Resampling in Sequential Monte Carlo

Suzie Brown

12 November 2019

Outline

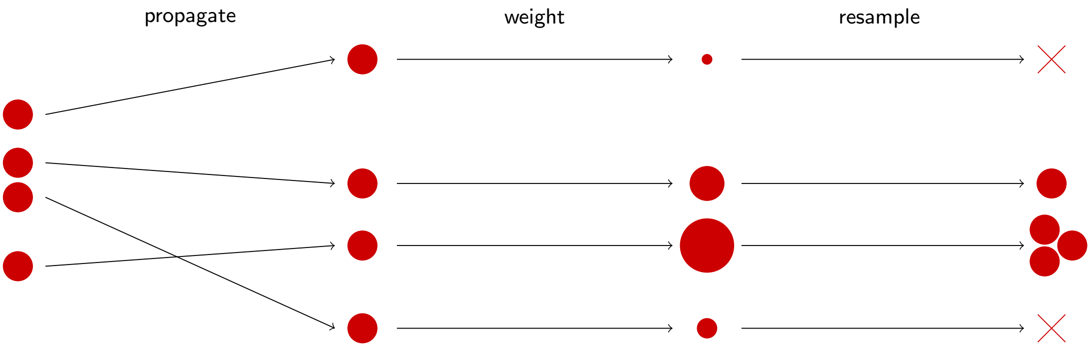
Sequential Monte Carlo

Motivation

- Want to approximate a sequence of target measures $(\eta_t)_{t\in\mathbb{N}}$
- Use a system of N particles with dynamics 'mimicking' the target
- A 'particle' consists of a position and a weight: $(x_t^{(i)}, w_t^{(i)}) = (x_i, w_i)$
- Approximate the measures η_t by (random) empirical measures η_t^N consisting of atoms at the particle positions

Sequential Monte Carlo

Illustration



Resampling

Motivation

- Resampling is necessary to prevent *weight degeneracy*
- But resampling causes *ancestral degeneracy*

Resampling

Motivation

- Resampling is necessary to prevent *weight degeneracy*
- But resampling causes *ancestral degeneracy*
- Strategy: resample in a way that minimises ‘unnecessary coalescences’

Resampling

Definition

We will take valid resampling schemes to be those satisfying

- The total number of particles N remains fixed
- The particles after resampling are equally weighted
- The scheme is unbiased: the expected number of offspring of particle i is equal to Nw_i for each i

Multinomial Resampling[1]

Definition

Parental indices $a_i \in \{1, \ldots, N\}$:

$$(a_i \mid w_{1:N}) \overset{iid}{\sim} \text{Categorical}(1:N, w_{1:N})$$

[1] Efron & Tibshirani (1994) 'An introduction to the bootstrap'

Multinomial Resampling[1]

Definition

Parental indices $a_i \in \{1, \ldots, N\}$:

$$(a_i \mid w_{1:N}) \overset{iid}{\sim} \text{Categorical}(1 : N, w_{1:N})$$

Offspring numbers $v_i \in \{0, \ldots, N\}$ such that $\sum v_i = N$:

$$(v_{1:N} \mid w_{1:N}) \sim \text{Multinomial}(N, w_{1:N})$$

[1] Efron & Tibshirani (1994) 'An introduction to the bootstrap'

Multinomial Resampling

Inversion Sampling

Draw uniform random variables

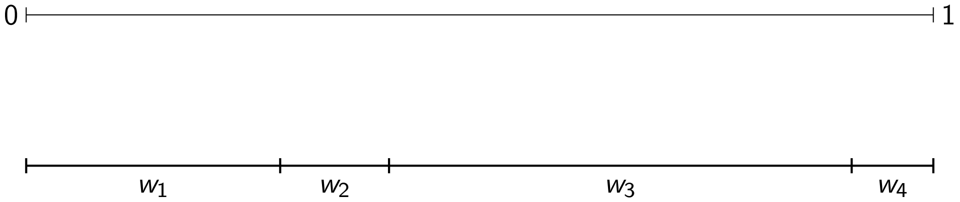
$$U_i \overset{iid}{\sim} \text{Uniform}(0,1); \qquad i = 1, \ldots, N$$

and determine the parental indices by inversion

$$a_i = \inf \left\{ k : \sum_{j=1}^{k} w_j \geq U_i \right\}$$

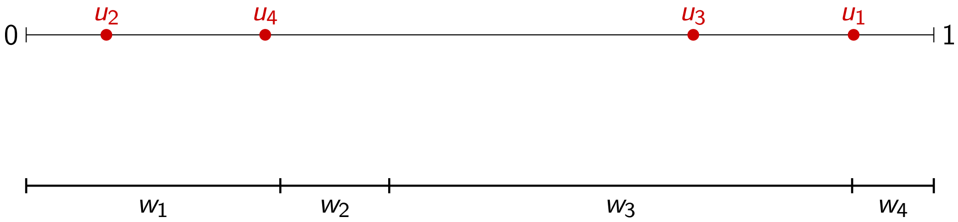
Multinomial Resampling

Inversion Sampling



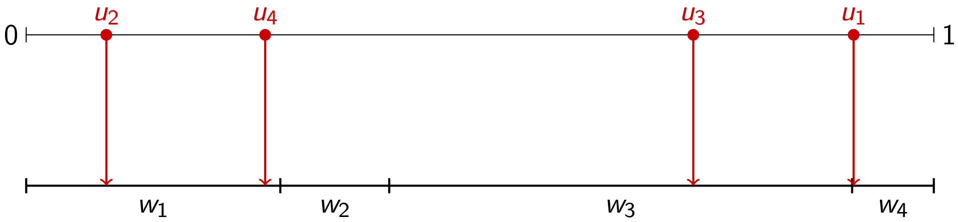
Multinomial Resampling

Inversion Sampling



Multinomial Resampling

Inversion Sampling



Residual Resampling[2,3]

Definition

1. Deterministically assign $\lfloor Nw_i \rfloor$ offspring to particle i; i=1,. . . , N
2. There are $R := N - \sum_{i=1}^{N} \lfloor Nw_i \rfloor$ offspring still to be assigned
3. Assign these randomly according to the residual weights $r_i := \frac{1}{R}(Nw_i - \lfloor Nw_i \rfloor)$

[2] Liu & Chen (1998) 'Sequential Monte Carlo methods for dynamic systems'
[3] Whitley (1994) 'A genetic algorithm tutorial'

Residual Resampling

Illustration

$w_1 = 0.28$

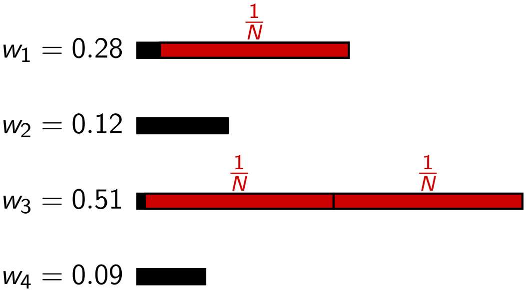
$w_2 = 0.12$

$w_3 = 0.51$

$w_4 = 0.09$

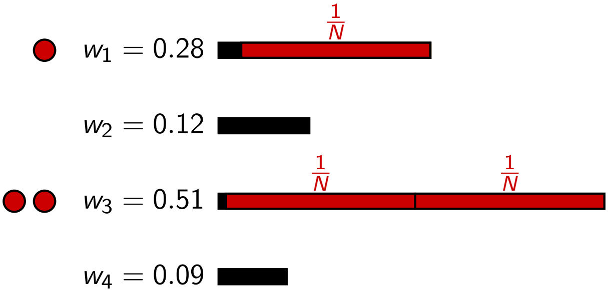
Residual Resampling

Illustration



Residual Resampling

Illustration



Residual Resampling

Illustration

$r_1 \propto 0.03$

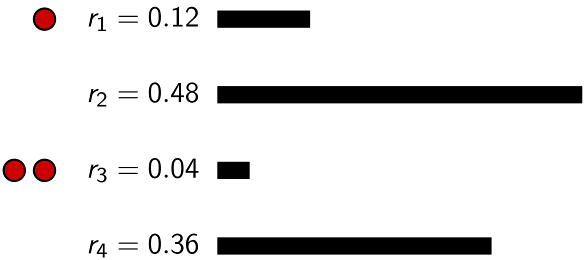
$r_2 \propto 0.12$

$r_3 \propto 0.01$

$r_4 \propto 0.09$

Residual Resampling

Illustration



Residual Resampling

Definition

If residuals are assigned using multinomial resampling, offspring counts are distributed

$$v_{1:N} \stackrel{d}{=} \lfloor N w_{1:N} \rfloor + \mathrm{Multinomial}(R, r_{1:N})$$

Stratified Resampling[4]

Definition

Draw uniformly from each stratified interval

$$U_i \sim \text{Uniform}\left(\frac{i-1}{N}, \frac{i}{N}\right); \qquad i = 1, \ldots, N$$

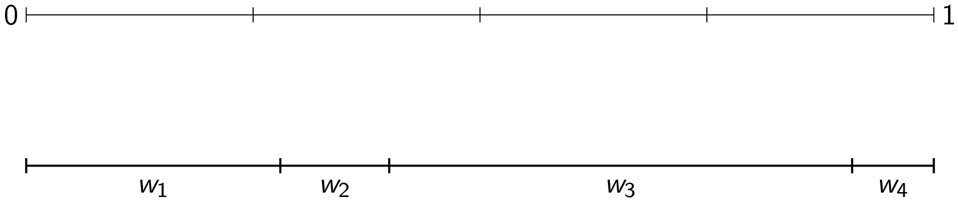
and determine the parental indices by inversion

$$a_i = \inf\left\{k : \sum_{j=1}^{k} w_j \geq U_i\right\}$$

[4] Kitagawa (1996) 'Monte Carlo filter and smoother for non-Gaussian nonlinear state space models'

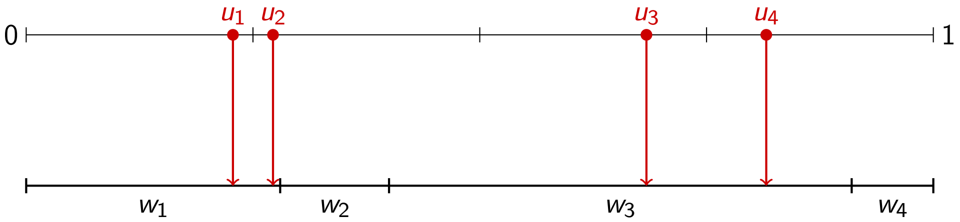
Stratified Resampling

Inversion Sampling



Stratified Resampling

Inversion Sampling



Systematic Resampling[5,6]

Definition

Draw uniformly from $[0, \frac{1}{N}]$, and add multiples of $\frac{1}{N}$

$$U_1 \sim \text{Uniform}\left(0, \frac{1}{N}\right)$$

$$U_i = U_1 + \frac{i-1}{N}; \qquad i = 2, \ldots, N$$

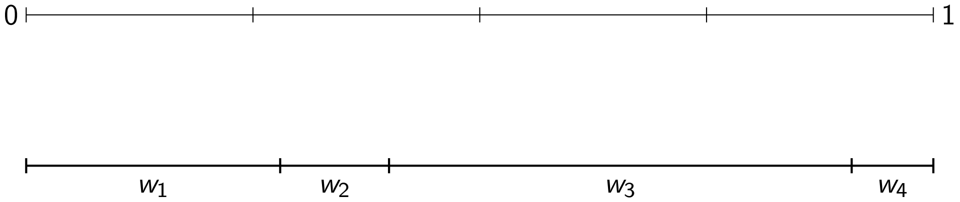
and determine the parental indices by inversion

$$a_i = \inf\left\{k : \sum_{j=1}^{k} w_j \geq U_i\right\}$$

[5] Carpenter, Clifford & Fearnhead (1999) 'Improved particle filter for nonlinear problems'
[6] Whitley (1994) 'A genetic algorithm tutorial'

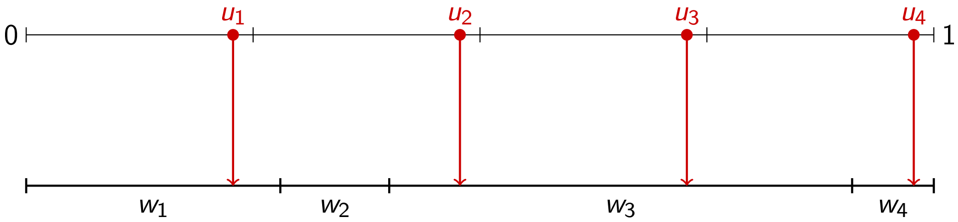
Systematic Resampling

Inversion Sampling



Systematic Resampling

Inversion Sampling



Properties of Resampling Schemes

Support of Offspring Counts

Suppose $w_i \in \left[\frac{k}{N}, \frac{k+1}{N}\right]$.
What are the possible values for v_i?

Multinomial: $v_i \in \{0, \dots, N\}$

Residual: $v_i \in \{k, \dots, k + R\}$

Stratified: $v_i \in \{k - 1, k, k + 1, k + 2\}$

Systematic: $v_i \in \{k, k + 1\}$

Properties of Resampling Schemes

One-Step Variance

Consider variance of our estimator, conditional on the previous step:

$$\mathrm{Var}\left[\frac{1}{N}\sum_{i=1}^{N}\varphi(X_t^{(i)}) \mid \mathcal{G}_{t-1}\right]$$

[7] Douc, Cappé & Moulines (2005) ‘Comparison of resampling schemes for particle filtering’

Properties of Resampling Schemes

One-Step Variance

Consider variance of our estimator, conditional on the previous step:

$$\mathrm{Var}\left[\frac{1}{N}\sum_{i=1}^{N}\varphi(X_t^{(i)}) \mid \mathcal{G}_{t-1}\right]$$

In this sense we have[7]

$$\begin{aligned} \mathrm{Var}[\text{stratified}] &\leq \mathrm{Var}[\text{multinomial}] \\ \mathrm{Var}[\text{residual-stratified}] \leq \mathrm{Var}[\text{residual-multinomial}] &\leq \mathrm{Var}[\text{multinomial}] \end{aligned}$$

[7] Douc, Cappé & Moulines (2005) 'Comparison of resampling schemes for particle filtering'

Properties of Resampling Schemes

Permutation Invariance

Stratified and systematic resampling are sensitive to the ordering of the particles.

Properties of Resampling Schemes

Permutation Invariance

Stratified and systematic resampling are sensitive to the ordering of the particles.

Example

$N = 6$

$w_{1:N} = \frac{1}{12}(3, 3, 2, 2, 1, 1)$

Is it possible to sample offspring counts $v_i = (1, 1, 1, 1, 1, 1)$?

Properties of Resampling Schemes

Permutation Invariance

Stratified and systematic resampling are sensitive to the ordering of the particles.

Example

$N = 6$

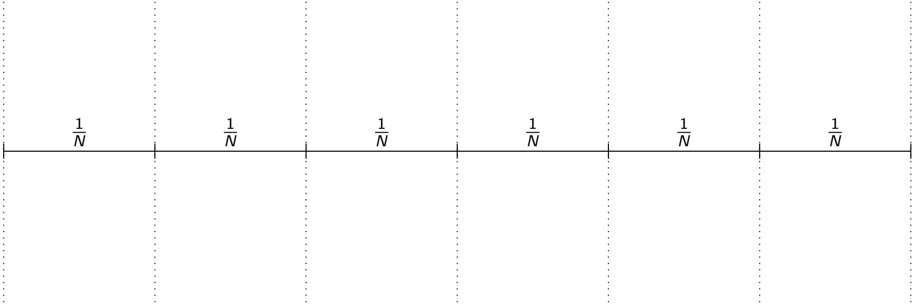
$w_{1:N} = \frac{1}{12}(3, 3, 2, 2, 1, 1)$

Is it possible to sample offspring counts $v_i = (1, 1, 1, 1, 1, 1)$?

Answer: it depends on the ordering!

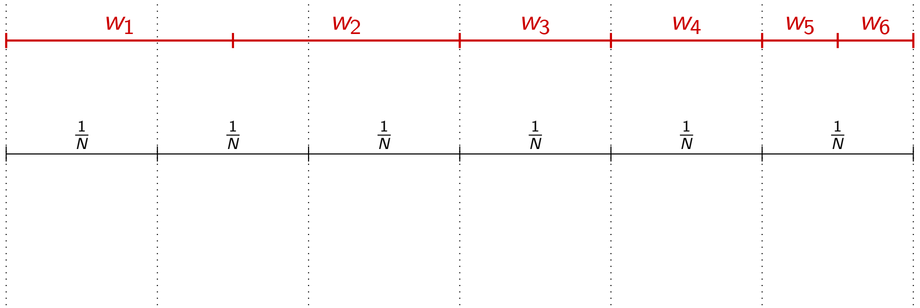
Properties of Resampling Schemes

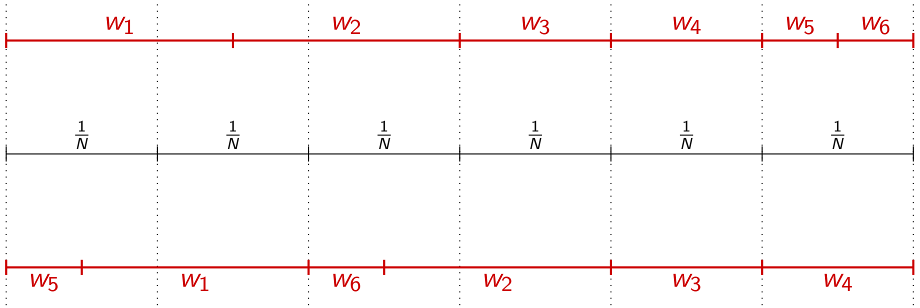
Permutation Invariance



Properties of Resampling Schemes

Permutation Invariance




w_1
w_2
w_3
w_4
w_5
w_6
$\frac{1}{N}$
$\frac{1}{N}$
$\frac{1}{N}$
$\frac{1}{N}$
$\frac{1}{N}$
$\frac{1}{N}$
w_5
w_1
w_6
w_2
w_3
w_4

Properties of Resampling Schemes

Permutation Invariance

- Kitagawa[8] suggested ordering the particles by their positions before resampling
- He ran an experiment suggesting that sorting reduces Monte Carlo variance
- This was later proved to be true[9]
- Sorting by position could be a sort of proxy for sorting by weight

[8] Kitagawa (1996) 'Monte Carlo filter and smoother for non-Gaussian nonlinear state space models'
[9] Gerber, Chopin & Whiteley (2018) 'Negative association, ordering and convergence of resampling methods'

Properties of Resampling Schemes

Degeneracy under Equal Weights

- Suppose all of the weights are multiples of $\frac{1}{N}$.

Properties of Resampling Schemes

Degeneracy under Equal Weights

- Suppose all of the weights are multiples of $\frac{1}{N}$.
- Then residual, stratified and systematic resampling all yield purely deterministic assignments of offspring.

Properties of Resampling Schemes

Degeneracy under Equal Weights

- Suppose all of the weights are multiples of $\frac{1}{N}$.
- Then residual, stratified and systematic resampling all yield purely deterministic assignments of offspring.
- In particular, if $w_{1:N} = \frac{1}{N}(1, \ldots, 1)$, these schemes do not resample at all (assigning exactly one offspring to each particle).

Properties of Resampling Schemes

Degeneracy under Equal Weights

- Suppose all of the weights are multiples of $\frac{1}{N}$.
- Then residual, stratified and systematic resampling all yield purely deterministic assignments of offspring.
- In particular, if $w_{1:N} = \frac{1}{N}(1, \ldots, 1)$, these schemes do not resample at all (assigning exactly one offspring to each particle).
- Under reasonable conditions, this situation has zero measure.

Properties of Resampling Schemes

Summary

	$\sup \lvert v_i - Nw_i \rvert$	low variance	invariant under permutations	degenerate if $w_{1:N} \propto (1, \dots, 1)$
multinomial	N	$\times$	$\checkmark$	$\times$
residual	R	$\checkmark$	$\checkmark$	$\checkmark$
stratified	2	$\checkmark$	$\times$	$\checkmark$
systematic	1	$\times$	$\times$	$\checkmark$

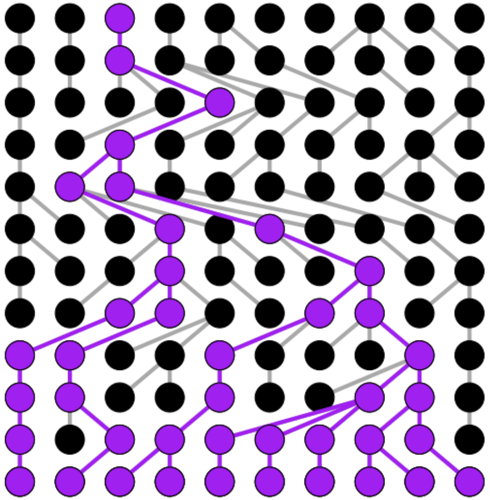
Properties of Resampling Schemes

Summary

	$\sup \lvert v_i - Nw_i \rvert$	low variance	invariant under permutations	degenerate if $w_{1:N} \propto (1, \dots, 1)$
multinomial	N	$\times$	$\checkmark$	$\times$
residual	R	$\checkmark$	$\checkmark$	$\checkmark$
stratified	2	$\checkmark$	$\times$	$\checkmark$
systematic	1	$\times$	$\times$	$\checkmark$
residual-strat	1	$\checkmark$	$\times$	$\checkmark$

Resampling and Genealogies

- Resampling creates a genealogy (family tree) of particles
- Properties of the genealogy affect performance of the SMC algorithm
- Different resampling schemes give different forms of genealogies
- Basic quantity for analysing genealogies is the pair coalescence probability



Coalescence Probability

Definition

The probability that a randomly chosen pair of particles at generation t share a common ancestor at generation $(t-1)$

$$c_N = \frac{1}{N(N-1)} \sum_{i=1}^{N} v_i(v_i - 1)$$

Coalescence Probability

Example

Consider the case where we have only two particles ($N = 2$)

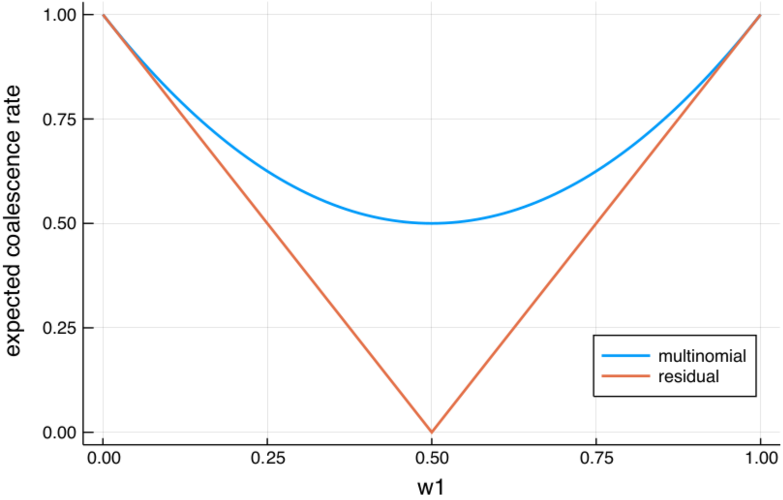
$$c_2 = \frac{1}{2}\left[v_1(v_1 - 1) + v_2(v_2 - 1)\right]$$

The expectation of c_2 conditional on knowing the weights (w_1, w_2) is

$$\begin{aligned} c_2 &= \frac{1}{2}\mathbb{E}[v_1(v_1 - 1) \mid w_{1:2}] + \frac{1}{2}\mathbb{E}[v_2(v_2 - 1) \mid w_{1:2}] \\ &= \mathbb{P}[v_1 = 2 \mid w_{1:2}] + \mathbb{P}[v_2 = 2 \mid w_{1:2}] \end{aligned}$$

Coalescence Probability

Example



- We proved that asymptotically (as $N \to \infty$) residual resampling dominates multinomial in terms of expected coalescence probability
- We also proved it in cases $N = 2$ and $N = 3$

- We proved that asymptotically (as $N \to \infty$) residual resampling dominates multinomial in terms of expected coalescence probability
- We also proved it in cases $N = 2$ and $N = 3$
- We conjecture that it holds for all finite N too
- It just remains to prove it for $N = 4, 5, \ldots$

- We proved that asymptotically (as $N \to \infty$) residual resampling dominates multinomial in terms of expected coalescence probability
- We also proved it in cases $N = 2$ and $N = 3$
- We conjecture that it holds for all finite N too
- It just remains to prove it for $N = 4, 5, \dots$
- We proved that systematic resampling (and some others) dominate multinomial in expected coalescence probability, for all N.

THE END