# 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：

# 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 $\sigma$-field of Borel sets. Let $\alpha$ 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 $\alpha$ 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 $\alpha$. 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 $P$ 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 $\alpha$ and let $X_{1}, \cdots, X_{n}$ be a sample of size $n$ from $P$. Then we can view $X_{1}$ has the distribution $Q$

[^0]and for $k=1, \cdots, n-1$ the conditional distribution $X_{k+1}$ given $X_{1}, \cdots, X_{k}$ is given by $\left(M Q(\cdot)+\sum_{j=1}^{k} \delta_{X_{j}}(\cdot)\right) /(M+k)$, where for $x \in X, \delta_{x}$ 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 $\sigma$-field of Borel sets. Let $\alpha$ 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 $\alpha$. 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 $\sum_{i=1}^{n} i m(i)=n$, let $\left(X_{1}, X_{2}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n))$ 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 $\left(X_{1}, \cdots, X_{n}\right)$ with $\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots$, $m(n)$ ) by ( $\left.X_{11}, \cdots, X_{1 m(1)}, X_{21}, X_{21}, \cdots, X_{2 m(2)}, X_{2 m(2)}, \cdots, X_{n 1}, \cdots, X_{n 1}\right)$. 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 $X_{s(1)}=\cdots=X_{t(1)}, \quad X_{s(2)}=\cdots=X_{t(2)}, \cdots, X_{s(m)}=\cdots=X_{t(m)}$ 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 $X_{j 1}, X_{j 2}, \cdots, X_{j m(j)}$ 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 $A_{i j} \in \boldsymbol{B}(i=1, \cdots, n, j=1, \cdots, m(i))$,

$$
\begin{align*}
& P\left(X_{i j} \in A_{i j}(i=1, \cdots, n, j=1, \cdots, m(i)),\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n))\right) \\
& \quad=n!M_{1}^{\Sigma_{1}^{m} m(i)} \prod_{i=1}^{n} \prod_{j=1}^{m(i)} Q\left(A_{i j}\right) / M^{(n)} \prod_{i=1}^{n}\left(m(i)!i^{m(i)}\right), \tag{2.1}
\end{align*}
$$

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 $\sum_{i=1}^{n} i m(i)=n$, let $\left(X_{n}, X_{n-1}, \cdots, X_{1}\right) \in C_{0}(m(1), \cdots, m(n))$ 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 $X_{n-m(1)}, \cdots, X_{n-(m(1)+2 m(2)-1)}$ occur twice each in the order $X_{n-m(1)}=X_{n-m(1)-1}$, $\cdots, X_{n-(m(1)+2 m(2)-2)}=X_{n-(m(1)+2 m(2)-1)}$ 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 $Y_{11}, \cdots, Y_{1 m(1)}, Y_{21}, Y_{21}, \cdots, Y_{2 m(2)}, Y_{2 m(2)}, \cdots$. Similarly we denote the realization of the above sample, $x_{n}, x_{n-1}, \cdots, x_{1}$, by $y_{11}, \cdots, y_{1 m(1)}, y_{21}, y_{21}, \cdots, y_{2 m(2)}$, $y_{2 m(2)}, \cdots$ Then we have the following

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

$$
\begin{align*}
& P\left(Y_{i j} \in A_{i j}(i=1, \cdots, n, j=1, \cdots, m(i)),\left(X_{n}, \cdots, X_{1}\right) \in C_{0}(m(1), \cdots, m(n))\right) \\
& \quad=\prod_{i=1}^{n}((i-1)!M)^{m(i)} \prod_{i=1}^{n} \prod_{j=1}^{m(i)} Q\left(A_{i j}\right) / M^{(n)} \tag{2.2}
\end{align*}
$$

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

For $\left(X_{2}, X_{1}\right) \in C_{0}(m(1), m(2))$ with $m(1)=2$ and $m(2)=0$, we have $Y_{11}=X_{2}, \quad Y_{12}=X_{1}$. For any $A_{1}, A_{2} \in \boldsymbol{B}$, from Lemma 2 we have

$$
\begin{aligned}
& P\left(Y_{11} \in A_{2}, Y_{12} \in A_{1},\left(X_{2}, X_{1}\right) \in C_{0}(2,0)\right) \\
& \quad=P\left(X_{2} \in A_{2}, X_{1} \in A_{1}, X_{2} \neq X_{1}\right) \\
& \quad=\int_{A_{1}} P\left(X_{2} \in A_{2}, X_{2} \neq x_{1} \mid x_{1}\right) d Q\left(x_{1}\right)
\end{aligned}
$$

Since from Lemma 2, given $X_{1}=x_{1}, X_{2}$ has the distribution $\left(\alpha(\cdot)+\delta_{x_{1}}(\cdot)\right) /(M+1)$ and $\alpha$ is nonatomic, we have

$$
\begin{aligned}
& P\left(Y_{11} \in A_{2}, Y_{12} \in A_{1},\left(X_{2}, X_{1}\right) \in C_{0}(2,0)\right) \\
& \quad=\int_{A_{1}} \alpha\left(A_{2}\right) /(M+1) d Q\left(x_{1}\right)=Q\left(A_{1}\right) \alpha\left(A_{2}\right) /(M+1) \\
& \quad=M^{m(1)} Q\left(A_{1}\right) Q\left(A_{2}\right) / M^{(2)} \quad \text { with } \quad m(1)=2, \quad m(2)=0
\end{aligned}
$$

For $\left(X_{2}, X_{1}\right) \in C_{0}(m(1), m(2))$ 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

$$
\begin{aligned}
& P\left(Y_{21} \in A,\left(X_{2}, X_{1}\right) \in C_{0}(0,1)\right) \\
& \quad=P\left(X_{2}=X_{1} \in A\right)=\int_{A} P\left(X_{2}=x_{1} \mid x_{1}\right) d Q\left(x_{1}\right)=\int_{A} 1 /(M+1) d Q\left(x_{1}\right) \\
& \quad=M^{m(2)} Q(A) / M^{(2)} \quad \text { with } \quad m(1)=0, \quad m(2)=1
\end{aligned}
$$

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 $\left(X_{n+1}, X_{n}, \cdots\right.$, $\left.X_{1}\right) \in C_{0}\left(m^{\prime}(1), \cdots, m^{\prime}(n+1)\right)$ and $\sum_{i=1}^{n+1} i m^{\prime}(i)=n+1$ by $Y_{11}^{\prime}, \cdots, Y_{1 m^{\prime}(1)}^{\prime}, Y_{21}^{\prime}, Y_{21}^{\prime}, \cdots, Y_{2 m(2)}^{\prime}$, $Y_{2 m^{\prime}(2)}^{\prime}, \cdots$ 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 $m^{\prime}(1) \geqq 1, m^{\prime}(n+1)=0$ and for $A_{i j} \in \boldsymbol{B}\left(i=1, \cdots, n, j=1, \cdots, m^{\prime}(i)\right)$

$$
\begin{aligned}
p_{1} & =P\left(Y_{i j}^{\prime} \in A_{i j}\left(i=1, \cdots, n, j=1, \cdots, m^{\prime}(i)\right),\left(X_{n+1}, \cdots, X_{1}\right) \in C_{0}\left(m^{\prime}(1), \cdots, m^{\prime}(n+1)\right)\right. \\
& =\int_{D_{1}} P\left(X_{n+1} \in A_{11}, X_{n+1} \neq x_{1}, \cdots, x_{n} \mid x_{1}, \cdots, x_{n}\right) d H\left(x_{1}, \cdots, x_{n}\right)
\end{aligned}
$$

where $H\left(x_{1}, \cdots, x_{n}\right)$ is the joint distribution of $X_{1}, \cdots, X_{n}$ and

$$
\begin{aligned}
D_{1}= & \left\{\left(x_{1}, \cdots, x_{n}\right) \mid\left(x_{n}, \cdots, x_{1}\right) \in C_{0}(m(1), \cdots, m(n)), m(1)=m^{\prime}(1)-1\right. \\
& m(i)=m^{\prime}(i)(i=2, \cdots, n), y_{1, j-1} \in A_{1 j} j\left(j=2, \cdots, m^{\prime}(1)\right)
\end{aligned}
$$

$$
\left.y_{i j} \in A_{i j}\left(i=2, \cdots, n, j=1, \cdots, m^{\prime}(i)\right)\right\} .
$$

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

$$
\begin{aligned}
p_{1}= & \int_{D_{1}} \alpha\left(A_{11}\right) /(M+n) d H\left(x_{1}, \cdots, x_{n}\right) \\
= & {\left[\alpha\left(A_{11}\right) /(M+n)\right] P\left(\left(X_{1}, \cdots, X_{n}\right) \in D_{1}\right) } \\
= & {\left[\alpha\left(A_{11}\right) /(M+n)\right] P\left(Y_{1, j-1} \in A_{1 j}\left(j=2, \cdots, m^{\prime}(1)\right),\right.} \\
& Y_{i j} \in A_{i j}\left(i=2, \cdots, n, j=1, \cdots, m^{\prime}(i)\right), \\
& \left.\left(X_{n}, \cdots, X_{1}\right) \in C_{0}\left(m^{\prime}(1)-1, m^{\prime}(2), \cdots, m^{\prime}(n)\right)\right) .
\end{aligned}
$$

Since we assume that the lemma holds for $n$ and $m^{\prime}(n+1)=0$,

$$
\begin{align*}
p_{1}= & {\left[\alpha\left(A_{11}\right) /(M+n)\right] M^{m^{\prime}(1)-1} \prod_{i=2}^{n}((i-1)!M)^{m^{\prime}(i)} } \\
& \times \prod_{j=2}^{m^{\prime}(1)} Q\left(A_{1 j}\right) \prod_{i=2}^{n} \prod_{j=1}^{m^{\prime}(i)} Q\left(A_{i j}\right) / M^{(n)} \\
= & \prod_{i=1}^{n+1}((i-1)!M)^{m^{\prime}(i)} \prod_{i=1}^{n+1} \prod_{j=1}^{m^{\prime}(i)} Q\left(A_{i j}\right) / M^{(n+1)} \tag{2.3}
\end{align*}
$$

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

$$
\begin{aligned}
p_{2} & =P\left(Y_{n+1,1}^{\prime} \in A_{n+1,1},\left(X_{n+1}, \cdots, X_{1}\right) \in C_{0}\left(m^{\prime}(1), \cdots, m^{\prime}(n+1)\right), m^{\prime}(n+1)=1\right) \\
& =\int_{D_{2}} P\left(X_{n+1}=x_{n} \mid x_{1}, \cdots, x_{n}\right) d H\left(x_{1}, \cdots, x_{n}\right),
\end{aligned}
$$

where $D_{2}=\left\{\left(x_{1}, \cdots, x_{n}\right) \mid x_{1}=\cdots=x_{n} \in A_{n+1,1}\right\}$. Since from Lemma 2 given $X_{1}, \cdots, X_{n}$, $X_{n+1}$ has the distribution $\left(\alpha(\cdot)+\sum_{i=1}^{n} \delta_{X_{i}}(\cdot)\right) /(M+n)$ and $\alpha$ is nonatomic,

$$
\begin{aligned}
p_{2} & =\int_{D_{2}} n /(M+n) d H\left(x_{1}, \cdots, x_{n}\right) \\
& =[n /(M+n)] P\left(X_{n}=\cdots=X_{1} \in A_{n+1,1}\right) \\
& =[n /(M+n)] P\left(Y_{n 1} \in A_{n+1,1},\left(X_{n}, \cdots, X_{1}\right) \in C_{0}(m(1), \cdots, m(n)), m(n)=1\right) .
\end{aligned}
$$

We assume that the lemma holds for $n$ and therefore

$$
\begin{align*}
p_{2} & =[n /(M+n)]((n-1)!) M Q\left(A_{n+1,1}\right) / M^{(n)} \\
& =(n!M)^{m^{\prime}(n+1)} Q\left(A_{n+1,1}\right) / M^{(n+1)} \quad \text { with } \quad m^{\prime}(n+1)=1 . \tag{2.4}
\end{align*}
$$

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

For $A_{i j} \in \boldsymbol{B}\left(i=k, \cdots, n, j=1, \cdots, m^{\prime}(i)\right)$, we have

$$
\begin{aligned}
p_{3} & =P\left(Y_{i j}^{\prime} \in A_{i j}\left(i=k, \cdots, n, j=1, \cdots, m^{\prime}(i)\right),\left(X_{n+1}, \cdots, X_{1}\right) \in C_{0}\left(m^{\prime}(1), \cdots, m^{\prime}(n+1)\right)\right) \\
& =\int_{D_{3}} P\left(X_{n+1}=x_{n}=\cdots=x_{n-k+2} \mid x_{1}, \cdots, x_{n}\right) d H\left(x_{1}, \cdots, x_{n}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
D_{3}= & \left\{\left(x_{1}, \cdots, x_{n}\right) \mid\left(x_{n}, \cdots, x_{1}\right) \in C_{0}(m(1), \cdots, m(n)), m(i)=0\right. \\
& (i=1, \cdots, k-2), m(k-1)=1, m(k)=m^{\prime}(k)-1 \\
& m(i)=m^{\prime}(i)(i=k+1, \cdots, n), y_{k-1,1} \in A_{k 1} \\
& \left.y_{k, j-1} \in A_{k j}\left(j=2, \cdots, m^{\prime}(k)\right), y_{i j} \in A_{i j}\left(i=k+1, \cdots, n, j=1, \cdots, m^{\prime}(i)\right)\right\} .
\end{aligned}
$$

By the similar argument to $p_{2}$, we have

$$
\begin{align*}
p_{3}= & \int_{D_{3}}(k-1) /(M+n) d H\left(x_{1}, \cdots, x_{n}\right) \\
= & {[(k-1) /(M+n)] P\left(\left(X_{1}, \cdots, X_{n}\right) \in D_{3}\right) } \\
= & {[(k-1) /(M+n)] P\left(Y_{k-1,1} \in A_{k 1}, Y_{k, j-1} \in A_{k j}\left(j=2, \cdots, m^{\prime}(k)\right),\right.} \\
& Y_{i j} \in A_{i j}\left(i=k+1, \cdots, n, j=1, \cdots, m^{\prime}(i)\right), \\
& \left.\left(X_{n}, \cdots, X_{1}\right) \in C_{0}\left(0, \cdots, 0,1, m^{\prime}(k)-1, m^{\prime}(k+1), \cdots m^{\prime}(n)\right)\right) \\
= & {\left[(k-1) /(M+n)\left[((k-2)!M)^{m^{\prime}(k)-1} \prod_{i=k+1}^{n}((i-1)!M)^{m^{\prime}(i)}\right.\right.} \\
& \times Q\left(A_{k 1}\right) \prod_{j=2}^{m^{\prime}(k)} Q\left(A_{k j}\right) \prod_{i=k+1}^{n} \prod_{j=1}^{m^{\prime}(i)} Q\left(A_{i j}\right) / M^{(n)} \\
= & \prod_{i=k}^{n}((i-1)!M)^{m^{\prime}(i)} \prod_{i=k}^{n} \prod_{j=1}^{m^{\prime}(i)} Q\left(A_{i j}\right) / M^{(n+1)} \tag{2.5}
\end{align*}
$$

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 ( $\left.X_{1}, \cdots, X_{n}\right) \in C_{0}(m(1), \cdots, m(n))$. The number of ways that $n$ observations $X_{1}, \cdots, X_{n}$ are permuted differently with $\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n))$ and $\sum_{i=1}^{n} i m(i)=n$ is $n!/ \sum_{i=1}^{n}\left[m(i)!(i!)^{m(i)}\right]$. To multiply the right-hand side of (2.2) with $\left(X_{1}, \cdots, X_{n}\right) \in C_{0}(m(1), \cdots, m(n))$ by this number yields (2.1).

If we take $A_{i j}=\boldsymbol{R}$ 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]).

$$
P\left(\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n))\right)=n!M^{\Sigma_{1}^{n} m(i)} / M^{(n)} \prod_{i=1}^{n}\left(m(i)!i^{m(i)}\right)
$$

The following theorem is essentially similar to Theorem 3.1 of Yamato [4].
Theorem 1. Given $\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n)), X_{11}, X_{12}, \cdots, X_{1 m(1)}, X_{21}, X_{22}, \cdots$, $X_{2 m(2)}, \cdots, X_{n 1}$ are independent and identically distributed with the distribution $Q$.

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

$$
\begin{aligned}
P\left(X_{i j} \in\right. & \left.A_{i j}(i=1, \cdots, n, j=1, \cdots, m(i)) \mid\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n))\right) \\
= & P\left(X_{i j} \in A_{i j}(i=1, \cdots, n, j=1, \cdots, m(i)),\right. \\
& \left.\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n))\right) / P\left(\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n))\right) \\
= & \sum_{i=1}^{n} \prod_{j=1}^{m(i)} Q\left(A_{i j}\right) .
\end{aligned}
$$

## 3. Expectation of Random Functionals

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

Theorem 2 (Yamato [4]). Let $h\left(x_{1}, \cdots, x_{n}\right)$ be a real-valued measurable function defined on $\left(\boldsymbol{R}^{n}, \boldsymbol{B}^{n}\right)$ and symmetric in $x_{1}, \cdots, x_{n}$. Let $\boldsymbol{P}$ be a Dirichlet process on $(\boldsymbol{R}, \boldsymbol{B})$ with parameter $\alpha$. Let $X_{1}, \cdots, X_{n}$ be a sample from $\boldsymbol{P}$. Then

$$
\begin{align*}
E h\left(X_{1}, \cdots, X_{n}\right)= & \Sigma^{*}\left[n!M^{\Sigma_{1}^{n} m(i)} / M^{(n)} \prod_{i=1}^{n}\left(m(i)!i^{m(i)}\right)\right] \\
& \int_{x^{\Sigma m(i)}} h\left(x_{11}, \cdots, x_{1 m(1)}, x_{21}, x_{21}, \cdots,\right. \\
& \left.x_{2 m(2)}, x_{2 m(2)}, \cdots, x_{n 1}, \cdots, x_{n 1}\right) \prod_{i=1}^{n} \prod_{j=1}^{m(i)} d Q\left(x_{i j}\right), \tag{3.1}
\end{align*}
$$

provided all integrals of the right-hand side exist. Where $\Sigma^{*}$ denotes the summation over all $n$ nonnegative integers $m(1), \cdots, m(n)$ satisfying $\sum_{i=1}^{n} i m(i)=n$ and in the arguments of the integrand of the right-hand side $x_{i s}$ 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 $\alpha$. From Theorem 1, for nonnegative intergers $m(1), \cdots, m(n)$ with $\sum_{i=1}^{n} i m(i)=n$, given $\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n)$ ), $X_{11}, \cdots, X_{1 m(n)}, X_{21}, \cdots, X_{2 m(2)}, \cdots, X_{n 1}$ are independent and identically distributed with the distribution $Q . \quad h$ is symmetric in $x_{1}, \cdots, x_{n}$. Therefore we have

$$
\begin{align*}
& E\left[h\left(X_{1}, \cdots, X_{n}\right) \mid\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n))\right] \\
&= E\left[h \left(X_{11}, \cdots, X_{1 m(1)}, X_{21}, X_{21}, \cdots, X_{2 m(2)}, X_{2 m(2)}, \cdots,\right.\right. \\
&\left.\left.X_{n 1}, \cdots, X_{n 1}\right) \mid\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n))\right]  \tag{3.2}\\
&= \int_{X^{\Sigma m(i)}} h\left(x_{11}, \cdots, x_{1 m(1)}, x_{21}, x_{21}, \cdots,\right. \\
&\left.x_{2 m(2)}, x_{2 m(2)}, \cdots, x_{n 1}, \cdots, x_{n 1}\right) \prod_{i=1}^{n} \prod_{j=1}^{m(i)} d Q\left(x_{i j}\right),
\end{align*}
$$

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

$$
P\left(\left(X_{1}, \cdots, X_{n}\right) \in C(m(1), \cdots, m(n))\right)=n!M^{\Sigma_{1}^{n} m(i)} / M^{(n)} \prod_{i=1}^{n}\left(m(i)!i^{m(i)}\right),
$$

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


[^0]:    * 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.

