A Sufficient Statistics Construction of Exponential Family Lévy Measure Densities for Nonparametric Conjugate Models: Supplementary Materail

Proof of Lemma 1: The result in Lemma 1 follows from the definition of the functional derivative and a modification of the proof in the case where the derivative is a standard partial derivative. For the sake of completeness we sketch the argument. A standard reference for the definitions that follow is [1].

Given a space $\mathcal{B}$ of functions, for example the space of piecewise continuous functions on $(0, \infty)$, and a functional $F : \mathcal{B} \longrightarrow \mathbb{R}$, the functional derivative of F with respect to $a(x) \in \mathcal{B}$ is defined as

$$\frac{\partial F[a(x)]}{\partial a(x)} = \lim_{\varepsilon \to 0} \frac{F[a(x) + \varepsilon \delta(x)] - F[a(x)]}{\varepsilon},$$

where $\delta(x)$ is an element of a class of test functions, usually taken to be a class of indicator functions or the class of bump functions on the domain of $\mathcal{B}$. Now, to prove the result in 1 from Lemma 1, it suffices to prove the result for the exponential family corresponding to the first component of $\eta(z)$, i.e. for $\eta_1(z)$. Thus, define the function $\psi(\eta_1(z))$ by

$$\psi(\eta_1(z)) = \int \exp(\eta_1(z) T_1(x)) \mu(dx).$$

Now assume $\eta_1(z)$ is such that $\psi(\eta_1(z))$ is finite. In addition assume $\exists\, \zeta > 0$ such that $\forall\, g(z) \in \mathcal{B}$ if $|\eta_1(z) - g(z)| < \zeta$ uniformly, then $\psi(g(z))$ exists and is finite. Let $0 < \delta(z) \leq 1$ and $\varepsilon_0 > 0$ be a pair such that $|\eta_1(z) - (\eta_1(z) + \varepsilon_0 \delta(z))| < \zeta$ uniformly. Define $\xi(z) = \eta_1(z) + \varepsilon_0 \delta(z)$ and note that

$$\frac{\psi(\xi(z)) - \psi(\eta_1(z))}{\varepsilon_0} = \int \frac{e^{\xi(z) T_1(x)} - e^{\eta_1(z) T_1(x)}}{\varepsilon_0} \mu(dx)$$

$$= \int e^{(\eta_1(z) T_1(x))} \frac{e^{(\xi(z) - \eta_1(z)) T_1(x)} - 1}{\varepsilon_0} \mu(dx) = \int e^{(\eta_1(z) T_1(x))} \frac{e^{\varepsilon_0 \delta(z) T_1(x)} - 1}{\varepsilon_0} \mu(dx)$$

$$\leq \int e^{(\eta_1(z) T_1(x))} \frac{e^{\zeta(\delta(z))|T_1(x)|}}{\zeta} \mu(dx)$$

$$\leq \int \frac{1}{\zeta} \Big| e^{(\eta_1(z) + \zeta \delta(z))|T_1(x)|} + e^{(\eta_1(z) - \zeta \delta(z))|T_1(x)|} \Big| \mu(dx).$$

Since the above integral is finite, it follows from the Lebesgue dominated convergence theorem that

$$\lim_{\varepsilon \to 0} \frac{\psi(\xi(z)) - \psi(\eta_1(z))}{\varepsilon_0} = \int \frac{\partial}{\partial(\eta_1(z))} \Big[exp(\eta_1(z) T_1(x)) \Big] \mu(dx).$$

Extension of the above to higher order functional derivatives proceeds by a standard induction argument. Finally, part 2. of Lemma 1 follows as both the usual chain rule and multiplication by a constant rule hold for functional derivatives.

Construction and Proof of Theorem 1: For every pair n, i define $\eta_{n,i} = \eta(\frac{i-\frac{1}{2}}{n})$, $A_{0,n,i} = A_0(\frac{i-1}{n}, \frac{i}{n}]$, and $T_{k,n,i} = A_{0,n,i} T_k(s)$, where the $T_{k,n,i}$ are independent random variables with $T_k(s)$ distributed according to a sufficient statistic from an exponential family with density

$$h(s) \exp\left\{ \Big\langle \eta_{n,i}, T(s) \Big\rangle - A(\eta_{n,i}) \right\}.$$

We use the notation $A_0(\frac{i-1}{n}, \frac{i}{n}]$ as shorthand for $A_0(\frac{i}{n}) - A_0(\frac{i-1}{n}+)$. Next define

$$T_{k,n}(0) = 0 \text{ and } T_{k,n}(t) = \sum_{\frac{i}{n} \le t} T_{k,n,i}(t) \text{ for } t \ge 0.$$

Note that with this definition, $T_{k,n}$ has independent increments. Next, we consider $T_k(s)$ from the exponential family $p_\theta(x|\eta(z))$ and perform a Taylor expansion on $e^{-\theta T_k(s)}$ yielding

$$\begin{aligned}
-\int (1 - e^{-\theta T_k(s)}) dL_t(s) &= -\int \left(1 - \left(\sum_{m=0}^{\infty} \frac{(-1)^m \theta^m T_k^m(s)}{m!} \right) \right) dL_t(s) \\
&= \sum_{m=1}^{\infty} \frac{(-1)^m \theta^m}{m!} \int T_k^m(s) dL_t(s) \\
&= \sum_{m=1}^{\infty} \frac{(-1)^m \theta^m}{m!} \int T_k^m(s) \left\{ \int_0^t e^{\langle \eta(z), T(s) \rangle - A(\eta(z))} dA_0(z) \right\} ds \\
&= \sum_{m=1}^{\infty} \frac{(-1)^m \theta^m}{m!} \int_0^t \left\{ \int T_k^m(s) e^{\langle \eta(z), T(s) \rangle - A(\eta(z))} ds \right\} dA_0(z) \\
&= \sum_{m=1}^{\infty} \frac{(-1)^m \theta^m}{m!} \int_0^t e^{-A(\eta(z))} \frac{\partial^m \left[e^{A(\eta(z))} \right]}{\partial \eta_k^m} dA_0(z).
\end{aligned}$$

In the above, the last equality holds by Lemma 1.

Next, in order to compute $\mathbb{E}[e^{-\theta T_{k,n}}]$ we must first compute $\mathbb{E}[T^m_{k,n,i}]$, where the expectation is with respect to the density of the random variable $T_{k,n,i}$. To that end we compute the density of $T_{k,n,i}$.

Since s is distributed according to the exponential family

$$h(x)e^{\eta(z),T(s)-A(\eta(z))},$$

as $T_k(s)$ satisfies the conditions of the theorem, by setting $u = T_k(s)$, and $v = A_{0,n,i}T_k(s)$, a simple calculation shows that $T_{k,n,i}$ has a density of the form

$$A^{-1}_{0,n,i}\left(\frac{dT_k^{-1}}{du}\Big|_{vA^{-1}_{0,n,i}}\right)h\Big(T_k^{-1}\Big(\frac{v}{A_{0,n,i}}\Big)\Big)\exp\big(\langle\eta_{n,i}, vA^{-1}_{0,n,i}\rangle - A(\eta_{n,i})\big).$$

As this integrates to 1 with respect to v, we have

$$\int\left\{A^{-1}_{0,n,i}\left(\frac{dT_k^{-1}}{du}\Big|_{vA^{-1}_{0,n,i}}\right)h\Big(T_k^{-1}\Big(\frac{v}{A_{0,n,i}}\Big)\Big)\exp\big(\langle\eta_{n,i}, vA^{-1}_{0,n,i}\rangle\big)\right\}dv = e^{A(\eta_{n,i})},$$

from which it follows that

$$\int\left\{A^{-1}_{0,n,i}\left(\frac{dT_k^{-1}}{du}\Big|_{vA^{-1}_{0,n,i}}\right)h\Big(T_k^{-1}\Big(\frac{v}{A_{0,n,i}}\Big)\Big)\exp\big(\langle\eta_{n,i}A^{-1}_{0,n,i}, v\rangle\big)\right\}dv = e^{A(\eta_{n,i})}.$$

Therefore, taking functional derivatives of both sides with respect to $\eta_{k,n,i}A^{-1}_{0,n,i}$, by Lemma 1 we have

$$\int\left\{T^m_{k,n,i}A^{-1}_{0,n,i}\left(\frac{dT_k^{-1}}{du}\Big|_{vA^{-1}_{0,n,i}}\right)h\Big(T_k^{-1}\Big(\frac{v}{A_{0,n,i}}\Big)\Big)\exp\big(\langle\eta_{n,i}A^{-1}_{0,n,i}, v\rangle\big)\right\}dv$$

$$= \frac{\partial^m}{\partial(\eta_{k,n,i}A^{-1}_{0,n,i})^m}\Big[e^{A(\eta_{n,i})}\Big] = A_{0,n,i}\frac{\partial^m}{\partial(\eta_{k,n,i})^m}\Big[e^{A(\eta_{n,i})}\Big]. \tag{1}$$

Multiplying both sides of (1) by $e^{-A(\eta_{n,i})}$, we conclude that

$$\mathbb{E}[T^m_{k,n,i}] = A_{0,n,i}\left(e^{-A(\eta_{n,i})}\frac{\partial^m}{\partial(\eta_{k,n,i})^m}\Big[e^{A(\eta_{n,i})}\Big]\right).$$

Thus, as $n \longrightarrow +\infty$

$$\mathbb{E}[T_{k.n}(t)] = \sum_{\frac{i}{n}\le t}\mathbb{E}[T_{k,n,i}(t)] =$$

$$\sum_{\frac{i}{n}\le t}A_{0,n,i}\left(e^{-A(\eta_{n,i})}\frac{\partial\big[e^{A(\eta_{n,i})}\big]}{\partial\eta_{k,n,i}}\right) \longrightarrow \int_0^t e^{-A(\eta(z))}\frac{\partial\big[e^{A(\eta(z))}\big]}{\partial\eta_k(z)}dA_0(z). \tag{2}$$

Now we consider the quantity $\mathbb{E}[e^{-\theta T_{k,n,i}}]$. Denoting the density of $T_{k,n,i}$ by $f_{k,n,i}(v)$ and performing a Taylor expansion on $e^{-\theta T_{k,n,i}}$ we have

$$\mathbb{E}[e^{-\theta T_{k,n,i}}] = \int e^{-\theta T_{k,n,i}(s)} f_{k,n,i}(v)dv = \sum_{m=0}^{\infty} \frac{(-1)^m \theta^m}{m!} \int T_{k,n,i}^m(s) f_{k,n,i}(v)dv$$

$$= 1 + \sum_{m=1}^{\infty} \frac{(-1)^m \theta^m}{m!} \mathbb{E}[T_{k,n,i}^m]$$

$$= 1 + \sum_{m=1}^{\infty} \frac{(-1)^m \theta^m}{m!} \left\{ A_{0,n,i} \left(e^{-A(\eta_{n,i})} \frac{\partial^m}{\partial(\eta_{k,n,i})^m} \left[e^{A(\eta_{n,i})} \right] \right) \right\},$$

where, similar to (2), as $n \longrightarrow +\infty$ we have

$$\sum_{\frac{i}{n} \leq t} A_{0,n,i} \left(e^{-A(\eta_{n,i})} \frac{\partial^m \left[e^{A(\eta_{n,i})} \right]}{\partial(\eta_{k,n,i})^m} \right) \longrightarrow \int_0^t e^{-A(\eta(z))} \frac{\partial^m \left[e^{A(\eta(z))} \right]}{\partial(\eta_k(z))^m} dA_0(z). \quad (3)$$

Defining

$$z_{n,i} = \sum_{m=1}^{\infty} \frac{(-1)^m \theta^m}{m!} \left\{ A_{0,n,i} \left(e^{-A(\eta_{n,i})} \frac{\partial^m}{\partial(\eta_{k,n,i})^m} \left[e^{A(\eta_{n,i})} \right] \right) \right\}$$

from (3) we conclude

$$\sum_{m=1}^{\infty} \frac{(-1)^m \theta^m}{m!} \sum_{\frac{i}{n} \leq t} \left\{ A_{0,n,i} \left(e^{-A(\eta_{n,i})} \frac{\partial^m}{\partial(\eta_{k,n,i})^m} \left[e^{A(\eta_{n,i})} \right] \right) \right\} =$$

$$\sum_{\frac{i}{n} \leq t} z_{n,i} \longrightarrow -\int (1 - e^{-\theta T_k(s)}) dL_t(s).$$

As

$$\mathbb{E}[e^{-\theta T_{k,n}}] = \mathbb{E}\left[\prod_{\frac{i}{n} \leq t} \exp(-\theta T_{k,n,i}) \right] = \prod_{\frac{i}{n} \leq t} (1 + z_{n,i}),$$

we may now invoke a Lemma A.1 from Appendix 1 of [2]. We state the result here for the sake of completeness:

Lemma 0.1 *Let $z_{n,i}$ be real numbers, for $n \geq 1$ and $i \geq 1$. Assume that, as $n \to \infty$, (i) $\sum_{a < \frac{i}{n} \leq b} z_{n.i} \longrightarrow z$, (ii) $\max_{a < \frac{i}{n} \leq b} |z_{n,i}| \longrightarrow 0$, (iii) $\limsup \sum_{a < \frac{i}{n} \leq b} z_{n.i} \leq M < +\infty$. Then $\prod_{a < \frac{i}{n} \leq b} (1 + z_{n.i}) \longrightarrow e^z$.*

From the above lemma and the preceding calculations, we conclude that

$$\mathbb{E}[e^{-\theta T_{k,n}(s)}] \longrightarrow \exp\left\{ -\int (1 - e^{\theta T_k(s)}) dL_t(s) \right\}.$$

As in [2] an analogous argument shows that the finite dimensional distributions of $\{T_{k,n}(s)\}$ converge properly as well. The fact that, for all $R > 0$, the sequence $\{T_{k,n}(s)\}_{n=1}^{\infty}$ is *tight* in the space $D([0, R])$ of all functions that are right continuous with left hand limits in the Skorohod topology follows from (2) and the proof of 15.6 in [3]. Hence, as in [2], the sequence $\{T_{k,n}(s)\}$ converges to a random element of $D([0, R])$ for every $R > 0$, and the process so defined may be taken to be the $[0, R]$ restriction of a Lévy process T_k on $[0, \infty)$ whose Lévy representation is given in Theorem 1. This completes the proof of the theorem.

References

[1] I. M. Gelfand and S. V. Fomin. *Calculus of Variations.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963.

[2] Nils Lid Hjort. Nonparametric Bayes estimators based on beta processes in models for life history data. *The Annals of Statistics*, 18:1259 – 1294, 1990.

[3] Patrick Billingsley. *Convergence of Probability Measures.* John Wiley and Sons, Inc., New York, 1968.