Bisection ideas in End-Point Conditioned Markov Process Simulation

Søren Asmussen & Asger Hobolth

Aarhus University, Denmark

http://home.imf.au.dk/asmus

RESIM2008

IRISA, Rennes, September 25, 2008

X finite Markov jump process on $[0, T]$

Aim:

simulate X conditioned on $X(0) = a$, $X(T) = b$

Applications:

statistical inference for Markov processes

finance

human genetics

genomics

Aim:

simulate X conditioned on
$X(0) = a$, $X(T) = b$

Naive algorithm:

Start from $X(0) = a$
First jump time exponential(q_a)
new state a' chosen w.p. $q_{aa'}/q_a$
($q_a = -q_{aa}$)
Repeat until T passed
Reject path if $X(T) \neq b$.
Inefficient if $P_{ab}(T)$ small
Otherwise hard to beat

Other algorithms:

Uniformization
Direct simulation

Uniformization

Unconditioned implementation:

Choose $\lambda > \max_i q_i$

Simulate Poisson(λ) process on $[0, T]$

$N(T)$ Poisson(λT)

Uniform epochs

State changes at epochs according to $p_{ij} = q_{ij}/\lambda$

Dummy transitions

Conditioned implementation:

$$\mathbb{P}(N(T) = n) \propto e^{-\lambda T}\frac{(\lambda T)^n}{n!}p_{ab}^n$$

Uniform epochs

State change at epoch k, state i according to $p_{ij}p_{jb}^{n-k-1}/p_{ib}^{n-k}$

Inefficient if q_i variable

Otherwise often good

Direct simulation

(Hobolth, 2008)

Start from $X(0) = a$

First jump to a'

with conditioned probability

First jump time from $f_{aa'b}(t)$

(conditioned density)

Repeat until T passed

Messy formulas and r.v. generation

Still often good

Comparisons:

Hobolth & Stone

Ann. Appl. Statist. (2009)

Bisection for BM in $[0, 1]$

$B(0) = 0,\ B(1) \sim N(0, 1)$

$B(1/2)$ from conditional normal given $B(0), B(1)$

$B(1/4)$ given $B(0), B(1/2)$

$B(3/4)$ given $B(1/2), B(1)$

Go on until desired resolution

SA-Hobolth:

Bisection algorithm for end-point conditioned CTMC (resolution not an issue)

Initialization, $X(0) = X(T) = a$

No jumps w.p. $\mathrm{e}^{-q_a T}$
W.p. $\mathrm{e}^{-q_a T}/P_{aa}(T)$,
take $X(0:T) \equiv a$
Otherwise at least two jumps,
go on to recursion

Initialization, $X(0) = a \neq b = X(T)$

One jump w.p.

$$R_{ab}(T) = \int_0^T q_a e^{-q_a t} \frac{q_{ab}}{q_a} e^{-q_b (T-t)} \, \mathrm{d}t$$

W.p. $R_{ab}(T)/P_{ab}(T)$,
take $X(0:T)$ with one jump.
Jump time density
prop. to integrand
exponential V truncated to $[0, T]$,
jump time V or $T - V$
Otherwise at least two jumps,
go on to recursion

Recursion, $a = b$

Know $[0, T]$ has ≥ 2 jumps

$X(T/2) =?$

$e_a = e^{-q_a T/2}$, $r_{ab} = R_{ab}(T/2)$,

$p_{ab} = P_{ab}(T/2)$

case	number of jumps in $[0, T/2]$	number of jumps in $[T/2, T]$	(unconditional) probability	not
1	0	0	$e_a e_a$	
2	0	≥ 2	$e_a(p_{aa} - e_a)$	
3	≥ 2	0	$(p_{aa} - e_a)e_a$	
4	≥ 2	≥ 2	$(p_{aa} - e_a)(p_{aa} - e_a)$	
5	1	1	$r_{ac} r_{ca}$	c
6	1	≥ 2	$r_{ac}(p_{ca} - r_{ca})$	c
7	≥ 2	1	$(p_{ac} - r_{ac})r_{ca}$	c
8	≥ 2	≥ 2	$(p_{ac} - r_{ac})(p_{ca} - r_{ca})$	c

Sample among cases 2-8

Finish intervals with 0 or 1

Go on with ≥ 2

$a \neq b$: very similar.

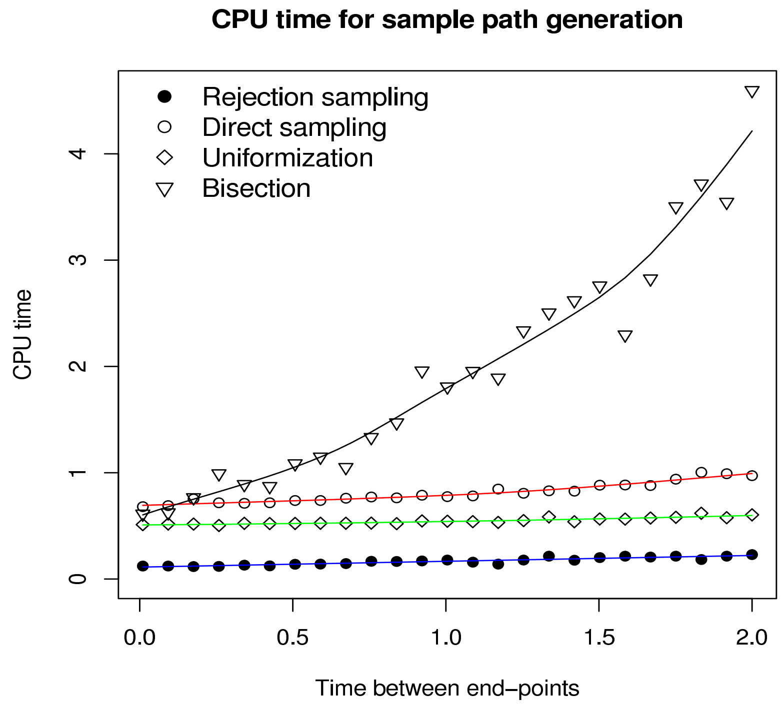
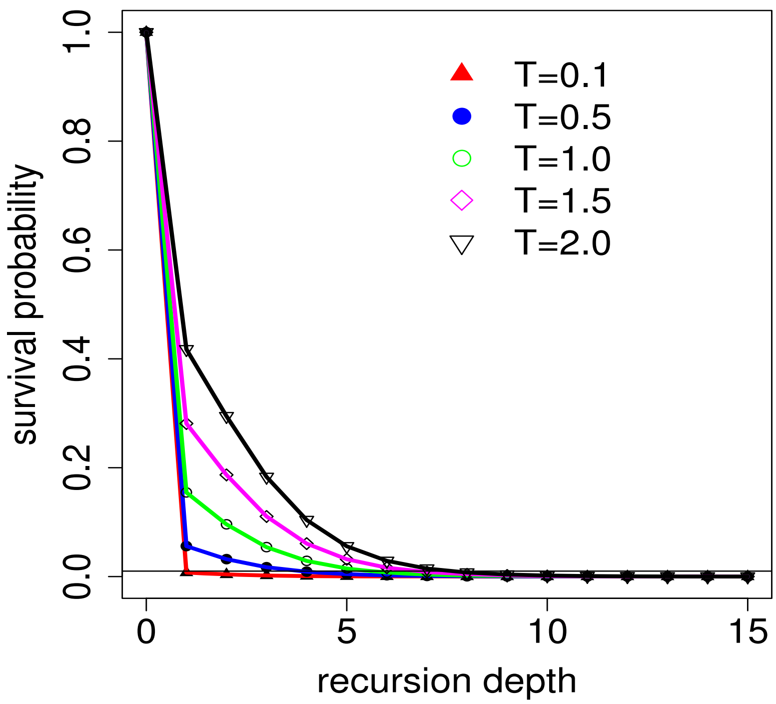
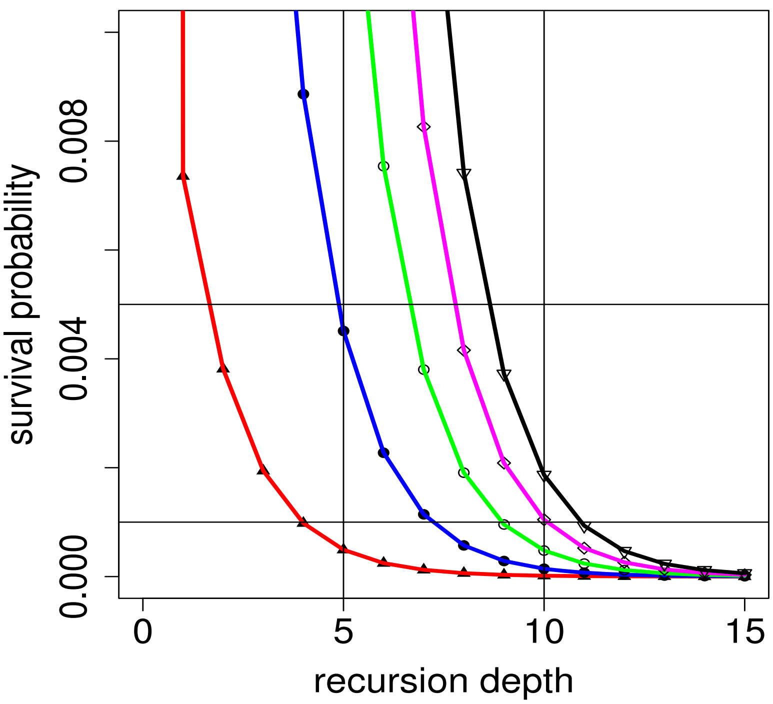

CPU time for sample path generation
Rejection sampling
Direct sampling
Uniformization
Bisection
CPU time
0
1
2
3
4
0.0
0.5
1.0
1.5
2.0
Time between end−points

Figure 1:



T=0.1
T=0.5
T=1.0
T=1.5
T=2.0
survival probability
recursion depth
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15

Figure 2:



survival probability
0.000
0.004
0.008
0
5
10
15
recursion depth

Figure 3: