# Stratified resampling: towards a finite time-scale

Suzie Brown

16th March 2021

The aim of this note is to calculate (bounds on) the probabilities of the four different outcomes that are possible for the marginal offspring count $\nu_t^{(i)}$ conditional on $w_t^{(i)}$, under stratified resampling. Ultimately, these bounds will be used to prove that the time-scale is finite (i.e. $\mathbb{P}[\tau_N(t)=\infty]=0$ for all finite $t$) under stratified resampling, assuming the additional constraints:

- that the weights are bounded away from the degenerate case $(1,\dots,1)/N$ in some way
- that the transition densities $q_t(x,x')$ are uniformly bounded above and away from zero.

Once the probabilities are calculated (in the following), a proof similar to the one used to prove the finite time-scale condition for stochastic rounding will prove the same for stratified resampling. The corresponding probabilities and hence finite time-scale proof for residual-stratified resampling should follow relatively easily by applying the bounds here within the residual resampling set-up.

## Let's do it!

Consider the marginal distribution of one offspring count $\nu_t^{(i)}$ conditional on the corresponding weight $w_t^{(i)}$. Henceforth we drop from the notation the dependence on $t$ and $i$, which are to be considered fixed throughout the following. As we have already seen, the possible values of $\nu$ are restricted conditional on $w$ to $\{\lfloor Nw\rfloor-1, \lfloor Nw\rfloor, \lfloor Nw\rfloor+1, \lfloor Nw\rfloor+2\}$. Denote by $p_i$ the conditional probability $\mathbb{P}[\nu=\lfloor Nw\rfloor+i \mid w]$, for $i=-1,0,1,2$.

We consider first a specific case (of particular interest for the finite time-scale proof) where $w=(1+\delta)/N$, with $\delta\in[0,1)$. First let's look at $p_0=\mathbb{P}[\nu=1\mid w=(1+\delta)/N]$. Thinking about the inversion sampling schematic, the resampling probabilities will depend upon where the length-$w$ interval falls with respect to the length-$(1/N)$ intervals for sampling. We split the possibilities into three cases (Figure 1).

### Case 1

In this case, one offspring is assigned almost surely from the interval that is entirely overlapping. Thus $p_0$ is just the probability that the partially overlapping interval does not contribute a second offspring to $\nu$. Hence,

$$p_0=\left(\frac{1}{N}-\frac{\delta}{N}\right)\div\frac{1}{N}=1-\delta. \tag{1}$$

### Case 2

In this case, one offspring is assigned almost surely from the interval that is entirely overlapping. Thus $p_0$ is just the probability that neither of the partially overlapping intervals contributes a second offspring to $\nu$. The lengths are such that $\delta_L+\delta_R=\delta$. We have

$$p_0=(1-\delta_L)(1-\delta_R)=1-\delta+\delta_L\delta_R. \tag{2}$$

Noting that $\delta_L\delta_R\le\delta^2/4\le\delta/4$, we conclude that

$$p_0\in\left[1-\delta,1-\frac{3\delta}{4}\right]. \tag{3}$$

When $\delta_L\in\{0,\delta\}$, this case collapses to Case 1, which is consistent with the bounds derived here.

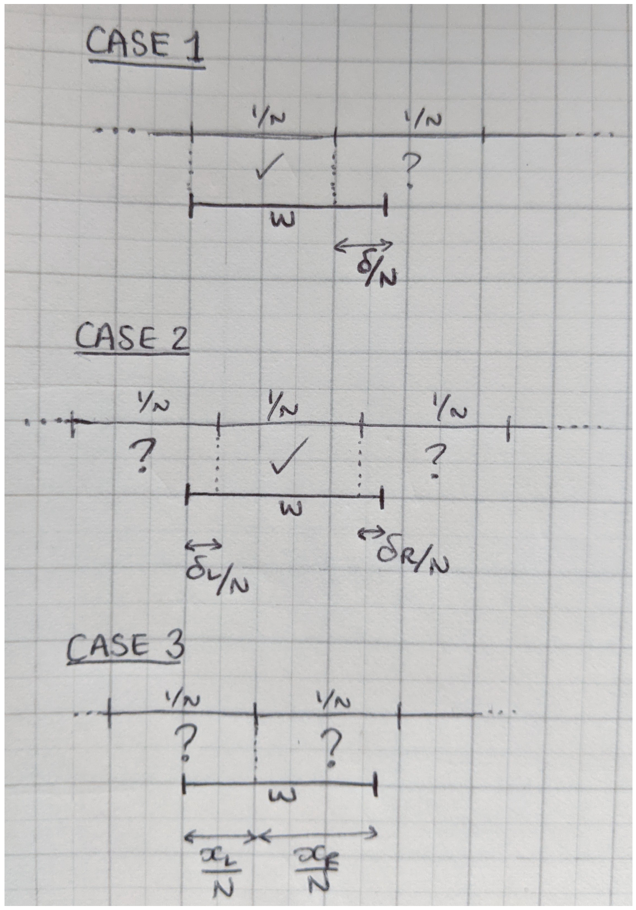


Figure 1: Sketch illustrating the difference between Cases 1–3.

## Case 3

Here $p_0$ is the probability that exactly one of the partially overlapping intervals contributes an offspring to $\nu$. The lengths are such that $x_L + x_R = 1 + \delta$, and also $x_L, x_R \in [\delta, 1]$ (otherwise we would be in Case 2). We have

$$p_0 = x_L(1 - x_R) + x_R(1 - x_L) = 1 + \delta - 2x_L x_R. \tag{4}$$

Notice that $\delta \le x_L x_R \le (1 + \delta)^2/4$, hence

$$p_0 \in \left[\frac{1 - \delta^2}{2}, 1 - \delta\right]. \tag{5}$$

When $\delta_L \in \{\delta, 1\}$, this case collapses to Case 1, which is consistent with the bounds derived here.

## Altogether

Overall, then, we have the bounds

$$p_0 \in \left[\frac{1 - \delta}{2}, 1 - \frac{3\delta}{4}\right]. \tag{6}$$

## Another thing

Now let $w = 2 + \delta$ for some $\delta \in [0, 1)$. Let's calculate $p_{-1} = \mathbb{P}[\nu = 1 \mid w = (2+\delta)/N]$. Again thinking about how the inversion sampling intervals could fall, there is only one case in which $p_{-1}$ is non-zero (Figure 2). In

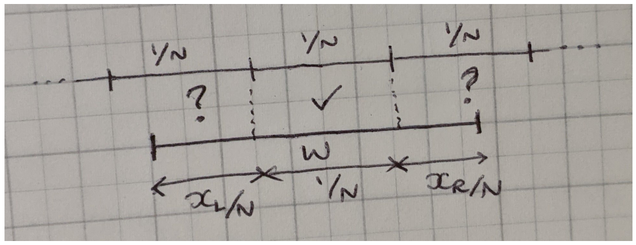


Figure 2: The only case giving non-zero probability $p_{-1}$.

this case,

$$p_{-1} = (1 - x_L)(1 - x_R) = (1 - x_L)(x_L - \delta). \tag{7}$$

This is a negative quadratic in $x_L$, with its unique maximum at $x_L = x_R = (1 + \delta)/2$. The maximum value of $p_{-1}$ is then

$$p_{-1} = \frac{1}{4}(1 - \delta)^2 \leq \frac{1}{4} \tag{8}$$

independently of $\delta$. So we have, for any $\delta$,

$$p_{-1} \in \left[0, \frac{1}{4}\right]. \tag{9}$$