Statistica Sinica **20**(2010)

Posterior consistency of species sampling priors

Gun ho Jang, Jaeyong Lee and Sangyeol Lee

University of Toronto and Seoul National University

Supplementary Material

This supplement contains technical details, proofs of lemmas and theorems in the main article. All section and equation numbers refer to the main article.

In the following, we prove Lemmas 1, 2 and 1, and which are needed in the proof of Theorem 1.

Proof of Lemma 1 (Necessity) Suppose that (i) and (ii) hold. By the Chebychev's inequality, for all P_0-continuity set U and $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}(|P(U) - P_0(U)| \geq \epsilon | X_1, \ldots, X_n) \\
&\quad \leq \quad \frac{1}{\epsilon^2}\mathbb{E}(|P(U) - P_0(U)|^2 | X_1, \ldots, X_n) \\
&\quad = \quad \frac{1}{\epsilon^2}[\mathbb{V}ar(P(U)|X_1, \ldots, X_n) + \{\mathbb{E}(P(U)|X_1, \ldots, X_n) - P_0(U)\}^2] \\
&\quad \to \quad 0 \qquad P_0^\infty - a.s.
\end{aligned}$$

Let O be a weak neighborhood of P_0. Then, there is a weak open set O_1 containing P_0 such that $P_0 \in O_1$ and

$$O_1 = \{Q \in \mathcal{M} : |Q(U_i) - P_0(U_i)| < \epsilon_i, i = 1, 2, \ldots, k\} \subset O$$

where U_is are P_0-continuity sets and ϵ_is are real numbers. Hence,

$$\begin{aligned}
&\mathbb{P}(O^c | X_1, \ldots, X_n) \leq \mathbb{P}(O_1^c | X_1, \ldots, X_n) \\
&\quad = \mathbb{P}[\{Q \in \mathcal{M} : |Q(U_i) - P_0(U_i)| < \epsilon_i, i = 1, 2, \ldots, k\}^c | X_1, \ldots, X_n] \\
&\quad \leq \sum_{i=1}^{k} \mathbb{P}[\{Q \in \mathcal{M} : |Q(U_i) - P_0(U_i)| \geq \epsilon_i\} | X_1, \ldots, X_n] \\
&\quad \to 0 \qquad P_0^\infty - a.s.
\end{aligned}$$

This completes the proof of the necessity.

(Sufficiency) Suppose that the posterior is consistent at P_0. Let U be a P_0-continuity set and $\epsilon > 0$ be a positive number. Then,

$$A = \{Q \in \mathcal{M} : |Q(U) - P_0(U)| < \epsilon\}$$

is a weak neighborhood of P_0.

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}[P(U)|X_1,\dots,X_n] &\geq \lim_{n\to\infty}\mathbb{E}[P(U)I(P\in A)|X_1,\dots,X_n]\\ &\geq \lim_{n\to\infty}(P_0(U)-\epsilon)\mathbb{P}(A|X_1,\dots,X_n)\\ &= P_0(U)-\epsilon, \quad P_0^\infty - a.s.,\end{aligned}$$

and

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}[P(U)|X_1,\dots,X_n] &\leq \lim_{n\to\infty}[(P_0(U)+\epsilon)\mathbb{P}(A|X_1,\dots,X_n)+\mathbb{P}(A^c|X_1,\dots,X_n)]\\ &= P_0(U)+\epsilon, \quad P_0^\infty - a.s.\end{aligned}$$

Since $\epsilon > 0$ is arbitrary,

$$\lim_{n\to\infty}\mathbb{E}[P(U)|X_1,\dots,X_n] = P_0(U), \quad P_0^\infty - a.s.$$

Similarly,

$$\lim_{n\to\infty}\mathbb{E}[P(U)^2|X_1,\dots,X_n] \geq \sup_{\epsilon\downarrow 0}\lim_{n\to\infty}(P_0(U)-\epsilon)^2\mathbb{P}(A|X_1,\dots,X_n) = P_0(U)^2, \quad P_0^\infty - a.s.$$

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}[P(U)^2|X_1,\dots,X_n] &\leq \sup_{\epsilon\downarrow 0}\lim_{n\to\infty}(P_0(U)+\epsilon)^2\mathbb{P}(A|X_1,\dots,X_n)+\mathbb{P}(A^c|X_1,\dots,X_n)\\ &= P_0(U)^2, \quad P_0^\infty - a.s..\end{aligned}$$

Thus, $\lim_{n\to\infty}\mathbb{E}[P(U)^2|X_1,\dots,X_n] = P_0(U)^2, \quad P_0^\infty - a.s.$, which implies (ii). ■

Proof of Lemma 2. (i). Since $X_1, X_2, \dots$ is iid sample from P_0, $\sum_{i=1}^n I(X_i \notin \mathcal{Z}) \sim Bin(n,\lambda)$, which implies $k^*/n = \frac{1}{n}\sum_{i=1}^n I(X_i \notin \mathcal{Z}) \to \lambda$, P_0^∞-a.s.

Fix $\epsilon > 0$. There is a sufficiently large l such that $q_1 + \cdots + q_l > 1 - \lambda - \epsilon/3$. By the strong law of large numbers, for all sufficiently large n with $n > 2l/\epsilon$ we have

$$|n_j/n - q_j| < \epsilon 3^{-1-j} \quad \text{for } j = 1,\dots,l \quad \text{and} \quad |k_n^*/n - \lambda| < \epsilon,$$

where $n_j = \sum_{i=1}^n I(X_i = z_j)$. Thus, for all sufficiently large n with $n > 2l/\epsilon$, we get

$$k_n/n \leq [l + n - (n_{n,1} + \cdots + n_{n,l})]/n \leq \epsilon/2 + 1 - (1-\lambda-\epsilon/2) \leq \lambda + \epsilon$$

and

$$k_n/n \geq k_n^*/n \geq \lambda - \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we get the result.
(ii). Suppose that $\lambda > 0$. Note

$$G_{k_n} = \frac{1}{k_n}\sum_{\tilde{X}_j\in\mathcal{Z}}\delta_{\tilde{X}_j} + \frac{1}{k_n}\sum_{\tilde{X}_j\notin\mathcal{Z}}\delta_{\tilde{X}_j} = \frac{k_n-k_n^*}{k_n}\frac{1}{k_n-k_n^*}\sum_{\tilde{X}_j\in\mathcal{Z}}\delta_{z_j} + \frac{k_n^*}{k_n}\frac{1}{k_n^*}\sum_{\tilde{X}_j\notin\mathcal{Z}}\delta_{\tilde{X}_j}.$$

From (i), $(k_n - k_n^*)/n \to 0$, P_0^∞-a.s. Combining these, we have

$$G_{k_n} \to \mu \quad P_0^\infty\text{-a.s.} \quad ■$$

Lemma 3. Under the assumptions in Theorem 1, the followings hold

(i) $\mathbb{E}[P(B)|X_1,\ldots,X_n] = [n/(n+b)]\tilde{F}_n(B)$

(ii) $\mathbb{E}[P(B)^2|X_1,\ldots,X_n] = [n^2\tilde{F}_n(B)^2 + n\tilde{F}_n(B) - a(b+ak_n)(1-\nu(B))]/[(n+b)(n+b+1)]$

where $\tilde{F}_n(B) = n^{-1}\sum_{i=1}^n I(X_i \in B) - ak_nG_{k_n}(B)/n + (b+ak_n)\nu(B)/n$.

Proof. From the posterior distribution (3.2), we have

$$\begin{aligned}
\mathbb{E}(P(B)|X_1,\ldots,X_n) &= \mathbb{E}[\sum_{j=1}^{k_n}\tilde{P}_jI(\tilde{X}_j \in B) + \tilde{R}_{k_n}PY(a,b+ak_n,\nu)(B)|X_1,\ldots,X_n] \\
&= \sum_{j=1}^{k_n}\frac{n_{jn}-a}{b+n}I(\tilde{X}_j \in B) + \frac{b+ak_n}{b+n}\nu(B) \\
&= \frac{n}{b+n}F_n(B) - \frac{ak_n}{b+n}G_{k_n}(B) + \frac{b+ak_n}{b+n}\nu(B)
\end{aligned}$$

where $F_n = n^{-1}\sum_{j=1}^n \delta_{X_j}$.

Simple algebra on the Dirichlet distribution and the Pitman-Yor process also give us

$$\begin{aligned}
&\mathbb{E}(P(B)^2|X_1,\ldots,X_n) = \mathbb{E}[(\sum_{j=1}^{k_n}\tilde{P}_jI(\tilde{X}_j \in B) + \tilde{R}_{k_n}PY(a,b+ak_n,\nu)(B))^2|X_1,\ldots,X_n] \\
&= \frac{n^2}{(b+n)(b+n+1)}\left(F_n(B) - \frac{ak_n}{n}G_{k_n}(B) + \frac{b+ak_n}{n}\nu(B)\right)^2 \\
&\quad + \frac{n}{(b+n)(b+n+1)}\left(F_n(B) - \frac{ak_n}{n}G_{k_n}(B) + \frac{b+ak_n}{n}\nu(B)(1-a+a\nu(B))\right) \\
&= \frac{n^2\tilde{F}_n(B)^2 + n\tilde{F}_n(B) - a(b+ak_n)(1-\nu(B))}{(n+b)(n+b+1)}. \quad \blacksquare
\end{aligned}$$

To prove Proposition 2 and Theorem 4, we need the following lemmas for the moments of the posterior.

Lemma 4. Suppose $P \sim SSP(p,\nu)$ and $X_1,\ldots,X_n$ given P is a random sample from P, where p is an EPPF and ν is a diffuse probability measure. Let $\mathbf{n}$ be the partition of n defined by $X_1,\ldots,X_n$, and $k = k(\mathbf{n})$. Then, for a measurable set B,

(i) $\mathbb{E}[P(B)|X_1,\ldots,X_n] = \sum_{i=1}^k p_i(\mathbf{n})I(\tilde{X}_j \in B) + p_{k+1}(\mathbf{n})\nu(B)$

(ii) $\mathbb{E}[P(B)^2|X_1,\ldots,X_n] = \sum_{i,j=1}^k p_i(\mathbf{n})p_j(\mathbf{n}^{i+})I(\tilde{X}_i \in B)I(\tilde{X}_j \in B) + 2\sum_{i=1}^k p_i(\mathbf{n})p_{k+1}(\mathbf{n}^{i+})I(\tilde{X}_i \in B)\nu(B) + p_{k+1}(\mathbf{n})p_{k+2}(\mathbf{n}^{(k+1)+})\nu^2(B) + p_{k+1}(\mathbf{n})p_{k+1}(\mathbf{n}^{(k+1)+})\nu(B)$.

Proof. Using the conditional distribution of the species sampling sequence, we obtain

$$
\begin{aligned}
\mathbb{E}[P(B)|X_1,\ldots,X_n] &= \mathbb{P}(X_{n+1}\in B|X_1,\ldots,X_n)\\
&= \sum_{i=1}^k p_i(\mathbf{n})I(\tilde{X}_i\in B)+p_{k+1}(\mathbf{n})\nu(B).
\end{aligned}
$$

To get (ii), first define $B_0 = B\setminus\{\tilde{X}_1,\ldots,\tilde{X}_k\}$. We expand $P(B)^2$ as follows.

$$
\begin{aligned}
P(B)^2 &= \left(\sum_{i=1}^k P(B\cap\{\tilde{X}_i\})+P(B_0)\right)^2\\
&= \sum_{i=1}^k P(B\cap\{\tilde{X}_i\})^2+\sum_{i\neq j}P(B\cap\{\tilde{X}_i\})P(B\cap\{\tilde{X}_j\})\\
&\quad +2P(B_0)\sum_{i=1}^k P(B\cap\{\tilde{X}_i\})+P(B_0)^2.
\end{aligned}
$$

We obtain the posterior expectation of the above expansion term by term.

$$
\begin{aligned}
&\mathbb{E}(P(B\cap\{\tilde{X}_j\})^2|X_1,\ldots,X_n)\\
=\;&\mathbb{E}[\mathbb{E}(P(B\cap\{\tilde{X}_j\})^2|X_1,\ldots,X_n,P)|X_1,\ldots,X_n]\\
=\;&\mathbb{E}[\mathbb{P}(X_{n+1}\in B\cap\{\tilde{X}_j\},X_{n+2}\in B\cap\{\tilde{X}_j\}|X_1,\ldots,X_n,F)|X_1,\ldots,X_n]\\
=\;&\mathbb{P}(X_{n+1}\in B\cap\{\tilde{X}_j\},X_{n+2}\in B\cap\{\tilde{X}_j\}|X_1,\ldots,X_n)\\
=\;&p_j(\mathbf{n})I(\tilde{X}_j\in B)p_j(\mathbf{n}^{j+})I(\tilde{X}_j\in B)\\
=\;&p_j(\mathbf{n})p_j(\mathbf{n}^{j+})I(\tilde{X}_j\in B).
\end{aligned}
$$

Similarly, we get for $i\neq j$,

$$
\begin{aligned}
&\mathbb{E}(P(B\cap\{\tilde{X}_i\})P(B\cap\{\tilde{X}_j\})|X_1,\ldots,X_n)\\
&\quad=\mathbb{P}(X_{n+1}\in B\cap\{\tilde{X}_i\},X_{n+2}\in B\cap\{\tilde{X}_j\}|X_1,\ldots,X_n)\\
&\quad=p_i(\mathbf{n})p_j(\mathbf{n}^{i+})I(\tilde{X}_i\in B)I(\tilde{X}_j\in B).
\end{aligned}
$$

For the third term,

$$
\begin{aligned}
&\mathbb{E}(P(B_0)P(B\cap\{\tilde{X}_j\})|X_1,\ldots,X_n)\\
&\quad=\mathbb{P}(X_{n+1}\in B\cap\{\tilde{X}_j\},X_{n+2}\in B_0|X_1,\ldots,X_n)\\
&\quad=p_j(\mathbf{n})p_{k+1}(\mathbf{n}^{j+})\nu(B)I(\tilde{X}_j\in B).
\end{aligned}
$$

Finally, the last term becomes

$$
\begin{aligned}
\mathbb{E}(P(B_0)^2|X_1,\ldots,X_n) &= \mathbb{P}(X_{n+1}\in B_0,X_{n+2}\in B_0|X_1,\ldots,X_n)\\
&= \int_{B_0} p_{k+1}(\mathbf{n})\Big(p_{k+2}(\mathbf{n}^{(k+1)+})\nu(B_0\setminus\{x_{n+1}\})\\
&\qquad +p_{k+1}(\mathbf{n}^{(k+1)+})I(x_{n+2}=x_{n+1})\Big)\,\nu(dx_{n+1})\\
&= p_{k+1}(\mathbf{n})p_{k+2}(\mathbf{n}^{(k+1)+})\nu(B_0)^2+p_{k+1}(\mathbf{n})p_{k+1}(\mathbf{n}^{(k+1)+})\nu(B_0).
\end{aligned}
$$

Using the fact that ν is a diffuse probability measure and collecting the above results, we obtain the result of the lemma. ■

Lemma 5. Suppose the same assumptions of Theorem 4. If the posterior is consistent, for any $B \subset \mathcal{Z}$,

$$\lim_{n\to\infty} \mathbb{E}[P(B)|X_1,\ldots,X_n] = P_0(B), \quad P_0^\infty - a.s.$$

Proof. We will show for $B = \{z\}$, $z \in \mathcal{Z}$. For an arbitrary $B \subset \mathcal{Z}$, a similar argument can be used to prove the lemma. An application of Theorem 1.2.2 of Ghosh and Ramamoorthi (2003) for a closed set B gives us

$$\limsup \mathbb{E}[P(B)|X_1,\ldots,X_n] \leq P_0(B), \quad P_0^\infty - a.s.$$

By the assumptions, for all sufficiently small $\epsilon > 0$, there exists an open set B_ϵ with $B \subset B_\epsilon$ such that $\nu(B_\epsilon) < \epsilon$ and $B_\epsilon \cap \mathcal{Z} = B$. Since $z \in \mathcal{Z}$, for all sufficiently large n, $z = \tilde{X}_j$ for some j. Thus,

$$\mathbb{E}[P(B_\epsilon)|X_1,\ldots,X_n] = p_j(\mathbf{n}) + p_{k+1}(\mathbf{n})\nu(B_\epsilon) \leq \mathbb{E}[P(B)|X_1,\ldots,X_n] + \epsilon.$$

This implies

$$P_0(B) \leq P_0(B_\epsilon) \leq \liminf_{n\to\infty} \mathbb{E}[P(B_\epsilon)|X_1,\ldots,X_n] \leq \liminf_{n\to\infty} \mathbb{E}[P(B)|X_1,\ldots,X_n] + \epsilon.$$

Since ϵ is arbitrary, this completes the proof. ■

Lemma 6. Assume the assumptions in Theorem 4 and P_0 is a mixture of discrete probability measure and ν with $\lambda > 0$. The following convergences hold $P_0^\infty - a.s.$

(i) $\displaystyle\sum_{i:\tilde{X}_i\in\mathcal{Z}} \sum_{j:\tilde{X}_j\in\mathcal{Z}} p_i(\mathbf{n})p_j(\mathbf{n}^{i+})I(\tilde{X}_i \in B)I(\tilde{X}_j \in B) \to P_0(B\cap\mathcal{Z})^2$,

(ii) $\displaystyle\sum_{i:\tilde{X}_i\in\mathcal{Z}} p_i(\mathbf{n})I(\tilde{X}_i \in B)\big[\sum_{j:\tilde{X}_j\notin\mathcal{Z}} p_j(\mathbf{n}^{i+})I(\tilde{X}_j \in B) + p_{k+1}(\mathbf{n}^{i+})\nu(B)\big] \to \lambda P_0(B\cap \mathcal{Z})\nu(B)$,

(iii) $\displaystyle\sum_{i:\tilde{X}_i\notin\mathcal{Z}} p_i(\mathbf{n})I(\tilde{X}_i \in B)\big[\sum_{j:\tilde{X}_j\notin\mathcal{Z}} p_j(\mathbf{n}^{i+})I(\tilde{X}_j \in B) + 2p_{k+1}(\mathbf{n}^{i+})\nu(B)\big]$
$+p_{k+1}(\mathbf{n})\nu(B)[p_{k+2}(\mathbf{n}^{(k+1)+})\nu(B) + p_{k+1}(\mathbf{n}^{(k+1)+})] \to (\lambda\nu(B))^2$.

Proof. Let the three left hand side terms in (i), (ii) and (iii) be I_1, I_2 and I_3 respectively.
(i) Using $p_j(\mathbf{n}^{i+}) = (n_j + I(i=j))/(n+1)$, I_1 is separated as follows.

$$\begin{aligned} I_1 &= \sum_{i:\tilde{X}_i\in\mathcal{Z}} p_i(\mathbf{n})I(\tilde{X}_i \in B)\frac{n}{n+1}F_{d,n}(B) \\ &\quad + \sum_{i:\tilde{X}_i\in\mathcal{Z}} p_i(\mathbf{n})I(\tilde{X}_i \in B)\Bigg[\sum_{j:\tilde{X}_j\in\mathcal{Z}} (p_j(\mathbf{n}^{i+}) - p_j^*(\mathbf{n}^{i+}))I(\tilde{X}_j \in B) + \frac{1}{n+1}I(\tilde{X}_i \in B)\Bigg] \\ &= I_{11} + I_{12} \end{aligned}$$

where $F_{d,n}(B) = n^{-1} \sum_{j:\tilde{X}_j \in \mathcal{Z}} n_j I(\tilde{X}_j \in B)$. By the strong law of large numbers, (4.5) and Proposition 3 (i),

$$|I_{11} - P_{0d}(B)^2| \leq \left| \frac{n}{n+1} F_{d,n}(B)^2 - P_{0d}(B)^2 \right| + \sum_{j:\tilde{X}_j \in \mathcal{Z}} |p_j(\mathbf{n}) - p_j^*(\mathbf{n})| \to 0, \quad P_0^\infty - a.s.$$

where $P_{0d}(B) = P_0(B \cap \mathcal{Z})$. From the assumption (4.5),

$$|I_{12}| \leq \sum_{i:\tilde{X}_i \in \mathcal{Z}} p_i(\mathbf{n}) \Bigg[\sum_{j:\tilde{X}_j \in \mathcal{Z}} |p_j(\mathbf{n}^{i+}) - p_j^*(\mathbf{n}^{i+})| + \frac{1}{n+1} I(\tilde{X}_i \in B) \Bigg] \to 0, \quad P_0^\infty - a.s.$$

(ii) I_2 is also separated into two groups. **Let** $H_{k^*(\mathbf{n})}(B) = (k^*(\mathbf{n}))^{-1} \sum_{i:\tilde{X}_i \notin B - \mathcal{Z}} p_i(\mathbf{n})$**.**

$$\begin{aligned} I_2 = & \sum_{i:\tilde{X}_i \in \mathcal{Z}} p_i(\mathbf{n}) I(\tilde{X}_i \in B)[\lambda\nu(B) \\ & \quad + (k^*(\mathbf{n}^{i+})p_+(\mathbf{n}^{i+}) + p_{k+1}(\mathbf{n}^{i+}) - \lambda)\nu(B) + k^*(\mathbf{n}^{i+})p_+(\mathbf{n}^{i+})(H_{k^*(\mathbf{n}^{i+})}(B) - \nu(B))] \\ = {} & I_{21} + I_{22}. \end{aligned}$$

By Proposition 3 (i), $I_{21} \to P_0(B \cap \mathcal{Z})\lambda\nu(B)$, P_0^∞-a.s. and

$$\begin{aligned} |I_{22}| \leq & \sum_{i:\tilde{X}_i \in \mathcal{Z}} p_i(\mathbf{n}) \big[|k^*(\mathbf{n}^{i+})p_+(\mathbf{n}^{i+}) + p_{k+1}(\mathbf{n}^{i+}) - \lambda| + |H_{k^*(\mathbf{n}^{i+})}(B) - \nu(B)| \big] \\ & \to 0, \quad P_0^\infty - a.s. \end{aligned}$$

(iii) Since $\sum_{i:\tilde{X}_i \notin \mathcal{Z}} p_i(\mathbf{n}) I(\tilde{X}_i \in B) = k^* p_+(\mathbf{n}) H_{k^*}(B)$, we have

$$\begin{aligned} I_3 = & (k^* p_+(\mathbf{n}) H_{k^*}(B) + p_{k+1}(\mathbf{n})\nu(B))\,\lambda\nu(B) \\ & + p_{k+1}(\mathbf{n})\nu(B) \big[(k^*(\mathbf{n}^{(k+1)+})p_+(\mathbf{n}^{(k+1)+}) + p_{k+2}(\mathbf{n}^{(k+1)+}) - \lambda)\nu(B) \\ & \qquad + p_{k+1}(\mathbf{n}^{(k+1)+}) + k^*(\mathbf{n}^{(k+1)+})p_+(\mathbf{n}^{(k+1)+})(H_{k^*(\mathbf{n}^{(k+1)+})}(B) - \nu(B)) \big] \\ & + \sum_{i:\tilde{X}_i \notin \mathcal{Z}} p_i(\mathbf{n}) I(\tilde{X}_i \in B) \Big[(k^*(\mathbf{n}^{i+})p_+(\mathbf{n}^{i+}) + p_{k+1}(\mathbf{n}^{i+}) - \lambda)\nu(B) \\ & \qquad + k^*(\mathbf{n}^{i+})p_+(\mathbf{n}^{i+})(H_{k^*(\mathbf{n}^{i+})}(B) - \nu(B)) \Big] \\ = {} & I_{31} + I_{32} + I_{33} + I_{34} + I_{35}. \end{aligned}$$

Note that by Proposition 3 (ii), $I_{31} \to \lambda\nu(B)\lambda\nu(B)$, P_0^∞-a.s. and

$$\begin{aligned} |I_{32}| & \leq |k^*(\mathbf{n}^{(k+1)+})p_+(\mathbf{n}^{(k+1)+}) + p_{k+2}(\mathbf{n}^{(k+1)+}) - \lambda| \to 0, \quad P_0^\infty - a.s. \\ |I_{33}| & \leq p_{k+1}(\mathbf{n}^{(k+1)+}) + |H_{k^*(\mathbf{n}^{(k+1)+})}(B) - \nu(B)| \to 0, \quad P_0^\infty - a.s. \\ |I_{34}| & \leq \sum_{i:\tilde{X}_i \notin \mathcal{Z}} p_i(\mathbf{n}) |k^*(\mathbf{n}^{i+})p_+(\mathbf{n}^{i+}) + p_{k+1}(\mathbf{n}^{i+}) - \lambda| \to 0, \quad P_0^\infty - a.s. \\ |I_{35}| & \leq \sum_{i:\tilde{X}_i \notin \mathcal{Z}} p_i(\mathbf{n}) |H_{k^*(\mathbf{n}^{i+})}(B) - \nu(B)| \to 0, \quad P_0^\infty - a.s. \end{aligned}$$

Thus, the lemma follows. ■

We now give the proofs of Propositions 2 and 3 and Theorem 4.

Proof of Proposition 2. From Lemma 2, we have

$$\begin{aligned}\mathbb{E}(P(B)|X_1,\ldots,X_n) &= \sum_{j=1}^{k}\frac{n_j}{n}I(\tilde{X}_j\in B)+\sum_{j=1}^{k}(p_j(\mathbf{n})-\frac{n_j}{n})I(\tilde{X}_j\in B)+p_{k+1}(\mathbf{n})\nu(B)\\ &= F_n(B)+\sum_{j=1}^{k}(p_j(\mathbf{n})-\frac{n_j}{n})I(\tilde{X}_j\in B)+p_{k+1}(\mathbf{n})\nu(B).\end{aligned}$$

By the assumptions and Proposition 3 (iii), the second and the third terms tend to 0. Thus, as $n\to\infty$, we get $|\mathbb{E}(P(B)|X_1,\ldots,X_n)-F_n(B)|\to 0$, $P_0^\infty - a.s.$

In (ii) of Lemma 2, $\mathbb{E}[P(B)^2|X_1,\ldots,X_n]$ is expressed as the sum of four terms, which we will call $I_1,\ldots,I_4$ in the order of appearance.

First term I_1 is separated into three parts as follows

$$\sum_{i,j=1}^{k}\left[p_i^*(\mathbf{n})p_j^*(\mathbf{n}^{i+})+p_i^*(\mathbf{n})(p_j(\mathbf{n}^{i+})-p_j^*(\mathbf{n}^{i+}))+(p_i(\mathbf{n})-p_i^*(\mathbf{n}))p_j(\mathbf{n}^{i+})\right]I(\tilde{X}_i,\tilde{X}_j\in B)$$

The second and third group converge to 0 $P_0^\infty - a.s.$ by (4.2) and (4.4). The first group becomes

$$\sum_{i,j=1}^{k}p_i^*(\mathbf{n})p_j^*(\mathbf{n}^{i+})I(\tilde{X}_i,\tilde{X}_j\in B)=\frac{n}{n+1}F_n(B)^2+\frac{1}{n+1}F_n(B).$$

Thus, $I_1\to P_0(B)^2$, $P_0^\infty - a.s.$ as $n\to\infty$.

The other terms are

$$\begin{aligned}|I_2+I_3+I_4| &\le p_{k+1}(\mathbf{n})\Big[2\sum_{i=1}^{k}p_i(\mathbf{n}^{(k+1)+})+p_{k+1}(\mathbf{n}^{(k+1)+})+p_{k+2}(\mathbf{n}^{(k+1)+})\Big]\\ &\le 2p_{k+1}(\mathbf{n})\to 0.\quad P_0^\infty - a.s.\end{aligned}$$

Thus, we get the desired result. ■

Proof of Proposition 3.
(i). Fix a Borel set B. Using $\sum_{i=1}^{k}(n_i/n)I(\tilde{X}_i\in\mathcal{Z}\cap B)=F_n(\mathcal{Z}\cap B)$,

$$\Big|\sum_{i=1}^{k}p_i(\mathbf{n})I(\tilde{X}_i\in\mathcal{Z}\cap B)-F_n(\mathcal{Z}\cap B)\Big|\le\sum_{i=1}^{k}|p_i(\mathbf{n})-n_i/n|.$$

Dominated convergence theorem and (4.5) give

$$\lim_{n\to\infty}\sum_{i=1}^{k}p_i(\mathbf{n})I(\tilde{X}_i\in\mathcal{Z}\cap B)=\lim_{n\to\infty}F_n(\mathcal{Z}\cap B)=P_0(\mathcal{Z}\cap B)=\sum_{i=1}^{\infty}q_iI(z_i\in B).$$

(ii). Apply (i) with $B = \mathcal{Z}$. Then,

$$\lim_{n\to\infty} \sum_{i:\tilde{X}_i \in \mathcal{Z}} p_i(\mathbf{n}) = \sum_{i=1}^{\infty} q_i = 1 - \lambda, \quad P_0^\infty - a.s.$$

By noting that $p_1(\mathbf{n}) + \cdots + p_{k+1}(\mathbf{n}) = 1$, we get

$$k^* p_+(\mathbf{n}) + p_{k+1}(\mathbf{n}) = \sum_{j:\tilde{X}_j \notin \mathcal{Z}} p_j(\mathbf{n}) + p_{k+1}(\mathbf{n}) = 1 - \sum_{j:\tilde{X}_j \in \mathcal{Z}} p_j(\mathbf{n}) \to \lambda, \quad P_0^\infty - a.s.$$

(iii). Note that (4.5) is assumed. The equation (4.4) is

$$\begin{aligned}
\lim_{n\to\infty} C_n &= \lim_{n\to\infty} \Big[\sum_{j:\tilde{X}_j \in \mathcal{Z}} |p_j(\mathbf{n}) - n_j/n| + \sum_{j:\tilde{X}_j \notin \mathcal{Z}} |p_j(\mathbf{n}) - n_j/n| \Big] \\
&= \lim_{n\to\infty} \Big[\sum_{j:\tilde{X}_j \in \mathcal{Z}} |p_j(\mathbf{n}) - n_j/n| + |(k^* p_+(\mathbf{n}) + p_{k+1}(\mathbf{n})) - k^*/n - p_{k+1}(\mathbf{n})| \Big] \\
&= \lim_{n\to\infty} \big[0 + |\lambda - \lambda - p_{k+1}(\mathbf{n})| \big] \\
&= \lim_{n\to\infty} p_{k+1}(\mathbf{n}), \quad P_0^\infty - a.s.
\end{aligned}$$

If $\lambda > 0$ is assumed, then P_0^∞-almost surely

$$0 = \lim_{n\to\infty} \sum_{j:\tilde{X}_j \notin \mathcal{Z}} |p_j(\mathbf{n}) - \frac{n_j}{n}| = \lim_{n\to\infty} (k^*/n)|np_+(\mathbf{n}) - 1| = \lambda \lim_{n\to\infty} |np_+(\mathbf{n}) - 1|. \quad \blacksquare$$

Proof of Theorem 4.
(Sufficiency)

Case A: when condition (i) holds. In this case, the posterior is consistent by the Proposition 3 (iii) and Proposition 2.

Case B: when condition (ii) holds. We first consider the case when P_0 is discrete, i.e., $\lambda = 0$. By noting that $p_{k+1}(\mathbf{n}) = 1 - \sum_{j=1}^{k} p_j(\mathbf{n}) \to 1 - (1 - \lambda) = 0$, $P_0^\infty - a.s.$, this case reduces to Case A and the posterior is consistent.

Now we consider the case when P_0 is a mixture of a discrete probability measure and ν with $\lambda > 0$. Note

$$\begin{aligned}
&\mathbb{E}[P(B)|X_1, \ldots, X_n] = \sum_{j=1}^{k} p_j(\mathbf{n}) I(\tilde{X}_j \in B) + p_{k+1}(\mathbf{n})\nu(B) \\
&= \sum_{j=1}^{k} p_j(\mathbf{n}) I(\tilde{X}_j \in \mathcal{Z} \cap B) + \sum_{j=1}^{k} p_j(\mathbf{n}) I(\tilde{X}_j \in B - \mathcal{Z}) + p_{k+1}(\mathbf{n})\nu(B).
\end{aligned}$$

The first term converges to $\sum_{j=1}^{\infty} q_j I(z_j \in B)$ by Proposition 3 (i). The fact that $H_{k^*}(B)$ converges to $\nu(B)$ together with Proposition 3 (ii) yields

$$k^* p_+(\mathbf{n})(H_{k^*}(B) - \nu(B)) + (k^* p_+(\mathbf{n}) + p_{k+1}(\mathbf{n}))\nu(B) \to \lambda\nu(B), \quad P_0^\infty - a.s,$$

where H_{k^*} is defined by $H_{k^*}(B) = (k^*)^{-1} \sum_{j=1}^{k} I(\tilde{X}_j \in B - \mathcal{Z})$ for all Borel set B. Thus,

$$\lim_{n\to\infty} \mathbb{E}[P(B)|X_1, \dots, X_n] = \sum_{j=1}^{\infty} q_j I(z_j \in B) + \lambda\nu(B) = P_0(B), \quad P_0^{\infty} - a.s.$$

Terms in Lemma 2 are regrouped as

$$\begin{aligned}
&\sum_{i:\tilde{X}_i\in\mathcal{Z}} \sum_{j:\tilde{X}_j\in\mathcal{Z}} p_i(\mathbf{n}) p_j(\mathbf{n}^{i+}) I(\tilde{X}_i \in B) I(\tilde{X}_j \in B) \\
&\quad + 2 \sum_{i:\tilde{X}_i\in\mathcal{Z}} p_i(\mathbf{n}) I(\tilde{X}_i \in B) \Bigg[\sum_{j:\tilde{X}_j\notin\mathcal{Z}} p_j(\mathbf{n}^{i+}) I(\tilde{X}_j \in B) + p_{k+1}(\mathbf{n}^{i+})\nu(B) \Bigg] \\
&\quad + \sum_{i:\tilde{X}_i\notin\mathcal{Z}} p_i(\mathbf{n}) I(\tilde{X}_i \in B) \Bigg[\sum_{j:\tilde{X}_j\notin\mathcal{Z}} p_j(\mathbf{n}^{i+}) I(\tilde{X}_j \in B) + 2p_{k+1}(\mathbf{n}^{i+})\nu(B) \Bigg] \\
&\quad + p_{k+1}(\mathbf{n})\nu(B)[p_{k+2}(\mathbf{n}^{(k+1)+})\nu(B) + p_{k+1}(\mathbf{n}^{(k+1)+})] \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

Lemma 4 shows that $I_1 \to P_0(B \cap \mathcal{Z})^2$, $I_2 \to 2\lambda P_0(B \cap \mathcal{Z})\nu(B)$ and $I_3 + I_4 \to (\lambda\nu(B))^2$, P_0^{∞}-a.s. Hence, we have

$$\lim_{n\to\infty} \mathbb{E}[P(B)^2|X_1, \dots, X_n] = (P_0(B \cap \mathcal{Z}) + \lambda\nu(B))^2 = (P_0(B))^2, \quad P_0^{\infty} - a.s.$$

Therefore, by Lemma 1, the posterior is consistent at P_0.

(Necessity) Suppose the posterior is consistent at P_0. By taking $B = \{z\}$ with $z \in \mathcal{Z}$ in Lemma 3, we have $p_j(\tilde{\mathbf{n}}) \to q_j$, $P_0^{\infty} - a.s$, for all j such that $\tilde{n}_j > 1$ and the decreasing ordering $\tilde{\mathbf{n}} = (\tilde{n}_1, \dots, \tilde{n}_k)$ of $n_1, \dots, n_k$. Also it is not hard to see $p_+(\mathbf{n}) - 1/n \to 0$.

Fix $\epsilon > 0$. There is L such that $\sum_{i>L} q_i < \epsilon/6$. By the strong law of large numbers and Proposition 3 (i), there is N_0 such that $|p_i^*(\tilde{\mathbf{n}}) - q_i| < \epsilon q_i/6$ and $|p_i(\tilde{\mathbf{n}}) - q_i| < \epsilon q_i/6$ for all $n \geq N_0$ and $i = 1, \dots, L$ where $p_i^*(\mathbf{n}) = n_i/n$. Again by the strong law of large numbers and Lemma 3 with $B = \mathcal{Z} - \{z_1, \dots, z_L\}$, we get

$$\sum_{j:\tilde{X}_j\in B} p_j^*(\mathbf{n}) \to \sum_{j>L} q_j \quad \text{and} \quad \sum_{j:\tilde{X}_j\in B} p_j(\mathbf{n}) \to \sum_{j>L} q_j, \quad P_0^{\infty} - a.s.$$

Also, there is N_1 such that for all $n \geq N_1$

$$\Bigg| \sum_{j:\tilde{X}_j\in B} p_j^*(\mathbf{n}) - \sum_{j>L} q_j \Bigg| < \epsilon/6 \quad \text{and} \quad \Bigg| \sum_{j:\tilde{X}_j\in B} p_j(\mathbf{n}) - \sum_{j>L} q_j \Bigg| < \epsilon/6.$$

By merging inequalities, for all $n \geq N_0 + N_1$

$$\sum_{j:\tilde{X}_j\in\mathcal{Z}} |p_j(\mathbf{n}) - n_j/n| \leq \epsilon.$$

The arbitrariness of $\epsilon > 0$ gives (4.5).

Now suppose that P_0 is not a mixture of a discrete probability measure and ν. Then, P_0 is a mixture of a discrete probability measure and a diffuse probability measure μ that is different from ν with $\lambda > 0$.

The predictive probability given by

$$\mathbb{E}(P(B)|X_1, \ldots, X_n) = \sum_{j=1}^{k} p_j(\mathbf{n}) I(\tilde{X}_j \in B) + p_{k+1}(\mathbf{n})\nu(B)$$

should converge to $P_0(B)$ for every P_0-continuity set B. Since $\mu \neq \nu$ and the predictive probability is eventually unaffected by ν, $p_{k+1}(\mathbf{n})$ should converge to 0.

Suppose that there is a sequence $\mathbf{n}_l$ such that $p_{k+1}(\mathbf{n}_l) \to \phi > 0$. Since $k^* p_+(\mathbf{n}) + p_{k+1}(\mathbf{n}) \to \lambda$, $k^* p_+(\mathbf{n}_l) \to \lambda - \phi \geq 0$. The predictive probability is

$$\sum_{j=1}^{k} p_j(\mathbf{n}_l) I(\tilde{X}_j \in \mathcal{Z} \cap B) + \sum_{j=1}^{k} p_j(\mathbf{n}_l) I(\tilde{X}_j \in B - \mathcal{Z}) + p_{k+1}(\mathbf{n}_l)\nu(B).$$

By the Proposition 3 (i), the first term converges to $\sum_{j=1}^{\infty} q_j I(z_j \in B), \quad P_0^{\infty} - a.s.$ The second term goes to

$$\begin{aligned} p_+(\mathbf{n}_l) \sum_{j=1}^{k} I(\tilde{X}_j \in B - \mathcal{Z}) &= k^* p_+(\mathbf{n}_l) \times \frac{1}{k^*} \sum_{j=1}^{k} I(\tilde{X}_j \in B - \mathcal{Z}) \\ &\to (\lambda - \phi)\mu(B - \mathcal{Z}) = (\lambda - \phi)\mu(B), \quad P_0^{\infty} - a.s.. \end{aligned}$$

Hence, the predictive probability $\mathbb{E}[P(B)|X_1, \ldots, X_{n_l}]$ converges to

$$\sum_{j=1}^{\infty} q_j I(z_j \in B) + (\lambda - \phi)\mu(B) + \phi\nu(B) = P_0(B) + \phi(\nu(B) - \mu(B)), \quad P_0^{\infty} - a.s.$$

It contradicts to the assumption that the posterior is consistent at P_0. Thus, $p_{k+1}(\mathbf{n}) \to 0, P_0^{\infty} - a.s.$. This completes the proof. ■

The following lemma is used to check the conditions of the posterior consistency for the N-IG process prior.

Lemma 7. When $0 \leq k \leq n$,

$$n w_{1,n} \to 1 \quad \text{as} \quad n \to \infty.$$

Proof. Since $y \geq 1$, obviously $1 - y^{-2} \leq 1$ and

$$n w_{1,n} = \frac{\int_1^{\infty} (1 - y^{-2})^n y^{k-1} e^{-ay} dy}{\int_1^{\infty} (1 - y^{-2})^{n-1} y^{k-1} e^{-ay} dy} \leq 1.$$

It is not hard to see that $(1 - y^{-2})^{n-1} y^{k-1} e^{-ay}$ on $(1, \infty)$ is unimodal and $\eta_{n,k}^2 = \text{argmax}_{y \geq 1} (1 - y^{-2})^{n-1} y^{k-1} e^{-ay}$ goes to infinity as n increases. Thus $\eta_{n,k}^2$ is unique and $(1 - y^{-2})^{n-1} y^{k-1} e^{-ay}$ increases on $(1, \eta_{n,k}^2)$ and decreases on $(\eta_{n,k}^2, \infty)$. For simplicity, let η be the positive root of $\eta^2_{n,k}$. By reducing the integration range of numerator as (η, ∞) and separating the integration range in denominator, we have

$$\begin{aligned}
nw_{1,n} &\geq \frac{\int_\eta^\infty (1-\eta^{-2})(1-y^{-2})^{n-1} y^{k-1} e^{-ay} dy}{\int_1^\eta (1-y^{-2})^{n-1} y^{k-1} e^{-ay} dy + \int_\eta^\infty (1-y^{-2})^{n-1} y^{k-1} e^{-ay} dy} \\
&\geq (1-\eta^{-2}) \left[1 + \frac{\int_1^\eta (1-y^{-2})^{n-1} y^{k-1} e^{-ay} dy}{\int_\eta^{\eta^2} (1-y^{-2})^{n-1} y^{k-1} e^{-ay} dy} \right]^{-1} \\
&\geq (1-\eta^{-2}) \left[1 + \frac{(\eta-1)(1-\eta^{-2})^{n-1} \eta^{k-1} e^{-a\eta}}{(\eta^2-\eta)(1-\eta^{-2})^{n-1} \eta^{k-1} e^{-a\eta}} \right]^{-1} \\
&= \frac{\eta - 1}{\eta} \to 1, \quad \text{as} \quad n \to \infty. \quad \blacksquare
\end{aligned}$$

The following lemma is used to check the conditions of the posterior consistency for the Poisson-Kingman process prior.

Lemma 8. When $0 \leq k \leq n$,

$$w(n,k) \to 1 \quad \text{as} \quad n \to \infty.$$

Proof. The upper bound is easily obtained by

$$w(n,k) = \frac{\int_0^\infty \frac{u}{b+u} u^{n-1} (b+u)^{ak-n} e^{-c\Gamma(1-a)(b+u)^a/a} du}{\int_0^\infty u^{n-1} (b+u)^{ak-n} e^{-c\Gamma(1-a)(b+u)^a/a} du} \leq 1.$$

The lower bound is quite similar to Lemma 5. For fixed n and k, it is easy to show that $u^{n-1}(b+u)^{ak-n} e^{-c\Gamma(1-a)(b+u)^a/a}$ is unimodal in u. Then, $v^2_{n,k} = \operatorname{argmax}_{u \geq 0} u^{n-1}(b+u)^{ak-n} e^{-c\Gamma(1-a)(b+u)^a/a}$ is unique and goes infinity as n goes to infinity. Besides, the argument function increases on $(0, v^2_{n,k})$ and decreases on $(v^2_{n,k}, \infty)$.

Let v be the positive square root of $v^2_{n,k}$. For the lower bound of $w(n,k)$, the integration range of numerator is reduced to (v, ∞). A similar calculation to Lemma 5 gives

$$\begin{aligned}
w(n,k) &\geq \frac{\int_v^\infty \frac{v}{b+v} u^{n-1} (b+u)^{ak-n} e^{-c\Gamma(1-a)(b+u)^a/a} du}{\int_0^\infty u^{n-1} (b+u)^{ak-n} e^{-c\Gamma(1-a)(b+u)^a/a} du} \\
&\geq \frac{v}{b+v} \left[1 + \frac{\int_0^v u^{n-1} (b+u)^{ak-n} e^{-c\Gamma(1-a)(b+u)^a/a} du}{\int_v^{v^2} u^{n-1} (b+u)^{ak-n} e^{-c\Gamma(1-a)(b+u)^a/a} du} \right]^{-1} \\
&\geq \frac{v}{b+v} \left[1 + \frac{v \cdot v^{n-1} (b+v)^{ak-n} e^{-c\Gamma(1-a)(b+v)^a/a}}{(v^2 - v) \cdot v^{n-1} (b+v)^{ak-n} e^{-c\Gamma(1-a)(b+v)^a/a}} \right]^{-1} \\
&= \frac{v-1}{b+v} \to 1 \quad \text{as} \quad n \to \infty. \quad \blacksquare
\end{aligned}$$

References

Ghosh, J. K. and Ramamoorthi, R. V. (2003). *Bayesian nonparametrics.* Springer-Verlag.