

THE EXPECTATION OF RANDOM FUNCTIONALS WITH THE DIRICHLET PROCESS AND ITS APPLICATIONS

Nomachi, Toshifumi
Graduate School of Science and Engineering, Kagoshima University

Yamato, Hajime
Faculty of Science, Kagoshima University

<https://doi.org/10.5109/13488>

出版情報 : Bulletin of informatics and cybernetics. 31 (2), pp.165-178, 1999-12. Research
Association of Statistical Sciences

バージョン :

権利関係 :



THE EXPECTATION OF RANDOM FUNCTIONALS WITH THE DIRICHLET PROCESS AND ITS APPLICATIONS

By

Toshifumi NOMACHI* and Hajime YAMATO†

Abstract

We deal with the expectation of random functionals with the Dirichlet process, using Sethuraman's representation. As an application of this expectation, we obtain the Bayes estimates of estimable parameters with the squared error loss based on the Dirichlet prior process. From the Bayes estimates, we obtain the limits of Bayes estimates, which may be used for the non-Bayesian inference. We see the differences of the limits of Bayes estimates from the corresponding U -statistics and the V -statistics, for five estimable parameters of degree two.

1. Introduction

We consider Bayes estimation of estimable parameters with squared error loss based on the Dirichlet process, which is introduced by Ferguson (1973). For some estimable parameters of degrees one and two, he gives the Bayes estimates and their limits. Using the results of Ferguson (1973) and Antoniak (1974), Yamato (1977b) gives the expectation of random functionals with the Dirichlet process. Using this expectation, he gives the Bayes estimates of estimable parameters of any degree. He also shows that for a sample of size n from the fixed distribution, the mean squared difference of the limit of Bayes estimate from the U -statistic is of order $O(n^{-2})$. In Section 2 we quote the definition and some properties of the Dirichlet process from Ferguson (1973) and Sethuraman (1994). We also quote the definition and some properties of the GEM (Generalized Engen-McCloskey) distribution, which is introduced by McCloskey (1965) and Engen (1978) (see Johnson et al. (1997)). In Section 3 we give the expectation of random functionals with the Dirichlet process using the result of Sethuraman (1994) and properties of the GEM distribution. As an application of this expectation, we can obtain Bayes estimates of estimable parameters with squared error loss based on the Dirichlet process. From the Bayes estimates, we obtain the limits of Bayes estimates, which may be used for the non-Bayesian inference. In Section 4 we consider the three estimators of estimable parameters of degree two, which are the limit of Bayes estimate, the U -statistic and the V -statistic for a sample from the fixed distribution. We give the mean

* Graduate School of Science and Engineering, Kagoshima University, Kagoshima 890-0065, Japan; Miyakonojo College of Technology, Miyazaki 885-8567, Japan

† Faculty of Science, Kagoshima University, Kagoshima 890-0065, Japan

squared difference of the limit of Bayes estimate from the corresponding U -statistic. We also give the mean squared errors of the limit of Bayes estimate and the corresponding V -statistic. Using these mean squared difference and the mean squared errors, we see the differences of the limit of Bayes estimate from the U -statistic and the V -statistic for the five estimable parameters of degree two. In Section 5 which is Appendix, two proofs are given for the Proposition 1 stated in Section 3.

2. The Dirichlet Process and the GEM Distribution

2.1. The Dirichlet process

Let $\alpha(\cdot)$ be a finite non-null measure on a measurable space $(\mathcal{X}, \mathcal{A})$ and $\mathcal{P}(\cdot)$ be a random probability on $(\mathcal{X}, \mathcal{A})$. We denote the k -dimensional product measurable space $(\mathcal{X} \times \cdots \times \mathcal{X}, \mathcal{A} \times \cdots \times \mathcal{A})$ by $(\mathcal{X}^k, \mathcal{A}^k)$.

DEFINITION 1. (Ferguson (1973)) We say $\mathcal{P}(\cdot)$ is the Dirichlet process on $(\mathcal{X}, \mathcal{A})$ with parameter α and write $\mathcal{P} \in \mathcal{D}(\alpha)$, if and only if for any m and every finite measurable partition (B_1, \dots, B_m) of \mathcal{X} (i.e. for every $m = 1, 2, \dots$ and $B_1, \dots, B_m \in \mathcal{A}$, $B_1 \cup \dots \cup B_m = \mathcal{X}$ and $B_i \cap B_j = \emptyset$ for $i \neq j$), the random vector $(\mathcal{P}(B_1), \dots, \mathcal{P}(B_m))$ has the Dirichlet distribution with parameter $(\alpha(B_1), \dots, \alpha(B_m))$, $D(\alpha(B_1), \dots, \alpha(B_m))$.

Ferguson (1973) gives the following representation of the Dirichlet process. Let J_1, J_2, \dots be random variables whose distributions are given by the following equations,

$$\begin{aligned} P\{J_1 \leq x_1\} &= \exp(N(x_1)) \quad \text{for } x_1 > 0 \\ P\{J_j \leq x_j | J_{j-1} = x_{j-1}, \dots, J_1 = x_1\} \\ &= \exp\{N(x_j) - N(x_{j-1})\} \quad \text{for } 0 < x_j < x_{j-1}; \quad j = 2, 3, \dots, \end{aligned}$$

where $N(x) = -\alpha(\mathcal{X}) \int_x^\infty e^{-y} y^{-1} dy$. Then the distributions of J_1, J_2, \dots depend on α only through $\alpha(\mathcal{X})$. For each $j = 1, 2, \dots$ we put $P_j = J_j / (\sum_{i=1}^\infty J_i)$. Then, $P_j \geq 0$ a.s. (almost surely) and $\sum_{j=1}^\infty P_j = 1$ a.s. Let V_1, V_2, \dots be a sequence of independent random variables on \mathcal{X} with the probability measure Q , where $Q(\cdot) = \alpha(\cdot) / \alpha(\mathcal{X})$. We assume that V_1, V_2, \dots are independent of P_1, P_2, \dots .

LEMMA 1. (Ferguson (1973)) Let $\mathcal{P}(\cdot)$ be a random probability on $(\mathcal{X}, \mathcal{A})$ defined by

$$\mathcal{P}(A) = \sum_{j=1}^{\infty} P_j \delta_{V_j}(A) \quad \text{for } A \in \mathcal{A}, \quad (2.1)$$

where δ_x is a measure with the total mass one on point x . Then $\mathcal{P}(\cdot)$ is the Dirichlet process with parameter α .

Sethuraman (1994) gives a new representation of the Dirichlet process. Let Y_1, Y_2, \dots be a sequence of independent random variables on \mathcal{X} with probability measure Q .

Let p_1, p_2, \dots be a sequence of random probabilities defined by

$$p_1 = \theta_1 \text{ and } p_j = (1 - \theta_1) \cdots (1 - \theta_{j-1})\theta_j \quad \text{for } j = 2, 3, \dots \quad (2.2)$$

where $\theta_1, \theta_2, \dots$ are independent and identically distributed random variables with a Beta distribution, $\mathcal{B}e(1, \alpha(\mathcal{X}))$.

LEMMA 2. (Sethuraman (1994)) *The random probability on $(\mathcal{X}, \mathcal{A})$ defined by*

$$\mathcal{P}(A) = \sum_{j=1}^{\infty} p_j \delta_{Y_j}(A) \quad \text{for } A \in \mathcal{A}, \quad (2.3)$$

is the Dirichlet process with parameter α .

From Lemma 2, we can get easily the following lemma.

LEMMA 3. (Sethuraman (1994)) *The random probability $\mathcal{P}(\cdot)$ defined by (2.3) satisfies the distributional equation,*

$$\mathcal{P}(A) \stackrel{d}{=} \theta_1 \delta_{Y_1}(A) + (1 - \theta_1) \mathcal{P}(A) \quad \text{for } A \in \mathcal{A}, \quad (2.4)$$

where $\stackrel{d}{=}$ means the stochastic equivalence relation and on the right-hand side \mathcal{P} is independent of θ_1 and Y_1 .

2.2. The GEM distribution

The joint distribution of $\mathbf{p} = (p_1, p_2, \dots)$ defined by (2.2) is called the GEM distribution. Now let $\mathbf{W} = (W_1, W_2, \dots)$ be a size-biased permutation of \mathbf{p} . That is, given \mathbf{p} , the conditional probability of \mathbf{W} is such that for $j = 1, 2, \dots$ and any distinct integers k_1, k_2, \dots, k_j ,

$$P[W_1 = p_{k_1}, \dots, W_j = p_{k_j} | \mathbf{p}] = p_{k_1} \frac{p_{k_2}}{1 - p_{k_1}} \cdots \frac{p_{k_j}}{1 - p_{k_1} - \cdots - p_{k_{j-1}}}.$$

Then \mathbf{W} has the same distribution as \mathbf{p} , since the GEM distribution is invariant under the size-biased permutation (see, for example, Ewens (1990)). Therefore from the lemmas 3 and 4 of Sibuya and Yamato (1995) we have the following lemma.

LEMMA 4. *For any positive integers c_1, c_2, \dots, c_s , we have*

$$E\left[\sum_{j_1 \neq \dots \neq j_s} p_{j_1}^{c_1} \cdots p_{j_s}^{c_s}\right] = \frac{\alpha^s}{\alpha^{[c_1 + \dots + c_s]}} \prod_{i=1}^s (c_i - 1)!,$$

where the above summation extends over all distinct positive integers j_1, \dots, j_s and $\alpha^{[s]} = \alpha(\alpha + 1) \cdots (\alpha + s - 1)$.

From Lemma 4 we have immediately the following lemma. The similar result based on (2.1) is presented by Yamato (1977b) using the result of Antoniak (1974).

LEMMA 5. For any combination (m_1, \dots, m_k) of k non-negative integers satisfying $\sum_{i=1}^k i \cdot m_i = k$, we have

$$E[\sum^* \prod_{i=1}^k \prod_{j=1}^{m_i} p_{ij}^i] = \frac{\alpha^{\sum_{j=1}^k m_j}}{\alpha^{[k]}} \prod_{i=1}^k \{(i-1)!\}^{m_i}, \quad (2.5)$$

where the summation \sum^* extends over every pair of all mutually distinct values p_{ij} ($j = 1, \dots, m_i$; $i = 1, \dots, k$) in (p_1, p_2, \dots) .

3. The Expectation of Random Functionals with the Dirichlet Process and Its Applications

The expectation of random functionals based on the Dirichlet process is given by Yamato (1977b) as follows:

PROPOSITION 1. (Yamato (1977b)) Let $\mathcal{P}(\cdot)$ be the Dirichlet process on $(\mathcal{X}, \mathcal{A})$ with parameter α and let g be a measurable real valued function on $(\mathcal{X}^k, \mathcal{A}^k)$ and symmetric in the arguments. If $\int_{\mathcal{X}^u} |g(x_1, \dots, x_1, \dots, x_u, \dots, x_u)| \prod_{i=1}^u d\alpha(x_i)$ exists for all combinations (r_1, r_2, \dots, r_u) satisfying with $r_1 + r_2 + \dots + r_u = k$ and $u = 1, 2, \dots, k$, then $\int_{\mathcal{X}^k} |g(x_1, \dots, x_k)| \prod_{i=1}^k d\mathcal{P}(x_i)$ is finite a.s. and we have

$$\begin{aligned} & E \int_{\mathcal{X}^k} g(x_1, \dots, x_k) \prod_{i=1}^k d\mathcal{P}(x_i) \\ &= \sum_{(\sum_{i=1}^k i \cdot m_i = k)} \frac{k! (\alpha(\mathcal{X}))^{\sum_{i=1}^k m_i}}{\prod_{i=1}^k [i^{m_i} (m_i)!] [\alpha(\mathcal{X})]^{[k]}} \int_{\mathcal{X}^{\sum_{i=1}^k m_i}} g(\underbrace{x_{11}, \dots, x_{1m_1}}_{m_1}, \\ & \quad \underbrace{x_{21}, x_{21}, \dots, x_{2m_2}, x_{2m_2}, \dots}_{2m_2}, \dots, \underbrace{x_{k1}, \dots, x_{k1}, \dots, x_{km_k}, \dots, x_{km_k}}_{km_k}) \prod_{i=1}^k \prod_{j=1}^{m_i} dQ(x_{ij}), \end{aligned} \quad (3.1)$$

where the above summation $\sum_{(\sum_{i=1}^k i \cdot m_i = k)}$ extends over the all combinations (m_1, \dots, m_k) of k non-negative integers satisfying $\sum_{i=1}^k i \cdot m_i = k$.

We give two different proofs of Proposition 1, which are presented in Appendix. In Proof 1, the proposition is proved by induction on the degree k , using the relation (2.4). In Proof 2, we use the expression (2.3) of the Dirichlet process and the property (2.5) of the GEM distribution.

Using Proposition 1, we may obtain the following Bayes estimate of estimable parameter $\theta(\mathcal{P})$ and its limit, based on the Dirichlet process with squared error loss. Let $\mathcal{P}(\cdot)$ be the Dirichlet process with parameter α on $(\mathcal{R}, \mathcal{B})$ and \mathcal{F} be a distribution function of \mathcal{P} . Let F_Q denote a distribution function of Q and X_1, X_2, \dots, X_n be a sample

from \mathcal{F} . Then Bayes estimate $\hat{\theta}(\mathcal{P})$ of $\theta(\mathcal{P})$ is given in (3.1) by replacing Q with \hat{P}_n , where $\hat{P}_n = q_n Q + (1 - q_n)P_n$, $q_n = \alpha(\mathcal{R})/(\alpha(\mathcal{R}) + n)$, and P_n is the empirical probability measure of the sample X_1, X_2, \dots, X_n .

Letting $\alpha(\mathcal{X})$ in $\hat{\theta}(\mathcal{P})$ tend to zero, we obtain the limit of Bayes estimate as follows:

$$\theta^* = \binom{n+k-1}{k}^{-1} \sum_{(r_1+\dots+r_n=k)} g(\underbrace{X_1, \dots, X_1}_{r_1}, \underbrace{X_2, \dots, X_2}_{r_2}, \dots, \underbrace{X_n, \dots, X_n}_{r_n}),$$

where the above summation $\sum_{(r_1+\dots+r_n=k)}$ extends over all non-negative integers r_1, \dots, r_n such that $r_1 + \dots + r_n = k$. (See Yamato (1977b).) Similar results have been obtained by Tiwari (1981). Limits of Bayes estimates are the average over the all k -subset of the sample (X_1, X_2, \dots, X_n) allowing the duplication. Let the kernel $g(x_1, \dots, x_k)$ depend only on the distinct values among x_1, \dots, x_k . For positive integers r_1, \dots, r_u satisfying $r_1 + \dots + r_u = k$, we put $g^*(x_1, \dots, x_u) = g(x_1, \dots, x_1, x_2, \dots, x_2, \dots, x_u, \dots, x_u)$, where on the right-hand side x_j appears r_j times for $j = 1, 2, \dots, u$. Then the limit of Bayes estimate may be represented as

$$\theta^* = \binom{n+k-1}{k}^{-1} \sum_{u=1}^k \binom{k-1}{k-u} \sum_{1 \leq i_1 < \dots < i_u \leq n} g^*(X_{i_1}, \dots, X_{i_u}). \quad (3.2)$$

We give its example in Subsection 4.3.

4. Limits of Bayes Estimates, U -statistics and V -statistics

4.1. Mean squared difference and mean squared errors

In this subsection, we fix the distribution \mathcal{P} and let X_1, X_2, \dots, X_n be a sample of size n from \mathcal{P} . We see the difference of the limit of Bayes estimate from the U -statistic by taking the mean squared difference. It is asymptotically $E[(U_n - \theta_n^*)^2] = O(n^{-2})$ for arbitrary degree (see Yamato (1977b)). We consider the case of degree two. The U -statistic with the symmetric kernel g of degree two is given by

$$U_n^{(2)} = \frac{1}{n(n-1)} \sum_{i \neq j} g(X_i, X_j),$$

which is the minimum variance unbiased estimator of estimable parameter $\theta(\mathcal{P}) = \int g(x, y) d\mathcal{P}(x) d\mathcal{P}(y)$ for all absolutely continuous distributions. Its variance is

$$\begin{aligned} \text{Var}(U_n^{(2)}) &= \frac{2}{n(n-1)} \{2(n-2)(Eg(X_1, X_2)g(X_1, X_3) - \theta^2) + E(g(X_1, X_2))^2 - \theta^2\}. \end{aligned} \quad (4.1)$$

(See Lee (1990) and Serfling (1980).) The squared difference of the limit of Bayes estimate from the U -statistic is written as

$$(\theta_n^* - U_n^{(2)})^2 = \frac{4}{n^2(n+1)^2} \left[\frac{1}{(n-1)^2} \left(\sum_{i \neq j} g(X_i, X_j) \right)^2 \right]$$

$$-\frac{2}{n-1} \sum_{i \neq j} g(X_i, X_j) \sum_{k=1}^n g(X_k, X_k) + \left(\sum_{i=1}^n g(X_i, X_i) \right)^2.$$

Taking the expectation of the above equation, we have

$$\begin{aligned} E(\theta_n^* - U_n^{(2)})^2 &= \frac{4}{n(n+1)^2} \left[\frac{1}{n-1} \{ (n-2)(n-3)\theta^2 + 2E(g(X_1, X_2))^2 \right. \\ &\quad + 4(n-2)Eg(X_1, X_2)g(X_1, X_3) \} + (n-1)(Eg(X_1, X_1))^2 \\ &\quad \left. - 2(n-2)\theta Eg(X_1, X_1) - 4Eg(X_1, X_2)g(X_1, X_1) + E(g(X_1, X_1))^2 \right]. \end{aligned} \quad (4.2)$$

The V -statistic associated with a symmetric kernel of degree k is the average over the all k -permutation of a sample allowing the duplication. The V -statistic with the symmetric kernel g of degree two is given by

$$V_n^{(2)} = \frac{1}{n^2} \sum_{i,j} g(X_i, X_j).$$

(See Lee (1990) and Serfling (1980).) The mean squared error of the limit of Bayes estimate is obtained as

$$\begin{aligned} \text{MSE}(\theta_n^*) & \\ &= \frac{(n-1)^2}{(n+1)^2} \text{Var}(U_n^{(2)}) + \frac{4(n-1)}{(n+1)^2} E(U_n^{(2)} - \theta)U_n^{(1)} + \frac{4}{(n+1)^2} E(U_n^{(1)} - \theta)^2, \end{aligned} \quad (4.3)$$

where $U_n^{(1)} = \sum_{i=1}^n g(X_i, X_i)/n$. The mean squared error of the V -statistic is obtained as

$$\begin{aligned} \text{MSE}(V_n^{(2)}) & \\ &= \frac{(n-1)^2}{n^2} \text{Var}(U_n^{(2)}) + \frac{2(n-1)}{n^2} E(U_n^{(2)} - \theta)U_n^{(1)} + \frac{1}{n^2} E(U_n^{(1)} - \theta)^2. \end{aligned} \quad (4.4)$$

The difference of these mean squared errors becomes

$$\begin{aligned} \text{MSE}(V_n^{(2)}) - \text{MSE}(\theta_n^*) & \\ &= \frac{n-1}{n^2(n+1)^2} \{ (n-1)(2n+1)\text{Var}(U_n^{(2)}) - (3n+1)E(U_n^{(1)} - \theta)^2 \\ &\quad - 2(n^2 - 2n - 1)E(U_n^{(2)} - \theta)U_n^{(1)} \}. \end{aligned} \quad (4.5)$$

Using the results of (4.1), (4.2), (4.3), (4.4) and (4.5), we see the differences of the limits of Bayes estimates from the U -statistics and the V -statistics in the following five subsections.

4.2. Estimation of $P\{X + Y \leq 0\}$

We consider the estimation of

$$\Delta = P\{X + Y \leq 0\} = \int_{\mathcal{R}^2} I(x + y \leq 0) dF(x) dF(y),$$

where $I(A)$ is an indicator function of A and F is the distribution function of \mathcal{P} . The Bayes estimate of Δ is

$$\hat{\Delta} = \frac{1}{(\alpha(\mathcal{R}) + n)(\alpha(\mathcal{R}) + n + 1)} \{ \alpha(\mathcal{R})^2 \Delta(Q) + \alpha(\mathcal{R}) F_Q(0) + 2\alpha(\mathcal{R}) \sum_{i=1}^n F_Q(-X_i) + 2 \sum_{1 \leq i \leq j \leq n} I(X_i + X_j \leq 0) \}.$$

The limit of Bayes estimate of Δ is

$$\Delta^* = \binom{n+1}{2}^{-1} \sum_{1 \leq i \leq j \leq n} I(X_i + X_j \leq 0).$$

It is well-known that the corresponding one-sample Wilcoxon's U -statistic and the V -statistic are given by

$$U_{\Delta} = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} I(X_i + X_j \leq 0), \quad V_{\Delta} = \frac{1}{n^2} \sum_{i,j=1}^n I(X_i + X_j \leq 0).$$

We assume that the underlying distribution \mathcal{P} is a fixed continuous distribution symmetric about zero. The mean squared difference of the limit of Bayes estimate Δ^* from the corresponding U -statistic U_{Δ} is obtained as

$$E[(\Delta^* - U_{\Delta})^2] = \frac{1}{3n(n-1)(n+1)}.$$

The variance of the U -statistic is given by $\text{Var}(U_{\Delta}) = (2n-1)/\{6n(n-1)\}$. The mean squared errors of the limit of Bayes estimate and the V -statistic are, respectively,

$$\text{MSE}(\Delta^*) = \frac{2n+1}{6n(n+1)}, \quad \text{MSE}(V_{\Delta}) = \frac{4n^2-1}{12n^3}.$$

The difference of these mean squared errors becomes

$$\text{MSE}(V_{\Delta}) - \text{MSE}(\Delta^*) = \frac{(n-1)(2n+1)}{12n^3(n+1)}.$$

Therefore we have

$$\text{MSE}(\Delta^*) < \text{MSE}(V_{\Delta}) < \text{Var}(U_{\Delta}) \quad \text{for } n \geq 2.$$

4.3. Estimation of probability weighted moments

We consider the estimation of probability weighted moments,

$$\beta_{r-1} = E[X\{F(X)\}^{r-1}] = \int_{\mathcal{R}} x\{F(x)\}^{r-1} dF(x), \quad r = 1, 2, \dots$$

The kernel of β_{r-1} is given by

$$g(x_1, \dots, x_r) = \frac{1}{r} \text{Max}\{x_1, \dots, x_r\}.$$

From (3.2), the limit of Bayes estimate of β_{r-1} is

$$\beta_{r-1}^* = \frac{1}{r} \binom{n+r-1}{r}^{-1} \sum_{u=1}^r \binom{r-1}{r-u} \sum_{1 \leq i_1 < i_2 < \dots < i_u \leq n} \text{Max}\{X_{i_1}, \dots, X_{i_u}\}.$$

The corresponding U -statistic is given by

$$U_{\beta_{r-1}} = \frac{1}{r} \binom{n}{r}^{-1} \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} \text{Max}\{X_{i_1}, X_{i_2}, \dots, X_{i_r}\}.$$

(See, for example, Lee (1990).) Especially for $r = 2$, the Bayes estimate of β_1 is

$$\begin{aligned} \hat{\beta}_1 &= \frac{1}{(\alpha(\mathcal{R}) + n)(\alpha(\mathcal{R}) + n + 1)} \{ \alpha(\mathcal{R})^2 \beta_1(Q) + \frac{\alpha(\mathcal{R})}{2} \int_{\mathcal{R}} x dQ(x) \\ &\quad + \alpha(\mathcal{R}) \sum_{i=1}^n \int_{\mathcal{R}} \text{Max}\{x, X_{(i)}\} dQ(x) + \sum_{i=1}^n i X_{(i)} \}, \end{aligned}$$

where $X_{(1)}, \dots, X_{(n)}$ denotes the order statistics for the sample X_1, \dots, X_n . The limit of Bayes estimate of β_1 is

$$\beta_1^* = \frac{1}{n(n+1)} \sum_{i=1}^n i X_{(i)}.$$

The corresponding U -statistic is given by

$$U_{\beta_1} = \frac{1}{2} \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \text{Max}\{X_i, X_j\} = \frac{1}{n(n-1)} \sum_{i=1}^n (i-1) X_{(i)}.$$

The corresponding V -statistic is given by

$$V_{\beta_1} = \frac{1}{2n^2} \sum_{i,j=1}^n \text{Max}\{X_i, X_j\}.$$

Assume that underlying distribution \mathcal{P} is a uniform distribution, $U(-\tau, \tau)$. The mean squared difference of the limit of Bayes estimate β_1^* from the corresponding U -statistic U_{β_1} is obtained as

$$E[(\beta_1^* - U_{\beta_1})^2] = \frac{\tau^2}{3n(n-1)(n+1)^2} (2n^2 - \frac{54}{5}n + \frac{73}{5}).$$

The variance of the U -statistic is given by $\text{Var}(U_{\beta_1}) = (4n-3)\tau^2/\{45n(n-1)\}$. The mean squared errors of the limit of Bayes estimate and the V -statistic are, respectively,

$$\text{MSE}(\beta_1^*) = \frac{4n^2 + 13n + 3}{45n(n+1)^2} \tau^2, \quad \text{MSE}(V_{\beta_1}) = \frac{16n^2 + 7n - 3}{180n^3} \tau^2.$$

The difference of the mean squared error from the variance becomes

$$\text{MSE}(\beta_1^*) - \text{Var}(U_{\beta_1}) = \frac{4n(n-2)}{45n(n-1)(n+1)^2}.$$

Therefore we have

$$\text{MSE}(V_{\beta_1}) < \text{Var}(U_{\beta_1}) < \text{MSE}(\beta_1^*) \quad \text{for } n > 2.$$

In the following three subsections, we also see the differences of the limits of Bayes estimates from the U -statistics and the V -statistics for the three parameters, for which Bayes estimates and their limits are given by Ferguson (1973) and Yamato (1977a).

4.4. Estimation of a variance

We consider the estimation of a variance of a distribution \mathcal{P} ,

$$\sigma^2 = \int_{\mathcal{R}^2} \frac{1}{2}(x-y)^2 dF(x)dF(y) = \int x^2 dF(x) - \left(\int x dF(x)\right)^2.$$

The limit of Bayes estimate of σ^2 is

$$\sigma^{2*} = \frac{1}{n+1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

(See Ferguson (1973).) It is well known that the corresponding U -statistic, sample variance, and the corresponding V -statistic are given by

$$U_{\sigma^2} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2, \quad V_{\sigma^2} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Assume that underlying distribution \mathcal{P} is a fixed continuous distribution, which is symmetric about zero and has the finite fourth moment. The mean squared difference of the limit of Bayes estimate σ^{2*} from the corresponding U -statistic U_{σ^2} is obtained as

$$\text{E}[(\sigma^{2*} - U_{\sigma^2})^2] = \frac{4}{n(n-1)(n+1)^2} \{(n^2 - 2n + 3)\mu_2^2 + (n-1)\mu_4\}, \quad (4.6)$$

where μ_i ($i = 2, 4$) is the i -th moment of the distribution \mathcal{P} . The variance of the U -statistic is given by $\text{Var}(U_{\sigma^2}) = \{((n-1)\mu_4 - (n-3)\mu_2^2) / \{n(n-1)\}\}$. The mean squared errors of the limit of Bayes estimate and the V -statistic are, respectively,

$$\begin{aligned} \text{MSE}(\sigma^{2*}) &= \frac{1}{n(n+1)^2} \{(\mu_4 - \mu_2^2)n^2 - 2(\mu_4 - 4\mu_2^2)n + \mu_4 - 3\mu_2^2\}, \\ \text{MSE}(V_{\sigma^2}) &= \frac{1}{n^3} \{(\mu_4 - \mu_2^2)n^2 - (2\mu_4 - 5\mu_2^2)n + \mu_4 - 3\mu_2^2\}. \end{aligned}$$

We consider the case that \mathcal{P} is a normal distribution, $N(0, 1)$. From (4.6), the mean

squared difference of the limit of Bayes estimate σ^{2*} from the corresponding U -statistic U_{σ^2} is obtained as

$$E[(\sigma^{2*} - U_{\sigma^2})^2] = \frac{4}{(n-1)(n+1)}.$$

The variance of the U -statistic is given by $\text{Var}(U_{\sigma^2}) = 2/(n-1)$. The mean squared errors of the limit of Bayes estimate and the V -statistic are, respectively,

$$\text{MSE}(\sigma^{2*}) = \frac{2}{n+1}, \quad \text{MSE}(V_{\sigma^2}) = \frac{2n-1}{n^2}.$$

Therefore we have

$$\text{MSE}(\sigma^{2*}) < \text{MSE}(V_{\sigma^2}) < \text{Var}(U_{\sigma^2}) \quad \text{for } n \geq 2.$$

Moreover in case that \mathcal{P} is a uniform distribution, $U(-\tau, \tau)$, we have

$$\begin{aligned} \text{MSE}(V_{\sigma^2}) &< \text{Var}(U_{\sigma^2}) < \text{MSE}(\sigma^{2*}) \quad \text{for } n \geq 12, \\ \text{MSE}(V_{\sigma^2}) &< \text{MSE}(\sigma^{2*}) \leq \text{Var}(U_{\sigma^2}) \quad \text{for } 3 \leq n \leq 11, \end{aligned}$$

where the last equality holds for $n = 11$.

4.5. Estimation of a squared mean

We consider the estimation of a squared mean of a distribution \mathcal{P} ,

$$\mu^2 = \left(\int_{\mathcal{R}} x dF(x) \right)^2.$$

The limit of Bayes estimate of μ^2 is

$$\mu^{2*} = \frac{n}{n+1}(\bar{X})^2 + \frac{1}{n+1}\bar{X}^2.$$

(See Yamato (1977a).) The corresponding U -statistic and the corresponding V -statistic are, respectively,

$$U_{\mu^2} = \frac{2}{n-1} \sum_{i < j} X_i X_j, \quad V_{\mu^2} = \frac{1}{n^2} \sum_{i,j} X_i X_j.$$

Assume that the underlying distribution \mathcal{P} is a fixed continuous distribution, which is symmetric about zero and has the finite fourth moment. The mean squared difference of the limit of Bayes estimate μ^{2*} from the corresponding U -statistic U_{μ^2} is obtained as

$$E[(\mu^{2*} - U_{\mu^2})^2] = \frac{4}{n(n-1)(n+1)^2} \{ (n^2 - 2n + 3)\mu_2^2 + (n+1)\mu_4 \}.$$

The variance of the U -statistic is given by $\text{Var}(U_{\mu^2}) = 2\mu_2^2/\{n(n-1)\}$. The mean squared errors of the limit of Bayes estimate and the V -statistic are, respectively,

$$\begin{aligned} \text{MSE}(\mu^{2*}) &= \frac{2}{n(n+1)^2} \{ 3\mu_2^2(n-1) + 2\mu_4 \}, \\ \text{MSE}(V_{\mu^2}) &= \frac{1}{n^3} \{ 3\mu_2^2(n-1) + \mu_4 \}. \end{aligned}$$

In case that \mathcal{P} is a normal distribution, $N(0, 1)$, the mean squared difference of the limit of Bayes estimate μ^{2*} from the corresponding U -statistic U_{μ^2} is obtained as $E[(\mu^{2*} - U_{\mu^2})^2] = 4/\{(n-1)(n+1)\}$. The variance of the U -statistic is given by $2/\{n(n-1)\}$. The mean squared errors of the limit of Bayes estimate and the V -statistic are obtained as $6/\{n(n+1)\}$ and $3/n^2$, respectively. Therefore we have

$$\text{Var}(U_{\mu^2}) < \text{MSE}(V_{\mu^2}) < \text{MSE}(\mu^{2*}) \quad \text{for } n \geq 4.$$

4.6. Estimation of a measure of concentration

We consider the estimation of a measure of concentration,

$$\gamma = E|X - Y| = \int_{\mathcal{R}^2} |x - y| dF(x) dF(y).$$

The limit of Bayes estimate of γ is

$$\gamma^* = \frac{2}{n(n+1)} \sum_{i < j} |X_i - X_j|.$$

(See Yamato (1977a).) The corresponding U -statistic and the corresponding V -statistic are, respectively,

$$U_\gamma = \frac{2}{n(n-1)} \sum_{i < j} |X_i - X_j|, \quad V_\gamma = \frac{2}{n^2} \sum_{i < j} |X_i - X_j|.$$

We consider the case that the underlying distribution \mathcal{P} is a normal distribution, $N(0, 1)$. The mean squared difference of the limit of Bayes estimate γ^* from the corresponding U -statistic U_γ is obtained as

$$E[(\gamma^* - U_\gamma)^2] = \frac{16}{n(n-1)(n+1)^2} \left\{ \frac{n^2}{\pi} + \left(\frac{2\sqrt{3}-5}{\pi} + \frac{1}{3} \right) n + \frac{6-4\sqrt{3}}{\pi} + \frac{1}{3} \right\}.$$

The variance of the U -statistic is given by $\text{Var}(U_\gamma) = 4\{(2\sqrt{3}-4)/\pi + 1/3\}n + (6-4\sqrt{3})/\pi + 1/3/\{n(n-1)\}$. The mean squared errors of the limit of Bayes estimate and the V -statistic are, respectively,

$$\begin{aligned} \text{MSE}(\gamma^*) &= \frac{4}{n(n+1)^2} \left\{ \left(\frac{2\sqrt{3}-4}{\pi} + \frac{1}{3} \right) n^2 + \frac{14-6\sqrt{3}}{\pi} n + \frac{4\sqrt{3}-6}{\pi} - \frac{1}{3} \right\}, \\ \text{MSE}(V_\gamma) &= \frac{4}{n^3} \left\{ \left(\frac{2\sqrt{3}-4}{\pi} + \frac{1}{3} \right) n^2 + \frac{11-6\sqrt{3}}{\pi} n + \frac{4\sqrt{3}-6}{\pi} - \frac{1}{3} \right\}. \end{aligned}$$

Therefore we have

$$\text{MSE}(V_\gamma) < \text{Var}(U_\gamma) < \text{MSE}(\gamma^*) \quad \text{for } n \geq 3.$$

In case that the underlying distribution is a uniform distribution, $U(-1/2, 1/2)$, the mean squared difference of the limit of Bayes estimate γ^* from the corresponding

U -statistic U_γ is obtained as $4(5n^2 - 4n + 3)/\{45n(n-1)(n+1)^2\}$. The variance of the U -statistic is given by $(n+3)/\{45n(n-1)\}$. The mean squared errors of the limit of Bayes estimate and the V -statistic are obtained as $(n^2 + 22n - 3)/\{45n(n+1)^2\}$ and $(n^2 + 7n - 3)/45n^3$, respectively. Therefore we have, $\text{MSE}(\gamma^*) = \text{Var}(U_\gamma) < \text{MSE}(V_\gamma)$ ($n = 2$), $\text{Var}(U_\gamma) < \text{MSE}(\gamma^*) = \text{MSE}(V_\gamma)$ ($n = 3$) and

$$\text{Var}(U_\gamma) < \text{MSE}(V_\gamma) < \text{MSE}(\gamma^*) \quad \text{for } n \geq 4.$$

5. Appendix

We give two different proofs of Proposition 1.

Proof 1. We prove Proposition 1 by induction. First we consider for $k = 1$. Using (2.4), we have

$$\int_{\mathcal{X}} g(x) d\mathcal{P}(x) \stackrel{d}{=} \theta_1 g(Y_1) + (1 - \theta_1) \int_{\mathcal{X}} g(x) d\mathcal{P}(x). \quad (5.1)$$

We take the expectations of the both sides of (5.1). Since \mathcal{P} is independent of (θ_1, Y_1) and θ_1 is independent of Y_1 on the right-hand side, we get,

$$\mathbb{E}\left[\int_{\mathcal{X}} g(x) d\mathcal{P}(x)\right] = \mathbb{E}[g(Y_1)] = \int_{\mathcal{X}} g(x) dQ(x).$$

Let $k \geq 2$. We assume that (3.1) holds for $2, 3, \dots, k-1$ and $g(x_1, \dots, x_k)$ is symmetric in the arguments. Then we have

$$\begin{aligned} & \int_{\mathcal{X}^k} g(x_1, \dots, x_k) \prod_{i=1}^k d\mathcal{P}(x_i) \\ & \stackrel{d}{=} \sum_{r=0}^{k-1} \binom{k}{r} \theta_1^{k-r} (1 - \theta_1)^r \int_{\mathcal{X}^r} g(\underbrace{Y_1, \dots, Y_1}_{k-r}, x_1, \dots, x_r) \prod_{i=1}^r d\mathcal{P}(x_i) \\ & \quad + (1 - \theta_1)^k \int_{\mathcal{X}^k} g(x_1, \dots, x_k) \prod_{i=1}^k d\mathcal{P}(x_i). \end{aligned} \quad (5.2)$$

We take the expectations of the both sides of (5.2). Note that on the right-hand side \mathcal{P} is independent of (θ_1, Y_1) and θ_1 is independent of Y_1 . For the first integral on the right-hand side of (5.2), we take the conditional expectation given $Y_1 = y_1$ and then take the expectation with respect to Y_1 having the distribution Q . Then we have

$$(1 - \mathbb{E}(1 - \theta_1)^k) \mathbb{E} \int_{\mathcal{X}^k} g(x_1, \dots, x_k) \prod_{i=1}^k d\mathcal{P}(x_i)$$

$$\begin{aligned}
&= \sum_{r=0}^{k-1} \binom{k}{r} \frac{\alpha(\mathcal{X})(\alpha(\mathcal{X}))^{[r]}}{(\alpha(\mathcal{X}))^{[k+1]}} (k-r)! \sum_{(\sum_{i=1}^r m_i = r)} \frac{r! (\alpha(\mathcal{X}))^{\sum_{i=1}^{k-r} m_i}}{\prod_{i=1}^r (i^{m_i} (m_i)! (\alpha(\mathcal{X}))^{[r]})} \\
&\quad \times \int_{\mathcal{X}^{1+\sum_{i=1}^r m_i}} g(\underbrace{y_1, \dots, y_1}_{k-r}, x_{11}, \dots, x_{1m_1}, \underbrace{x_{21}, x_{21}}_2, \dots, \underbrace{x_{2m_2}, x_{2m_2}}_2, \\
&\quad \dots, \underbrace{x_{r1}, \dots, x_{r1}}_r, \dots, \underbrace{x_{rm_r}, \dots, x_{rm_r}}_r) \prod_{i=1}^r \prod_{j=1}^{m_i} dQ(x_{ij}) dQ(y_1).
\end{aligned}$$

For each $r \in \{0, 1, \dots, k-1\}$, y_1 appears $k-r$ times. We put $m'_{k-r} = m_{k-r} + 1$ and $m'_j = m_j$ for $j \neq k-r$. From the combination (m_1, m_2, \dots, m_k) satisfying $\sum_{i=1}^r m_i = r$, we have the combination $(m'_1, m'_2, \dots, m'_{k+1})$ which satisfies $\sum_{i=1}^r m'_i = k$. Thus we have

$$\begin{aligned}
&\mathbb{E} \left[\int_{\mathcal{X}^k} g(x_1, \dots, x_k) \prod_{i=1}^k d\mathcal{P}(x_i) \right] \\
&= \sum_{(\sum_{i=1}^k m'_i = k)} \frac{(k)! [\alpha(\mathcal{X})]^{\sum_{i=1}^k m'_i}}{\prod_{i=1}^k (i^{m'_i} (m'_i)! (\alpha(\mathcal{X}))^{[k]})} \int_{\mathcal{X}^{\sum_{i=1}^k m'_i}} g(x_{11}, \dots, x_{1m'_1}, x_{21}, x_{21}, \\
&\quad \dots, x_{2m'_2}, x_{2m'_2}, \dots, x_{k1}, \dots, x_{k1}, \dots, x_{km'_k}, \dots, x_{km'_k}) \prod_{i=1}^k \prod_{j=1}^{m'_i} dQ(x_{ij}).
\end{aligned}$$

Therefore Proposition 1 is proved. \square

Proof 2. Since Y_1, Y_2, \dots are independent of (p_1, p_2, \dots) and by using (2.3), we have

$$\begin{aligned}
&\mathbb{E} \left[\int_{\mathcal{X}^k} g(x_1, \dots, x_k) \prod_{i=1}^k d\mathcal{P}(x_i) \right] \tag{5.3} \\
&= \sum_{(\sum_{i=1}^k m_i = k)} \frac{k!}{\prod_{i=1}^k [(i!)^{m_i} (m_i)!]} \mathbb{E}[g(Y_{11}, \dots, Y_{1m_1}, Y_{21}, Y_{21}, \dots, \\
&\quad Y_{2m_2}, Y_{2m_2}, \dots, Y_{k1}, \dots, Y_{k1}, \dots, Y_{km_k}, \dots, Y_{km_k})] \mathbb{E} \left[\sum_{i=1}^k \prod_{j=1}^{m_i} p_{ij}^i \right].
\end{aligned}$$

Y_1, Y_2, \dots are mutually independent random variables with Q and we apply the relation (2.4) to the right-hand side of (5.3). Thus the expectation of random functional is equal to

$$\begin{aligned}
&\sum_{(\sum_{i=1}^k m_i = k)} \frac{k! (\alpha(\mathcal{X}))^{\sum_{i=1}^k m_i}}{\prod_{i=1}^k [i^{m_i} (m_i)! (\alpha(\mathcal{X}))^{[k]}]} \int_{\mathcal{X}^{\sum_{i=1}^k m_i}} g(x_{11}, \dots, x_{1m_1}, x_{21}, \\
&\quad x_{21}, \dots, x_{2m_2}, x_{2m_2}, \dots, x_{k1}, \dots, x_{k1}, \dots, x_{km_k}, \dots, x_{km_k}) \prod_{i=1}^k \prod_{j=1}^{m_i} dQ(x_{ij}).
\end{aligned}$$

Thus the proof of Proposition 1 is completed. \square

Acknowledgments

The authors would like to express gratitude to Professor Takashi Yanagawa who gave us the valuable comments and the suggestions. The authors also thank the referee for his useful comments.

References

- Antoniak, C. E. (1974), Mixtures of Dirichlet process with applications to Bayesian nonparametric problems. *The Annals of Statistics*, **2**, 1152–1174.
- Engen, S. (1978), *Stochastic Abundance Models with Emphasis on Biological Communities and Species Diversity*. Chapman & Hall, London.
- Ewens, W. J. (1990), Population genetics theory — the past and the future. in: S. Lessard, ed., *Mathematical and Statistical Developments of Evolutionary Theory*. (Kluwer Academic Publishers, Dordrecht) 177–227.
- Ferguson, T. S. (1973), A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, **1**, 209–230.
- Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997), *Discrete Multivariate Distributions*. John Wiley & Sons, New York.
- Lee, A. J. (1990), *U-statistics: Theory and Practice*. Dekker, New York.
- McCloskey, J. W. (1965), A model for the distribution of individuals by species in an environment. *Ph.D. thesis*, Michigan State University.
- Serfling, R. J. (1980), *Approximation theorems of mathematical statistics*. John Wiley & Sons, New York.
- Sethuraman, J. (1994), A constructive definition of Dirichlet priors. *Statistica Sinica*, **4**, 639–650.
- Sibuya, M. and Yamato, H. (1995), Ordered and unordered random partitions of an integer and the GEM distribution. *Statistics & Probability Letters*, **25**, 177–183.
- Tiwari, R. C. (1981), A mathematical study of the Dirichlet process. *Ph.D. thesis*, Florida State University.
- Yamato, H. (1977a), Relations between limiting Bayes estimates and the U-statistics for estimable parameters of degrees 2 and 3. *Communications in Statistics*, **A6**, 55–66.
- Yamato, H. (1977b), Relations between limiting Bayes estimates and U-statistics for estimable parameters. *Journal of the Japan Statistical Society*, **7**, 57–66.

Received March 24, 1999

Revised June 10, 1999

Re-revised September 9, 1999