Non-triviality condition

Suzie Brown

April 27, 2020

Multinomial resampling: neutral case

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

$$\mathbb{E}\left[c_N(t) \middle| \mathbf{w} = \left(\frac{1}{N}, \dots, \frac{1}{N}\right)\right] = \frac{1}{N}.$$

Proof.

$$\begin{aligned}
\mathbb{E}\left[c_N(t) | \mathbf{w} = (1/N, \dots, 1/N)\right] &= \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}\left[(\nu_t^{(i)})_2 \mid \mathbf{w} = (1/N, \dots, 1/N)\right] \\
&= \frac{1}{(N)_2} \sum_{i=1}^{N} (N)_2 \left(\frac{1}{N}\right)^2 = \sum_{i=1}^{N} \frac{1}{N^2} = \frac{1}{N}
\end{aligned}$$

□

Lemma 2. *For all $N \geq 4$, for all t,*

$$\mathbb{E}\left[(c_N(t))^2 \middle| \mathbf{w} = \left(\frac{1}{N}, \dots, \frac{1}{N}\right)\right] = \frac{N+2}{N^3}.$$

Proof.

$$\begin{aligned}
&\mathbb{E}\left[(c_N(t))^2 \middle| \mathbf{w} = (1/N, \dots, 1/N)\right] = \frac{1}{(N)_2^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbb{E}\left[(\nu_t^{(i)})_2 (\nu_t^{(j)})_2 \middle| \mathbf{w} = (1/N, \dots, 1/N)\right] \\
&= \frac{1}{(N)_2^2} \left\{ \sum_{i=1}^{N} \sum_{j \neq i}^{N} \mathbb{E}\left[(\nu_t^{(i)})_2 (\nu_t^{(j)})_2 \middle| \mathbf{w} = (1/N, \dots, 1/N)\right] + \sum_{i=1}^{N} \mathbb{E}\left[(\nu_t^{(i)})_2^2 \middle| \mathbf{w} = (1/N, \dots, 1/N)\right] \right\} \\
&= \frac{1}{(N)_2^2} \left\{ \sum_{i=1}^{N} \sum_{j \neq i}^{N} \mathbb{E}\left[(\nu_t^{(i)})_2 (\nu_t^{(j)})_2 \middle| \mathbf{w} = (1/N, \dots, 1/N)\right] + \sum_{i=1}^{N} \mathbb{E}\left[(\nu_t^{(i)})_4 + 4(\nu_t^{(i)})_3 + 2(\nu_t^{(i)})_2 \middle| \mathbf{w} = (1/N, \dots, 1/N)\right] \right\} \\
&= \frac{1}{(N)_2^2} \left\{ \sum_{i=1}^{N} \sum_{j \neq i}^{N} (N)_4 \left(\frac{1}{N}\right)^2 \left(\frac{1}{N}\right)^2 + \sum_{i=1}^{N} \left((N)_4 \left(\frac{1}{N}\right)^4 + 4(N)_3 \left(\frac{1}{N}\right)^3 + 2(N)_2 \left(\frac{1}{N}\right)^2 \right) \right\} \\
&= \frac{1}{(N)_2^2} \left\{ N(N-1)(N)_4 \frac{1}{N^4} + N(N)_4 \frac{1}{N^4} + 4N(N)_3 \frac{1}{N^3} + 2N(N)_2 \frac{1}{N^2} \right\} \\
&= \frac{(N-2)(N-3)}{N^4} + \frac{(N-2)(N-3)}{N^4(N-1)} + \frac{4(N-2)}{N^3(N-1)} + \frac{2}{N^2(N-1)} \\
&= \frac{1}{N^4(N-1)} \left[(N-2)(N-3)(N-1+1) + 4N(N-2) + 2N^2\right] \\
&= \frac{1}{N^3(N-1)} \left[N^2 - 5N + 6 + 4N - 8 + 2N\right] = \frac{N^2 + N - 2}{N^3(N-1)} = \frac{(N+2)(N-1)}{N^3(N-1)} = \frac{N+2}{N^3}.
\end{aligned}$$

□

Lemma 3. *For all $N \geq 4$, for all t,*

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

Proof. We apply the Paley-Zygmund inequality,

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

Setting $\theta = 2/N$ and using Lemmata 1–2,

$$\mathbb{P}\left[c_N(t) > \frac{2}{N^2}\middle|\mathbf{w} = (1/N, \dots, 1/N)\right] \geq \left(1 - \frac{2}{N}\right)^2 \frac{(1/N)^2}{(N+2)/N^3} = \left(1 - \frac{2}{N}\right)^2 \frac{N}{N+2}.$$

□

NB: We actually have an exact expression for the above probability, which is $1 - N!N^{-N}$ (see for example the proof of Lemma 4 below). Perhaps it would be better to use that... although the asymptotics seem a bit more obscure then? This could save a page of workings though so probably a good idea in the end.

Theorem 1. *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 $\sum_{r=1}^{\infty} \mathbb{P}(E_r) = \infty$. A lower bound on $\mathbb{P}(E_r)$ is given in Lemma 3. For any fixed $N \geq 4$, the bound 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 $N_0 = 3$, we have that $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all $N > N_0$ and all finite t, as required. □

Multinomial resampling: non-neutral case

Lemma 4. *For all $N \geq 2$, for all t, for any weight vector $(w_1, \dots, w_N)$,*

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

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

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

$$\begin{aligned}
\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \mid \mathbf{w} = (w_1, \dots, w_N)\right] &= 1 - \mathbb{P}[c_N(t) = 0 \mid \mathbf{w} = (w_1, \dots, w_N)] \\
&= 1 - \mathbb{P}[\nu_t^{(1:N)} = (1, \dots, 1) \mid \mathbf{w} = (w_1, \dots, w_N)].
\end{aligned}$$

Since, conditional on the weights, $\nu_t^{(1:N)} \sim \text{Multinomial}(N, (w_1, \dots, w_N))$, the probability of interest is

$$\mathbb{P}[\nu_t^{(1:N)} = (1, \dots, 1) \mid \mathbf{w} = (w_1, \dots, w_N)] = N! \prod_{i=1}^{N} w_i. \tag{1}$$

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 $\sum w_i = 1$ 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 reduces 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 calculate 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 2. *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 Theorem 1, denote the sequence of events $E_r := \{c_N(r) > 2/N^2\}$ for $r \in \mathbb{N}$. We know from Theorem 1 that, in the neutral case, E_r occurs infinitely often. Lemma 4 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: optimal weights

NB: The exposition below is more explicit than necessary, in order to reduce dependencies between sections. The expectations under CSMC-mn do not really need to be calculated directly, as they are equal to the expectations under standard-mn, where $(N-1)$ replaces N everywhere except in the leading $(N)_2$ factors. It is probably also possible to infer Theorem 3 by a direct modification of Theorem 1, without the need to calculate moments and apply the PZ inequality again.

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

$$\mathbb{E}\left[c_N(t) \mid \mathbf{w} = \mathbf{w}^*\right] = \frac{N-2}{N(N-1)}.$$

Proof. Since the immortal particle has weight zero, the remaining offspring counts are distributed as Multinomial$(N-1, (1/(N-1), \ldots, 1/(N-1)))$. We can apply the usual formula for factorial moments of the Multinomial distribution:

$$\mathbb{E}\left[c_N(t) \mid \mathbf{w} = \mathbf{w}^*\right] = \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}\left[(\nu_t^{(i)})_2 \mid \mathbf{w} = \mathbf{w}^*\right] = \frac{1}{(N)_2} \sum_{i \neq i^*}^{N} (N-1)_2 \left(\frac{1}{N-1}\right)^2 = \frac{N-2}{N(N-1)}.$$

$\square$

Lemma 6. *For all $N \geq 4$, for all t,*

$$\mathbb{E}\left[(c_N(t))^2 \middle| \mathbf{w} = \mathbf{w}^*\right] = \frac{(N+1)(N-2)^2}{N^2(N-1)^3}.$$

Proof.

$$\begin{aligned}
\mathbb{E}\left[(c_N(t))^2 \mid \mathbf{w} = \mathbf{w}^*\right] &= \frac{1}{(N)_2^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbb{E}\left[(\nu_t^{(i)})_2 (\nu_t^{(j)})_2 \middle| \mathbf{w} = \mathbf{w}^*\right] \\
&= \frac{1}{(N)_2^2} \sum_{i \neq i^*}^{N} \sum_{j \neq i^*}^{N} \mathbb{E}\left[(\nu_t^{(i)})_2 (\nu_t^{(j)})_2 \middle| \mathbf{w} = \mathbf{w}^*\right] \\
&= \frac{1}{(N)_2^2} \left\{ \sum_{i \neq i^*}^{N} \sum_{j \neq i, i^*}^{N} \mathbb{E}\left[(\nu_t^{(i)})_2 (\nu_t^{(j)})_2 \middle| \mathbf{w} = \mathbf{w}^*\right] + \sum_{i \neq i^*}^{N} \mathbb{E}\left[(\nu_t^{(i)})_2^2 \middle| \mathbf{w} = \mathbf{w}^*\right] \right\} \\
&= \frac{1}{(N)_2^2} \left\{ \sum_{i \neq i^*}^{N} \sum_{j \neq i, i^*}^{N} \mathbb{E}\left[(\nu_t^{(i)})_2 (\nu_t^{(j)})_2 \middle| \mathbf{w} = \mathbf{w}^*\right] + \sum_{i \neq i^*}^{N} \mathbb{E}\left[(\nu_t^{(i)})_4 + 4(\nu_t^{(i)})_3 + 2(\nu_t^{(i)})_2 \middle| \mathbf{w} = \mathbf{w}^*\right] \right\} \\
&= \frac{1}{(N)_2^2} \left\{ \sum_{i \neq i^*}^{N} \sum_{j \neq i, i^*}^{N} (N-1)_4 \left(\frac{1}{N-1}\right)^4 + \sum_{i \neq i^*}^{N} \left((N-1)_4 \left(\frac{1}{N-1}\right)^4 + 4(N-1)_3 \left(\frac{1}{N-1}\right)^3 + 2(N-1)_2 \left(\frac{1}{N-1}\right)^2 \right) \right\} \\
&= \frac{1}{(N)_2^2} \left\{ \frac{(N-1)_2 (N-1)_4}{(N-1)^4} + \frac{(N-1)(N-1)_4}{(N-1)^4} + \frac{4(N-1)(N-1)_3}{(N-1)^3} + \frac{2(N-1)(N-1)_2}{(N-1)^2} \right\} \\
&= \frac{(N+1)(N-2)^2}{N^2(N-1)^3}.
\end{aligned}$$

$\square$

Lemma 7. *For all $N \geq 4$, for all t,*

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

Proof. We apply the Paley-Zygmund inequality, with $\theta = \frac{2(N-1)}{N(N-2)}$:

$$\begin{aligned}
\mathbb{P}\left[c_N(t) > \theta\, \mathbb{E}[c_N(t) | \mathbf{w} = \mathbf{w}^*] | \mathbf{w} = \mathbf{w}^*\right] &\geq (1-\theta)^2 \frac{\mathbb{E}[c_N(t) | \mathbf{w} = \mathbf{w}^*]^2}{\mathbb{E}[(c_N(t))^2 | \mathbf{w} = \mathbf{w}^*]} \\
&= \left(1 - \frac{2(N-1)}{N(N-2)}\right)^2 \frac{(N-2)^2}{N^2(N-1)^2} \frac{N^2(N-1)^3}{(N+1)(N-2)^2} = \left(1 - \frac{2(N-1)}{N(N-2)}\right)^2 \frac{N-1}{N+1}.
\end{aligned}$$

$\square$

Theorem 3. *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 1; Lemma 7 provides the bound on $P(E_r)$ which is strictly positive and constant in r. $\square$

Conditional SMC with multinomial resampling: general weights

Lemma 8. *For all $N \geq 2$, for all t, for any weight vector $(w_1, \ldots, w_N)$,*

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

Proof.

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

since the immortal particle i^* is automatically assigned one offspring. This is equivalent to the expression we had in the standard case (1), except with $N-1$ particles rather than N. As we saw in Lemma 4, this function is maximised at the vector of equal weights, in this case $\mathbf{w}_{-i^*} = \frac{1}{N-1}(1, \ldots, 1)$. This leaves zero weight for the immortal particle, so overall the maximum is attained at $\mathbf{w}^* = \frac{1}{N-1}\{(1, \ldots, 1) - \mathbf{e}_{i^*}\}$ as required. □

Theorem 4. *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 in Theorem 1, denote the sequence of events $E_r := \{c_N(r) > 2/N^2\}$ for $r \in \mathbb{N}$. We know from the argument behind Theorem 3 (which is completely analogous to Theorem 1) that, in the neutral case, E_r occurs infinitely often. Lemma 8 tells us that $\mathbb{P}[E_r \mid \mathbf{w} = (w_1, \ldots, w_N)] \geq \mathbb{P}[E_r \mid \mathbf{w} = \mathbf{w}^*]$ for all r. Therefore, by a coupling argument, we conclude that E_r occurs infinitely often in the general case as well. □