Supplementary Material for

AN ASYMPTOTIC ANALYSIS OF A CLASS OF DISCRETE NONPARAMETRIC PRIORS

Pierpaolo De Blasi[1], Antonio Lijoi[2] and Igor Prünster[1]

[1] *University of Torino and Collegio Carlo Alberto, Italy*
[2] *University of Pavia and Collegio Carlo Alberto, Italy*

Abstract: This supplementary materical contains: (i) the proof of Proposition 1 stated in the main manuscript; (ii) an auxiliary technical lemma, which is used for proving Theorem 1 within the main manuscript; (iii) details on the asymptotic derivations concerning the specific examples considered in Section 4. Equations that are specific to this supplementary file are labeled as (A$\star$). All other referenced equations correspond to those of the main paper.

1. Proof of Proposition 1.

Without loss of generality we assume that the support of the prior guess $\mathrm{E}[\tilde{p}(\,\cdot\,)] = P^*$ coincides with $\mathbb{X}$. Let us start by considering the case of $\sigma < 0$. Let $d_{\mathbb{X}}$ be the distance on $\mathbb{X}$ and let d_w denote the Prokhorov distance on $\mathbf{P}_{\mathbb{X}}$. We wish to show that any weak-neighborhood of G_0 has positive Q mass for any probability measure $G_0 \in \mathbf{P}_X$. Since $\mathbb{X}$ is separable, it is well-known that the set of discrete distributions with a finite number of point masses is dense in $\mathbf{P}_{\mathbb{X}}$, with respect to d_w. Hence, for any $\epsilon > 0$ there exists a positive integer k_0, vector of weights $(p_1^0, \ldots, p_{k_0}^0)$ in the k_0-dimensional simplex Δ_{k_0} and points $x_1^0, \ldots, x_{k_0}^0$ in $\mathbb{X}$ such that $d_w(G_{\mathbf{p}^0,\mathbf{x}^0}, G_0) < \epsilon/2$, where $G_{\mathbf{p}^0,\mathbf{x}^0} = \sum_{i=1}^{k_0} p_i^0 \delta_{x_i^0}$. For any $\eta, \delta > 0$ introduce the sets

$$U_0(\eta) = \{\mathbf{p} = (p_1, \ldots, p_{k_0}) \in \Delta_{k_0} :\ |p_i - p_i^0| < \eta \,\text{for any}\ i = 1, \ldots, k_0\}$$
$$V_0(\eta) = \{\mathbf{x} = (x_1, \ldots, x_{k_0}) \in \mathbb{X}^{k_0} :\ d_{\mathbb{X}}(x_i, x_i^0) < \delta \,\text{for any}\ i = 1, \ldots, k_0\}$$

and $W_0(\eta, \delta)$ stands for the set of discrete probability distributions $G_{\mathbf{p},\mathbf{x}} = \sum_{i=1}^{k_0} p_i \delta_{x_i}$ for $\mathbf{p} \in U_0(\eta)$ and $\mathbf{x} \in V_0(\delta)$. Recall that, conditionally on $K = k_0$, the vector $(p_1, \ldots, p_{k^0})$ has symmetric Dirichlet distribution with parameter $|\sigma|$. This fact, combined with the assumptions on π and on P^*, entails $Q(W_0(\eta,\delta)) > 0$. The proof is completed by showing that, for appropriate choices of η and δ, any $G_{\mathbf{p},\mathbf{x}}$ in $W_0(\eta,\delta)$ is such that $d_w(G_{\mathbf{p},\mathbf{x}}, G_0) < \epsilon$. But this follows by standard arguments. Since

$$d_w(G_{\mathbf{p},\mathbf{x}}, G_0) \le d_w(G_{\mathbf{p},\mathbf{x}}, G_{\mathbf{p}^0,\mathbf{x}^0}) + \frac{\epsilon}{2},$$

we next show that $\eta = \delta/k_0$ implies that $d_w(G_{\mathbf{p},\mathbf{x}}, G_{\mathbf{p}^0,\mathbf{x}^0}) < \delta$ so that $\delta = \epsilon/2$ would work. For $A \in \mathscr{X}$, the set A^ρ stands for A enlarged by its $d_{\mathbb{X}}$-neighbourhood with radius ρ, $A^\rho = \{x : d_{\mathbb{X}}(x, A) < \rho\}$. When $\rho > \delta$, it is obvious that $x_i^0 \in A$ implies that $x_i \in A^\rho$ whenever $\mathbf{x} = (x_1, \ldots, x_{k_0})$ is in $V_0(\delta)$. One can equivalently say that if $I_0 = \{i : x_i^0 \in A\}$ and $I = \{i : x_i \in A^\rho\}$, then $I \supset I_0$ and

$$\begin{aligned} G_{\mathbf{p}^0,\mathbf{x}^0}(A) - G_{\mathbf{p},\mathbf{x}}(A^\rho) &= \sum_{i \in I_0 \cap I} (p_i^0 - p_i) - \sum_{i \in I \setminus I_0} p_i \\ &\le \sum_{i \in I} (p_i^0 - p_i) \le \eta \, \mathrm{card}(I) = \eta k_0 = \delta < \rho. \end{aligned}$$

On the other hand, if $\rho < \delta$, then there exists some set A in $\mathscr{X}$ such that $x_i^0 \in A$ and $x_i \notin A^\rho$ so that is not possible to bound $G_{\mathbf{p}^0,\mathbf{x}^0}(A) - G_{\mathbf{p},\mathbf{x}}(A^\rho)$ by ρ. This completes the proof of the case $\sigma < 0$. The Dirichlet case is well known (Ferguson, 1973; Majumdar, 1992) and the general $\sigma = 0$ case follows by direct extension of the results concerning the Dirichlet process. The case of $\sigma > 0$ follows immediately from the representation of Gibbs-type partitions with $\sigma > 0$ in terms of stable completely random measures (Gnedin and Pitman, 2005, Theorem 12 (iii)). $\square$

2. An auxiliary lemma.

Lemma 1. *Let $I(n,k)$ be defined as in (3.3). Then*

$$I(n,k) = \left(\frac{V_{n+2,k}}{V_{n+1,k}} - \frac{V_{n+2,k+1}}{V_{n+1,k+1}}\right)(n - \sigma k) + \frac{V_{n+2,k}}{V_{n+1,k}} \tag{A1}$$

$$I(n,k) = \frac{V_{n+2,k+2}}{V_{n+1,k+1}} - \frac{V_{n+2,k+1}}{V_{n+1,k}} + \frac{V_{n+2,k+1}}{V_{n+1,k+1}}(1-\sigma) \tag{A2}$$

$$\begin{aligned} &\frac{V_{n+2,k}}{V_{n+1,k}} - \frac{V_{n+2,k+1}}{V_{n+1,k+1}} \\ &\quad \underset{(>)}{\le} [n+1-\sigma(k+1)]\left(\frac{V_{n+2,k+2}}{V_{n+1,k+1}} - \frac{V_{n+2,k+1}}{V_{n+1,k}}\right) \quad \textit{for} \quad \begin{array}{l} 0 \le \sigma < 1 \\ (\sigma < 0) \end{array} \end{aligned} \tag{A3}$$

Proof. The proof relies on the backward recursion defining the weights of Gibbs-type priors, which is stated in (1.4). As for equation (A1),

$$\begin{aligned}\left(\frac{V_{n+2,k}}{V_{n+1,k}} - \frac{V_{n+2,k+1}}{V_{n+1,k+1}}\right)(n-\sigma k) + \frac{V_{n+2,k}}{V_{n+1,k}} \\ &= \frac{V_{n+2,k}}{V_{n+1,k}}(n+1-\sigma k) - \frac{V_{n+2,k+1}}{V_{n+1,k+1}}(n-\sigma k) \\ &= 1 - \frac{V_{n+2,k+1}}{V_{n+1,k}} - \frac{V_{n+2,k+1}}{V_{n+1,k+1}}(n-\sigma k) \\ &= 1 - \frac{V_{n+2,k+1}}{V_{n+1,k}}\left(1 + \frac{V_{n+1,k}}{V_{n+1,k+1}}(n-\sigma k)\right) \\ &= 1 - \frac{V_{n+2,k+1}}{V_{n+1,k}}\frac{V_{n,k}}{V_{n+1,k+1}} = I(n,k)\end{aligned}$$

where we used the backward recursion (1.4) for $(n+1,k)$ in the second equality and for (n,k) in the last equality.

As for equation (A2), we use the backward recursion (1.4) for $(n+1,k+1)$ to get $V_{n+2,k+2} + V_{n+2,k+1}(1-\sigma) = V_{n+1,k+1} - (n-\sigma k)V_{n+2,k+1}$. Then

$$\begin{aligned}\frac{V_{n+2,k+2}}{V_{n+1,k+1}} - \frac{V_{n+2,k+1}}{V_{n+1,k}} + \frac{V_{n+2,k+1}}{V_{n+1,k+1}}(1-\sigma) \\ &= \frac{V_{n+1,k+1} - V_{n+2,k+1}(n-\sigma k)}{V_{n+1,k+1}} - \frac{V_{n+2,k+1}}{V_{n+1,k}} \\ &= 1 - \frac{V_{n+2,k+1}}{V_{n+1,k+1}}(n-\sigma k) - \frac{V_{n+2,k+1}}{V_{n+1,k}} \\ &= 1 - \frac{V_{n+2,k+1}}{V_{n+1,k+1}}\left((n-\sigma k) + \frac{V_{n+1,k+1}}{V_{n+1,k}}\right) \\ &= 1 - \frac{V_{n+2,k+1}}{V_{n+1,k}}\frac{V_{n,k}}{V_{n+1,k}} = I(n,k)\end{aligned}$$

where we used again the backward recursion for (n,k) in the last equality.

As for equation (A3), use the backward recursion (1.4) for $(n+1,k+1)$ and $(n+1,k)$ on the right hand side, respectively, to get

$$\begin{aligned}\frac{V_{n+2,k+2}}{V_{n+1,k+1}} - \frac{V_{n+2,k+1}}{V_{n+1,k}} \\ &= \left(1 - \frac{V_{n+2,k+1}}{V_{n+1,k+1}}(n+1-\sigma(k+1))\right) - \left(1 - \frac{V_{n+2,k}}{V_{n+1,k}}(n+1-\sigma k)\right) \\ &= (n+1-\sigma(k+1))\left(\frac{V_{n+2,k}}{V_{n+1,k}}\frac{n+1-\sigma k}{n+1-\sigma(k+1)} - \frac{V_{n+2,k+1}}{V_{n+1,k+1}}\right).\end{aligned}$$

Finally, consider that $\frac{n+1-\sigma k}{n+1-\sigma(k+1)} \geq 1$ for $0 \leq \sigma \leq 1$ implies the first inequality and that $\frac{n+1-\sigma k}{n+1-\sigma(k+1)} < 1$ for $\sigma < 0$ implies the second inequality. $\square$

3. Details for the determination of (4.3) and (4.5).

Determination of (4.3). Consider (4.2) with $\kappa_n = n$ for any n a.s.-P_0^∞, which corresponds to the case of diffuse P_0. By virtue of Eq. (17) of Erdélyi, Magnus, Oberhettinger and Tricomi (1953, Section 6.13.2), the functions ${}_1F_1(n; 2n; \lambda)$ and ${}_1F_1(n+1; 2n+2; \lambda)$ have the same asymptotic expansion as $n \to \infty$, namely

$$\frac{\sqrt{2\pi}\Gamma(2n)}{\sqrt{n/2}\Gamma(n)\Gamma(n)}\, \mathrm{e}^{\lambda/2}\left(\frac{1}{2}\right)^{2n}\big[1+O(1/n)\big].$$

This means that

$$\frac{{}_1F_1(n; 2n; \lambda)}{{}_1F_1(n+1; 2n+2; \lambda)} \to 1$$

as $n \to \infty$, and therefore (4.3) follows. $\square$

Determination of (4.5). A diffuse P_0 implies $\kappa_n = n$ for any n a.s.-P_0^∞ in (4.4). Then, by Eq. (16) in Erdélyi, Magnus, Oberhettinger and Tricomi (1953, Section 2.3.2), one obtains the following asymptotic expansions, as $n \to \infty$,

$${}_2F_1(n+1, n+2; 2n+2; \eta) \sim \left(\frac{2}{\eta}\right)^{4+2n}\big(2-\eta-2\sqrt{1-\eta}\big)^{n+2}C(\eta)$$

$${}_2F_1(n, n+1; 2n; \eta), \sim \left(\frac{2}{\eta}\right)^{2+2n}\big(2-\eta-2\sqrt{1-\eta}\big)^{n+1}C(\eta),$$

where $C(\eta) = \big[(1+2\sqrt{1-\eta}/\eta)^2 - ((2-\eta)/\eta)^2\big]^{-\frac{3}{2}}$. These asymptotic equivalences immediately yield (4.5). $\square$

References

Erdélyi, A., Magnus, W., Oberhettinger, F. and Tricomi, F.G. (1953). *Higher Trascendental Functions*, Vol. 1,McGraw-Hill, New York.

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

Gnedin, A. and Pitman, J. (2005). Exchangeable Gibbs partitions and Stirling triangles. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **325**, 83-102 (translation in *Journal of Mathematical Sciences* **138** (2006), 5674–5685).

Majumdar, S. (1992). On topological support of Dirichlet prior. *Statist. Probab. Lett.* **15**, 385–388.