Conditions for convergence

$$X^{(N)}(t) = \frac{Y^{(N)}_{\lfloor t/h_N \rfloor} - a_N}{b_N} \xrightarrow{d} X(t)$$

diffusion

1) Asymmetric RW

$Y^{(N)}$ a RW on $\mathbb{Z}$ with

$$p_{i,i+1} = \frac{1}{2} + \frac{\mu}{2\sqrt{N}} \qquad p_{i,i-1} = \frac{1}{2} - \frac{\mu}{2\sqrt{N}}$$

for N large enough so they are in $[0,1]$.

Define

$$X^{(N)}(t) := \frac{Y^{(N)}_{\lfloor Nt \rfloor}}{\sqrt{N}}$$

→ Time rescaling by $h_N = \frac{1}{N}$

→ space rescaling by $b_N = \sqrt{N}$

→ no centering

$$Y^{(N)}_{\lfloor Nt \rfloor} = i \in \mathbb{Z} \Rightarrow X^{(N)}(t) = \frac{i}{\sqrt{N}}$$

$$\Rightarrow \Delta X^{(N)}(t) = \frac{1}{\sqrt{N}} \left(Y^{(N)}_{\lfloor Nt \rfloor + 1} - Y_{\lfloor Nt \rfloor} \right)$$

For $p \in \mathbb{N}$

$$E_x\left[\left(\Delta X^{(N)}(t)\right)^p\right] =$$

$$= \frac{1}{N^{p/2}}\left[(+1)^p\left(\frac{1}{2}+\frac{\mu}{2\sqrt{N}}\right)+(-1)^p\left(\frac{1}{2}-\frac{\mu}{2\sqrt{N}}\right)\right]$$

- $p = 1$

$$E_x\left(\Delta X^{(N)}(t)\right) = \frac{1}{N^{1/2}}\left[\cancel{\frac{1}{2}}+\frac{\mu}{2\sqrt{N}}-\cancel{\frac{1}{2}}+\frac{\mu}{2\sqrt{N}}\right]$$

$$= \underbrace{\frac{1}{N}}_{h_N}\ \underbrace{\mu}_{\mu(x)\equiv\mu\in\mathbb{R}} \qquad \left(E_x(\Delta X(t)) = \mu(x)h + o\right)$$

- $p = 2$

$$E_x\left[\left(\Delta X^{(N)}(t)\right)^2\right] = \frac{1}{N}\left[\frac{1}{2}+\cancel{\frac{\mu}{2\sqrt{N}}}+\frac{1}{2}-\cancel{\frac{\mu}{2\sqrt{N}}}\right] = \underbrace{\frac{1}{N}}_{h_N}$$

$$\Rightarrow \qquad \sigma^2(x) \equiv 1$$

- $p = 4$

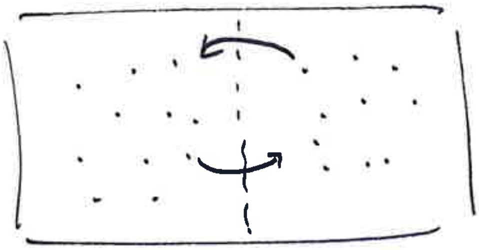
$$E_x\left[\left(\Delta X^{(N)}(t)\right)^4\right] = \frac{1}{N^2} = o(h_N)$$

Then $X^{(N)} \xrightarrow{d} X$ where X solves

$$dX(t) = \mu\, dt + dB(t) \quad , \quad X(t)\in\mathbb{R}$$

2) EHRENFEST URN

Total $2N$ balls
separated by membrane



↑ membrane

1 ball selected at
random and moved to other space

$Y^{(N)}$ = no. balls in first space

$$S = \{0, \dots, 2N\}$$

$$p_{i,i-1} = \frac{i}{2N} = \frac{1}{2} - \frac{N-i}{2N}$$

$$p_{i,i+1} = 1 - \frac{i}{2N} = \frac{1}{2} + \frac{N-i}{2N}$$

spatially inhomogeneous RW on
S finite. Define

$$X^{(N)}(t) := \frac{Y^{(N)}_{\lfloor Nt \rfloor} - N}{\sqrt{N}}$$

$\Rightarrow \; Y^{(N)} = i \quad \Rightarrow \quad i = x\sqrt{N} + N$
when $X^{(N)}(t) = x$

$$P_x\left(\Delta X^{(N)}(t) = \pm \frac{1}{\sqrt{N}}\right) = P\left(\Delta Y^{(N)}_{\lfloor Nt \rfloor} = \pm 1 \;\middle|\; i = x\sqrt{N} + N\right)$$

$$= \frac{1}{2} \pm \frac{\cancel{N} - (x\sqrt{N} + \cancel{N})}{2N} = \frac{1}{2} \mp \frac{x}{2\sqrt{N}}$$

spatial inhomogeneity

$$E_x\left(\Delta X^{(N)}(t)\right) = \frac{1}{\sqrt{N}}\left(\cancel{\frac{1}{2}} - \frac{x}{2\sqrt{N}}\right) - \frac{1}{\sqrt{N}}\left(\cancel{\frac{1}{2}} + \frac{x}{2\sqrt{N}}\right)$$

$$= -\frac{x}{2N} - \frac{x}{2N} = -\frac{x}{N} \qquad \left(h_N = \frac{1}{N};\ \mu(x) = -x\right)$$

$$E_x\left[\left(\Delta X^{(N)}(t)\right)^2\right] = \frac{1}{N}\left(\frac{1}{2} - \cancel{\frac{x}{2\sqrt{N}}}\right) + \frac{1}{N}\left(\frac{1}{2} + \cancel{\frac{x}{2\sqrt{N}}}\right) = \frac{1}{N}$$

$$\sigma^2(x) \equiv 1$$

$$E_x\left[\left(\Delta X^{(N)}(t)\right)^4\right] = O\left(\frac{1}{N^2}\right) = o\left(\frac{1}{N}\right) = o(h_N)$$

$$\Longrightarrow \quad X^{(N)} \xrightarrow{d} X \quad \text{s.t.}$$

$$dX(t) = -X(t)\,dt + dB(t), \quad X(t) \in \mathbb{R}$$

called ORNSTEIN-UHLENBECK diffusion, stationary w.r.t.

$$N\left(0, \frac{1}{2}\right)$$

applications in math finance and biology.

It can be seen as a continuous-time analog of an $AR(1)$.

Discretize over Δt interval to get

$$\underbrace{\Delta X_k}_{X_{k+1} - X_k} = - X_k \Delta t + \sqrt{\Delta t}\, \varepsilon_k$$

$$\varepsilon_k \overset{iid}{\sim} N(0, \tfrac{1}{2})$$

$$\Rightarrow \quad E(\Delta X_k) = - X_k \Delta t$$

$$Var(\Delta X_k) = \frac{\Delta t}{2}$$

and we write

$$X_{k+1} = X_k - X_k \Delta t + \sqrt{\Delta t}\, \varepsilon_k$$

$$\overset{!}{=} \underbrace{(1 - \Delta t)}_{a} X_k + \underbrace{\sqrt{\Delta t}}_{b}\, \varepsilon_k$$

---

3) BRANCHING PROCESSES

$Y_n^{(N)}$ a BP $\quad Y_n^{(N)} = \sum_{i=1}^{Y_{n-1}^{(N)}} Z_i^{(N)}$

$Z_i^{(N)} \overset{iid}{\sim}$ with mean $\mu^{(N)} \in \mathbb{R}$

variance $\sigma^2 > 0$

$$E(Y_n^{(N)} \mid Y_{n-1}^{(N)} = y) = \mu^{(N)} y$$

$$\Rightarrow E_y(\Delta Y_n^{(N)}) = \mu^{(N)} y - y = y(\mu^{(N)} - 1)$$

$$\mathrm{Var}_y(Y_m^{(N)}) = \sigma^2 y$$

Define $\quad X^{(N)}(t) := \dfrac{Y^{(N)}_{\lfloor Nt \rfloor}}{N}$

$$\Rightarrow \quad X^{(N)}(t) = x \quad \Rightarrow \quad y = Nx$$

$$E_x[\Delta X^{(N)}(t)] = \frac{1}{N} E_y[\Delta Y^{(N)}_{\lfloor Nt \rfloor}] = \frac{1}{N} y(\mu^{(N)} - 1)$$

$$E_x[(\Delta X^{(N)}(t))^2] = \mathrm{Var}_x(\Delta X^{(N)}(t)) + (E_x[\Delta X(t)])^2$$

$$= \frac{1}{N^2} \underbrace{\mathrm{Var}_y(\Delta Y_m)}_{\mathrm{Var}_y(Y_{m+1}) + 0} + \frac{1}{N^2} y^2 (\mu^{(N)} - 1)^2$$

$$= \frac{1}{N^2} \sigma^2 y + \frac{1}{N^2} y^2 (\mu^{(N)} - 1)^2$$

if we let $\quad \mu^{(N)} = 1 + \dfrac{m}{N}$

$$\mu^{(N)} - 1 = \frac{m}{N}$$

$$\Rightarrow \quad E_x[\Delta x] = \frac{1}{N} y \frac{m}{N} = \frac{my}{N^2} = \frac{mx}{N}$$

since $y = Nx$

$$E_x[(\Delta x)^2] = \frac{\sigma^2 x}{N} + \underbrace{\frac{1}{N^2} N^2 x^2 \frac{m^2}{N^2}}_{o(\frac{1}{N})}$$

$h_N = \frac{1}{N}$, $\mu(x) = mx$

$\sigma^2(x) = \sigma^2 \cdot x$

$\Rightarrow X^{(N)} \xrightarrow{d} X$ s.t.

$$dX(t) = m\, X(t)\, dt + \sigma \sqrt{X(t)}\, dB(t)$$

$X(t) \in \mathbb{R}_+$

called

- COX - INGERSOLL - ROSS diffusion in math finance

- CONTINOUST - STATE branching process in math biology

0 is an absorbing state

4) <u>Wright - Fisher</u>

$$Y^{(N)}_{m+1} \mid Y_m = i \sim \text{Binom}\left(N, \frac{i}{N}\right)$$

$$X^{(N)}(t) := \frac{Y^{(N)}_{\lfloor Nt \rfloor}}{N} \quad \text{\% Type 0}$$

$$Z_x := \left(X^{(N)}(t+\tfrac{1}{N}) \mid X^{(N)}(t) = \underbrace{x}_{i/N}\right) \sim \frac{1}{N}\,\text{Binom}(N, x)$$

$$E_x\left[\Delta X^{(N)}(t)\right] = E(Z_x) - x = \frac{1}{N}\, Nx - x = 0$$

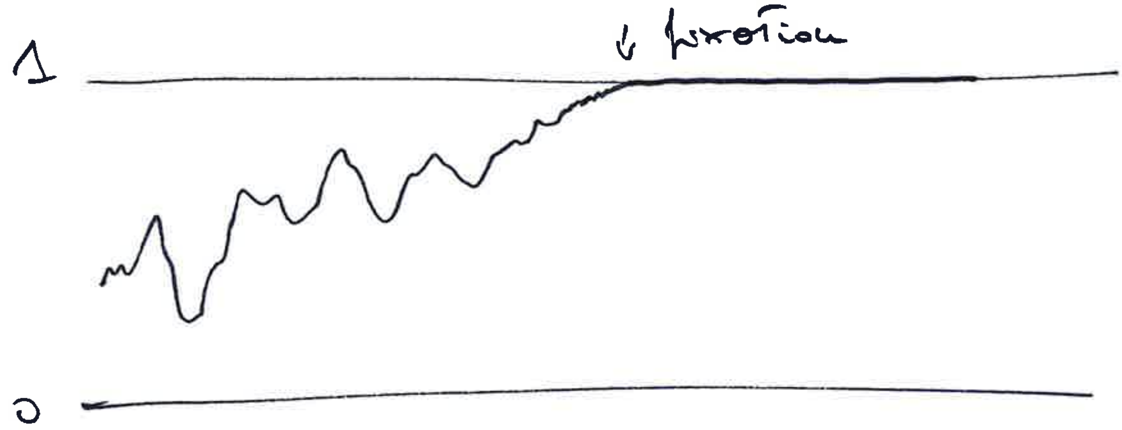
$$E_x\left[\left(\Delta X^{(N)}(t)\right)^2\right] =$$

$$= \text{Var}_x\left[\underbrace{\Delta X(t)}_{Z_x - x}\right] + \left(\underbrace{E_x\left[\Delta X(t)\right]}_{=0}\right)^2$$

$$= \frac{1}{N^2} N x(1-x) = \underbrace{\frac{1}{N}}_{h_N} \underbrace{x(1-x)}_{\sigma^2(x)}$$

$X^{(N)} \xrightarrow{d} X$ s.t.

$$dX(t) = \sqrt{X(t)(1-X(t))}\, dB(t)$$



Add mutations

$$Y_{n+1} | Y_n = i \sim \text{Binom}(N, p_i)$$

$$p_i = \alpha\left(1 - \frac{i}{N}\right) + (1-\beta)\frac{i}{N}$$

$$\alpha = P(1 \to 0)$$
$$\beta = P(0 \to 1)$$

$$\tilde{Z}_x := \left(X^{(N)}\left(t + \tfrac{1}{N}\right) \,\middle|\, X^{(N)}(t) = x \right) \sim$$

$$\sim \frac{1}{N} \text{Binom}(N, p_x) \qquad x = \frac{i}{N}$$

$$p_x = \alpha(1-x) + (1-\beta)x$$

$$E_x\left[\Delta X^{(N)}(t)\right] = E_x(\tilde{Z}_x) - x$$

$$= \frac{1}{N} N p_x - x = \dots = \frac{1}{N} \overbrace{\left[\alpha(1-x) - \beta x\right]}^{\mu(x)}$$

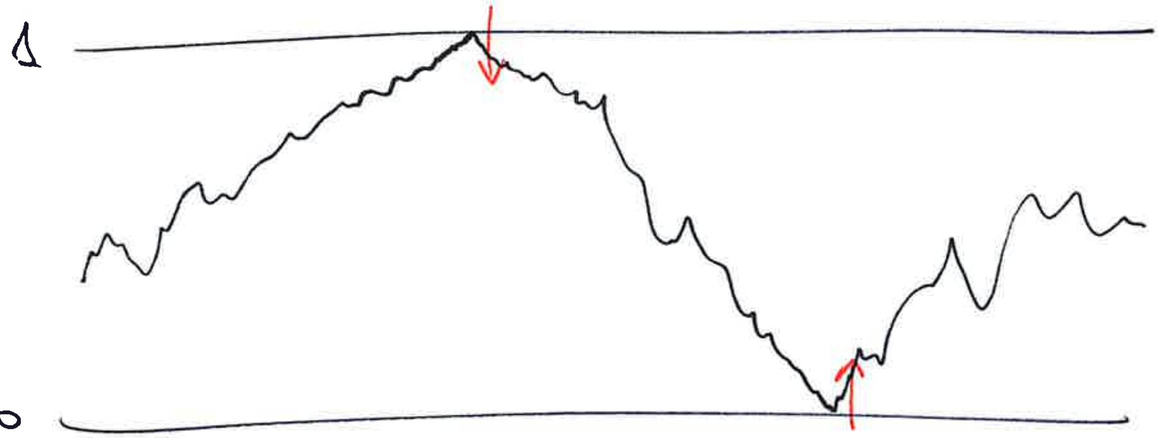
$$E_x\left[\left(\Delta X^{(N)}(t)\right)^2\right] = \dots = \frac{1}{N}\underbrace{x(1-x)}_{\sigma^2(x)} + o\left(\frac{1}{N}\right)$$

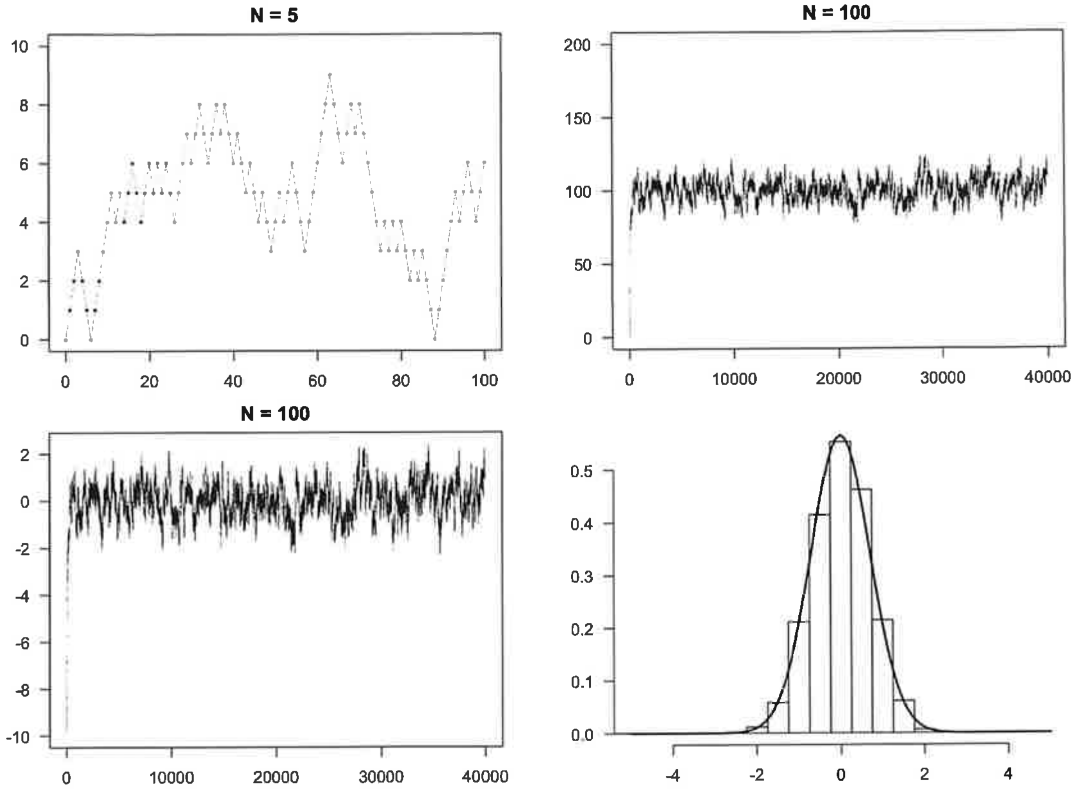
$$\Longrightarrow \quad X^{(N)} \xrightarrow{d} X \quad \text{s.t.}$$

$$dX(t) = \left[\alpha(1-X(t)) - \beta X(t)\right] dt + \underbrace{\sqrt{X(t)(1-X(t))}}_{=0 \text{ if } X(t)=0,1}\, dB(t)$$

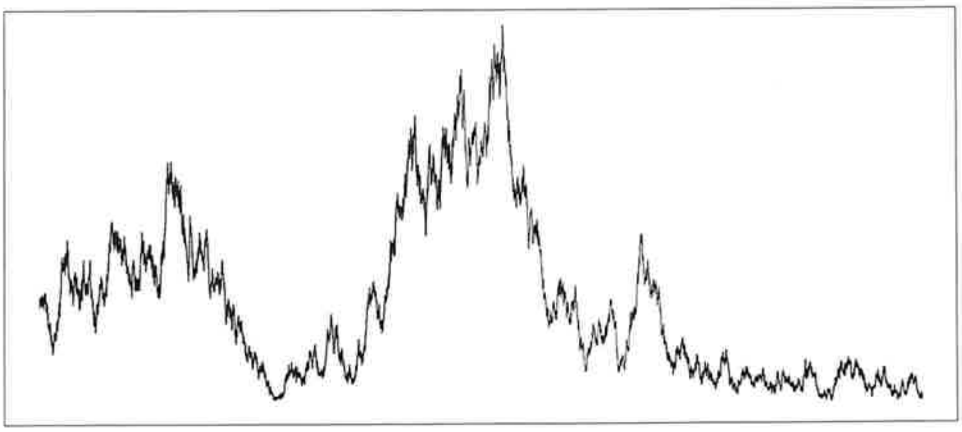
$$\text{at } \quad X(t)=0 \qquad dX(t) = \alpha\, dt \quad \uparrow$$

$$\text{at } \quad X(t)=1 \qquad dX(t) = -\beta\, dt \quad \downarrow$$

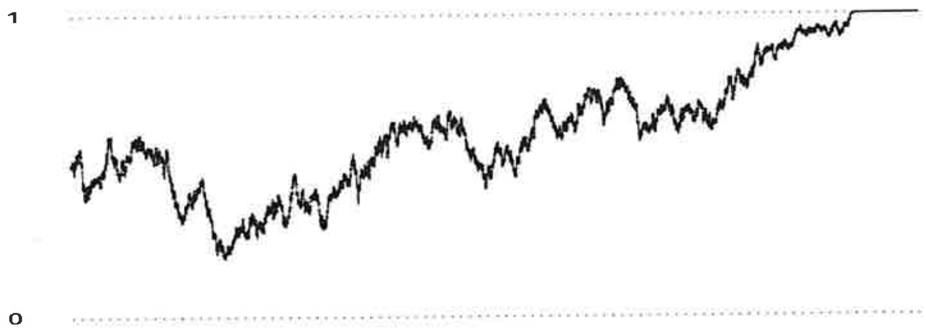




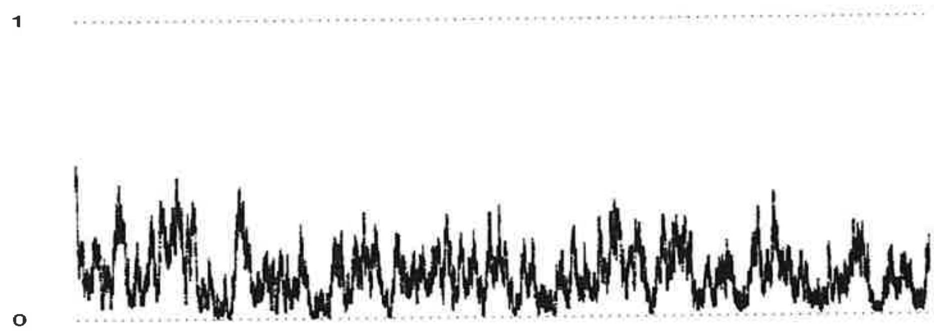
Top: two paths of the Ehrenfest urn for $N = 5, 100$.
Bottom: centered and rescaled process (left); ergodic frequencies vs. $N(0, 1/2)$ (right).



Sample path of a CIR diffusion on $\mathbb{R}_+$.

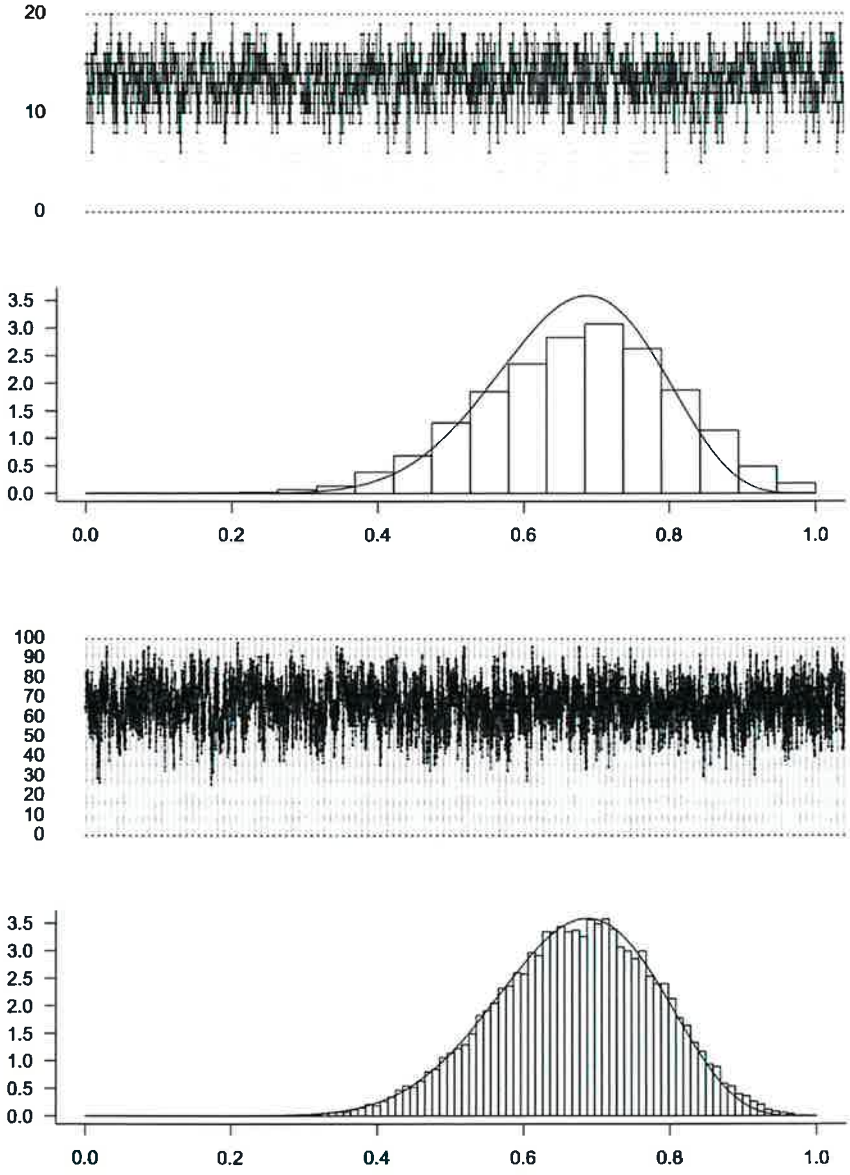


Sample path of a WF diffusion *without* mutation, exhibiting fixation at 1.



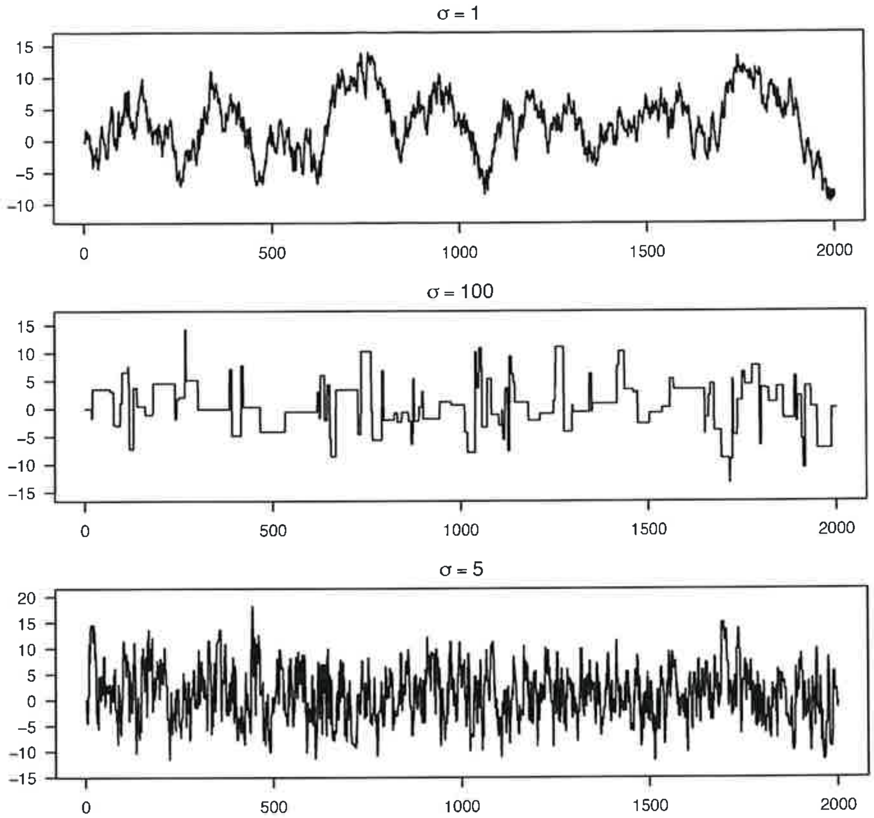
Sample path of a WF diffusion *with* mutation, with $a = 1$ and $b = 6$.

Application 1: approximation of stationary distribution

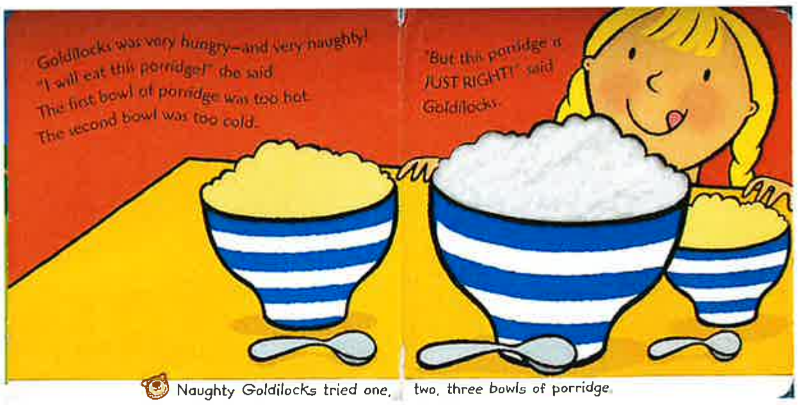


WF paths and ergodic frequencies (histograms) against Beta$(2a, 2b)$ density (solid) for different values of N.

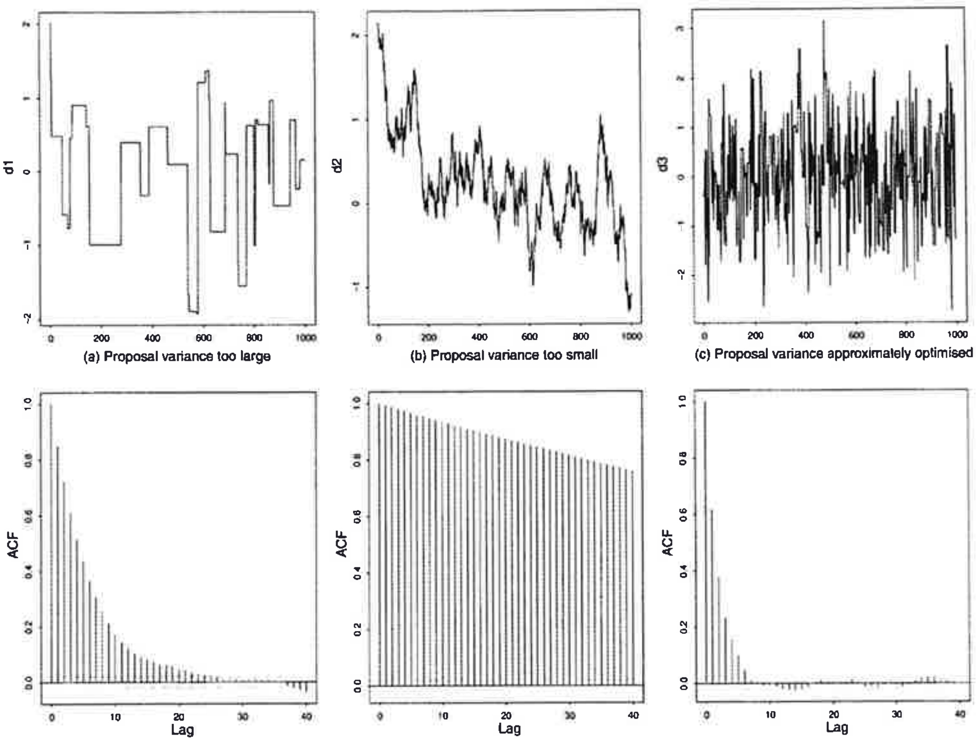
Application 2: Tuning of variance in the RW-MH (Goldilocks principle) through scaling limits



If the proposal variance is *too small*, the space is not explored efficiently (top); if the proposal variance is *too large*, the algorithm gets stuck in the same state for long periods (middle); if the proposal variance is *just right*, the space is explored efficiently. The right tuning can be investigated through the scaling limit of the MCMC (more at the end of the course).



This has been called the *Goldilocks principle* for RW-MH, terminology which recalls the famous fairy tale for kids.



Paths of a RW-MH algorithm for different choices of the scaling σ. Figure from ROBERTS, G.O. and ROSENTHAL, J.S. (2001). Optimal scaling for various Metropolis–Hastings algorithms. *Statist. Sci.*, **16**, 351–367.

They study a rescaling of a class of RW-MH and how to tune the variance of the proposal accordingly.