The Annals of Statistics
1981, Vol. 9, No. 4, 803-811

POSTERIOR DISTRIBUTION OF A DIRICHLET PROCESS FROM QUANTAL RESPONSE DATA[1]

BY P. K. BHATTACHARYA

University of Arizona

The posterior distribution of a Dirichlet process from N quantal responses at r dosage levels has been recognized by Antoniak as a mixture of Dirichlet processes. The purpose of this paper is to develop a systematic procedure for computing finite-dimensional distributions of such mixtures which can be equivalently expressed as multivariate beta distributions with random parameter vectors. It is shown that the sequence of random parameter vectors of the updated beta posteriors from observations at increasing dosage levels evolves in a manner which is described by r separate Markov chains. This description is then used to derive the asymptotic posterior distribution. The weak limits of the relevant Markov chains are shown to be solutions of certain stochastic differential equations and the random parameter vector of the posterior beta distribution is shown to be asymptotically normal, the mean vector and covariance matrix of which are given by recursion formulas.

1. Introduction. Let X be the tolerance level of an individual to a drug and let F denote the population cdf of X. Instead of observing a random sample of X, fix r doses $x_1 < \cdots < x_r$ and assign $N\alpha_j$ subjects to x_j for making inference about F, where $\alpha_1, \cdots, \alpha_r > 0$ and $N\alpha_1, \cdots, N\alpha_r$ are integers. The observations U_{ji}, $1 \le i \le N\alpha_j$, $1 \le j \le r$, taking values 1 or 0 according as the tolerance of the ith individual assigned to the dose x_j is less than or equal to x_j, or greater than x_j, are called quantal responses.

Ayer, et al. (1955), proposed an estimate of F based on such data. Subsequently, Kraft and van Eeden (1964) established some general properties of Bayesian methods using Dirichlet process (DP) priors, while Ramsey (1972) and Wesley (1976), using an algorithm of Turnbull (1976), developed methods for computing the posterior mode. However, the problem of evaluating the posterior distribution remained intractable, because these posteriors are complicated mixtures of DP's as observed by Antoniak (1974). Ferguson (1973) showed that with each observation X_i the DP parameter α gains a unit mass at X_i. However, with the corresponding quantal response U_{ji}, the DP parameter gains a unit mass at a random point whose distribution is easily determined, giving rise to a mixture of DP's. With several U_{ji} the posterior becomes a very complicated mixture. This is also typical for randomly censored data as shown by Susarla and Van Ryzin (1976) and Blum and Susarla (1977).

In this paper we develop a systematic procedure for computing finite-dimensional distributions of the posterior of a DP from quantal data for which it is enough to work with multinomial distributions. For making inference about F at the r dosage levels define $\theta_1 + \cdots + \theta_j = F(x_j)$, $1 \le j \le r$ and $\theta_{r+1} = 1 - F(x_r)$. Corresponding to a prior DP (α) for F, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_{r+1})$ follows a multivariate beta distribution Be($\mathbf{a}$) with parameter vector $\mathbf{a} = (a_1, \cdots, a_{r+1})$ where $a_1 + \cdots + a_j = \alpha(-\infty, x_j]$, $1 \le j \le r$ and $a_{r+1} = \alpha(x_{r+1}, \infty)$. Similarly for the posterior, corresponding to a mixture of DP for F, $\boldsymbol{\theta}$ will have a multivariate beta distribution with a random parameter vector. In Section 2, the distri-

Received August 1978; revised June 1980.

[1] This research was supported by NSF Grant MCS 77-01663.

AMS 1970 *Subject Classifications.* Primary 62E20, 62C10; secondary 62G99, 62P10.

Key words and phrases. Dirichlet process mixtures, quantal response data, asymptotic posterior distribution, Markov chain, weak convergence, stochastic differential equation.

bution of this random vector is described in Theorem 1 in terms of r separate Markov chains (MC's).

This result is then used in Section 3 to derive the asymptotic posterior distribution when the information content of the prior as well as the number of observations at each dose is large, as would be the case if for large N, $N\alpha_1, \cdots, N\alpha_r$ quantal responses are observed at the r doses along with $N\alpha_0$ complete observations. The asymptotic posterior distribution is determined by the asymptotic distribution of the random parameter vector of the multivariate beta posterior. Using a theorem of Strook and Varadhan (1969), the weak limits of the relevant MC's are shown to be solutions of certain stochastic differential equations (Theorem 2), from which the asymptotic distribution of the random parameter vector is obtained (Theorem 3). The slightly more general problem of determining the posterior of F on an arbitrary finite set is discussed in Remarks 1 and 2 following Theorems 1 and 3.

2. Exact distribution of the random parameter vectors of the successive beta posteriors. U_{ji} are independent Bernoulli variables taking values 1 and 0 with probabilities $\theta_1 + \cdots + \theta_j$ and $\theta_{j+1} + \cdots + \theta_{r+1}$ respectively, $1 \le i \le N\alpha_j$, $1 \le j \le r$, where $\alpha_j > 0$, $N\alpha_j$ are integers, and $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_{r+1})$ is an unknown probability vector following a multivariate beta prior Be($\mathbf{c}$) with parameter vector $\mathbf{c} = (c_1, \cdots, c_{r+1})$, $c_j > 0$. Let $m_j = \sum_{i=1}^{N\alpha_j} U_{ji}$ and $n_j = N\alpha_j - m_j$. Clearly m_j is sufficient for $\boldsymbol{\theta}$ in $U_{j1}, \cdots, U_{j,N\alpha_j}$. For each j, rearrange the U_{ji}'s so that the first m_j are 1's and the last n_j are 0's.

The posterior distribution of $\boldsymbol{\theta}$ given $m_1, \cdots, m_j$ is a beta distribution with a random parameter vector $\boldsymbol{\nu}_j$. In this section we shall express $\boldsymbol{\nu}_j$, $1 \le j \le r$, in terms of a sequence of artificial random vectors forming r distinct MC's $\{\hat{\boldsymbol{\nu}}_j = \boldsymbol{\nu}_j(0), \boldsymbol{\nu}_j(1), \cdots, \boldsymbol{\nu}_j(N\alpha_j) = \boldsymbol{\nu}_j\}$, $1 \le j \le r$. These chains are described as follows. For $j = 1$, $\hat{\boldsymbol{\nu}}_1 \equiv \mathbf{c}$ and $\boldsymbol{\nu}_1(k)$ represents the random parameter vector of the posterior given $U_{11}, \cdots, U_{1k}$, for $0 \le k \le N\alpha_1$. For $j > 1$, the jth chain is initiated by choosing $\hat{\boldsymbol{\nu}}_j$ from the conditional distribution of $\boldsymbol{\nu}_{j-1}$ given m_j; subsequently, $\boldsymbol{\nu}_j(k)$ represents the random parameter vector of the posterior given $U_{j1}, \cdots, U_{jk}$ obtained from the prior Be($\hat{\boldsymbol{\nu}}_j$) conditioned on $\hat{\boldsymbol{\nu}}_j$.

We begin with a random variable Y taking values in $\{1, \cdots, r+1\}$ with probability vector $\boldsymbol{\theta}$ following a prior distribution Be($\mathbf{a}$), $\mathbf{a}$ nonrandom, of which the pdf with respect to $d\boldsymbol{\theta} = d\theta_1 \ldots d\theta_r$ is

$$g_{\mathbf{a}}(\boldsymbol{\theta}) = \{\Gamma(\textstyle\sum_1^{r+1} a_l)/\prod_1^{r+1} \Gamma(a_l)\} \prod_1^{r+1} \theta_l^{a_l - 1}, \qquad \theta_{r+1} = 1 - \sum_1^r \theta_l.$$

Let $\mathbf{e}_l$ be the $(r+1)$-dimensional lth coordinate vector, J a proper subset of $\{1, \cdots, r+1\}$ and $\boldsymbol{\Delta}_J$ a random vector with

$$P[\boldsymbol{\Delta}_J = \mathbf{e}_l] = p(\mathbf{a}, \mathbf{a} + \mathbf{e}_l) = a_l/\textstyle\sum_{l \in J} a_l, \qquad l \in J. \tag{1}$$

Since $g_{\mathbf{a}}(\boldsymbol{\theta}) \cdot \theta_l = (a_l/\sum_1^{r+1} a_l) g_{\mathbf{a}+\mathbf{e}_l}(\boldsymbol{\theta}) = E_{\mathbf{a}}(\theta_l) g_{\mathbf{a}+\mathbf{e}_l}(\boldsymbol{\theta})$, the posterior of $\boldsymbol{\theta}$ given $Y \in J$ is

$$dP(\boldsymbol{\theta} \mid J) = \textstyle\sum_{l \in J} E_{\mathbf{a}}(\theta_l) g_{\mathbf{a}+\mathbf{e}_l}(\boldsymbol{\theta})\, d\boldsymbol{\theta} / \sum_{l \in J} E_{\mathbf{a}}(\theta_l) = g_{\mathbf{a}+\boldsymbol{\Delta}_J}(\boldsymbol{\theta})\, d\boldsymbol{\theta}.$$

Now suppose the prior of $\boldsymbol{\theta}$ is Be($\boldsymbol{\nu}_0$) where $\boldsymbol{\nu}_0$ is a random vector with $P[\boldsymbol{\nu}_0 = \mathbf{a}] = \Pi_0(\mathbf{a})$, $\mathbf{a} \in S_0$, S_0 being a finite (or countable) set with $\sum_1^{r+1} a_l = c$ for all $\mathbf{a} \in S_0$. Then the posterior of $\boldsymbol{\theta}$ given $Y \in J$ is

$$\begin{aligned} dP(\boldsymbol{\theta} \mid J) &= \textstyle\sum_{a \in S_0} P[\boldsymbol{\nu}_0 = \mathbf{a} \mid Y \in J]\, dP(\boldsymbol{\theta} \mid \boldsymbol{\nu}_0 = \mathbf{a}, Y \in J) \\ &= \textstyle\sum_{a \in S_0} P[\boldsymbol{\nu}_0 = \mathbf{a} \mid Y \in J] g_{\mathbf{a}+\boldsymbol{\Delta}_J}(\boldsymbol{\theta})\, d\boldsymbol{\theta} \end{aligned}$$

by the previous case, and

$$\begin{aligned} P[\boldsymbol{\nu}_0 = \mathbf{a} \mid Y \in J] &= \Pi_0(\mathbf{a}) E_{\mathbf{a}}(\textstyle\sum_{l \in J} \theta_l) / \sum_{a \in S_0} \Pi_0(\mathbf{a}) E_{\mathbf{a}}(\sum_{l \in J} \theta_l) \\ &= \Pi_0(\mathbf{a})(\textstyle\sum_{l \in J} a_l) / \sum_{\mathbf{a} \in S_0} \Pi_0(\mathbf{a})(\sum_{l \in J} a_l) = \hat{\Pi}_0(\mathbf{a}), \quad \text{say}. \end{aligned}$$

Thus letting $S_1 = \{\mathbf{a} + \mathbf{e}_l \mid \mathbf{a} \in S_0, l \in J\}$ and $\Pi_1(\mathbf{b}) = \sum_{\mathbf{a}\in S_0} \hat{\Pi}_0(\mathbf{a})p(\mathbf{a}, \mathbf{b})$ for $b \in S_1$, where $p(\mathbf{a}, \mathbf{b})$ is given by (1), we have

$$\begin{aligned} dP(\theta \mid J) &= \textstyle\sum_{\mathbf{a}\in S_0} \hat{\Pi}_0(\mathbf{a}) g_{\mathbf{a}+\Delta_J}(\boldsymbol{\theta})\, d\boldsymbol{\theta} \\ &= \textstyle\sum_{\mathbf{a}\in S_0} \hat{\Pi}_0(\mathbf{a}) \sum_{l\in J} p(\mathbf{a}, \mathbf{a} + \mathbf{e}_l) g_{\mathbf{a}+\mathbf{e}_l}(\boldsymbol{\theta})\, d\boldsymbol{\theta} \\ &= \textstyle\sum_{b\in S_1} \Pi_1(\mathbf{b}) g_{\mathbf{b}}(\boldsymbol{\theta})\, d\boldsymbol{\theta}, \end{aligned}$$

which shows that the posterior of $\boldsymbol{\theta}$ is $\mathrm{Be}(\boldsymbol{\nu}_1)$, where $\boldsymbol{\nu}_1$ is a random vector with $P[\boldsymbol{\nu}_1 = \mathbf{b}] = \Pi_1(\mathbf{b})$ for $\mathbf{b} \in S_1$. Moreover, $\sum_1^{r+1} b_l = c + 1$ for all $b \in S_1$, which allows the above argument to be extended in the following manner.

Suppose the prior of $\boldsymbol{\theta}$ is $\mathrm{Be}(\boldsymbol{\nu}_0)$ as above and let $Y_1, Y_2, \cdots$ be i.i.d. as Y. If $J_1, J_2, \cdots$ are proper subsets of $\{1, \cdots, r + 1\}$, then we use the above argument repeatedly to see that the posterior of $\boldsymbol{\theta}$ given $Y_1 \in J_1, \cdots, Y_i \in J_i$ is $\mathrm{Be}(\boldsymbol{\nu}_i)$ where the distributions of the random vectors $\boldsymbol{\nu}_i$ are described recursively by the formulas:

$$\begin{aligned} &P[\boldsymbol{\nu}_i = \mathbf{b}] = \Pi_i(\mathbf{b}) = \textstyle\sum_{a\in S_{i-1}} \hat{\Pi}_{i-1}(\mathbf{a})p_{i-1}(\mathbf{a}, \mathbf{b}), \qquad \mathbf{b} \in S_i, \\ &S_i = \{\mathbf{a} + \mathbf{e}_l \mid \mathbf{a} \in S_{i-1}, l \in J_i\}, \\ &p_{i-1}(\mathbf{a}, \mathbf{a} + \mathbf{e}_l) = a_l / \textstyle\sum_{l\in J_i} a_l, \qquad l \in J_i, \\ &\hat{\Pi}_{i-1}(\mathbf{a}) = \Pi_{i-1}(\mathbf{a})(\textstyle\sum_{l\in J_i} a_l)/\sum_{a\in S_{i-1}} \Pi_{i-1}(\mathbf{a})(\sum_{l\in J_i} a_l), \qquad \mathbf{a} \in S_{i-1}. \end{aligned}$$

In particular, if $\boldsymbol{\nu}_0 \equiv \mathbf{a}^*$, i.e. $S_0 \equiv \{\mathbf{a}^*\}$ and if $J_1, J_2, \ldots$ are such that $\sum_{l\in J_{i+1}} a_l$ are constants for all $\mathbf{a} \in S_i$, $i = 0, 1, \cdots$, then $\hat{\Pi}_i = \Pi_i$, $i = 0, 1, \ldots$ in the above formulas. In such a case, $\boldsymbol{\nu}_i$ has the same distribution as $\boldsymbol{\nu}_i^*$, where $\{\boldsymbol{\nu}_i^*, i = 0, 1, \ldots\}$ is a MC with transition probabilities $p_i(\mathbf{a}, \mathbf{b}) = P[\boldsymbol{\nu}_{i+1}^* = \mathbf{b} \mid \boldsymbol{\nu}_i^* = \mathbf{a}]$ starting at $\mathbf{a}^*$. We use this idea in the following lemma.

LEMMA 1. *The posterior of $\boldsymbol{\theta}$ from the prior* $\mathrm{Be}(\mathbf{a}^*)$ *and $N\alpha_j$ quantal responses at the* jth *dose of which the first m_j are* 1 *and the last $n_j = N\alpha_j - m_j$ are* 0 *is* $\mathrm{Be}(\boldsymbol{\nu}_{j,\mathbf{a}^*}(N\alpha_j))$, *where* $\{\boldsymbol{\nu}_{j,\mathbf{a}^*}(k), 0 \le k \le N\alpha_j\}$ *is a* MC *starting at* $\mathbf{a}^*$ *and having transition probabilities*

$$\begin{aligned} &\Pi_{jk}(\mathbf{a}, \mathbf{a} + (\mathbf{e}_{jl}; \mathbf{0})) = a_l / \textstyle\sum_{l=1}^{j} a_l, && 1 \le l \le j, 0 \le k \le m_j - 1, \\ &\Pi_{jk}(\mathbf{a}, \mathbf{a} + (\mathbf{0}; \mathbf{f}_{jl})) = a_l / \textstyle\sum_{l=j+1}^{r+1} a_l, && j + 1 \le l \le r + 1, m_j \le k \le N\alpha_j - 1, \end{aligned} \tag{2}$$

where $\mathbf{e}_{jl}$ and $\mathbf{f}_{jl}$ are the j- and $(r + 1 - j)$-dimensional lth *coordinate vectors.*

PROOF. The observations U_{ji} with $\sum_{i=1}^{N\alpha_j} U_{ji} = m_j$ are equivalent to $Y_i \in J_i$, $1 \le i \le N\alpha_j$, with $J_i = \{1, \cdots, j\}$, $1 \le i \le m_j$ and $J_i = \{j + 1, \cdots, r + 1\}$, $m_j + 1 \le i \le N\alpha_j$. Clearly $\sum_{l=1}^{r+1} a_l$ and $\sum_{l\in J_{i+1}} a_l$ are constants for all $\mathbf{a} \in S_i$, $0 \le i \le N\alpha_j - 1$ and the argument of the above paragraph becomes applicable. This proves the lemma.

We now extend this lemma to the case when the prior distribution of $\boldsymbol{\theta}$ is $\mathrm{Be}(\boldsymbol{\nu}_{j-1})$ where $\boldsymbol{\nu}_{j-1}$ is a random vector taking values in S_{j-1}^* with probabilities $\Pi_{j-1}(\mathbf{a}) = P[\boldsymbol{\nu}_{j-1} = \mathbf{a}]$, where $S_0^* = \{\mathbf{c}\}$ and for $j \ge 1$,

$$S_j^* = \{\mathbf{c} + \boldsymbol{\Delta} \mid \Delta_l \quad \text{are nonnegative integers}, \quad \textstyle\sum_{l=1}^{r+1} \Delta_l = N \sum_{l=1}^{j} \alpha_l\}.$$

LEMMA 2. *If in the hypothesis of Lemma* 1 *the prior distribution of $\boldsymbol{\theta}$ is* $\mathrm{Be}(\boldsymbol{\nu}_{j-1})$ *as above, then in the conclusion, the posterior of $\boldsymbol{\theta}$ is* $\mathrm{Be}(\boldsymbol{\nu}_j(N\alpha_j))$, *where* $\{\boldsymbol{\nu}_j(k), 0 \le k \le N\alpha_j\}$ *is a* MC *with the same transition probabilities as in Lemma* 1, *but having initial distribution $\hat{\Pi}_j$ given by*

$$\begin{aligned} \hat{\Pi}_j(\mathbf{a}) &= P[\boldsymbol{\nu}_{j-1} = \mathbf{a} \mid \textstyle\sum_{i=1}^{N\alpha_j} U_{ji} = m_j] \\ &= \Pi_{j-1}(\mathbf{a})\Psi_j(\mathbf{a}, m_j)/\textstyle\sum_{\mathbf{a}\in S_{j-1}^*} \Pi_{j-1}(\mathbf{a})\Psi_j(\mathbf{a}, m_j), \end{aligned} \tag{3}$$

$$\Psi_j(\mathbf{a}, m_j) = \{\Gamma(\textstyle\sum_{l=1}^{j} a_l + m_j)/\Gamma(\sum_{l=1}^{j} a_l)\} \cdot \{\Gamma(\sum_{l=1}^{r+1-j} a_{j+l} + n_j)/\Gamma(\sum_{l=1}^{r+1-j} a_{j+l})\}.$$

PROOF. The posterior of $\boldsymbol{\theta}$ given m_j is

$$dP(\boldsymbol{\theta} \mid \sum_{i=1}^{N\alpha_j} U_{ji} = m_j) = \sum_{\mathbf{a} \in S_{j-1}^*} P[\boldsymbol{\nu}_{j-1} = \mathbf{a} \mid \sum_{i=1}^{N\alpha_j} U_{ji} = m_j]\, dP(\boldsymbol{\theta} \mid \boldsymbol{\nu}_{j-1} = \mathbf{a}, \sum_{i=1}^{N\alpha_j} U_{ji} = m_j).$$

For arbitrary $\mathbf{a}^* \in S_{j-1}^*$, $dP(\boldsymbol{\theta} \mid \boldsymbol{\nu}_{j-1} = \mathbf{a}^*, \sum_{i=1}^{N\alpha_j} U_{ji} = m_j)$ is $\mathrm{Be}(\boldsymbol{\nu}_{j,\mathbf{a}^*}(N\alpha_j))$ by Lemma 1, while straightforward calculations reduce

$$P[\boldsymbol{\nu}_{j-1} = \mathbf{a} \mid \sum_{i=1}^{N\alpha_j} U_{ji} = m_j] = \Pi_{j-1}(\mathbf{a}) \int g_{\mathbf{a}}(\boldsymbol{\theta})(\sum_{l=1}^{j} \theta_l)^{m_j}0(\sum_{l=j+1}^{r+1} \theta_l)^{n_j}\, d\boldsymbol{\theta} \Big/$$

$$\sum_{\mathbf{a} \in S_{j-1}^*} \Pi_{j-1}(\mathbf{a}) \int g_{\mathbf{a}}(\boldsymbol{\theta})(\sum_{l=1}^{j} \theta_l)^{m_j}(\sum_{l=j+1}^{r+1} \theta_l)^{n_j}\, d\boldsymbol{\theta}$$

to the form given in (3). This completes the proof.

We now arrive at the main result of this section.

THEOREM 1. *If the prior distribution of* $\boldsymbol{\theta}$ *is* $\mathrm{Be}(\mathbf{c})$, $\mathbf{c}$ *nonrandom, then the posterior distribution of* $\boldsymbol{\theta}$ *given* $m_1, \cdots, m_j$ *is* $\mathrm{Be}(\boldsymbol{\nu}_j)$, *where* $\boldsymbol{\nu}_j$, $1 \le j \le r$, *are described by r distinct* MC's

$$\{\hat{\boldsymbol{\nu}}_j = \boldsymbol{\nu}_j(0), \boldsymbol{\nu}_j(1), \cdots, \boldsymbol{\nu}_j(N\alpha_j) = \boldsymbol{\nu}_j\}, \qquad 1 \le j \le r.$$

Let $\hat{\Pi}_j$ *and* Π_j *denote respectively the initial distribution and the distribution of* $\boldsymbol{\nu}_j(N\alpha_j)$ *at the end of the* jth *chain, and let* Π_{jk} *denote the transition probability of the* jth *chain at its* kth *step. Then* $\hat{\Pi}_1(\mathbf{c}) = 1$, $\hat{\Pi}_j$ *is obtained from* Π_{j-1} *for* $j \ge 2$ *by* (3), *and* Π_{jk} *are given by* (2).

PROOF. The theorem follows from Lemma 2 by induction.

REMARK 1. Taking $\sum_{l=1}^{j} \theta_l = F(x_j)$, $1 \le j \le r$ and $\theta_{r+1} = 1 - F(x_r)$ where $x_1 < \cdots < x_r$ are the dosage levels, we can find the posterior joint distribution of $F(x_1), \cdots, F(x_r)$ from Theorem 1. To find the joint distribution of F on $y_1 < \cdots < y_s$ in general, suppose $\{y_1, \cdots, y_s\}$ includes $\{x_1, \cdots, x_r\}$ as a subset and let $1 \le \gamma(1) < \cdots < \gamma(r) \le s$ be such that $x_j = y_{\gamma(j)}$ and define $\theta_1^* + \cdots + \theta_l^* = F(y_l)$, $1 \le l \le s$ and $\theta_{s+1}^* = 1 - F(y_s)$. As before, let m_j denote the number of positive responses among $N\alpha_j$ observations at the jth dosage level. Then Theorem 1 extends as follows. Starting with a $(s+1)$-dimensional prior $\mathrm{Be}(\mathbf{c}^*)$ for $\boldsymbol{\theta}^*$, let $\alpha_{\gamma(j)}^* = \alpha_j$, $m_{\gamma(j)}^* = m_j$, $n_{\gamma(j)}^* = n_j$, $1 \le j \le r$, and for all l other than $\gamma(1), \cdots, \gamma(r)$, set $\alpha_l^* = m_l^* = n_l^* = 0$. Now Theorem 1 applies to the posterior of $\boldsymbol{\theta}^*$ given $m_1^*, \cdots, m_l^*$. Of course, some of the MC's will be nonexistent and if the lth chain is such, then $\Pi_{l-1} = \hat{\Pi}_l = \Pi_l$.

3. Asymptotics.

For the asymptotics, we shall take $\mathbf{c} = N\alpha_0\mathbf{p} = N\alpha_0(p_1, \cdots, p_{r+1})$ where $p_l > 0$ with $\sum_1^{r+1} p_l = 1$ and write

$$\begin{aligned} &A_j = \alpha_0 + \alpha_1 + \cdots + \alpha_j, \\ (4) \qquad &P_j = \sum_{l=1}^{j} p_l, \qquad \boldsymbol{\lambda}_j = (\lambda_{j1}, \cdots, \lambda_{jj}), \qquad \lambda_{jl} = p_l/P_j, \\ &Q_j = 1 - P_j, \qquad \boldsymbol{\mu}_j = (\mu_{j1}, \cdots, \mu_{j,r+1-j}), \qquad \mu_{jl} = p_{j+l}/Q_j, \end{aligned}$$

for $1 \le j \le r$. Assume that $m_j = \sum_{i=1}^{N\alpha_j} U_{ji}$ differs from $N\alpha_jP_j$ by an order of $\sqrt{N}$. For example, this will be the case when the prior summarizes the information of $N\alpha_0$ complete observations. Specifically,

$$(5) \qquad m_j = m_j^{(N)} = N\alpha_jP_j + N^{1/2}V_j \quad \text{and} \quad n_j = n_j^{(N)} = N\alpha_jQ_j - N^{1/2}V_j.$$

For deriving the asymptotic distributions, normalize the first j coordinates of $\boldsymbol{\nu}_j(k)$ as

$$\mathbf{X}_j(k) = \mathbf{X}_j^{(N)}(k) = N^{-1/2}[\boldsymbol{\nu}_j'(k) - (NA_{j-1}P_j + k)\boldsymbol{\lambda}_j], \qquad 0 \le k \le m_j,$$

(6a)

$$\mathbf{X}_j(m_j + k) = \mathbf{X}_j(m_j), \qquad 1 \le k \le n_j$$

and the last $r + 1 - j$ coordinates of $\boldsymbol{\nu}_j(k)$ as

$$\mathbf{Y}_j(k) = \mathbf{Y}_j^{(N)}(k) = N^{-1/2}[\boldsymbol{\nu}_j''(o) - NA_{j-1}Q_j\boldsymbol{\mu}_j], \qquad 0 \le k \le m_j,$$

(6b)

$$\mathbf{Y}_j(m_j + k) = N^{-1/2}[\boldsymbol{\nu}_j''(m_j + k) - (NA_{j-1}Q_j + k)\boldsymbol{\mu}_j], \qquad 1 \le k \le n_j,$$

where $\boldsymbol{\nu}_j'(k) = (\nu_{j1}(k), \cdots, \nu_{jj}(k)), \boldsymbol{\nu}_j''(k) = (\nu_{j,j+1}(k), \cdots, \nu_{j,r+1}(k))$ and A_j, P_j, Q_j, $\boldsymbol{\lambda}_j$, $\boldsymbol{\mu}_j$, m_j and n_j are given by (4) and (5). The reason for such normalization is that according to our arrangement of the U_{ji}, $\boldsymbol{\nu}_j'(m_j + k)$ remains unchanged for $0 \le k \le n_j$ and $\boldsymbol{\nu}_j''(k)$ remains unchanged for $0 \le k \le m_j$. Putting $\mathbf{X}_j(k)$ and $\mathbf{Y}_j(k)$ together, the $(r + 1)$-dimensional MC's

(7a) $$\mathbf{Z}_j(k) = \mathbf{Z}_j^{(N)}(k) = (\mathbf{X}_j(k), \mathbf{Y}_j(k)), \qquad 0 \le k \le N\alpha_j, 1 \le j \le r,$$

are obtained as the normalized version of $\{\boldsymbol{\nu}_j(k)\}$. Finally, $\boldsymbol{\nu}_j$ and $\hat{\boldsymbol{\nu}}_j$ are normalized as

(7b) $$\mathbf{Z}_j = \mathbf{Z}_j^{(N)} = N^{-1/2}(\boldsymbol{\nu}_j - NA_j\mathbf{p}), \qquad \hat{\mathbf{Z}}_j = \hat{\mathbf{Z}}_j^{(N)} = N^{-1/2}(\hat{\boldsymbol{\nu}}_j - NA_{j-1}\mathbf{p}).$$

As in the case of $\boldsymbol{\nu}_j^{(N)}$, $\hat{\boldsymbol{\nu}}_j^{(N)}$ and $\{\boldsymbol{\nu}_j^{(N)}(k)\}$, denote the distributions of $\mathbf{Z}_j^{(N)}$, $\hat{\mathbf{Z}}_j^{(N)}$ and the transition probabilities of $\{\mathbf{Z}_j^{(N)}(k)\}$ as $\Pi_j^{(N)}$, $\hat{\Pi}_j^{(N)}$ and $\Pi_{jk}^{(N)}$ respectively. Observe that though $\boldsymbol{\nu}_j = \boldsymbol{\nu}_j(N\alpha_j)$, their normalized versions $\mathbf{Z}_j$ and $\mathbf{Z}_j(N\alpha_j)$ are related by the formula

(8) $$\mathbf{Z}_j = \mathbf{Z}_j(N\alpha_j) + V_j(\boldsymbol{\lambda}_j; -\boldsymbol{\mu}_j),$$

which is easily obtained by using (4)–(7). The connection between the distributions Π_{j-1} of $\mathbf{Z}_{j-1}$ and $\hat{\Pi}_j$ of $\hat{\mathbf{Z}}_j$ will be obtained by transforming formula (3) to obtain

(9)
$$\hat{\Pi}_j(\mathbf{z}) = \Pi_{j-1}(\mathbf{z})\Psi_j^*(\mathbf{z}, m_j) \textstyle\sum_{\mathbf{z}\in\mathscr{Z}_{j-1}} \Pi_{j-1}(\mathbf{z})\Psi_j^*(\mathbf{z}, m_j),$$
$$\Psi_j^*(\mathbf{z}, m_j) = \Psi_j(NA_{j-1}\mathbf{p} + N^{1/2}\mathbf{z}, m_j),$$
$$\mathscr{Z}_{j-1} = \{N^{-1/2}(\mathbf{a} - NA_{j-1}p) \,|\, \mathbf{a} \in S_{j-1}^*\}.$$

To study the weak convergence of the MC's $\{\mathbf{Z}_j^{(N)}(k)\}$, transform $\{\mathbf{X}_j(k)\}$ and $\{\mathbf{Y}_j(k)\}$ to continuous time $0 \le t \le 1$ as

$$\xi_j(t) = \xi_j^{(N)}(t) = ([m_jt] + 1 - m_jt)\mathbf{X}_j([m_jt]) + (m_jt - [m_jt])\mathbf{X}_j([m_jt] + 1)$$
$$\eta_j(t) = \eta_j^{(N)}(t) = ([n_jt] + 1 - n_jt)\mathbf{Y}_j(m_j + [n_jt]) + (n_jt - [n_jt])\mathbf{Y}_j(m_j + [n_jt] + 1).$$

It will be shown that the mutually independent processes $\xi_j^{(N)}(t)$ and $\eta_j^{(N)}(t)$ starting at $\mathbf{x}^*$ and $\mathbf{y}^*$, converge weakly to $\xi_j^*(t)$ and $\eta_j^*(t)$ respectively, which are solutions of appropriate stochastic differential equations. Hence conditionally, given $\mathbf{Z}_j^{(N)}(0) = (\mathbf{x}^*; \mathbf{y}^*)$, $\mathbf{Z}_j^{(N)}(N\alpha_j) \to_w (\xi_j^*(1), \eta_j^*(1))$. The asymptotic posterior distribution will be derived from the solutions of these stochastic differential equations. Since $\{\xi_j^{(N)}(t)\}$ and $\{\eta_j^{(N)}(t)\}$ are independent, we shall treat them separately.

Weak convergence of MC's to diffusion has been studied by Skorokhod (1965) and later (under simpler conditions) by Strook and Varadhan (1969). We shall use Theorem 10.3 of Strook and Varadhan for our purpose. Since

$$|\mathbf{X}_j^{(N)}(k + 1) - \mathbf{X}_j^{(N)}(k)| \le N^{-1/2} \max_{1\le j\le l} |\mathbf{e}_{jl} - \boldsymbol{\lambda}_j| \le N^{-1/2}(1 + |\boldsymbol{\lambda}_j|^2)^{1/2}$$

with probability 1, it is enough to analyze the limiting behavior of the conditional mean vector and product-moment matrix of the increment $\mathbf{X}_j^{(N)}(k + 1) - \mathbf{X}_j^{(N)}(k)$ given $\mathbf{X}_j^{(N)} = \mathbf{x}$. Moreover, the sum of the j coordinates of $\mathbf{X}_j^{(N)}(k)$ remains constant for all k, so for $\xi_j^{(N)}(0) = \mathbf{X}_j^{(N)}(0)$ starting at $\mathbf{x}^*$, it is enough to restrict attention to those $\mathbf{x}$ for which

$\sum_{l=1}^{j} x_l = \sum_{l=1}^{j} x_l^*$. For these $\mathbf{x}$ it can be verified that

$$E[\mathbf{X}_j^{(N)}(k+1) - \mathbf{x} \mid \mathbf{X}_j^{(N)}(k) = \mathbf{x}] = \mathbf{b}(k/m_j, \mathbf{x})/m_j + \boldsymbol{\beta}(N; k, \mathbf{x}),$$

$$E[(\mathbf{X}_j^{(N)}(k+1) - \mathbf{x}) \otimes (\mathbf{X}_j^{(N)}(k+1) - \mathbf{x}) \mid \mathbf{X}_j^{(N)}(k) = \mathbf{x}] = A(k/m_j, \mathbf{x})/m_j + R(N; k, \mathbf{x}),$$

with the vector $\mathbf{b}(t, \mathbf{x})$ and the matrix $A(t, \mathbf{x})$ given by

$$\mathbf{b}(t, \mathbf{x}) = \alpha_j(\mathbf{x} - \boldsymbol{\lambda}_j \textstyle\sum_{l=1}^{j} x_l^*)/(A_{j-1} + \alpha_j t), \; A(t, \mathbf{x}) = \alpha_j P_j K(\boldsymbol{\lambda}_j), \tag{10}$$

where $K(\boldsymbol{\lambda}_j)$ is the covariance matrix of a multinomial distribution with probability vector $\boldsymbol{\lambda}_j$, i.e., $K(\boldsymbol{\lambda}_j)$ has $\lambda_{jl}(1 - \lambda_{jl})$ on its diagonal and $-\lambda_{jl}\lambda_{jl'}$ for its off-diagonal elements, while $m_j^{(N)} \mid \boldsymbol{\beta}(N; k, \mathbf{x}) \mid$ and $m_j^{(N)} \mid R(N; k, \mathbf{x}) \mid$ are uniformly bounded and tend to 0 as $N \to \infty$ uniformly for $k \geq 0$ and $\mathbf{x}$ in compact sets. Theorem 10.3 of Strook and Varadhan (1969) thus becomes applicable and $\xi_j^{(N)}(t) \to_w \xi_j^*(t)$, where $\{\xi_j^*(t), 0 \leq t \leq 1\}$ is a diffusion starting at $\mathbf{x}^*$ with drift coefficient vector $\mathbf{b}(t, \mathbf{x})$ and diffusion matrix $A(t, \mathbf{x})$ given by (10).

Let $\varepsilon_{jl}(\boldsymbol{\lambda}_j)$ and $\phi_{jl}(\boldsymbol{\lambda}_j)$, $1 \leq l \leq j - 1$ denote the normalized eigenvectors and the corresponding eigenvalues of the positive semidefinite matrix $K(\boldsymbol{\lambda}_j)$ of rank $j - 1$. Write

$$K^{1/2}(\boldsymbol{\lambda}_j) = \varepsilon_j(\boldsymbol{\lambda}_j)\mathrm{Diag}(\sqrt{\phi_{jl}(\boldsymbol{\lambda}_j)}), \tag{11}$$

where $\varepsilon_j(\boldsymbol{\lambda}_j)$ is the $j \times (j-1)$ matrix whose lth column is $\varepsilon_{jl}(\boldsymbol{\lambda}_j)$. Then the diffusion $\xi_j^*(t)$ can be expressed as the solution of the stochastic differential equation

$$d\xi_j^*(t) = \alpha_j(A_{j-1} + \alpha_j t)^{-1}\{\xi_j^*(t) - (\textstyle\sum_{l=1}^{j} x_l^*)\boldsymbol{\lambda}_j\}\, dt + (\alpha_j P_j)^{1/2} K_j^{1/2}(\boldsymbol{\lambda}_j)\, d\mathbf{W}(t) \tag{12a}$$

starting at $\mathbf{x}^*$, where $\mathbf{W}(t)$ is a $(j-1)$-dimensional standard Brownian motion (see Gikhman and Skorokhod (1969), pages 402–403). In the same manner, $\eta_j^{(N)}(t) \to_w \eta_j^*(t)$ which can be expressed as the solution of

$$d\eta_j^*(t) = \alpha_j(A_{j-1} + \alpha_j t)^{-1}\{\eta_j^*(t) - (\textstyle\sum_{l=1}^{r+1-j} y_l^*)\boldsymbol{\mu}_j\}\, dt - (\alpha_j Q_j)^{1/2} K_j^{1/2}(\boldsymbol{\mu}_j)\, d\mathbf{W}'(t) \tag{12b}$$

starting at $\mathbf{y}^*$, where $K^{1/2}(\boldsymbol{\mu}_j)$ is a $(r + 1 - j) \times (r - j)$ matrix defined analogously as $K^{1/2}(\boldsymbol{\lambda}_j)$ and $\mathbf{W}'(t)$ *is a* $(r - j)$-dimensional standard Brownian motion independent of $\mathbf{W}(t)$.

The stochastic d.e.'s (12a) and (12b) are linear equations of a particularly simple form for which explicit solutions can be obtained by standard methods (see Theorem 8.2.2 of Arnold (1974)). The solution of (12a) starting at $\mathbf{x}^*$ is

$$\xi_j^*(t) = (1 + \alpha_j A_{j-1}^{-1} t)\mathbf{x}^* - \alpha_j A_{j-1}^{-1} t(\textstyle\sum_{l=1}^{j} x_l^*)\boldsymbol{\lambda}_j$$
$$+ (A_{j-1} + \alpha_j t)(\alpha_j P_j)^{1/2} K^{1/2}(\boldsymbol{\lambda}_j) \int_0^t (A_{j-1} + \alpha_j s)^{-1}\, d\mathbf{W}(s).$$

Hence $\xi_j^*(t)$ follows a j-dimensional (singular) normal distribution with mean vector

$$(1 + \alpha_j A_{j-1}^{-1} t)\mathbf{x}^* - \alpha_j A_{j-1}^{-1} t(\textstyle\sum_{l=1}^{j} x_l^*)\boldsymbol{\lambda}_j$$

and covariance matrix $\alpha_j P_j(1 + \alpha_j A_{j-1}^{-1} t)tK(\boldsymbol{\lambda}_j)$. Similarly, $\eta_j^*(t)$ follows a $(r + 1 - j)$-dimensional normal distribution with mean vector and covariance matrix given by replacing P_j, $\boldsymbol{\lambda}_j$, $\mathbf{x}^*$ and $\sum_{l=1}^{j} x_l^*$ by Q_j, $\boldsymbol{\mu}_j$, $\mathbf{y}^*$ and $\sum_{l=1}^{r+1-j} y_l^*$ respectively in the above formulas. Furthermore, $\xi_j^*(t)$ and $\eta_j^*(t)$ are independent. Using these results in conjunction with (8), we have the following theorem.

THEOREM 2. *Conditionally, given* $\hat{\mathbf{Z}}_j^{(N)} = \mathbf{z}^* = (\mathbf{x}^*, \mathbf{y}^*)$, $\mathbf{Z}_j^{(N)} \to_w \mathbf{Z}_j^* = (\xi_j^*(1) + V_j\boldsymbol{\lambda}_j; \eta_j^*(1) - V_j\boldsymbol{\mu}_j)$, *which follows a* $(r+1)$*-dimensional singular normal distribution with*

$$E(\mathbf{Z}_j^*) = (1 + \alpha_j A_{j-1}^{-1})(\mathbf{x}^*; \mathbf{y}^*) - \alpha_j A_{j-1}^{-1}((\textstyle\sum_{l=1}^{j} x_l^*)\boldsymbol{\lambda}_j; (\sum_{l=1}^{r+1-j} y_l^*)\boldsymbol{\mu}_j) + V_j(\boldsymbol{\lambda}_j; -\boldsymbol{\mu}_j),$$

$$\mathrm{Cov}(\mathbf{Z}_j^*) = \alpha_j(1 + \alpha_j A_{j-1}^{-1}) \begin{bmatrix} p_j K(\boldsymbol{\lambda}_j) & 0 \\ 0 & Q_j K(\boldsymbol{\mu}_j) \end{bmatrix}.$$

To find the limiting distribution of $\mathbf{Z}_r^{(N)}$ (which will give us the limiting posterior distribution), these conditional weak convergences have to be pieced together. For this we need to know (a) how the distribution of $\mathbf{Z}_{j-1}^{(N)}$ changes to that of $\hat{\mathbf{Z}}_j^{(N)}$ asymptotically and (b) how the weak convergence of the distribution of $\hat{\mathbf{Z}}_j^{(N)}$ and that of the conditional distribution of $\mathbf{Z}_j^{(N)}$ given $\hat{\mathbf{Z}}_j^{(N)}$ combine to yield the weak convergence of the marginal distribution of $\mathbf{Z}_j^{(N)}$. These questions are dealt with below.

For $j \geq 2$, extend $\Psi_j^*(\mathbf{z}, m_j)$ to all $\mathbf{z} \in \mathbb{R}^{r+1}$ for which its formula is meaningful and set $\Psi_j^*(\mathbf{z}, m_j) = 0$ elsewhere in $\mathbb{R}^{r+1}$. Rewrite (9) as

$$d\hat{\Pi}_j^{(N)}(\mathbf{z}) = g_j^{(N)}(\mathbf{z})\, d\Pi_{j-1}^{(N)}(\mathbf{z}) \Big/ \int_{\mathbb{R}^{r+1}} g_j^{(N)}(\mathbf{z})\, d\Pi_{j-1}^{(N)}(\mathbf{z}),$$

$$g_j^{(N)}(\mathbf{z}) = e^{N\alpha_j} N^{-N\alpha_j} (A_{j-1}^{A_{j-1}} A_j^{-A_j})^N P_j^{-m_j} Q_j^{-n_j} \Psi_j^*(\mathbf{z}, m_j).$$

The sequence of nonnegative functions $\{g_j^{(N)}\}$ is uniformly bounded by $K_j \exp[V_j^2/(2\alpha_j P_j Q_j)]$ where K_j is a constant not depending on N or $\mathbf{z}$, and $g_j^{(N)}(\mathbf{z})$ converges to the bounded continous function

$$g_j(\mathbf{z}) = \frac{A_{j-1}}{A_j} \exp\left[\frac{V_j^2}{2\alpha_j P_j Q_j} - \frac{\alpha_j}{2A_j A_{j-1} P_j Q_j}\left(\textstyle\sum_{l=1}^{j} z_j - \frac{A_{j-1} V_j}{\alpha_j}\right)^2\right] \tag{13}$$

uniformly for $\mathbf{z}$ in compact sets. Hence $\hat{\Pi}_j^{(N)} \to_w \hat{\Pi}_j$ defined by

$$d\hat{\Pi}_j(\mathbf{z}) = g_j(\mathbf{z})\, d\Pi_{j-1}(\mathbf{z}) \Big/ \int_{\mathbb{R}^{r+1}} g_j(\mathbf{z})\, d\Pi_{j-1}(\mathbf{z}), \tag{14}$$

where Π_{j-1} is the weak limit of $\Pi_{j-1}^{(N)}$. This answers question (a).

To answer (b), denote the conditional weak convergence of Theorem 2 by $\tilde{\Pi}_j^{(N)}(\cdot \mid \mathbf{z}) \to_w \tilde{\Pi}_j(\cdot \mid \mathbf{z})$ which is uniform for $\mathbf{z}$ in compact sets by Corollary 10.2 of Strook and Varadhan (1969) and note that $\tilde{\Pi}_j(\cdot \mid \mathbf{z})$ is continuous a.e. $(\hat{\Pi}_j)$ in $\mathbf{z}$. This conditional weak convergence can, therefore, be combined with the weak $\hat{\Pi}_j^{(N)} \to_w \hat{\Pi}_j$ for $\hat{\mathbf{Z}}_j^{(N)}$ to yield the weak limit of the marginal distribution $\Pi_j^{(N)}$ of $\mathbf{Z}_j^{(N)}$ as Π_j defined by

$$\Pi_j(B) = \int_{\mathbb{R}^{r+1}} \tilde{\Pi}_j(B \mid \mathbf{z})\, d\hat{\Pi}_j(\mathbf{z}) \tag{15}$$

for $(r+1)$-dimensional Borel sets B.

Finally we need the following lemma to carry out the computation required for formulas (13) and (14).

LEMMA 3. *Let Π denote a d-dimensional (possibly singular) normal distribution with mean vector $\mathbf{M} = (M_1, \cdots, M_d)$ and covariance matrix $\boldsymbol{\Sigma} = ((\sigma_{ii'}))$. For $-\infty < a < \infty$, $b > 0$ and $1 \leq k \leq d$, define*

$$d\hat{\Pi}(\mathbf{z}) = \exp[-(2b)^{-1}(\textstyle\sum_{l=1}^{k} z_l - a)^2]\, d\Pi(\mathbf{z}) \Big/ \int_{R^d} \exp[-(2b)^{-1}(\sum_{l=1}^{k} z_l - a)^2]\, d\Pi(\mathbf{z}).$$

Then $\hat{\Pi}$ is a d-dimensional (possibly singular) normal distribution with mean vector $\hat{\mathbf{M}} = (\hat{M}_1, \cdots, \hat{M}_d)$ and coveriance matrix $\hat{\boldsymbol{\Sigma}} = ((\hat{\sigma}_{ii'}))$ given by

$$\hat{M}_i = M_i + (\textstyle\sum_{l=1}^{k} \sigma_{il})(b + \sum_{l=1}^{k} \sum_{l'=1}^{k} \sigma_{ll'})^{-1}(a - \sum_{l=1}^{k} M_l),$$

$$\hat{\sigma}_{ii'} = \sigma_{ii'} - (\textstyle\sum_{l=1}^{k} \sigma_{il})(\sum_{l=1}^{k} \sigma_{i'l})(b + \sum_{l=1}^{k} \sum_{l'=1}^{k} \sigma_{ll'})^{-1}.$$

PROOF. $\hat{\Pi}$ is the conditional distribution of $\mathbf{Z}$ given $X = a$ where the marginal distribution of $\mathbf{Z}$ is Π and the conditional distribution of X given $\mathbf{Z}$ is univariate normal with mean $\sum_{l=1}^{k} Z_l$ and variance b. The lemma now follows from the properties of multivariate normal distributions.

We now state and prove our final result.

THEOREM 3. *For $1 \le j \le r$, the posterior distribution of $\boldsymbol{\theta}$ given $\{U_{j'i}, 1 \le i \le N\alpha_{j'}, 1 \le j' \le j\}$ with respect to* $\text{Be}(N\alpha_0\mathbf{p})$ *is* $\text{Be}(\boldsymbol{\nu}_j^{(N)})$ *where $\boldsymbol{\nu}_j^{(N)}$ is an $(r+1)$-dimensional random vector and $\mathbf{Z}_j^{(N)} = N^{-1/2}(\boldsymbol{\nu}_j^{(N)} - NA_j\mathbf{p})$ is asymptotically normally distributed with mean vector $\mathbf{M}_j = (M_{j,1}, \cdots, M_{j,r+1})$ and covariance matrix $\boldsymbol{\Sigma}_j = ((\sigma_{j,ii'}))$ given by the recursion formulas*:

$$M_{1,i} = \begin{cases} V_1, & i = 1 \\ -V_1\mu_{1i}, & 2 \le i \le r+1 \end{cases}, \qquad \alpha_0^2\boldsymbol{\Sigma}_1 = \alpha_0\alpha_1A_1\begin{bmatrix} 0 & 0 \\ 0 & Q_1K(\boldsymbol{\mu}_1)\end{bmatrix},$$

$$\hat{M}_{j,i} = M_{j-1,i} + \frac{\sum_{l=1}^{j}\sigma_{j-1,il}}{A_jA_{j-1}P_jQ_j + \alpha_j\sum_{l=1}^{j}\sum_{l'=1}^{j}\sigma_{j-1,ll'}}\cdot\left(A_{j-1}V_j - \alpha_j\sum_{l=1}^{j}M_{j-1,l}\right)$$

$$\hat{\sigma}_{j,ii'} = \sigma_{j-1,ii'} - \frac{\alpha_j\left(\sum_{l=1}^{j}\sigma_{j-1,il}\right)\left(\sum_{l=1}^{j}\sigma_{j-1,i'l}\right)}{A_jA_{j-1}P_jQ_j + \alpha_j\sum_{l=1}^{j}\sum_{l'=1}^{j}\sigma_{j-1,ll'}}$$

$$A_{j-1}\mathbf{M}_j = A_j\hat{\mathbf{M}}_j + \left(A_{j-1}V_j - \alpha_j\sum_{l=1}^{j}\hat{M}_{j,l}\right)(\boldsymbol{\lambda}_j; -\boldsymbol{\mu}_j)$$

$$A_{j-1}^2\boldsymbol{\Sigma}_j = \alpha_jA_jA_{j-1}\begin{bmatrix} P_jK(\boldsymbol{\lambda}_j) & 0 \\ 0 & Q_jK(\boldsymbol{\mu}_j)\end{bmatrix} + A_j^2\hat{\boldsymbol{\Sigma}}_j - \alpha_jA_jS_j + \alpha_j^2T_j,$$

where $S_j = ((s_{j,ii'}))$, $T_j = ((t_{j,ii'}))$ are given by

$$s_{j,ii'} = \begin{cases} \lambda_{ji'}\sum_{l=1}^{j}\hat{\sigma}_{j,il} + \lambda_{ji}\sum_{l=1}^{j}\hat{\sigma}_{j,i'l}, & 1 \le i, i' \le j \\ \mu_{j,i'-j}\sum_{l=1}^{j}\hat{\sigma}_{j,il} + \lambda_{ji}\sum_{l=j+1}^{r+1}\hat{\sigma}_{j,i'l}, & 1 \le i \le j,\ j+1 \le i' \le r+1 \\ \mu_{j,i'-j}\sum_{l=j+1}^{r+1}\hat{\sigma}_{j,il} + \lambda_{j,i-j}\sum_{l=j+1}^{r+1}\hat{\sigma}_{j,i'l}, & j+1 \le i, i' \le r+1 \end{cases}$$

$$t_{j,ii'} = \begin{cases} \lambda_{ji}\lambda_{ji'}\sum_{l=1}^{j}\sum_{l'=1}^{j}\sigma_{j,ll'}, & 1 \le i, i' \le j \\ \lambda_{ji}\mu_{j,i'-j}\sum_{l=1}^{j}\sum_{l'=j+1}^{r+1}\sigma_{j,ll'}, & 1 \le i \le j,\ j+1 \le i' \le r+1 \\ \mu_{j,i-j}\mu_{j,i'-j}\sum_{j+1}^{r+1}\sum_{l'=j+1}^{r+1}\sigma_{j,ll'}, & j+1 \le i, i' \le r+1, \end{cases}$$

where $K(\boldsymbol{\lambda}_j)$ and $K(\boldsymbol{\mu}_j)$ are given by (11) *and* $V_j = N^{-1/2}\left(\sum_{i=1}^{N\alpha_j}U_{ji} - N\alpha_jP_j\right)$.

PROOF. The asymptotic normality of $\mathbf{Z}_1^{(N)}$ and the formulas for $\mathbf{M}_1$ and $\boldsymbol{\Sigma}_1$ are obtained by an application of Theorem 2 to the MC $\{\mathbf{Z}_1^{(N)}(k), 0 \le k \le N\alpha_j\}$ starting at $\mathbf{0}$. Having established the asymptotic distribution of $\mathbf{Z}_{j-1}^{(N)}$, apply (13), (14) and Lemma 3 to show that $\hat{\mathbf{Z}}_j^{(N)}$ is asymptotically normal with mean vector $\hat{\mathbf{M}}_j = (\hat{M}_{j,1}, \cdots, \hat{M}_{j,r+1})$ and $\hat{\boldsymbol{\Sigma}}_j = ((\hat{\sigma}_{j,ii'}))$ as given by this theorem. Now applying Theorem 2 in conjunction with these formulas for $\hat{\mathbf{M}}_j$ and $\hat{\boldsymbol{\Sigma}}_j$ and (15), the asymptotic normality of $\mathbf{Z}_j^{(N)}$ and the formulas for $\mathbf{M}_j$ and $\boldsymbol{\Sigma}_j$ follow after a little simplification. This completes the proof.

REMARK 2. As we mentioned at the end of Remark 1, from Theorem 3 we may also find the asymptotic posterior joint distribution of $F(y_1), \cdots, F(y_s)$ for a set of points $y_1 < \cdots < y_s$ of which the dosage levels $x_1 < \cdots < x_r$ form a proper subset. For this all we have to do is to take $\alpha_l = 0$ and $V_l = 0$ for all l for which y_l is not one of the dosage levels.

4. Acknowledgement. The suggestions made by a referee were very helpful in clarifying some misleading statements in the original version which improved the presentation of Section 2.

REFERENCES

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

ARNOLD, L. (1974). *Stochastic Differential Equations: Theory and Applications.* Wiley, New York.

AYER, M., BRUNK, H. D., EWING, G. M., REID, W. T. and SILVERMAN, E. (1955). An empirical distribution function for sampling with incomplete information. *Ann. Math. Statist.* **26** 641–647.

BLUM, J. and SUSARLA, V. (1977). On the posterior distribution of a Dirichlet process given randomly right censored observations. *Stoch. Processes Appl.* **5** 207–211.

FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230.

GIKHMAN, I. I. and SKOROKHOD, A. V. (1969). *Introduction to the Theory of Random Processes.* Saunders, Philadelphia.

KRAFT, C. H. and VAN EEDEN, C. (1964). Bayesian bio-assay. *Ann. Math. Statist.* **35** 886–890.

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

SKOROKHOD, A. V. (1965). *Studies in the Theory of Random Processes.* Addison-Wesley, Reading.

STROOK, D. W. and VARADHAN, S. R. S. (1969). Diffusion processes with continuous coefficients II. *Comm. Pure Appl. Math.* **22** 479–530.

SUSARLA, V. and VAN RYZIN, J. (1976). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **61** 897–902.

TURNBULL, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *J. Roy. Statist. Soc. B* **38** 290–295.

WESLEY, M. N. (1976). Bioassay: estimating the mean of the tolerance distribution. Tech. Rep. No. 17, Division of Biostatistics, Stanford Univ.

DIVISION OF STATISTICS
UNIVERSITY OF CALIFORNIA
DAVIS, CA 95616