PROPERTIES OF SAMPLES FROM DISTRIBUTIONS CHOSEN FROM A DIRICHLET PROCESS

Yamato, Hajime
Department of Mathematics, Faculty of Science, Kagoshima University

https://doi.org/10.5109/13358

出版情報：Bulletin of informatics and cybernetics. 21 (1/2), pp.77-83, 1984-03. Research Association of Statistical Sciences
バージョン：
権利関係：

Bulletin of Informatics and Cybernetics Vol. 21, No. 1~2, 1984

PROPERTIES OF SAMPLES FROM DISTRIBUTIONS CHOSEN FROM A DIRICHLET PROCESS*

By

Hajime YAMATO**

Abstract

The joint distributions of samples from distributions chosen from a Dirichlet process with nonatomic parameter are given and the conditional distributions of the samples are derived, by the method different from Yamato [4]. By making use of the above result, the expectations of functions of the samples are evaluated.

1. Introduction

The Dirichlet process was introduced by Ferguson [2] for Bayesian nonparametric inference. It is well-known that a distribution chosen from a Dirichlet process is discrete with probability one. The purpose of this paper is to show properties of samples from distributions chosen from a Dirichlet process with nonatomic parameter by the method different from Yamato [4] and to give its application. The author assumes that readers are familiar with the Dirichlet process. For the definition of the Dirichlet process, see Ferguson [2].

Let $\boldsymbol{R}$ be the real line and let $\boldsymbol{B}$ be the σ-field of Borel sets. Let α be a nonnull finite measure on $(\boldsymbol{R}, \boldsymbol{B})$. $Q(\cdot)$ denotes $\alpha(\cdot)/\alpha(\boldsymbol{R})$ and M denotes $\alpha(\boldsymbol{R})$. We list some properties of the Dirichlet process for the later use.

LEMMA 1 (*Ferguson* [2]). *Let* $\boldsymbol{P}$ *be a Dirichlet process on* $(\boldsymbol{R}, \boldsymbol{B})$ *with parameter* α *and let* X *be a sample of size* 1 *from* $\boldsymbol{P}$. *Then for* $A \in \boldsymbol{B}$

$$P(X \in A) = Q(A)\,.$$

Let $X_1, \cdots, X_n$ be a sample of size n from a distribution $\boldsymbol{P}$ chosen from a Dirichlet process on $(\boldsymbol{R}, \boldsymbol{B})$ with parameter α. Then, as stated in Korwar and Hollander [3], we can view the observations $X_1, \cdots, X_n$ as being obtained sequentially as follows: Let X_1 be a sample of size 1 from $\boldsymbol{P}$; having obtained X_1, let X_2 be a sample of size 1 from the conditional distribution $\boldsymbol{P}$ given X_1; and so on until $X_1, \cdots, X_n$ are obtained. Thus by Lemma 1 we have the following lemma, which is essentially similar to the statement of Zehnwirth [5, p. 16].

LEMMA 2. *Let* $\boldsymbol{P}$ *be a Dirichlet process on* $(\boldsymbol{R}, \boldsymbol{B})$ *with parameter* α *and let* $X_1, \cdots, X_n$ *be a sample of size* n *from* P. *Then we can view* X_1 *has the distribution* Q

* This research was partly supported by the Grant-in-Aid for Scientific Research Project No. 58540117 from the Ministry of Education.

** Department of Mathematics, Faculty of Science, Kagoshima University, Kagoshima 890, Japan.

and for $k=1, \cdots, n-1$ *the conditional distribution* X_{k+1} *given* $X_1, \cdots, X_k$ *is given by* $\left(MQ(\cdot)+\sum_{j=1}^{k}\delta_{X_j}(\cdot)\right)\Big/(M+k)$, *where for* $x\in X$, δ_x *denotes the measure on* $(\boldsymbol{R}, \boldsymbol{B})$ *giving the mass one to the point* x.

In Section 2, we shall give the joint distribution of samples from distributions chosen from a Dirichlet process with nonatomic parameter, by the method different from Yamato [4]. Furthermore, we shall derive the conditional distribution of the samples, which is essentially similar to Theorem 3.1 of Yamato [4].

We shall use the above result to evaluate expectation of functions of the samples for nonatomic parameter in Section 3.

2. Properties of Samples

Let $\boldsymbol{R}$ be the real line and let $\boldsymbol{B}$ be the σ-field of Borel sets. Let α be a nonnull finite measure on $(\boldsymbol{R}, \boldsymbol{B})$ and nonatomic. $Q(\cdot)$ denotes $\alpha(\cdot)/\alpha(\boldsymbol{R})$ and M denotes $\alpha(\boldsymbol{R})$. Let $X_1, \cdots, X_n$ be a sample of size n from a distribution $\boldsymbol{P}$ chosen from a Dirichlet process on $(\boldsymbol{R}, \boldsymbol{B})$ with parameter α. We can consider that the sample $X_1, \cdots, X_n$ is obtained sequentially, as stated in Section 1. For nonnegative integers $m(1), \cdots, m(n)$ with $\sum_{i=1}^{n} im(i)=n$, let $(X_1, X_2, \cdots, X_n)\in C(m(1), \cdots, m(n))$ be the event that there are $m(1)$ distinct values of X that occur only once, $m(2)$ that occur exactly twice, $\cdots$, $m(n)$ that occur exactly n times. We denote the sample $(X_1, \cdots, X_n)$ with $(X_1, \cdots, X_n)\in C(m(1), \cdots, m(n))$ by $(X_{11}, \cdots, X_{1m(1)}, X_{21}, X_{21}, \cdots, X_{2m(2)}, X_{2m(2)}, \cdots, X_{n1}, \cdots, X_{n1})$. Note that if $m(n)\geqq 1$ then $m(1)=\cdots=m(n-1)=0$ and $m(n)=1$. If $m(1)=2$ and $X_s\neq X_t$ with $s<t$ are different from the remainders, then $X_{11}=X_s$, $X_{12}=X_t$. Suppose that $m(j)=m(1<j<m)$. If $X_{s(1)}=\cdots=X_{t(1)}$, $X_{s(2)}=\cdots=X_{t(2)}, \cdots, X_{s(m)}=\cdots=X_{t(m)}$ with $s(1)<s(2)<\cdots<s(m)$ and $s(i)=\min(s(i), \cdots, t(i))$ $(i=1, \cdots, m)$ and the number of each equal X's are j, then $X_{j1}, X_{j2}, \cdots, X_{jm(j)}$ are equal to $X_{s(1)}, X_{s(2)}, \cdots, X_{s(m)}$ in that order. The following lemma is essentially similar to Proposition 3.2 of Yamato [4].

LEMMA 3. *For any* $A_{ij}\in\boldsymbol{B}(i=1, \cdots, n, j=1, \cdots, m(i))$,

$$P(X_{ij}\in A_{ij}(i=1, \cdots, n, j=1, \cdots, m(i)), (X_1, \cdots, X_n)\in C(m(1), \cdots, m(n)))$$
$$=n!\,M^{\Sigma_1^m m(i)}\prod_{i=1}^{n}\prod_{j=1}^{m(i)}Q(A_{ij})/M^{(n)}\prod_{i=1}^{n}(m(i)!\,i^{m(i)}), \tag{2.1}$$

where $M^{(n)}=M(M+1)\cdots(M+n-1)$.

Before proving Lemma 3 we shall prepare Lemma 4. For nonnegative integers $m(1), \cdots, m(n)$ with $\sum_{i=1}^{n} im(i)=n$, let $(X_n, X_{n-1}, \cdots, X_1)\in C_0(m(1), \cdots, m(n))$ be the event that $X_n, X_{n-1}, \cdots, X_{n-(m(1)-1)}$, in that order, are unique in the sample and occur only once; that $X_{n-m(1)}, \cdots, X_{n-(m(1)+2m(2)-1)}$ occur twice each in the order $X_{n-m(1)}=X_{n-m(1)-1}$, $\cdots$, $X_{n-(m(1)+2m(2)-2)}=X_{n-(m(1)+2m(2)-1)}$ and etc. We use the similar notations to Antoniak [1] with respect to C and C_0. We denote the sample $X_n, X_{n-1}, \cdots, X_1$ with $(X_n, \cdots, X_1)\in C_0(m(1), \cdots, m(n))$ by $Y_{11}, \cdots, Y_{1m(1)}, Y_{21}, Y_{21}, \cdots, Y_{2m(2)}, Y_{2m(2)}, \cdots$. Similarly we denote the realization of the above sample, $x_n, x_{n-1}, \cdots, x_1$, by $y_{11}, \cdots, y_{1m(1)}, y_{21}, y_{21}, \cdots, y_{2m(2)}, y_{2m(2)}, \cdots$ Then we have the following

LEMMA 4. *For any* $A_{ij}\in \boldsymbol{B}(i=1, \cdots, n,\ j=1, \cdots, m(i))$

$$P(Y_{ij}\in A_{ij}(i=1, \cdots, n,\ j=1, \cdots, m(i)),\ (X_n, \cdots, X_1)\in C_0(m(1), \cdots, m(n)))$$
$$=\prod_{i=1}^{n}((i-1)!\,M)^{m(i)}\prod_{i=1}^{n}\prod_{j=1}^{m(i)}Q(A_{ij})/M^{(n)}\,. \tag{2.2}$$

PROOF. At first we shall prove the lemma for $n=2$. Two non-negative integers $(m(1), m(2))$ with $m(1)+2m(2)=2$ are $(2, 0)$ and $(0, 1)$. Let X_1, X_2 be a sample of size 2.

For $(X_2, X_1)\in C_0(m(1), m(2))$ with $m(1)=2$ and $m(2)=0$, we have $Y_{11}=X_2$, $Y_{12}=X_1$. For any A_1, $A_2\in \boldsymbol{B}$, from Lemma 2 we have

$$\begin{aligned}&P(Y_{11}\in A_2,\ Y_{12}\in A_1,\ (X_2, X_1)\in C_0(2, 0))\\ &=P(X_2\in A_2,\ X_1\in A_1,\ X_2\neq X_1)\\ &=\int_{A_1}P(X_2\in A_2,\ X_2\neq x_1\,|\,x_1)dQ(x_1)\,.\end{aligned}$$

Since from Lemma 2, given $X_1=x_1$, X_2 has the distribution $(\alpha(\cdot)+\delta_{x_1}(\cdot))/(M+1)$ and α is nonatomic, we have

$$\begin{aligned}&P(Y_{11}\in A_2,\ Y_{12}\in A_1,\ (X_2, X_1)\in C_0(2, 0))\\ &=\int_{A_1}\alpha(A_2)/(M+1)dQ(x_1)=Q(A_1)\alpha(A_2)/(M+1)\\ &=M^{m(1)}Q(A_1)Q(A_2)/M^{(2)}\qquad \text{with}\quad m(1)=2\,,\quad m(2)=0\,.\end{aligned}$$

For $(X_2, X_1)\in C_0(m(1), m(2))$ with $m(1)=0$ and $m(2)=1$, we have $Y_{21}=X_2=X_1$. For any $A\in \boldsymbol{B}$, from Lemma 2 we have

$$\begin{aligned}&P(Y_{21}\in A,\ (X_2, X_1)\in C_0(0, 1))\\ &=P(X_2=X_1\in A)=\int_A P(X_2=x_1\,|\,x_1)dQ(x_1)=\int_A 1/(M+1)dQ(x_1)\\ &=M^{m(2)}Q(A)/M^{(2)}\qquad \text{with}\quad m(1)=0\,,\quad m(2)=1\,.\end{aligned}$$

Thus the lemma holds for $n=2$. Next we assume that the lemma holds for $n\geqq 2$ and show that it holds for $n+1$. We denote the sample $X_{n+1}, X_n, \cdots, X_1$ with $(X_{n+1}, X_n, \cdots, X_1)\in C_0(m'(1), \cdots, m'(n+1))$ and $\sum_{i=1}^{n+1}im'(i)=n+1$ by $Y'_{11}, \cdots, Y'_{1m'(1)}, Y'_{21}, Y'_{21}, \cdots, Y'_{2m(2)}, Y'_{2m'(2)}, \cdots$ For a sample of size $n+1$ we have two cases: The one is that X_{n+1} occurs only once and the other is that X_{n+1} equals to the previous observation.

For the case that X_{n+1} occurs only once, we have $m'(1)\geqq 1$, $m'(n+1)=0$ and for $A_{ij}\in \boldsymbol{B}(i=1, \cdots, n,\ j=1, \cdots, m'(i))$

$$\begin{aligned}p_1&=P(Y'_{ij}\in A_{ij}(i=1, \cdots, n,\ j=1, \cdots, m'(i)),\ (X_{n+1}, \cdots, X_1)\in C_0(m'(1), \cdots, m'(n+1))\\ &=\int_{D_1}P(X_{n+1}\in A_{11},\ X_{n+1}\neq x_1, \cdots, x_n\,|\,x_1, \cdots, x_n)dH(x_1, \cdots, x_n)\,,\end{aligned}$$

where $H(x_1, \cdots, x_n)$ is the joint distribution of $X_1, \cdots, X_n$ and

$$\begin{aligned}D_1=\{&(x_1, \cdots, x_n)\,|\,(x_n, \cdots, x_1)\in C_0(m(1), \cdots, m(n)),\ m(1)=m'(1)-1\,,\\ &m(i)=m'(i)(i=2, \cdots, n),\ y_{1,j-1}\in A_{1j}(j=2, \cdots, m'(1)),\end{aligned}$$

$$y_{ij}\in A_{ij}(i=2, \cdots, n, j=1, \cdots, m'(i))\} .$$

Since from Lemma 2, given $X_1, \cdots, X_n$, X_{n+1} has the distribution $\left(\alpha(\cdot)+\sum_{i=1}^{n}\delta_{X_i}(\cdot)\right)\Big/(M+n)$ and α is nonatomic,

$$\begin{aligned}p_1&=\int_{D_1}\alpha(A_{11})/(M+n)dH(x_1, \cdots, x_n)\\&=[\alpha(A_{11})/(M+n)]P((X_1, \cdots, X_n)\in D_1)\\&=[\alpha(A_{11})/(M+n)]P(Y_{1,j-1}\in A_{1j}(j=2, \cdots, m'(1)),\\&\quad Y_{ij}\in A_{ij}(i=2, \cdots, n, j=1, \cdots, m'(i)),\\&\quad (X_n, \cdots, X_1)\in C_0(m'(1)-1, m'(2), \cdots, m'(n))) .\end{aligned}$$

Since we assume that the lemma holds for n and $m'(n+1)=0$,

$$\begin{aligned}p_1&=[\alpha(A_{11})/(M+n)]M^{m'(1)-1}\prod_{i=2}^{n}((i-1)!\,M)^{m'(i)}\\&\quad\times\prod_{j=2}^{m'(1)}Q(A_{1j})\prod_{i=2}^{n}\prod_{j=1}^{m'(i)}Q(A_{ij})/M^{(n)}\\&=\prod_{i=1}^{n+1}((i-1)!\,M)^{m'(i)}\prod_{i=1}^{n+1}\prod_{j=1}^{m'(i)}Q(A_{ij})/M^{(n+1)} .\end{aligned}\tag{2.3}$$

For the case that X_{n+1} equals to the previous observation, at first we consider the case of $m'(n+1)=1$ and next the case of $m'(n+1)=0$. In case of $m'(n+1)=1$ where $X_1, \cdots, X_{n+1}$ are all equal, for $A_{n+1,1}\in\boldsymbol{B}$ we have

$$\begin{aligned}p_2&=P(Y'_{n+1,1}\in A_{n+1,1}, (X_{n+1}, \cdots, X_1)\in C_0(m'(1), \cdots, m'(n+1)), m'(n+1)=1)\\&=\int_{D_2}P(X_{n+1}=x_n\,|\,x_1, \cdots, x_n)dH(x_1, \cdots, x_n) ,\end{aligned}$$

where $D_2=\{(x_1, \cdots, x_n)\,|\,x_1=\cdots=x_n\in A_{n+1,1}\}$. Since from Lemma 2 given $X_1, \cdots, X_n$, X_{n+1} has the distribution $\left(\alpha(\cdot)+\sum_{i=1}^{n}\delta_{X_i}(\cdot)\right)\Big/(M+n)$ and α is nonatomic,

$$\begin{aligned}p_2&=\int_{D_2}n/(M+n)dH(x_1, \cdots, x_n)\\&=[n/(M+n)]P(X_n=\cdots=X_1\in A_{n+1,1})\\&=[n/(M+n)]P(Y_{n1}\in A_{n+1,1}, (X_n, \cdots, X_1)\in C_0(m(1), \cdots, m(n)), m(n)=1) .\end{aligned}$$

We assume that the lemma holds for n and therefore

$$\begin{aligned}p_2&=[n/(M+n)]((n-1)!)MQ(A_{n+1,1})/M^{(n)}\\&=(n!\,M)^{m'(n+1)}Q(A_{n+1,1})/M^{(n+1)} \qquad \text{with} \quad m'(n+1)=1 .\end{aligned}\tag{2.4}$$

Finally we consider the case that X_{n+1} equals to the previous observation and $m'(n+1)=0$. Since $m'(1)=0$, we suppose that there exists an integer k such that $2\leqq k\leqq n$, $m'(1)=\cdots=m'(k-1)=0$, $m'(k)\geqq 1$ and $m'(n+1)=0$.

For $A_{ij}\in\boldsymbol{B}(i=k, \cdots, n, j=1, \cdots, m'(i))$, we have

$$p_3=P(Y'_{ij}\in A_{ij}(i=k,\cdots,n,\ j=1,\cdots,m'(i)),\ (X_{n+1},\cdots,X_1)\in C_0(m'(1),\cdots,m'(n+1)))$$
$$=\int_{D_3}P(X_{n+1}=x_n=\cdots=x_{n-k+2}\mid x_1,\cdots,x_n)dH(x_1,\cdots,x_n),$$

where

$$D_3=\{(x_1,\cdots,x_n)\mid(x_n,\cdots,x_1)\in C_0(m(1),\cdots,m(n)),\ m(i)=0$$
$$(i=1,\cdots,k-2),\ m(k-1)=1,\ m(k)=m'(k)-1$$
$$m(i)=m'(i)(i=k+1,\cdots,n),\ y_{k-1,1}\in A_{k1}$$
$$y_{k,j-1}\in A_{kj}(j=2,\cdots,m'(k)),\ y_{ij}\in A_{ij}(i=k+1,\cdots,n,\ j=1,\cdots,m'(i))\}.$$

By the similar argument to p_2, we have

$$p_3=\int_{D_3}(k-1)/(M+n)dH(x_1,\cdots,x_n)$$
$$=[(k-1)/(M+n)]P((X_1,\cdots,X_n)\in D_3)$$
$$=[(k-1)/(M+n)]P(Y_{k-1,1}\in A_{k1},\ Y_{k,j-1}\in A_{kj}(j=2,\cdots,m'(k)),$$
$$Y_{ij}\in A_{ij}(i=k+1,\cdots,n,\ j=1,\cdots,m'(i)),$$
$$(X_n,\cdots,X_1)\in C_0(0,\cdots,0,1,m'(k)-1,m'(k+1),\cdots m'(n)))$$
$$=[(k-1)/(M+n)[((k-2)!M)^{m'(k)-1}\prod_{i=k+1}^{n}((i-1)!M)^{m'(i)}$$
$$\times Q(A_{k1})\prod_{j=2}^{m'(k)}Q(A_{kj})\prod_{i=k+1}^{n}\prod_{j=1}^{m'(i)}Q(A_{ij})/M^{(n)}$$
$$=\prod_{i=k}^{n}((i-1)!M)^{m'(i)}\prod_{i=k}^{n}\prod_{j=1}^{m'(i)}Q(A_{ij})/M^{(n+1)} \tag{2.5}$$

From the evaluations of p_1, p_2, p_3, we know that the lemma holds for $n+1$ and thus proved it by induction.

PROOF OF LEMMA 3. Lemma 4 also holds for $(X_1,\cdots,X_n)\in C_0(m(1),\cdots,m(n))$. The number of ways that n observations $X_1,\cdots,X_n$ are permuted differently with $(X_1,\cdots,X_n)\in C(m(1),\cdots,m(n))$ and $\sum_{i=1}^{n}im(i)=n$ is $n!\Big/\sum_{i=1}^{n}[m(i)!(i!)^{m(i)}]$. To multiply the right-hand side of (2.2) with $(X_1,\cdots,X_n)\in C_0(m(1),\cdots,m(n))$ by this number yields (2.1).

If we take $A_{ij}=\boldsymbol{R}$ for $i=1,\cdots,n$, $j=1,\cdots,m(i)$ in Lemma 3, then we have the following lemma which is found in Antoniak [1].

LEMMA 5. (*Antoniak* [1]).

$$P((X_1,\cdots,X_n)\in C(m(1),\cdots,m(n)))=n!\,M^{\Sigma_1^n m(i)}/M^{(n)}\prod_{i=1}^{n}(m(i)!\,i^{m(i)}).$$

The following theorem is essentially similar to Theorem 3.1 of Yamato [4].

THEOREM 1. *Given* $(X_1,\cdots,X_n)\in C(m(1),\cdots,m(n))$, $X_{11},X_{12},\cdots,X_{1m(1)},X_{21},X_{22},\cdots,X_{2m(2)},\cdots,X_{n1}$ *are independent and identically distributed with the distribution* Q.

PROOF. For any $A_{ij}\in\boldsymbol{B}(i=1,\cdots,n,\ j=1,\cdots,m(i))$, by Lemma 3 and 5 we have

$$P(X_{ij}\in A_{ij}(i=1, \cdots, n, j=1, \cdots, m(i))\,|\,(X_1, \cdots, X_n)\in C(m(1), \cdots, m(n)))$$
$$=P(X_{ij}\in A_{ij}(i=1, \cdots, n, j=1, \cdots, m(i)),$$
$$(X_1, \cdots, X_n)\in C(m(1), \cdots, m(n)))/P((X_1, \cdots, X_n)\in C(m(1), \cdots, m(n)))$$
$$=\sum_{i=1}^{n}\prod_{j=1}^{m(i)} Q(A_{ij}).$$

3. Expectation of Random Functionals

By the use of Theorem 1 we shall prove the following theorem (Yamato [4]) for nonatomic parameter α. Our method of proof is different from Yamato [4]. $\boldsymbol{R}^n$ is the n-dimensional Euclidean space and $\boldsymbol{B}^n$ is the σ-field of Borel subsets of $\boldsymbol{R}^n$ for $n=2, 3, \cdots$.

THEOREM 2 (*Yamato* [4]). *Let $h(x_1, \cdots, x_n)$ be a real-valued measurable function defined on $(\boldsymbol{R}^n, \boldsymbol{B}^n)$ and symmetric in $x_1, \cdots, x_n$. Let $\boldsymbol{P}$ be a Dirichlet process on $(\boldsymbol{R}, \boldsymbol{B})$ with parameter α. Let $X_1, \cdots, X_n$ be a sample from $\boldsymbol{P}$. Then*

$$Eh(X_1, \cdots, X_n)=\sum{}^*\Big[n\,!\,M^{\Sigma_1^n m(i)}/M^{(n)}\prod_{i=1}^{n}(m(i)\,!\,i^{m(i)})\Big]$$
$$\int_{\boldsymbol{X}^{\Sigma m(i)}} h(x_{11}, \cdots, x_{1m(1)}, x_{21}, x_{21}, \cdots,$$
$$x_{2m(2)}, x_{2m(2)}, \cdots, x_{n1}, \cdots, x_{n1})\prod_{i=1}^{n}\prod_{j=1}^{m(i)} dQ(x_{ij}), \tag{3.1}$$

provided all integrals of the right-hand side exist. Where $\sum^*$ denotes the summation over all n nonnegative integers $m(1), \cdots, m(n)$ satisfying $\sum_{i=1}^{n} im(i)=n$ and in the arguments of the integrand of the right-hand side x_{is} appears at exactly i times for $i=1, 2, \cdots, n$ and $s=1, \cdots, m(i)$.

PROOF. We give the proof for nonatomic parameter α. From Theorem 1, for nonnegative intergers $m(1), \cdots, m(n)$ with $\sum_{i=1}^{n} im(i)=n$, given $(X_1, \cdots, X_n)\in C(m(1), \cdots, m(n))$, $X_{11}, \cdots, X_{1m(n)}, X_{21}, \cdots, X_{2m(2)}, \cdots, X_{n1}$ are independent and identically distributed with the distribution Q. h is symmetric in $x_1, \cdots, x_n$. Therefore we have

$$E[h(X_1, \cdots, X_n)\,|\,(X_1, \cdots, X_n)\in C(m(1), \cdots, m(n))]$$
$$=E[h(X_{11}, \cdots, X_{1m(1)}, X_{21}, X_{21}, \cdots, X_{2m(2)}, X_{2m(2)}, \cdots,$$
$$X_{n1}, \cdots, X_{n1})\,|\,(X_1, \cdots, X_n)\in C(m(1), \cdots, m(n))] \tag{3.2}$$
$$=\int_{\boldsymbol{X}^{\Sigma m(i)}} h(x_{11}, \cdots, x_{1m(1)}, x_{21}, x_{21}, \cdots,$$
$$x_{2m(2)}, x_{2m(2)}, \cdots, x_{n1}, \cdots, x_{n1})\prod_{i=1}^{n}\prod_{j=1}^{m(i)} dQ(x_{ij}),$$

which exists for each n nonnegative integers $m(1), \cdots, m(n)$ with $\sum_{i=1}^{n} im(i)=n$ by the assumption. Since by Lemma 5 for each n nonnegative integers $m(1), \cdots, m(n)$ with

$\sum_{i=1}^{n} im(i)=n,$

$$P((X_1, \cdots, X_n)\in C(m(1), \cdots, m(n)))=n\,!\,M^{\Sigma_1^n m(i)}/M^{(n)}\prod_{i=1}^{n}(m(i)\,!\,i^{m(i)}),$$

taking expectation of (3.2) we have (3.1).

References

[1] ANTONIAK, C. A.: *Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems.* Ann. Statist. 2 (1974), 1152-1174.

[2] FERGUSON, T. S.: *A Bayesian analysis of some nonparametric problems.* Ann. Statist. 1 (1973), 209-230.

[3] KORWAR, R. M. and HOLLANDER, M.: *Contributions to the theory of Dirichlet processes.* Ann. Probability, 1 (1973), 705-711.

[4] YAMATO, H.: *Expectation of functions of samples from distributions chosen from a Dirichlet process,* (in preparation).

[5] ZEHNWIRTH, B.: *Credibility and the Dirichlet process.* Scan. Actuarial J., (1979), 13-23.

Communicated by S. Kano
Received September 21, 1983