Non-triviality condition (shortened version)

Suzie Brown

April 28, 2020

Multinomial resampling

Lemma 1. *For all $N \geq 2$, for all t,*

$$\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathbf{w} = \left(\frac{1}{N}, \dots, \frac{1}{N}\right)\right] = 1 - \frac{N!}{N^N}.$$

Proof. Fix arbitrary t and $N \geq 2$. Since $2/(N)_2 > 2/(N^2)$ is the smallest possible non-zero value for $c_N(t)$,

$$\begin{aligned}\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \mid \mathbf{w} = (1/N, \dots, 1/N)\right] &= 1 - \mathbb{P}[c_N(t) = 0 \mid \mathbf{w} = (1/N, \dots, 1/N)] \\ &= 1 - \mathbb{P}[\nu_t^{(1:N)} = (1, \dots, 1) \mid \mathbf{w} = (1/N, \dots, 1/N)].\end{aligned}$$

Conditional on the weights, $\nu_t^{(1:N)} \sim \text{Multinomial}(N, (1/N, \dots, 1/N))$, so the probability of interest is

$$\mathbb{P}[\nu_t^{(1:N)} = (1, \dots, 1) \mid \mathbf{w} = (1/N, \dots, 1/N)] = N! \prod_{i=1}^{N} \frac{1}{N} = \frac{N!}{N^N}.$$

□

Lemma 2. *In the neutral case (i.e. when all weights are equal at every time step) with multinomial resampling, there exists N_0 such that for all $N > N_0$, for all finite t, $\mathbb{P}[\tau_N(t) = \infty] = 0$.*

Proof. Let us rewrite the event of interest in a different way.

$$\begin{aligned}\mathbb{P}[\tau_N(t) = \infty] = 0 &\Leftrightarrow \mathbb{P}[\tau_N(t) < \infty] = 1 \\ &\Leftrightarrow \mathbb{P}\left[\min\left\{s > 1 : \sum_{r=1}^{s} c_N(r) < t\right\} < \infty\right] = 1 \\ &\Leftrightarrow \mathbb{P}\left[\exists s < \infty : \sum_{r=1}^{s} c_N(r) < t\right] = 1\end{aligned}$$

It is sufficient to show that, for all $N > N_0$, $c_N(r)$ is bounded away from zero infinitely often in r. We consider the sequence of events $E_r := \{c_N(r) > 2/N^2\}$ for $r \in \mathbb{N}$. In the neutral case, the resampled family sizes at each generation are independent, hence the events E_r are independent. By the second Borel-Cantelli lemma, E_r occurs infinitely often if $\sum_{r=1}^{\infty} \mathbb{P}(E_r) = \infty$. An expression for $\mathbb{P}(E_r)$ is given in Lemma 1. For any fixed $N \geq 2$, the probability is strictly positive and constant in r, so the Borel-Cantelli condition is satisfied, thus we conclude that E_r occurs infinitely often. Hence, taking $N_0 = 1$, we have that $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all $N > N_0$ and all finite t, as required. □

Lemma 3. *For all $N \geq 2$, for all t, for any weight vector $(w_1, \dots, w_N) \in \mathcal{S}_{N-1}$,*

$$\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathbf{w} = (w_1, \dots, w_N)\right] \geq \mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathbf{w} = \left(\frac{1}{N}, \dots, \frac{1}{N}\right)\right].$$

That is, the probability of having at least one merger is minimised by the vector of equal weights.

Proof. Fix arbitrary t and $N \geq 2$. Recall that

$$1 - \mathbb{P}\left[c_N(t) > \frac{2}{N^2} \mid \mathbf{w} = (w_1, \dots, w_N)\right] = N! \prod_{i=1}^{N} w_i. \tag{1}$$

We will show that the global maximum of this function on the simplex $\mathcal{S}_{N-1}$ is attained at $\mathbf{w} = (1/N, \dots, 1/N)$. This weight vector will therefore minimise the probability of the complementary event, implying the desired result.

First, since we are working on the simplex, we encode the constraint $\sum w_j = 1$ by rewriting the function to optimise as

$$f(\mathbf{w}) := \prod_{i=1}^{N} w_i = \left(1 - \sum_{j=1}^{N-1} w_j\right) \prod_{i=1}^{N-1} w_i$$

where we have also dropped the constant positive factor $N!$. Now, for every $k \in \{1, \dots, N-1\}$, we solve

$$\frac{\partial f(\mathbf{w})}{\partial w_k} = \left(1 - w_k - \sum_{j=1}^{N-1} w_j\right) \prod_{i \neq k}^{N-1} w_i = 0.$$

The product over $i \neq k$ is constant for each k, so this amounts to solving

$$w_k = 1 - \sum_{j=1}^{N-1} w_j = w_N$$

for all k. The unique solution is $w_1 = w_2 = \cdots = w_N = 1/N$.

To verify that this critical point is a maximum, we can evaluate the Hessian H:

$$H_{kl}(\mathbf{w}) = \begin{cases} -2 \prod_{i \neq k}^{N-1} w_i & k = l \\ \left(1 - w_k - w_l - \sum_{j=1}^{N-1} w_j\right) \prod_{i \neq k,l}^{N-1} w_i & k \neq l \end{cases}$$

$$H_{kl}((1/N, \dots, 1/N)) = \begin{cases} -2 \left(\frac{1}{N}\right)^{N-2} & k = l \\ -\left(\frac{1}{N}\right)^{N-2} & k \neq l \end{cases}$$

and show that H is negative semi-definite: for any $\mathbf{x} \in \mathbb{R}^{N-1}$,

$$\begin{aligned} \mathbf{x}^T H \mathbf{x} &= \sum_{k=1}^{N-1} \left[-2 \left(\frac{1}{N}\right)^{N-2} x_k^2 - \sum_{l \neq k}^{N-1} \left(\frac{1}{N}\right)^{N-2} x_k x_l \right] = \left(\frac{1}{N}\right)^{N-2} \left[-\sum_{k=1}^{N-1} 2x_k^2 - \sum_{k=1}^{N-1} \sum_{l \neq k}^{N-1} x_k x_l \right] \\ &= \left(\frac{1}{N}\right)^{N-2} \left[-\sum_{k=1}^{N-1} x_k^2 - \sum_{k=1}^{N-1} \sum_{l=1}^{N-1} x_k x_l \right] = \left(\frac{1}{N}\right)^{N-2} \left[-\sum_{k=1}^{N-1} x_k^2 - \left(\sum_{k=1}^{N-1} x_k\right)^2 \right] \leq 0. \end{aligned}$$

□

Theorem 1. *With multinomial resampling, conditional on any sequence of weight vectors* $\mathbf{w}_r^{(1:N)} \in \mathcal{S}_{N-1}; r \in \mathbb{N}$, *there exists* N_0 *such that for all* $N > N_0$, *for all finite* t, $\mathbb{P}[\tau_N(t) = \infty] = 0$.

Proof. As in Lemma 2, denote the sequence of events $E_r := \{c_N(r) > 2/N^2\}$ for $r \in \mathbb{N}$. We know from Lemma 2 that, in the neutral case, E_r occurs infinitely often. Lemma 3 tells us that $\mathbb{P}[E_r \mid \mathbf{w} = (w_1, \dots, w_N)] \geq \mathbb{P}[E_r \mid \mathbf{w} = (1/N, \dots, 1/N)]$ for all r. Therefore, by a coupling argument, we conclude that E_r occurs infinitely often in the non-neutral case as well. □

Conditional SMC with multinomial resampling

Define $\mathbf{w}^* := \frac{1}{N-1} [(1, \dots, 1) - \mathbf{e}_{i^*}]$, where i^* is the immortal index at generation t, and $\mathbf{e}_i$ denotes a 1-hot vector.

Lemma 4. *For all* $N \geq 4$, *for all* t,

$$\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathbf{w} = \mathbf{w}^*\right] = 1 - \frac{(N-1)!}{(N-1)^{N-1}}$$

Proof. Under $\mathbf{w}^*$, the immortal parent has zero weight and is therefore assigned exactly one offspring (the immortal particle). The remaining $N-1$ offspring are assigned to the remaining $N-1$ parents according to a Multinomial distribution with equal weights. We therefore have

$$\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \mid \mathbf{w} = \mathbf{w}^*\right] = 1 - \mathbb{P}[\nu_t^{(1:N)} = (1, \dots, 1) \mid \mathbf{w} = \mathbf{w}^*] = 1 - (N-1)! \prod_{i \neq i^*}^{N} \frac{1}{N-1} = 1 - \frac{(N-1)!}{(N-1)^{N-1}}.$$

□

Lemma 5. *In conditional SMC with multinomial resampling, in the optimal case where the weight vector is equal to $\mathbf{w}^*$ at every time step, there exists N_0 such that for all $N > N_0$, for all finite t, $\mathbb{P}[\tau_N(t) = \infty] = 0$.*

Proof. The proof is exactly the same as for Theorem 2; Lemma 4 provides the expression for $P(E_r)$ which is strictly positive and constant in r. □

Lemma 6. *For all $N \geq 2$, for all t, for any weight vector $(w_1, \dots, w_N) \in \mathcal{S}_{N-1}$,*

$$\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathbf{w} = (w_1, \dots, w_N)\right] \geq \mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathbf{w} = \mathbf{w}^*\right].$$

Proof. Recall that

$$1 - \mathbb{P}\left[c_N(t) > \frac{2}{N^2} \mid \mathbf{w} = (w_1, \dots, w_N)\right] = (N-1)! \prod_{i \neq i^*}^{N} w_i.$$

Ignoring the immortal particles, this is equivalent to multinomial resampling in the standard case (1), only with $N-1$ particles rather than N. As we saw in Lemma 3, this function is maximised at the vector of equal weights, in this case $\mathbf{w}_{-i^*} = \frac{1}{N-1}(1, \dots, 1)$. This leaves zero weight for the immortal particle, so overall the maximum is attained at $\mathbf{w}^* = \frac{1}{N-1}\left[(1, \dots, 1) - \mathbf{e}_{i^*}\right]$ as required. □

Theorem 2. *In conditional SMC with multinomial resampling, conditional on any sequence of weight vectors $\mathbf{w}_r^{(1:N)} \in \mathcal{S}_{N-1}; r \in \mathbb{N}$, there exists N_0 such that for all $N > N_0$, for all finite t, $\mathbb{P}[\tau_N(t) = \infty] = 0$.*

Proof. As before, consider the sequence of events $E_r := \{c_N(r) > 2/N^2\}$ for $r \in \mathbb{N}$. We know from the argument behind Lemma 5 (which is completely analogous to Lemma 2) that, in the case $\mathbf{w} = \mathbf{w}^*$, E_r occurs infinitely often. Lemma 6 tells us that $\mathbb{P}[E_r \mid \mathbf{w} = (w_1, \dots, w_N)] \geq \mathbb{P}[E_r \mid \mathbf{w} = \mathbf{w}^*]$ for all r. Therefore, by a coupling argument, we conclude that E_r occurs infinitely often in the general case as well. □