26.11.2012 (1)
5.12.2012

DICHOTOM - PR.	
GENETIC — mymon.	
Noise-ind. transition in genetic models	
Chandler — the concept of friction	

STOCHASTIC DIFFUSION FROM SCALING OF DISCRETE TIME PROCESSES,

Brownian motion is a regular diffusion process on the interval $(-\infty, +\infty)$ with a drift term $A(x)=0$ and a diffusion term $B(x)=\sigma^2$

$$\Delta_h \cancel{X} = X(h) - X(0)$$ normally distributed with $E[\Delta_h X]=0$
$Var[\Delta_h X] = \sigma^2 h$

$$E[\Delta_h X / X(0)=x] = 0$$
$$E[(\Delta_h X)^2 / X(0)=x] = \sigma^2 h$$

$$E\left[(\Delta_h X)^4 / X(0)=x\right] =$$

$$= \frac{1}{\sqrt{2\pi h}\,\sigma} \int_{-\infty}^{+\infty} y^4 \exp\left(-\frac{1}{2}\frac{y^2}{\sigma^2 h}\right) dy$$

$$= \left(\frac{y^2}{\sigma^2 h} = 2z, \quad \frac{2y\,dy}{\sigma^2 h} = 2dz \to dy = \frac{\sigma^2 h\, dz}{y} \right.$$

$$y = (2\sigma^2 h z)^{1/2}$$

$$= \frac{1}{(2\pi h)^{1/2}\sigma} 2 \cdot \int_{0}^{+\infty} (2\sigma^2 h z)^2 \frac{\sigma^2 h\, dz}{(2\sigma^2 h z)^{1/2}} e^{-z}$$

26.11. (2)

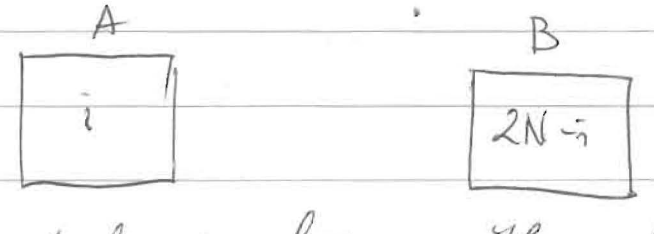
$$= \frac{1}{(2\pi h)^{1/2}\sigma}\, 2\int_{+0}^{+\infty} \frac{4\sigma^4 h^2 \cdot \sigma^2 \cdot h}{\sigma (2h)^{1/2}}\, z^{3/2} e^{-z}\, dz$$

$$= \frac{1}{\sqrt{\pi}}\, 2 \int_{-0}^{+\infty} 2\sigma^4 h^2 \, z^{3/2} e^{-z}\, dz$$

$$= \frac{4\sigma^4 h^2}{\sqrt{\pi}} \int_0^{\infty} z^{3/2} e^{-z}\, dz = \frac{3\cancel{\sqrt{\pi}}}{\cancel{4}} \cdot \frac{\cancel{4}\sigma^4 h^2}{\cancel{\sqrt{\pi}}}$$

O-U process from the Ehrenfest urn model

$$\left.\begin{array}{lll} i & \rightarrow & \text{urn } A \\ 2N-i & \rightarrow & \quad B \end{array}\right\} 2N \text{ balls}$$



Any of boxes chosen with prob $1/2$, then

$$P_{ij} = \begin{cases} \frac{2N-i}{N}\cdot\frac{1}{2} & j = i+1 \\ \frac{i}{2N} & j = i-1 \\ 0 & \text{otherwise} \end{cases}$$

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

$$\frac{1}{2}\,\frac{i}{N} \qquad i \to i-1$$

Let us define

$$\text{Prob}\{\Delta X = \pm 1 \mid X_N(t) = x\} = \frac{1}{2} + \frac{N-x}{2N} =$$

$X_N(t)$ = number of particles in box A at time t

$$\Delta X = X_N(t+\Delta t) - X_N(t)$$

$x \equiv i$

$$\frac{1}{2} \pm \frac{N-x}{2N} \begin{cases} = \frac{1}{2} + \frac{1}{2} - \frac{x}{2N} = 1 - \frac{x}{2N} & (\Delta X = +1) \\ = \frac{1}{2} - \frac{1}{2} + \frac{x}{2N} = \frac{x}{2N} & (\Delta X = -1) \end{cases}$$

Let N increase and Δt decrease while $N\Delta t = 1 = \text{const}$

Define $\Rightarrow$ $$\boxed{Y_N(\tau) \doteq \frac{X_N([N\tau]) - N}{\sqrt{N}}}$$

approximating process

Hence: $X_N([N\tau]) = \sqrt{N} \cdot Y_N(\tau) + N$

Now

$$\Pr\left\{\Delta Y = \pm \frac{1}{\sqrt{N}} \mid Y_N(\tau) = y\right\} = \Pr\{\Delta X = \pm 1 \mid X_N(N\tau) =$$

$$x = N + y\sqrt{N}\} = \frac{1}{2} \pm \frac{N-(N+y\sqrt{N})}{2N} \quad \text{(same as above)}$$

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

We would like to estimate

?

$$\lim_{h\to 0} \frac{1}{h} E\left[\Delta Y(h) \mid Y(t)=y\right] \to A(y)$$

$$\lim_{h\to 0} \frac{1}{h} E\left[\Delta Y^2(h) \mid Y(t)=y\right] \to B(y)$$

$h = \frac{1}{N}$

$Pr\{\Delta Y = \pm\frac{1}{\sqrt{N}}$ (when ΔX changes by 1, ΔY as a rescaled variable changes by $\frac{1}{\sqrt{N}}$)$\mid$

$$Y(\tau)=y\} = Pr\{\Delta X = \pm 1 \mid X[N\tau]=x = N + y\sqrt{N}\}$$

$$= \frac{1}{2} \pm \frac{N-X}{2N} = \frac{1}{2} \pm \frac{N-(N+y\sqrt{N})}{2N}$$

$$= \frac{1}{2} \mp \frac{y}{2\sqrt{N}}$$

$$E\left[\Delta Y(h) \mid Y(t)=y\right] = \frac{1}{\sqrt{N}}\left(\frac{1}{2} - \frac{y}{2\sqrt{N}}\right) +$$

$$+ (-)\frac{1}{\sqrt{N}} \cdot \left(\frac{1}{2} + \frac{y}{2\sqrt{N}}\right) = -\frac{y}{N} = -y\cdot h$$

hence $\lim_{h\to 0} \frac{1}{h} E\left[\Delta Y(h) \mid Y(t)=y\right] = -y$!

5

$$\Delta Y^2(h) = \frac{1}{N} \rightarrow$$

$$E\left[\Delta Y^2(h) \mid Y(t)=y\right] = \frac{1}{N}\cdot\left(\frac{1}{2}-\frac{y}{2\sqrt{N}}\right)$$

$$+\frac{1}{N}\left(\frac{1}{2}+\frac{y}{2\sqrt{N}}\right) = \frac{1}{N}$$

$$\lim_{h\to 0}\frac{1}{h}E\left[\Delta Y^2(h) \mid Y(t)=y\right] = N\cdot\frac{1}{N}=1$$

similar computation leads to

$$\frac{1}{h}E\left[\Delta Y^4(h) \mid Y(t)=y\right] = \frac{1}{N}\rightarrow 0$$

hence the process scales to a diffusion process

$$\dot{y} = -y + (\sigma)\,\xi(t) \quad LE_q$$

KBE $$\frac{\partial p(y,t)}{\partial t} = -y\,\partial_y p(y,t) + \frac{1}{2}\partial_{yy}p(y,t)$$

FPE $$\frac{\partial p(y,t)}{\partial t} = \partial_y\, y\, p(y,t) + \frac{1}{2}\partial_{yy}p(y,t)$$

Ito transformation rules

How to determine the form of infinitesimal parameters of new processes built from certain transformations applied to original processes?

$$Y(t) = g\big(X(t)\big) \quad g \text{ - continuous and monotone}$$

$\{X(t), t>0\} \longrightarrow \mu(x), \; \sigma^2(x)$
$x \in [l, r]$ — drift, diffusion

Assumption:
Two uniformly continuous $g'(x)$ and $g''(x)$ exist.

$\{Y(X(t)), t>0\}$
$y \in [g(l), g(r)]$

$$\mu_Y(y) = \tfrac{1}{2}\sigma^2(x)g''(x) + \mu(x)g'(x)$$

$$\sigma_Y^2(y) = \sigma^2(x)\left[g'(x)\right]^2$$

PROOF:

$$g(x+\Delta x) = g(x) + \Delta x\, g'(x) + \frac{1}{2}(\Delta x)^2 g''(x) +$$
$$+ \frac{1}{2}(\Delta x)^2 \left[g''(\xi) - g''(x) \right]$$

where $\xi \in [x, x+\Delta x]$

$$\Downarrow$$

$$g(X(t+h)) = g(X(t)) + \Delta X g'(X) + \frac{1}{2}(\Delta X)^2 g''(X)$$
$$+ \frac{1}{2}(\Delta X)^2 \left[g''(\xi(X)) - g''(X(t)) \right]$$

or otherwise

$$Y(t+h) - Y(t) = \Delta X\, g'(X) + \frac{1}{2}(\Delta X)^2 g''(X)$$
$$+ \frac{1}{2}(\Delta X)^2 \left[g''(\xi(X)) - g''(X(t)) \right]$$

$$\lim_{h\to 0} \frac{1}{h} E\left[Y(t+h) - Y(t) \mid Y(t) = y \right] =$$

$$\underbrace{\lim_{h\to 0} E\left[\frac{\Delta X}{h} g'(X) + \frac{1}{2} \frac{1}{h}(\Delta X)^2 g''(X) + \frac{1}{2} \underbrace{\frac{(\Delta X)^2}{h}}_{\text{convergent}} \underbrace{\left[g''(\xi(x)) - g''(X(t)) \right]}_{= 0 \text{ by continuity of } g''!} \right]}_{}$$

$$\equiv \quad \mu(x) g'(x) + \frac{1}{2}\underbrace{\sigma^2(x) g''(x)}_{}$$

$= 0$ (written above the last line)

second moment:

$$\left[Y(t+h)-Y(t)\right]^2 = \Delta X^2\left[g'(x)\right]^2 + R$$

where R contain $(\Delta X)^3$ and higher order terms

Since $\lim_{h\to 0} \frac{1}{h} E\left[|\Delta X|^r \,|\, X(t)=x\right]=0$ for $r \geq 3$

(by definition, for any stochastic diffusion process $X(t)$)

$$\Rightarrow \lim_{h\to 0} \frac{1}{h} E\left[\left\{Y(t+h)-Y(t)\right\}^2 / \, Y(t)=y\right] =$$

$$= \sigma^2(x)\left[g'(x)\right]^2$$

EXAMPLE:

The Euclidean distance from the origin of an n-dim Brownian motion $\sqrt{Z}$, where

$$Z(t) = X_1^2(t) + \ldots + X_n^2(t)$$

Let

$$X_1(t+\Delta t) = x_1 + \Delta X_1 \,, \quad Z(t+\Delta t) = z + \Delta Z$$

$$= x_1^2 + \ldots + x_n^2 + \Delta Z$$

$$\downarrow$$

$$\Rightarrow \Delta Z = \left[X_1(t+\Delta t)\right]^2 - x_1^2 + \ldots + \left[X_n(t+\Delta t)\right]^2 - x_n^2$$

$$= 2\left(x_1\Delta X_1 + \ldots + x_n \Delta X_n\right) + \left[(\Delta X_1)^2 + \ldots + (\Delta X_n)^2\right]$$

ΔX_i are independent and normally distributed with zero mean ($\mu = 0$) and variance Δt

$$\Downarrow$$

$$E[\Delta Z \mid Z(t) = z] = n \cdot \Delta t$$

$$E[(\Delta Z)^2 \mid Z(t) = z] = 4(x_1^2 + \ldots + x_n^2)\Delta t + o(\Delta t)$$
$$= 4 \cancel{z} \Delta t$$

$$E[(\Delta Z)^4 \mid Z(t) = z] = o((\Delta t)^2)$$

$$\Downarrow$$

$$\mu(z) = n$$
$$\sigma^2(z) = 4z$$

$Y(t) = g(Z) = \sqrt{Z}$ defines a Bessel process

$g'(z) = \frac{1}{2} z^{-1/2}$; $g'' = -\frac{1}{4} z^{-3/2}$

$\Rightarrow \quad \mu(z) = n \;, \; \sigma^2(z) = 4z \Rightarrow \sigma_Y^2(y) = ?$

$$\mu_Y(y) = \frac{1}{2}\sigma^2(z) \cdot g''(z) + \mu(z) g'(z) \Big|_{y^2 = z}$$

$$= \frac{1}{2} \cdot 4z \cdot \left(-\frac{1}{4}\right) z^{-3/2} + n \frac{1}{2} z^{-1/2}$$

$$= -\frac{1}{2} z^{-1/2} + \frac{n}{2} z^{-1/2} = \frac{n-1}{2} z^{-1/2}$$

$$= \frac{n-1}{2} \frac{1}{y}$$

$$\sigma_Y^2(y) = \sigma^2(z)\left[g'(z)\right]^2\Big|_{y^2=z}$$

$$\sigma_Y^2(y) = 1 \quad \begin{cases} \sigma^2(z) = 4z = 4y^2 \\ g'(z) = \frac{1}{2} z^{-1/2} = \frac{1}{2} y^{-1} \\ \left[g'(z)\right]^2 = \frac{1}{4} y^{-2} \end{cases}$$

The Bessel process is governed by a Langevin equation:

$$dY_t = \frac{n-1}{2Y} dt + dW_t$$

FPEq $\quad \frac{\partial}{\partial t} p(y,t) = -\frac{\partial}{\partial y}\left(\frac{n-1}{2y} p(y,t)\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2} p(y,t)$

$$p_s(y) = N \exp\left(2\int^y \frac{n-1}{2y'} dy'\right) \sim C|y|^{n-1}$$

WRIGHT-FISHER GENETIC MODEL

$$P_{ij} = \binom{N}{j} p_i^j (1-p_i)^{N-j}$$

where $p_i = \dfrac{(1+s)[i(1-\alpha)+(N-i)\beta]}{(1+s)[i(1-\alpha)+(N-i)\beta] + [i\alpha+(N-i)(1-\beta)]}$

PARENTAL POPULATION

i A types $\quad$ $N-i$ a types

is subject to mutation, selection + sampling forces

$\boxed{N \to \infty \text{, diffusion approximation}}$

(composition of the next generation is determined through N binomial trials)

let $\alpha = \dfrac{r_1}{N}$ $\quad r_1, r_2, s > 0$

$\beta = \dfrac{r_2}{N}$

$$Y_N(\tau) \equiv \frac{X([N\tau])}{N}, \quad h = \frac{1}{N}$$

1) We analyze changes in the population of genes with mutations only.

$$\Delta Y_N(\tau, h) = Y_N\left(\tau + \frac{1}{N}\right) - Y_N(\tau) =$$

$$\left\{X\left([N\tau]+1\right) - X\left([N\tau]\right)\right\} N^{-1}$$

$$E\left[\Delta Y_N(\tau,h) \mid Y_N(\tau) = \xi = \frac{i}{N}\right] = E\left[\left\{Y_N(\tau+\tfrac{1}{N}) - Y_N(\tau)\right\} \mid Y_N(\tau) = \frac{i}{N}\right] =$$

$= \left(NY_N(\tau + \frac{1}{N}) \text{ is distributed binomially!}\right) =$

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

$$= \frac{1}{N}\left[-\gamma_1 \frac{i}{N} + \left(1 - \frac{i}{N}\right)\gamma_2\right]$$

hence $\lim\limits_{h \to 0} \frac{1}{h} E\left[\Delta Y_N(\tau, h) \mid Y_N(\tau) = \xi\right] =$

$$= -\gamma_1 \xi + (1-\xi)\gamma_2$$

where $\xi = \lim\limits_{N \to \infty} \frac{i}{N}$

$$\lim_{h \to 0} \frac{1}{h} E\left[\Delta^2 Y_N(\tau, h) \mid Y_N(\tau) = \xi\right] =$$

$$= \lim_{N \to \infty} N\, E\left[\Delta^2 Y_N(\tau, h) \mid Y_N(\tau) = \xi\right] \quad \left(h = \frac{1}{N}\right)$$

$$\left(Y_N^2(\tau + \tfrac{1}{N}) - 2Y_N(\tau + \tfrac{1}{N}) Y_N(\tau) + Y_N^2(\tau) \right)$$

$$= \lim_{N \to \infty} N\, E\left[\frac{N^2 Y_N^2(\tau + \frac{1}{N})}{N^2} - \frac{2N Y_N(\tau + \frac{1}{N})}{N}(\xi) + \xi^2\right]$$

Recall, that for binomial PDF

$$\phi(s) = \langle s^n \rangle = \sum_{n=0}^{N} \binom{N}{p} p^n s^n q^{N-n} = (ps+q)^N$$

hence,

$$\phi''(s)\Big|_{s=1} = N(N-1)p^2 = \langle n^2 \rangle - \langle n \rangle \text{ and}$$

$$\langle n \rangle = Np$$

and

$$\Downarrow \qquad \langle n^2 \rangle = N(N-1)p^2 + Np$$

$$E_i\left[\frac{N^2 Y_N^2(\tau + \frac{1}{N})}{N^2}\right] = \frac{1}{N^2}\left[Np_i(1-p_i) + N^2 p_i^2\right]$$

$$\lim_{N\to\infty} N E_i\left[\frac{1}{N^2}\left\{Np_i(1-p_i) + N^2 p_i^2\right\} - 2p_i \xi + \xi^2\right] \overset{N\to\infty}{=}$$

where $\xi = \frac{i}{N}$

$$\alpha = \frac{r_1}{N}, \quad \beta = \frac{r_2}{N} \qquad = \xi(1-\xi)$$

and $p_i = \dfrac{i(1-\alpha) + (N-i)\beta}{i(1-\alpha) + (N-i)\beta + i\alpha + (N-i)(1-\beta)} =$

$$= \frac{i - \alpha i + N\beta - i\beta}{N}$$

Resulting diffusion process is described by

$$\mu(x) = -\gamma_1 x + (1-x)\gamma_2$$

$$= -(\gamma_1 + \gamma_2)x + \gamma_2$$

$$\sigma^2(x) = x(1-x)$$

$$\frac{\partial p(x,t)}{\partial t} = \frac{\partial}{\partial x}\left\{\left[\gamma_1 + \gamma_2\right]x - \gamma_2\right]p(x,t)$$

$$+ \frac{\partial^2}{\partial x^2} x(1-x)\, p(x,t)$$

$$\lim_{t \to \infty} p(x,t) = p_s(x) = \frac{\Gamma(2(\gamma_1 + \gamma_2))}{\Gamma(2\gamma_1)\Gamma(2\gamma_2)} x^{2\gamma_1 - 1} x^{2\gamma_2 - 1}$$