Comparing expected coalescence rates for multinomial & residual resampling

Suzie Brown

July 25, 2019

Case $N = 2$

Lemma 1. *For all weight vectors* $w_t^{(1:2)}$, $\mathbb{E}[c_2^m(t)|w_t^{(1:2)}] \geq \mathbb{E}[c_2^r(t)|w_t^{(1:2)}]$.

Proof. With only $N = 2$ particles, the coalescence rate becomes

$$\mathbb{E}[c_N(t)|w_t^{(1:2)}] = \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}\left[(v_t^{(i)})_2 | w_t^{(1:2)}\right] = \mathbb{P}[v_t^{(1)} = 0] + \mathbb{P}[v_t^{(1)} = 2].$$

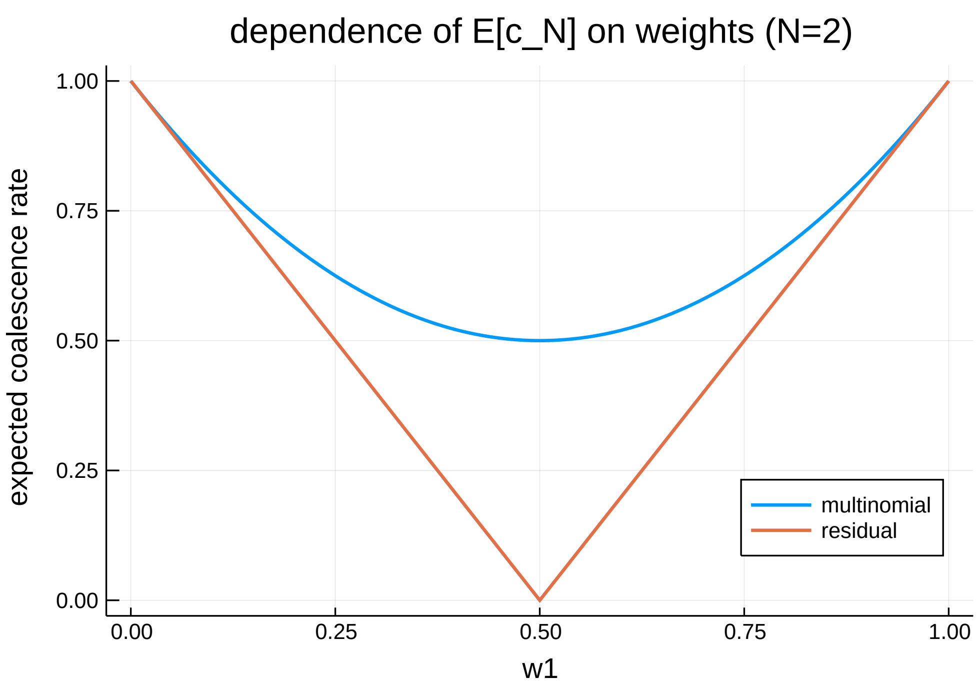
For residual resampling,

$$\mathbb{E}[c_2^r(t)|w_t^{(1:2)}] = \mathbb{I}_{\{w_t^{(1)} \geq 1/2\}}(2w_t^{(1)} - 1) + \mathbb{I}_{\{w_t^{(1)} < 1/2\}}(2w_t^{(2)} - 1)$$

And for multinomial resampling,

$$\begin{aligned}
\mathbb{E}[c_2^m(t)|w_t^{(1:2)}] &= (w_t^{(1)})^2 + (w_t^{(2)})^2 \\
&= \mathbb{I}_{\{w_t^{(1)} \geq 1/2\}}((w_t^{(1)})^2 + (w_t^{(2)})^2) + \mathbb{I}_{\{w_t^{(1)} < 1/2\}}((w_t^{(1)})^2 + (w_t^{(2)})^2) \\
&\geq \mathbb{I}_{\{w_t^{(1)} \geq 1/2\}}(w_t^{(1)})^2 + \mathbb{I}_{\{w_t^{(1)} < 1/2\}}(w_t^{(2)})^2
\end{aligned}$$

Then since $(w_t^{(i)} - 1)^2 = (w_t^{(i)})^2 - 2w_t^{(i)} + 1 \geq 0$, we have that $(w_t^{(i)})^2 \geq 2w_t^{(i)} - 1$ and hence we conclude the proof. $\square$



Case $N = 3$

Given a weight vector $(w_t^{(1)}, w_t^{(2)}, w_t^{(3)})$, let $w_{(1)} \geq w_{(2)} \geq w_{(3)}$ denote the weights sorted from high to low. With $N = 3$ there are many more cases than with $N = 2$, and these are described below, using the sorted weights.

In each case for the conditions on the sorted weights, the possible offspring count vectors (sorted in the same order as the weights) are listed, along with the probability of each (conditional on the given case). Finally, using these outcomes and associated probabilities, the conditional expectation of interest is calculated.

Case	Weights	Offspring counts	Conditional probabilities	$\mathbb{E}[c_2^r(t) \| w_t^{(1:3)}]$
(A)	$w_{(1)} = 1$	$(3,0,0)$	1	1
(B)	$2/3 < w_{(1)} < 1$	$(3,0,0)$ $(2,1,0)$ $(2,0,1)$	$3w_{(1)} - 2$ $3w_{(2)}$ $3w_{(3)}$	$2w_{(1)} - 1$
(C)	$w_{(1)} = 2/3$	$(2,1,0)$ $(2,0,1)$	$3w_{(2)}$ $3w_{(3)}$	$1/3$
(D1)	$1/3 < w_{(1)} < 2/3$ and $1/3 \leq w_{(2)} < 2/3$	$(2,1,0)$ $(1,2,0)$ $(1,1,1)$	$3w_{(1)} - 1$ $3w_{(2)} - 1$ $3w_{(3)}$	$1/3 - w_{(3)}$
(D2)	$1/3 < w_{(1)} < 2/3$ and $w_{(2)} < 1/3$	$(3,0,0)$ $(2,1,0)$ $(2,0,1)$ $(1,2,0)$ $(1,0,2)$ $(1,1,1)$	$(3/2)^2(w_{(1)} - 1/3)^2$ $(3/2)^2 2(w_{(1)} - 1/3)w_{(2)}$ $(3/2)^2 2(w_{(1)} - 1/3)w_{(3)}$ $(3/2)^2 w_{(2)}^2$ $(3/2)^2 w_{(3)}^2$ $(3/2)^2 2w_{(2)}w_{(3)}$	$(1/4)(3w_{(1)} - 1)(w_{(1)} + 1)$
(E)	$w_{(1)} = 1/3$	$(1,1,1)$	1	0

Lemma 2. *For all weight vectors $w_t^{(1:3)}$, $\mathbb{E}[c_3^r(t) | w_t^{(1:3)}] \leq \mathbb{E}[c_3^m(t) | w_t^{(1:3)}]$.*

Proof. We have the following expression in the case of residual resampling:

$$\begin{aligned}\mathbb{E}[c_3^r(t)|w_t^{(1:3)}] = &(2w_{(1)} - 1)\mathbb{I}_{\{2/3 \leq w_{(1)} \leq 1\}} \\ &+ (1/3 - w_{(3)})\mathbb{I}_{\{1/3 \leq w_{(1)} < 2/3\}}\mathbb{I}_{\{1/3 \leq w_{(2)} < 2/3\}} \\ &+ (1/4)(3w_{(1)} - 1)(w_{(1)} + 1)\mathbb{I}_{\{1/3 \leq w_{(1)} < 2/3\}}\mathbb{I}_{\{w_{(2)} < 1/3\}}\end{aligned}$$

compared to the following in the case of multinomial resampling:

$$\begin{aligned}\mathbb{E}[c_3^m(t)|w_t^{(1:3)}] &= w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2 \\ &= (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)\mathbb{I}_{\{2/3 \leq w_{(1)} \leq 1\}} \\ &\quad + (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)\mathbb{I}_{\{1/3 \leq w_{(1)} < 2/3\}}\mathbb{I}_{\{1/3 \leq w_{(2)} < 2/3\}} \\ &\quad + (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)\mathbb{I}_{\{1/3 \leq w_{(1)} < 2/3\}}\mathbb{I}_{\{w_{(2)} < 1/3\}}.\end{aligned}$$

Hence it suffices to show the following:

(i) $(2w_{(1)} - 1) \leq (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)$

(ii) $(1/3 - w_{(3)}) \leq (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)$

(iii) $(1/4)(3w_{(1)} - 1)(w_{(1)} + 1) \leq (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)$

First consider (ii). Since $w_{(3)}$ is defined as the smallest of the three weights, we know that $w_{(3)} \in [0, 1/3]$. Meanwhile, the RHS is the sum of the squared weights, which is always between $1/3$ and 1. Therefore (ii) is true.

Now let us consider (i). We have the identity $(w_{(1)}-1)^2 = w_{(1)}^2 - 2w_{(1)} + 1$, which implies that $w_{(1)}^2 \geq 2_w(1) - 1$. Since $w_{(2)}^2 + w_{(3)}^2 \geq 0$ we can therefore conclude that (i) is true.

Finally consider (iii). In this case it is sufficient to show that $(1/4)(3w_{(1)} - 1)(w_{(1)} + 1) \leq w_{(1)}^2$. Note that

$$
\begin{aligned}
& w_{(1)}^2 - \frac{1}{4}(3w_{(1)} - 1)(w_{(1)} + 1) \\
& = \frac{1}{4}w_{(1)}^2 - \frac{1}{2}w_{(1)} + \frac{1}{4} \\
& = \frac{1}{4}(w_{(1)} - 1)^2 \geq 0.
\end{aligned}
$$

Therefore (iii) is also true, concluding the proof. □

Remark 1. *Examining the table, we can see that the function* $\mathbb{E}[c_3^r(t)|w_t^{(1:3)}]$ *is continuous in* $w_{(1)}$*. However, the plot clearly shows discontinuities. These occur at the boundaries between different orderings when we sort the weights from high to low.*

General N

Lemma 3. *In the case* $w_{(1)} \geq \frac{N-1}{N}$*, we have*

$$\mathbb{E}[c_N^r(t)|w_t^{(1:N)}] = 2w_{(1)} - 1.$$

Proof. In this case, (N-1) offspring are deterministically assigned to particle 1. The one remaining offspring can either be assigned to particle 1 or to some other particle.

- The first option yields offspring vector $(N, 0, 0, \dots)$ and occurs with probability $(w_{(1)} - \frac{N-1}{N})/(1/N) = Nw_{(1)} - (N-1)$. The resulting value of c_N^r is $N(N-1)/(N)_2 = 1$.
- The second option yields an offspring vector that is some permutation of $(N-1, 1, 0, 0, \dots)$, and occurs with probability $1 - Nw_{(1)} - (N-1) = N - Nw_{(1)}$. The resulting value of c_N^r is $(N-1)(N-2)/(N)_2 = (N-2)/N$.

So, under this constraint on the weights, we have the following expectation:

$$\mathbb{E}[c_N^r(t)|w_t^{(1:N)}] = Nw_{(1)} - (N-1) + \frac{N-2}{N}(N - Nw_{(1)}) = 2w_{(1)} - 1.$$

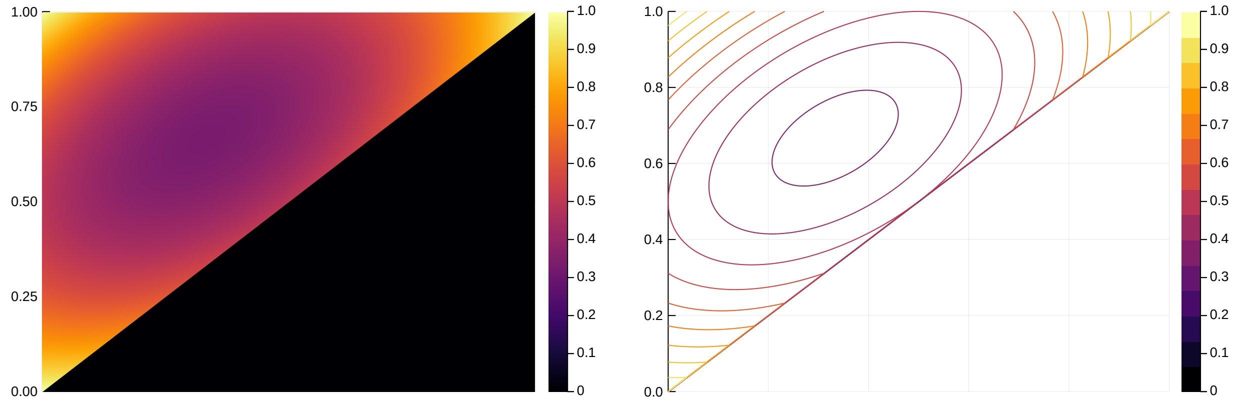
□

Lemma 4. *For* $N \geq 3$*, in the case* $w_{(1)}, w_{(2)}, \dots, w_{(N-1)} \in \left[\frac{1}{N}, \frac{2}{N}\right]$*, we have*

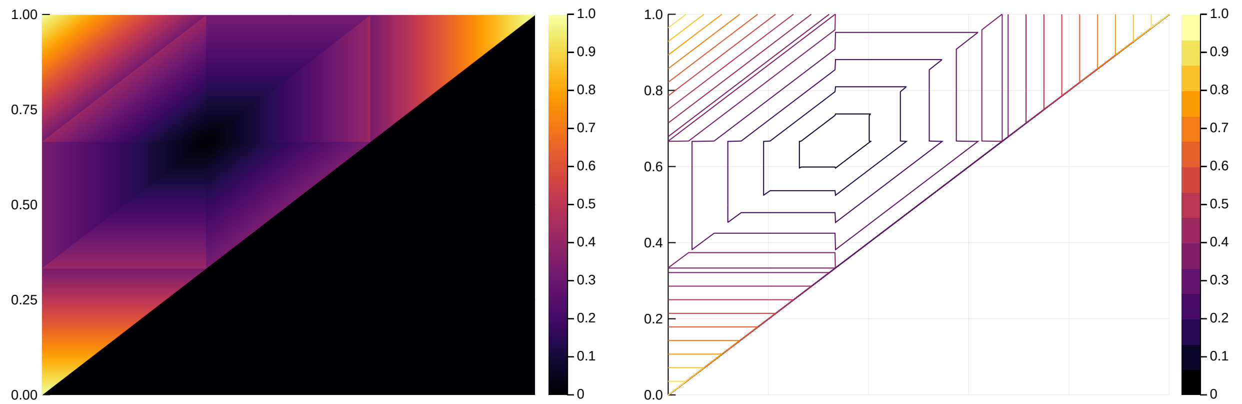
$$\mathbb{E}[c_N^r(t)|w_t^{(1:N)}] = \frac{2}{N-1}(1/N - w_{(N)}).$$

Proof. In this case, one offspring is deterministically assigned to each of the particles $1, 2, \dots, N-1$. The one remaining offspring can be assigned either to particle N or to one of the others.

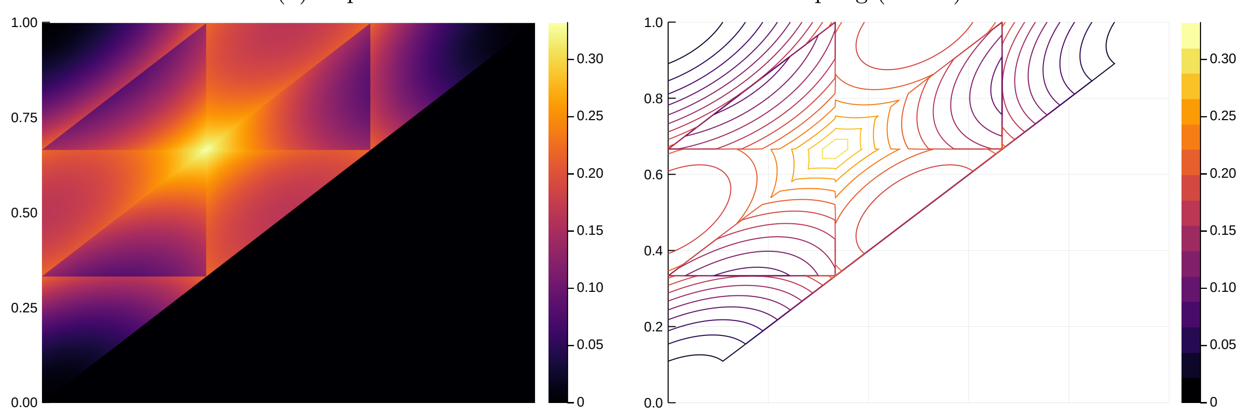
- The first option yields offspring vector $(1, 1, \dots, 1)$ and occurs with probability $w_{(N)}/(1/N) = Nw_{(N)}$. The resulting value of c_N^r is 0.
- The second option yields an offspring vector that is some permutation of $(2, 0, 1, 1, \dots, 1)$ and occurs with probability $1 - Nw_{(N)}$. The resulting value of c_N^r is 2.



(a) Expected coalescence rate with multinomial resampling ($N = 3$)



(b) Expected coalescence rate with residual resampling ($N = 3$)



(c) Difference between expected coalescence rate with multinomial or residual resampling ($N = 3$)

So, under this constraint on the weights, we have the following expectation:

$$\mathbb{E}[c_N^r(t)|w_t^{(1:N)}] = 2(1 - Nw_{(N)})/(N)_2 = \frac{2}{N-1}(1/N - w_{(N)}).$$

□

Conjecture 1. *The number of cases to consider for population size N (excluding zero-measure cases which can be included in other cases) is*

$$p(N-1) + \frac{(N-1)(N-2)}{2}$$

where $p(\cdot)$ denotes the partition function (number of integer partitions).

This expression was determined by pattern-spotting in low-N cases. I believe a combinatorial proof could be found but, since it is not a particularly useful result, I shan't bother.