Simplex Diffusion HMM

Stephan Schiffels

May 2021

Dirichlet distribution

$$p(\{x_i\}|\{\alpha_i\}) = \frac{\Gamma(\mathrm{A})}{\prod_{i=1}^K \Gamma(\alpha_i)} \prod_{i=1}^K x_i^{\alpha_i - 1} \tag{1}$$

with $\mathrm{A} = \sum_i \alpha_i$.

The likelihood of a set of observations $\{n_1 \dots n_k\}$ given a Dirichlet prior represented by $\{\alpha_1, \dots, \alpha_k\}$ is described by the so-called Dirichlet-Multinomial distribution, which is a so-called compound distribution:

$$p(\{n_i\}|\{\alpha_i\}) = \frac{\Gamma(A)\Gamma(N+1)}{\Gamma(N+A)} \prod_{i=1}^k \frac{\Gamma(n_k + \alpha_k)}{\Gamma(\alpha_k)\Gamma(n_k + 1)} \tag{2}$$

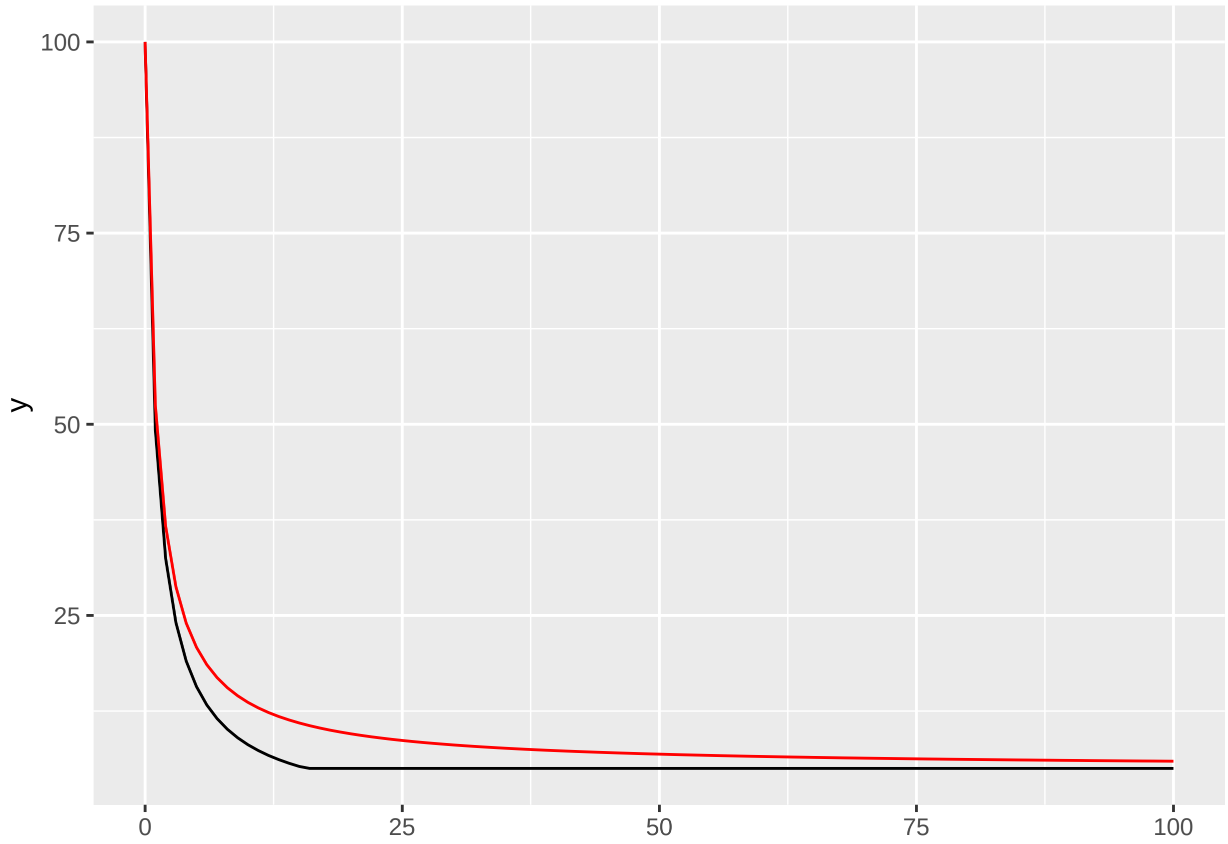
where $N = \sum_i n_i$.

Diffusion step

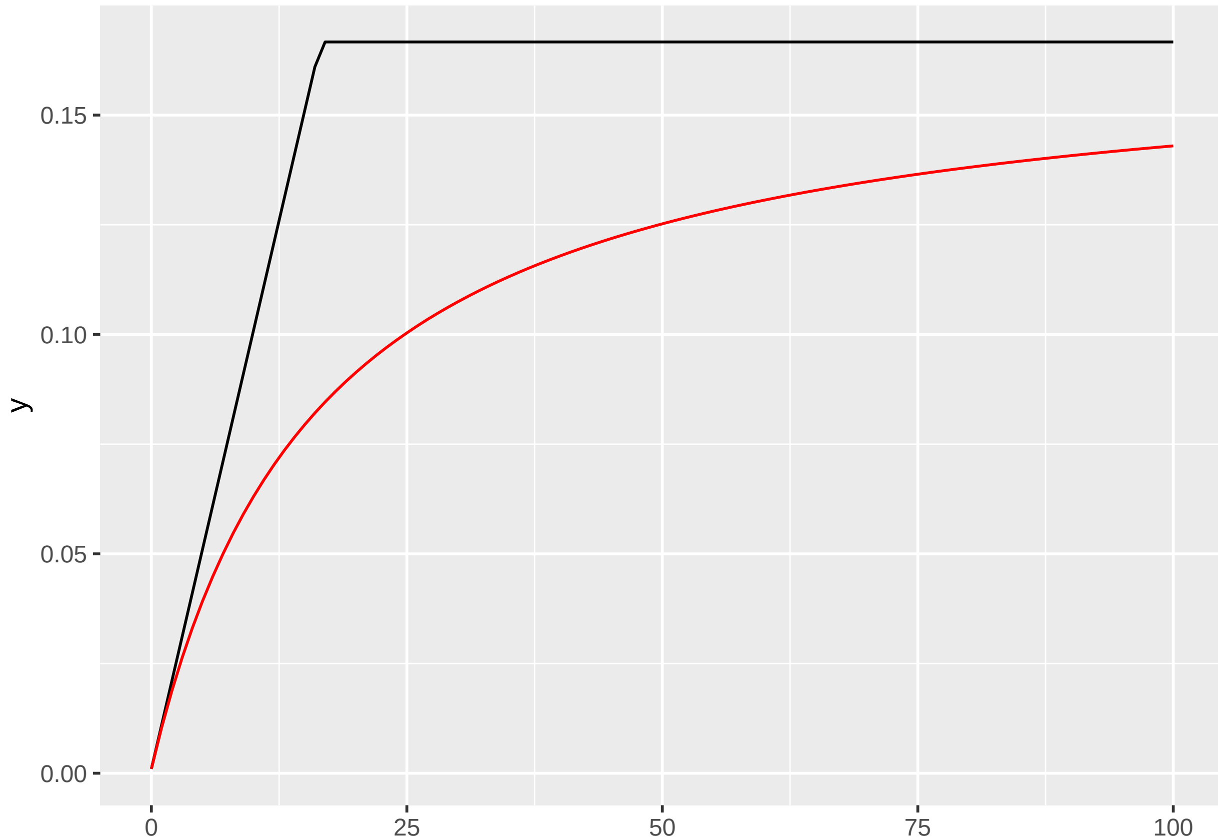
```
a_diffusion <- function(t, a0, diffusion_constant, k)
  pmax(k, (a0 - diffusion_constant * t * (a0 + 1)) / (1 + diffusion_constant * t * (a0 + 1)))

a_diffusion2 <- function(t, a0, diffusion_constant, k)
  a0 * (1.0 + k * diffusion_constant * t) / (1.0 + a0 * diffusion_constant * t)

ggplot() + geom_function(fun = ~a_diffusion(., 100, 0.01, 5)) + xlim(0, 100) +
  geom_function(fun = ~a_diffusion2(., 100, 0.01, 5), col="red")
```




```r
var_part <- function(a) (1 / (a + 1))
ggplot() + geom_function(fun = ~var_part(a_diffusion(., 1000, 0.01, 5))) + xlim(0, 100) +
  geom_function(fun = ~var_part(a_diffusion2(., 1000, 0.01, 5)), col="red")
```


y
0.15
0.10
0.05
0.00
0
25
50
75
100