SUPPLEMENT TO "INCONSISTENCY OF PITMAN–YOR PROCESS MIXTURES FOR THE NUMBER OF COMPONENTS"

By Jeffrey W. Miller* and Matthew T. Harrison*

Brown University

This supplementary document contains various proofs that were excluded from the main document (Miller and Harrison, 2013) in the interest of space.

S1. Miscellaneous proofs.

Proof of Proposition 6.3. (1) For any $\theta \in \Theta$ and any $j \in \{1, \ldots, k\}$,

$$\int_{\mathcal{X}} s_j(x)^2 p_\theta(x)\, d\lambda(x) = \exp(-\kappa(\theta)) \frac{\partial^2}{\partial \theta_j^2} \int_{\mathcal{X}} \exp(\theta^{\mathtt{T}} s(x))\, d\lambda(x) < \infty$$

(Hoffmann-Jørgensen, 1994, 8.36.1). Since P has density $f = \sum \pi_i p_{\theta(i)}$ with respect to λ, then

$$\mathbb{E} s_j(X)^2 = \int_{\mathcal{X}} s_j(x)^2 f(x)\, d\lambda(x) = \sum_{i=1}^{t} \pi_i \int_{\mathcal{X}} s_j(x)^2 p_{\theta(i)}(x)\, d\lambda(x) < \infty,$$

and hence

$$(\mathbb{E}|s(X)|)^2 \le \mathbb{E}|s(X)|^2 = \mathbb{E}s_1(X)^2 + \cdots + \mathbb{E}s_k(X)^2 < \infty.$$

(2) Note that $S_P(s) \subset S_\lambda(s)$ (in fact, they are equal since P_θ and λ are mutually absolutely continuous for any $\theta \in \Theta$), and therefore

$$S_P(s) \subset S_\lambda(s) \subset C_\lambda(s) = \overline{\mathcal{M}}$$

by Proposition B.1(8). Hence,

$$\mathbb{P}(s(X) \in \overline{\mathcal{M}}) \ge \mathbb{P}(s(X) \in S_P(s)) = Ps^{-1}(\mathrm{support}(Ps^{-1})) = 1.$$

(3) Suppose λ is absolutely continuous with respect to Lebesgue measure, $\mathcal{X}$ is open and connected, and s is real analytic. Let $L \subset \mathbb{R}^k$ be a hyperplane, and write $L = \{z \in \mathbb{R}^k : z^{\mathtt{T}} y = b\}$ where $y \in \mathbb{R}^k \smallsetminus \{0\}$, $b \in \mathbb{R}$. Define $g : \mathcal{X} \to \mathbb{R}$ by $g(x) = s(x)^{\mathtt{T}} y - b$. Then g is real analytic on $\mathcal{X}$, since a finite sum of real analytic functions is real analytic. Since $\mathcal{X}$ is connected, it follows that either g is

*Supported in part by NSF grant DMS-1007593 and DARPA contract FA8650-11-1-715.

identically zero, or the set $V = \{x \in \mathcal{X} : g(x) = 0\}$ has Lebesgue measure zero (Krantz, 1992). Now, g cannot be identically zero, since for any $\theta \in \Theta$, letting $Z \sim P_\theta$, we have

$$0 < y^{\mathtt{T}}\kappa''(\theta)y = y^{\mathtt{T}}(\operatorname{Cov} s(Z))y = \operatorname{Var}(y^{\mathtt{T}}s(Z)) = \operatorname{Var} g(Z)$$

by Proposition B.1(2) and (3). Consequently, V must have Lebesgue measure zero. Hence, $P(V) = 0$, since P is absolutely continuous with respect to λ, and thus, with respect to Lebesgue measure. Therefore,

$$\mathbb{P}(s(X) \in L) = \mathbb{P}(g(X) = 0) = P(V) = 0. \qquad \square$$

Proof of Lemma C.1. By Taylor's theorem, for any $x \in B_\varepsilon(x_0)$ there exists z_x on the line between x_0 and x such that, letting $y = x - x_0$,

$$f(x) = f(x_0) + y^{\mathtt{T}}f'(x_0) + \tfrac{1}{2}y^{\mathtt{T}}f''(z_x)y = f(x_0) + \tfrac{1}{2}y^{\mathtt{T}}f''(z_x)y.$$

Since $z_x \in B_\varepsilon(x_0)$, and thus $A \trianglelefteq f''(z_x) \trianglelefteq B$,

$$\tfrac{1}{2}y^{\mathtt{T}}Ay \leq f(x) - f(x_0) \leq \tfrac{1}{2}y^{\mathtt{T}}By.$$

Hence,

$$\begin{aligned} e^{tf(x_0)}\int_{B_\varepsilon(x_0)} \exp(-tf(x))\,dx &\leq \int_{B_\varepsilon(x_0)} \exp(-\tfrac{1}{2}(x-x_0)^{\mathtt{T}}(tA)(x-x_0))\,dx \\ &= (2\pi)^{k/2}|(tA)^{-1}|^{1/2}\ \mathbb{P}\big(|(tA)^{-1/2}Z| \leq \varepsilon\big). \end{aligned}$$

Along with a similar argument for the lower bound, this implies

$$\Big(\frac{2\pi}{t}\Big)^{k/2}\frac{C(t,\varepsilon,B)}{|B|^{1/2}} \leq e^{tf(x_0)}\int_{B_\varepsilon(x_0)} \exp(-tf(x))\,dx \leq \Big(\frac{2\pi}{t}\Big)^{k/2}\frac{C(t,\varepsilon,A)}{|A|^{1/2}}.$$

Considering the rest of the integral, outside of $B_\varepsilon(x_0)$, we have

$$0 \leq \int_{E\smallsetminus B_\varepsilon(x_0)} \exp(-tf(x))\,dx \leq \exp\big(-(t-s)(f(x_0)+\delta)\big)\,g(s).$$

Combining the preceding four inequalities yields the result. $\square$

Although we do not need it (and thus, we omit the proof), the following corollary gives the well-known asymptotic form of the Laplace approximation. (As usual, $g(t) \sim h(t)$ as $t \to \infty$ means that $g(t)/h(t) \to 1$.)

Corollary S1.1. *Let $E \subset \mathbb{R}^k$ be open. Let $f : E \to \mathbb{R}$ be C^2 smooth such that for some $x_0 \in E$ we have that $f'(x_0) = 0$, $f''(x_0)$ is positive definite, and $f(x) > f(x_0)$ for all $x \in E \smallsetminus \{x_0\}$. Suppose there exists $\varepsilon > 0$ such that $B_\varepsilon(x_0) \subset E$ and $\inf\{f(x) - f(x_0) : x \in E \smallsetminus B_\varepsilon(x_0)\}$ is positive, and suppose there is some $s > 0$ such that $\int_E e^{-sf(x)}\,dx < \infty$. Then*

$$\int_E \exp(-tf(x))\,dx \ \sim\ \Big(\frac{2\pi}{t}\Big)^{k/2}\frac{\exp(-tf(x_0))}{|f''(x_0)|^{1/2}}$$

as $t \to \infty$.

S2. Capture lemma. In this section, we prove Lemma 8.1, which is restated here for the reader's convenience.

The following definitions are standard. Let $\mathcal{S}$ denote the unit sphere in $\mathbb{R}^k$, that is, $\mathcal{S} = \{x \in \mathbb{R}^k : |x| = 1\}$. We say that $H \subset \mathbb{R}^k$ is a *halfspace* if $H = \{x \in \mathbb{R}^k : x^{\mathsf{T}} u \prec b\}$, where $\prec$ is either $<$ or $\leq$, for some $u \in \mathcal{S}$, $b \in \mathbb{R}$. We say that $L \subset \mathbb{R}^k$ is a *hyperplane* if $L = \{x \in \mathbb{R}^k : x^{\mathsf{T}} u = b\}$ for some $u \in \mathcal{S}$, $b \in \mathbb{R}$. Given $U \subset \mathbb{R}^k$, let ∂U denote the *boundary of* U, that is, $\partial U = \bar{U} \setminus U^\circ$. So, for example, if H is a halfspace, then ∂H is a hyperplane. The following notation is also useful: given $x \in \mathbb{R}^k$, we call the set $R_x = \{ax : a > 0\}$ the *ray through* x.

We give the central part of the proof first, postponing some plausible intermediate results for the moment.

LEMMA S2.1 (Capture lemma). *Let $V \subset \mathbb{R}^k$ be open and convex. Let P be a probability measure on $\mathbb{R}^k$ such that:*

(1) *$\mathbb{E}|X| < \infty$ when $X \sim P$,*
(2) *$P(\bar{V}) = 1$, and*
(3) *$P(L) = 0$ for any hyperplane L that does not intersect V.*

If $X_1, X_2, \ldots \overset{\text{iid}}{\sim} P$, then for any $\beta \in (0, 1]$ there exists $U \subset V$ compact such that $\mathcal{I}_\beta(X_{1:n}, U) \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$.

PROOF. Without loss of generality, we may assume $0 \in V$ (since otherwise we can translate to make it so, obtain U, and translate back). Let $\beta \in (0, 1]$. By Proposition S2.3 below, for each $u \in \mathcal{S}$ there is a closed halfspace H_u such that $0 \in H_u^\circ$, R_u intersects $V \cap \partial H_u$, and $\mathcal{I}_\beta(X_{1:n}, H_u) \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$. By Proposition S2.5 below, there exist $u_1, \ldots, u_r \in \mathcal{S}$ (for some $r > 0$) such that the set $U = \bigcap_{i=1}^r H_{u_i}$ is compact and $U \subset V$. Finally,

$$\mathcal{I}_\beta(X_{1:n}, U) = \prod_{i=1}^r \mathcal{I}_\beta(X_{1:n}, H_{u_i}) \xrightarrow[n \to \infty]{\text{a.s.}} 1. \qquad \square$$

The main idea of the lemma is exhibited in the following simpler case, which we will use to prove Proposition S2.3.

PROPOSITION S2.2. *Let $V = (-\infty, c)$, where $-\infty < c \leq \infty$. Let P be a probability measure on $\mathbb{R}$ such that:*

(1) *$\mathbb{E}|X| < \infty$ when $X \sim P$, and*
(2) *$P(V) = 1$.*

If $X_1, X_2, \ldots \overset{\text{iid}}{\sim} P$, then for any $\beta \in (0, 1]$ there exists $b < c$ such that $\mathcal{I}_\beta(X_{1:n}, (-\infty, b]) \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$.

PROOF. Let $\beta \in (0, 1]$. By continuity from above, there exists $a < c$ such that $\mathbb{P}(X > a) < \beta$. If $\mathbb{P}(X > a) = 0$ then the result is trivial, taking $b = a$.

Suppose $\mathbb{P}(X > a) > 0$. Let b such that $\mathbb{E}(X \mid X > a) < b < c$, which is always possible, by a straightforward argument (using $\mathbb{E}|X| < \infty$ in the $c = \infty$ case). Let $B_n = B_n(X_1, \dots, X_n) = \{i \in \{1, \dots, n\} : X_i > a\}$. Then

$$\begin{aligned}\frac{1}{|B_n|}\sum_{i \in B_n} X_i &= \frac{1}{\frac{1}{n}|B_n|}\frac{1}{n}\sum_{i=1}^{n} X_i\, I(X_i > a)\\ &\xrightarrow[n\to\infty]{\text{a.s.}} \frac{\mathbb{E}(X\, I(X > a))}{\mathbb{P}(X > a)} = \mathbb{E}(X \mid X > a) < b.\end{aligned}$$

Now, fix $n \in \{1, 2, \dots\}$, and suppose $0 < |B_n| < \beta n$ and $\frac{1}{|B_n|}\sum_{i \in B_n} X_i < b$, noting that with probability 1, this happens for all n sufficiently large. We show that this implies $\mathcal{I}_\beta(X_{1:n}, (-\infty, b]) = 1$. This will prove the result.

Let $A \subset \{1, \dots, n\}$ such that $|A| \geq \beta n$. Let $M = \{\pi_1, \dots, \pi_{|A|}\}$ where π is a permutation of $\{1, \dots, n\}$ such that $X_{\pi_1} \geq \cdots \geq X_{\pi_n}$ (that is, $M \subset \{1, \dots, n\}$ consists of the indices of $|A|$ of the largest entries of $(X_1, \dots, X_n)$). Then $|M| = |A| \geq \beta n \geq |B_n|$, and it follows that $B_n \subset M$. Therefore,

$$\frac{1}{|A|}\sum_{i \in A} X_i \leq \frac{1}{|M|}\sum_{i \in M} X_i \leq \frac{1}{|B_n|}\sum_{i \in B_n} X_i \leq b,$$

as desired. □

The first of the two propositions used in Lemma S2.1 is the following.

PROPOSITION S2.3. *Let V and P satisfy the conditions of Lemma S2.1, and also assume $0 \in V$. If $X_1, X_2, \dots \overset{\text{iid}}{\sim} P$ then for any $\beta \in (0, 1]$ and any $u \in \mathcal{S}$ there is a closed halfspace $H \subset \mathbb{R}^k$ such that*

1. *$0 \in H^\circ$,*
2. *R_u intersects $V \cap \partial H$, and*
3. *$\mathcal{I}_\beta(X_{1:n}, H) \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$.*

PROOF. Let $\beta \in (0, 1]$ and $u \in \mathcal{S}$. Either (a) $R_u \subset V$, or (b) R_u intersects ∂V.

(Case (a)) Suppose $R_u \subset V$. Let $Y_i = X_i^{\mathsf{T}} u$ for $i = 1, 2, \dots$. Then $\mathbb{E}|Y_i| \leq \mathbb{E}|X_i||u| = \mathbb{E}|X_i| < \infty$, and thus, by Proposition S2.2 (with $c = \infty$) there exists $b \in \mathbb{R}$ such that $\mathcal{I}_\beta(Y_{1:n}, (-\infty, b]) \xrightarrow{\text{a.s.}} 1$. Let us choose this b to be positive, which is always possible since $\mathcal{I}_\beta(Y_{1:n}, (-\infty, b])$ is nondecreasing as a function of b. Let $H = \{x \in \mathbb{R}^k : x^{\mathsf{T}} u \leq b\}$. Then $0 \in H^\circ$, since $b > 0$, and R_u intersects $V \cap \partial H$ at bu, since $R_u \subset V$ and $bu^{\mathsf{T}} u = b$. And since $\frac{1}{|A|}\sum_{i \in A} Y_i \leq b$ if and only if $\frac{1}{|A|}\sum_{i \in A} X_i \in H$, we have $\mathcal{I}_\beta(X_{1:n}, H) \xrightarrow{\text{a.s.}} 1$.

(Case (b)) Suppose R_u intersects ∂V at some point $z \in \mathbb{R}^k$. Note that $z \neq 0$ since $0 \notin R_u$. Since $\bar{V}$ is convex, it has a supporting hyperplane at z, and thus, there exist $v \in \mathcal{S}$ and $c \in \mathbb{R}$ such that $G = \{x \in \mathbb{R}^k : x^{\mathsf{T}} v \leq c\}$ satisfies $\bar{V} \subset G$ and

$z \in \partial G$ (Rockafellar, 1970, 11.2). Note that $c > 0$ and $V \cap \partial G = \varnothing$ since $0 \in V$ and V is open. Letting $Y_i = X_i^{\mathsf{T}} v$ for $i = 1, 2, \ldots$, we have

$$\mathbb{P}(Y_i \le c) = \mathbb{P}(X_i^{\mathsf{T}} v \le c) = \mathbb{P}(X_i \in G) \ge \mathbb{P}(X_i \in \bar{V}) = P(\bar{V}) = 1,$$

and hence,

$$\mathbb{P}(Y_i \ge c) = \mathbb{P}(Y_i = c) = \mathbb{P}(X_i^{\mathsf{T}} v = c) = \mathbb{P}(X_i \in \partial G) = P(\partial G) = 0,$$

by our assumptions on P, since ∂G is a hyperplane that does not intersect V. Consequently, $\mathbb{P}(Y_i < c) = 1$. Also, as before, $\mathbb{E}|Y_i| < \infty$. Thus, by Proposition S2.2, there exists $b < c$ such that $\mathcal{I}_\beta(Y_{1:n}, (-\infty, b]) \xrightarrow{\text{a.s.}} 1$. Since $c > 0$, we may choose this b to be positive (as before). Letting $H = \{x \in \mathbb{R}^k : x^{\mathsf{T}} v \le b\}$, we have $\mathcal{I}_\beta(X_{1:n}, H) \xrightarrow{\text{a.s.}} 1$. Also, $0 \in H^\circ$ since $b > 0$.

Now, we must show that R_u intersects $V \cap \partial H$. First, since $z \in R_u$ means $z = au$ for some $a > 0$, and since $z \in \partial G$ means $z^{\mathsf{T}} v = c > 0$, we find that $u^{\mathsf{T}} v > 0$ and $z = cu/u^{\mathsf{T}} v$. Therefore, letting $y = bu/u^{\mathsf{T}} v$, we have $y \in R_u \cap V \cap \partial H$, since

(i) $b/u^{\mathsf{T}} v > 0$, and thus $y \in R_u$,
(ii) $y^{\mathsf{T}} v = b$, and thus $y \in \partial H$,
(iii) $0 < b/u^{\mathsf{T}} v < c/u^{\mathsf{T}} v$, and thus y is a (strict) convex combination of $0 \in V$ and $z \in \bar{V}$, hence $y \in V$ (Rockafellar, 1970, 6.1). □

To prove Proposition S2.5, we need the following geometrically intuitive facts.

PROPOSITION S2.4. *Let $V \subset \mathbb{R}^k$ be open and convex, with $0 \in V$. Let H be a closed halfspace such that $0 \in H^\circ$. Let $T = \{x/|x| : x \in V \cap \partial H\}$. Then*

(1) *T is open in $\mathcal{S}$,*
(2) *$T = \{u \in \mathcal{S} : R_u$ intersects $V \cap \partial H\}$, and*
(3) *if $x \in H$, $x \ne 0$, and $x/|x| \in T$, then $x \in V$.*

PROOF. Write $H = \{x \in \mathbb{R}^k : x^{\mathsf{T}} v \le b\}$, with $v \in \mathcal{S}$, $b > 0$. Let $\mathcal{S}_+ = \{u \in \mathcal{S} : u^{\mathsf{T}} v > 0\}$. (1) Define $f : \partial H \to \mathcal{S}_+$ by $f(x) = x/|x|$, noting that $0 \notin \partial H$. It is easy to see that f is a homeomorphism. Since V is open in $\mathbb{R}^k$, then $V \cap \partial H$ is open in ∂H. Hence, $T = f(V \cap \partial H)$ is open in $\mathcal{S}_+$, and since $\mathcal{S}_+$ is open in $\mathcal{S}$, then T is also open in $\mathcal{S}$. Items (2) and (3) are easily checked. □

PROPOSITION S2.5. *Let $V \subset \mathbb{R}^k$ be open and convex, with $0 \in V$. If $(H_u : u \in \mathcal{S})$ is a collection of closed halfspaces such that for all $u \in \mathcal{S}$,*

(1) *$0 \in H_u^\circ$ and*
(2) *R_u intersects $V \cap \partial H_u$,*

then there exist $u_1, \ldots, u_r \in \mathcal{S}$ (for some $r > 0$) such that the set $U = \bigcap_{i=1}^r H_{u_i}$ is compact and $U \subset V$.

PROOF. For $u \in \mathcal{S}$, define $T_u = \{x/|x| : x \in V \cap \partial H_u\}$. By part (1) of Proposition S2.4, T_u is open in $\mathcal{S}$, and by part (2), $u \in T_u$, since R_u intersects $V \cap \partial H_u$. Thus, $(T_u : u \in \mathcal{S})$ is an open cover of $\mathcal{S}$. Since $\mathcal{S}$ is compact, there is a finite subcover: there exist $u_1, \ldots, u_r \in \mathcal{S}$ (for some $r > 0$) such that $\bigcup_{i=1}^r T_{u_i} \supset \mathcal{S}$, and in fact, $\bigcup_{i=1}^r T_{u_i} = \mathcal{S}$. Let $U = \bigcap_{i=1}^r H_{u_i}$. Then U is closed and convex (as an intersection of closed, convex sets). Further, $U \subset V$ since for any $x \in U$, if $x = 0$ then $x \in V$ by assumption, while if $x \neq 0$ then $x/|x| \in T_{u_i}$ for some $i \in \{1, \ldots, r\}$ and $x \in U \subset H_{u_i}$, so $x \in V$ by Proposition S2.4(3).

In order to show that U is compact, we just need to show it is bounded, since we already know it is closed. Suppose not, and let $x_1, x_2, \ldots \in U \smallsetminus \{0\}$ such that $|x_n| \to \infty$ as $n \to \infty$. Let $v_n = x_n/|x_n|$. Since $\mathcal{S}$ is compact, then (v_n) has a convergent subsequence such that $v_{n_i} \to u$ for some $u \in \mathcal{S}$. Then for any $a > 0$, we have $av_{n_i} \in U$ for all i sufficiently large (since av_{n_i} is a convex combination of $0 \in U$ and $|x_{n_i}|v_{n_i} = x_{n_i} \in U$ whenever $|x_{n_i}| \geq a$). Since $av_{n_i} \to au$, and U is closed, then $au \in U$. Thus, $au \in U$ for all $a > 0$, i.e. $R_u \subset U$. But $u \in T_{u_j}$ for some $j \in \{1, \ldots, r\}$, so R_u intersects ∂H_{u_j} (by Proposition S2.4(2)), and thus $au \notin H_{u_j} \supset U$ for all $a > 0$ sufficiently large. This is a contradiction. Therefore, U is bounded. $\square$

References.

HOFFMANN-JØRGENSEN, J. (1994). *Probability with a view toward statistics* **2**. Chapman & Hall.

KRANTZ, S. G. (1992). *Function theory of several complex variables*. AMS Chelsea Publishing, Providence.

MILLER, J. W. and HARRISON, M. T. (2013). Inconsistency of Pitman–Yor process mixtures for the number of components.

ROCKAFELLAR, R. T. (1970). *Convex Analysis*. Princeton University Press.

DIVISION OF APPLIED MATHEMATICS
BROWN UNIVERSITY
PROVIDENCE, RI 02912
E-MAIL: Jeffrey_Miller@Brown.edu
Matthew_Harrison@Brown.edu