc. While Bayesian nonparametric problems are typically quite hard, the fact that the bioassay setting provides a binomial likelihood enables simplification particularly with regard to prior specification. Even so, previous attacks have enjoyed limited success in developing desired marginal posterior distributions and their features due to computational difficulties. However, the Gibbs sampler approach is shown to be well–suited for the general handling of such "likelihood × prior" forms. Moreover, a potentially useful extension can be handled as well

through this sampling–based approach. Other extensions, for instance, to multivariate dichotomous response and to nominal polytomous response are worthy of further investigation.

ACKNOWLEDGEMENT

In conducting this research, the authors are grateful for the financial support of the National Science Foundation.

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.

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

Cox, D.R. and Snell, E.J. (1989). Analysis of Bioassay Data, 2nd. Edition, 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.

Finney, D.J. (1978). Statistical Methods In Biological Assay, New York: Hafner.

Gelfand, A.E., Hills, S.E., Racine–Poon, A. and Smith, A.F.M. (1990).

Illustration of Bayesian inference in normal data models using Gibbs sampling. J. Am. Statist. Assoc. (To appear).

Gelfand, A.E. and Smith, A.F.M. (1990). Sampling based approaches to calculating marginal densities. J. Am. Statist. Assoc., 85, 398–409.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans. on Pattern Analysis and Machine Intelligence, 6, 721–741.

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

Kuo, L. (1988). Linear Bayes estimators of the potency curve in bioassay. Biometrika, 75, 91–96.

Maxwell, D. (1961). Analyzing Qualitative Data, London: Methuen.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953). Equations of state space calculations by fast computing machines, J. Chem. Phys., 21, 1087–1091.

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

Robertson, T., Wright, F.T. and Dykstra, R.L. (1988). Order Restricted Statistical Inference, Chichester: J. Wiley & Sons.

Tanner, M. and Wong, W. (1987). The calculation of posterior distributions by data augmentation (with discussion), J. Am. Statist. Assoc., 82, 528–550.

$t_i[\log_2(\text{concentration})]$	0	1	2	3	4
No. of Deaths	2	8	15	23	27
MLE	0.067	0.267	0.500	0.767	0.900
$F_o(t_i)$	0.000	0.023	0.500	0.977	1.000

Bayes Estimate (S.D.)
[95% Interval Estimate]

M					
1	0.102(0.068) [0.019,0.267]	0.265(0.112) [0.129,0.454]	0.505(0.131) [0.297,0.698]	0.759(0.129) [0.585,0.898]	0.869(0.098) [0.712,0.961]
30	0.063(0.028) [0.023,0.123]	0.136(0.039) [0.092,0.185]	0.502(0.038) [0.439,0.567]	0.872(0.039) [0.826,0.917]	0.919(0.031) [0.863,0.959]
100	0.029(0.009) [0.014,0.049]	0.083(0.012) [0.069,0.099]	0.501(0.016) [0.474,0.526]	0.919(0.016) [0.904,0.933]	0.962(0.011) [0.944,0.976]
300	0.010(0.003) [0.006,0.016]	0.049(0.004) [0.042,0.055]	0.500(0.010) [0.485,0.514]	0.952(0.010) [0.946,0.958]	0.986(0.004) [0.981,0.991]

		Age Groups				
		5–7	8–9	10–11	12–13	14–15
Disturbed Dream Rating	2	7	13	7	10	3
	0 or 1	7	26	20	21	9
	−1	7	10	23	28	32

Age Catetories	0	1	2	3	4
Data:					
n	21	59	50	59	44
Y	14	36	43	49	41
X	7	10	23	28	32
Prior Guess:					
$F_{C,0}$	0.309	0.500	0.692	0.841	0.933
$F_{A,0}$	0.159	0.309	0.500	0.691	0.841
α	5.840	4.500	3.160	2.111	1.468
β	2.049	2.340	2.340	2.049	1.643
δ	2.111	3.160	4.500	5.840	6.889
Bayes Estimates:					
$\hat{F}_C$	0.505(0.029)	0.631(0.022)	0.797(0.020)	0.849(0.017)	0.924(0.015)
	[0.446,0.562]	[0.590,0.671]	[0.760,0.832]	[0.820,0.877]	[0.899,0.945]
$\hat{F}_A$	0.192(0.023)	0.236(0.023)	0.426(0.022)	0.520(0.022)	0.726(0.023)
	[0.151,0.241]	[0.195,0.283]	[0.384,0.469]	[0.478,0.559]	0.681,0.769]
MLE:					
$\hat{F}_C$	0.667	0.610	0.860	0.831	0.932
$\hat{F}_A$	0.333	0.170	0.460	0.475	0.727
Isotonic Regression of MLE:					
$\hat{F}_C$	0.625	0.625	0.844	0.844	0.932
$\hat{F}_A$	0.213	0.213	0.460	0.475	0.727

Fig. 1. Bayes estimates of the potency curve. – – – = Bayes estimate for M=1, ··· = Bayes estimate for M=100, —— = Prior curve, F_0, while squares denote maximum likelihood estimates at t=0,1,2,3,4.

Fig. 2. The Bayes estimate of F for M = 1 and its 95% confidence band.

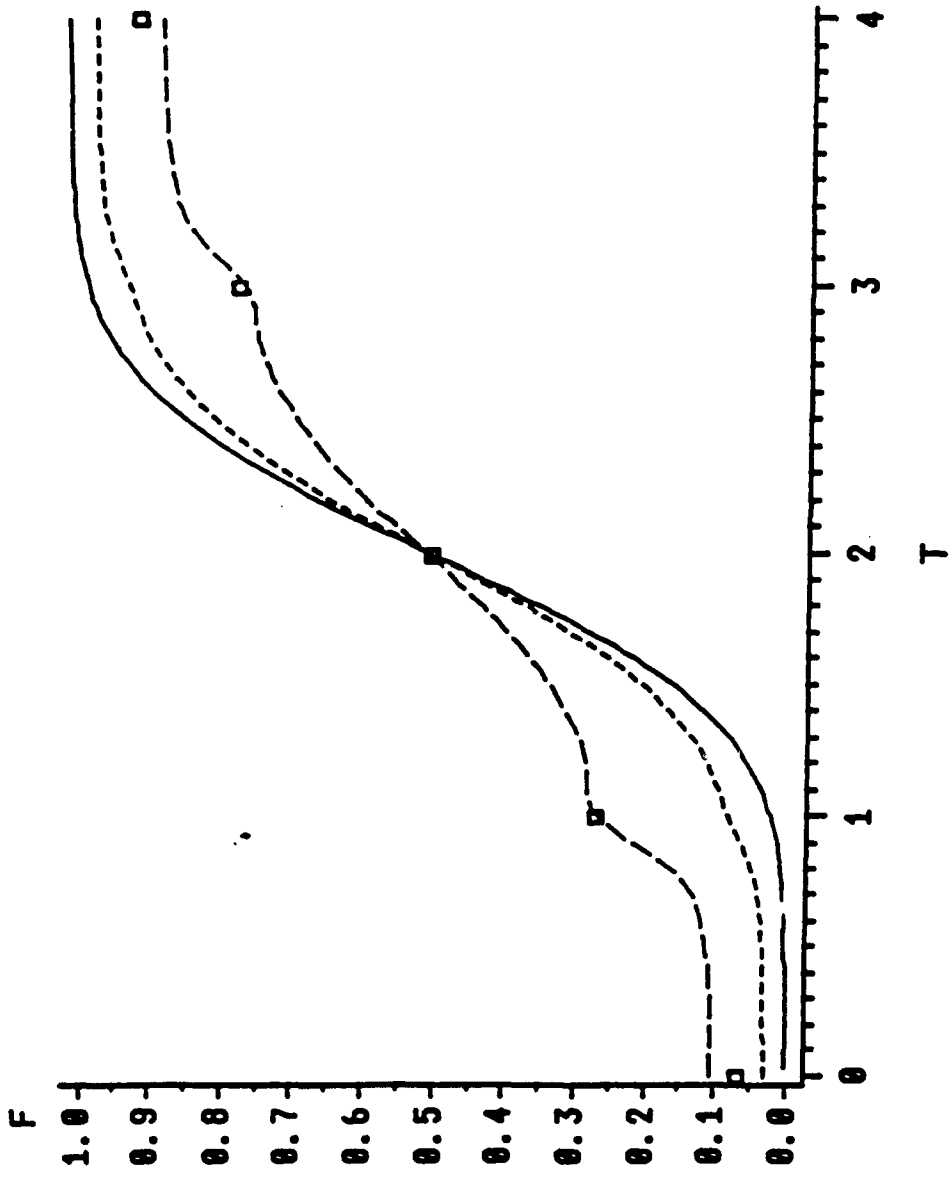


Figure 1

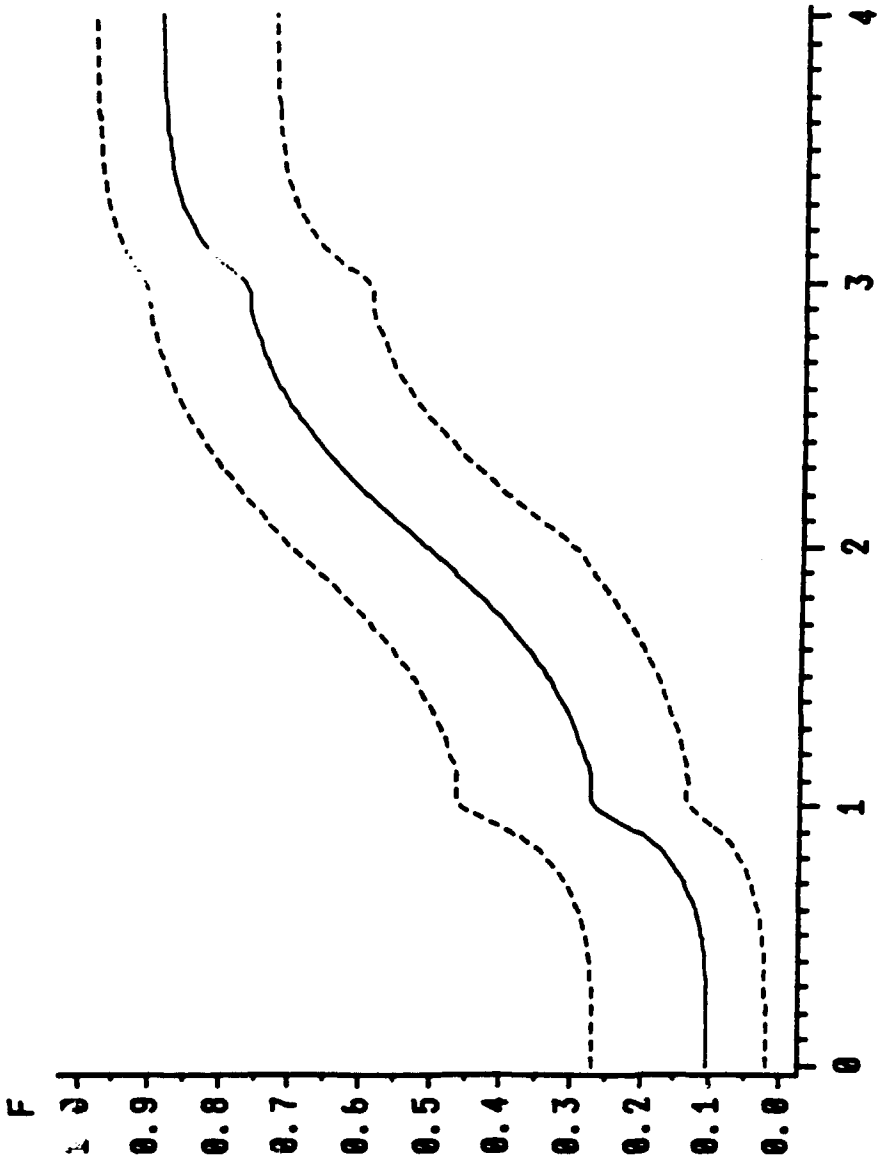


Figure 2

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 458	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) NONPARAMETRIC BAYESIAN BIOASSAY INCLUDING ORDERED POLYTOMOUS RESPONSE		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Alan E. Gelfand and Lynn Kuo		8. CONTRACT OR GRANT NUMBER(s) N0025-92-J-1264
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305-4065		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 111		12. REPORT DATE September 14, 1992
		13. NUMBER OF PAGES_
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Key Words: bioassay, Dirichlet process prior, Gibbs sampler, potency curve, product Beta prior.

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

See Reverse Side

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Previous attempts at implementing fully Bayesian nonparametric bioassay have enjoyed limited success due to computational difficulties. We show here how this problem may be generally handled using a sampling based approach to develop desired marginal posterior distributions and their features. A useful extension is presented which treats the case of ordered polytomous response. Illustrative examples are provided.