Filtering and Particle Systems

Pierre Del Moral

L.S.P-C.N.R.S, F-Toulouse III

Contents

$\longrightarrow$ **Simulation et algorithmes stochastiques**
Nathalie Baroli, P. DM, Cepadues Editions 2001.

1 Stochastic modeling

$\longrightarrow$ **Probability**
A.N. Shiryaev, *Second edition, Springer*
Graduate texts in Mathematics.

1.1 Elementary probability theory

1.1.1 Space of elementary events

$\longrightarrow$ An experiment with finite number of outcomes

$$\Omega = \{\omega_1, \ldots, \omega_N\}$$

Example 1.1 (Single toss of a coin) $\Omega = \{H, T\}$ *(H=Head, T=Tail)*

Example 1.2 (n-tosses of a coin)

$$\Omega = \{\omega \,:\, \omega = (a_1, \ldots, a_n)\,,\; a_i = H \text{ or } T\}$$

Example 1.3 (Sampling with replacement) $\sim$*[selection n balls from an M-urn]*

$$\begin{aligned} \Omega^r_o &= \{\omega \,:\, \omega = (a_1, \ldots, a_n)\,,\; a_i = 1, \ldots, M\} \quad \text{ordered samples} \\ \Omega^r_{uo} &= \{\omega \,:\, \omega = \{a_1, \ldots, a_n\}\,,\; a_i = 1, \ldots, M\} \quad \text{un-ordered samples} \end{aligned}$$

$$\Longrightarrow \qquad 1) \qquad |\Omega^r_o| = M^n \qquad \text{and} \qquad 2) \qquad |\Omega^r_{uo}| = N(M, n) = C^n_{M+n-1}$$

(Proof 2)) by induction: first $N(k,1) = k = C_k^1$ *and if* $N(k,n) \overset{\forall k \leq M}{\equiv} C_{k+n-1}^n$ *then*
with $a_1 \leq a_2 \leq \ldots \leq a_{n+1}$ *($\uparrow$ arrangement) we have*

$$\begin{aligned} a_1 &= 1 \longrightarrow N(M,n) \\ a_1 &= 2 \longrightarrow N(M-1,n) \\ &\vdots \\ a_1 &= M \longrightarrow N(1,n) \end{aligned}$$

Using $C_k^{l-1} = C_{k+1}^l - C_k^l$ *we obtain*

$$N(M,n+1) = \sum_{k=1}^{M} N(k,n) = \sum_{k=1}^{M} C_{k+n-1}^n = \sum_{k=1}^{M} (C_{k+n}^{n+1} - C_{k+n-1}^{n+1}) = C_{M+n}^{n+1}$$

Example 1.4 (Sampling without replacement)

$$\begin{aligned} \Omega_o^{ur} &= \{\omega : \omega = (a_1, \ldots, a_n)\,,\ a_i \neq a_j\,,\ \ a_i = 1, \ldots, M\} \quad \text{ordered samples} \\ \Omega_{uo}^{ur} &= \{\omega : \omega = \{a_1, \ldots, a_n\}\,,\ a_i \neq a_j\,,\ a_i = 1, \ldots, M\} \quad \text{un-ordered samples} \end{aligned}$$

$$\Longrightarrow \qquad |\Omega_o^{ur}| = M(M-1)\ldots(M-n+1) = A_M^n \qquad \text{and} \qquad |\Omega_{uo}^{ur}| = C_M^n$$

Note $\{a_1, \ldots, a_n\} \longrightarrow n!$ *associated ordered samples*
$\Longrightarrow A_M^n = C_M^n \times n!$

$\longrightarrow$ Experiment with uncountable number of outcomes
$\sim$ Lebesgue measure theory.

1.1.2 Events

- $A \subset \Omega \longrightarrow$ outcome $\quad \omega \in A\,, \quad \omega \notin A$

Example 1.5 (3-tosses)

$$\begin{aligned} \Omega &= \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\} \quad (\Rightarrow |\Omega| = 2^3) \\ A &= \text{"appearance of at least two Heads"} = \{HHH, HHT, HTH, THH\} \end{aligned}$$

- Elementary events $\longrightarrow$ new events

Logical connectives:	"or", "and", "not".
Set theory operations:	union ($\cup$), intersection ($\cap$), complement $(.)^c$.
	certain event$= \Omega$, impossible event $\emptyset$.

- $\mathcal{A}$ collection of sets $A \subset \Omega \longrightarrow \sigma(\mathcal{A})$ algebra generated by $\mathcal{A}$

$$(\Omega \in \mathcal{A} \quad \text{and} \quad A \in \mathcal{A}\,,\ B \in \mathcal{A} \implies A \cup B,\ A \cap B\,,\ A - B \in \mathcal{A})$$

Example 1.6 (3-tosses)

$$\begin{aligned} \mathcal{A} &= \{\{\underline{HH}T\}, \{\underline{HH}H\}\} \text{ (first two tails)} \\ \sigma(\mathcal{A}) &= \{\Omega, \emptyset, \{\underline{HH}T\}, \{\underline{HH}H\}, \{\underline{HH}T, \underline{HH}H\}\} \end{aligned}$$

- $\sigma -$ algebra $\mathcal{A}$ over an uncountable set Ω

$$A_n \in \mathcal{A} \implies A = \cup_n A_n \in \mathcal{A} \quad \text{(countable union stability)}$$

1.1.3 Probability of an outcome

$\longrightarrow$ finite number of outcomes

- $\omega \in \Omega = \{\omega_1, \dots, \omega_N\} \longrightarrow p(\omega) \in [0,1] \quad \text{s.t} \quad \sum_{i=1}^{N} p(\omega_i) = 1$

$$A \in \mathcal{A} \ (\subset \mathcal{P}(\Omega)) \longrightarrow \mathbb{P}(A) \stackrel{\text{def.}}{=} \sum_{i\,:\,\omega_i \in A} p(\omega_i)$$

$$\downarrow$$

$$(\Omega, \mathcal{A}, \mathbb{P}) \ \text{ Probability space}$$

Properties:

- $\mathbb{P}(\Omega) = 1$, $\mathbb{P}(\emptyset) = 0$.
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B)$
  $= \mathbb{P}(A - B) + \mathbb{P}(B - A) + \mathbb{P}(A \cup B)$.
- $A \cap B = \emptyset \Longrightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
- $\mathbb{P}(\Omega - A) = 1 - \mathbb{P}(A)$.

$\longrightarrow$ uncountable number of outcomes $\sim$ countably additive probability

$$A_n \in \mathcal{A} \ (\sigma - \text{algebra}) \longrightarrow \mathbb{P}(\cup_n A_n) = \sum_n \mathbb{P}(A_n)$$

$$\downarrow$$

Caratheodory extension theo.

$\mathbb{P}$ countably additive on an algebra $\mathcal{B}$
$\Longrightarrow \exists!$ countably additive extension on the σ-algebra $\mathcal{A}$ generated by $\mathcal{B}$

$$\downarrow \ \text{(Lebesgue integration)}$$

$$(\Omega, \mathcal{A}, \mathbb{P}) \ \text{ Probability space}$$

1.1.4 Classical models

Example 1.7 (equiprobable outcomes)

$$\begin{aligned}\Omega &= \{\omega_1, \ldots, \omega_N\} \\ p(\omega_1) &= \ldots = p(\omega_N) = \frac{1}{N} \Longrightarrow \mathbb{P}(A) = \frac{|\{i \ : \ \omega_i \in A\}|}{N}\end{aligned}$$

Suppose

$$\Omega = \Omega_o^r = \{\omega \ : \ \omega = (a_1, \ldots, a_n)\,,\ a_i = 1, \ldots, M\} \quad \text{and} \quad N = |\Omega| = M^n.$$

If

$$A = \text{"no repetition"} = \{\omega \ : \ \omega = (a_1, \ldots, a_n)\,,\ a_i \neq a_j\}$$

then

$$\begin{aligned}\mathbb{P}(A) = \frac{|\{i \ : \ \omega_i \in A\}|}{N} &= \frac{M(M-1)\ldots(M-n+1)}{M^n} \\ &= \left(1 - \frac{1}{M}\right)\left(1 - \frac{2}{M}\right)\ldots\left(1 - \frac{n-1}{M}\right)\end{aligned}$$

Example 1.8 (Uniform outcomes) (**on** $[0,1]$)

$$\Omega = [0,1]\,, \qquad \mathcal{A} = \mathbf{B}([0,1])\,, \qquad \mathbb{P}(du) = 1_{[0,1]}(u)\ du.$$

$$0 \le a \le b \le 1 \implies \mathbb{P}([a,b]) = \int_a^b \mathbb{P}(du) = b - a$$

Example 1.9 (Gaussian outcomes) $(m \in \mathbb{R}, \sigma > 0)$

$$\Omega = \mathbb{R}\,, \qquad \mathcal{A} = \mathbf{B}(\mathbb{R})\,, \qquad \mathbb{P}(du) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(u-m)^2}{2\sigma^2}}\; du.$$

Example 1.10 (Exponential outcomes) $(\lambda > 0)$

$$\Omega = \mathbb{R}_+ = [0, \infty)\,, \qquad \mathcal{A} = \mathbf{B}(\mathbb{R}_+)\,, \qquad \mathbb{P}(du) = 1_{\mathbb{R}_+}(u)\; \lambda\; e^{-\lambda u}\; du.$$

Example 1.11 (Poissonnian outcomes) $(\lambda > 0)$

$$\Omega = \mathbb{N}\,, \qquad \mathcal{A} = \mathcal{P}(\mathbb{N})\,, \qquad \mathbb{P}(n) = \frac{\lambda^n}{n!}\; e^{-\lambda}.$$

Binomial $\Omega = \{\omega \ : \ \omega = (a_1, \ldots, a_n)\,,\ a_i \in \{0,1\}\}$ [failure/success]

$$\omega = (a_1, \ldots, a_n) \in \Omega \ \longrightarrow \ p(\omega) = p^{\sum_{i=1}^n a_i} \ q^{n - \sum_{i=1}^n a_i} \in [0,1]\,, \qquad p + q = 1$$

For $k = 0, \ldots, n$ we have that

$$A_k = \left\{ \omega = (a_1, \ldots, a_n) \ : \ \sum_{i=1}^n a_i = k \right\} \implies \mathbb{P}(A_k) = C_n^k \ p^k \ q^{n-k}$$

Multinomial $\Omega = \{\omega \ : \ \omega = (a_1, \ldots, a_n)\,,\ a_i \in \{b_1, \ldots, b_r\}\}$

$$\begin{aligned} \omega = (a_1, \ldots, a_n) \in \Omega \ \longrightarrow \ & \nu_i(\omega) = |\{j \ : \ a_j = b_i\}| \\ p(\omega) \ = \ & p_1^{\nu_1(\omega)} \ldots \ p_r^{\nu_r(\omega)} \in [0,1]\,, \qquad p_1 + \ldots + p_r = 1 \end{aligned}$$

For $n_1 + \ldots + n_r = n$ we set

$$C_n(n_1, \ldots, n_r) = \{\omega \ : \ \nu_1(\omega) = n_1, \ldots, \nu_r(\omega) = n_r\}\,.$$

Then we have that

$$\begin{aligned} |C_n(n_1, \ldots, n_r)| &= C_n^{n_1} \times C_{n-n_1}^{n_2} \times \ldots \times C_{n-(n_1+\ldots+n_{r-1})}^{n_r} \\ &= \frac{n!}{n_1! \ldots n_r!} \end{aligned}$$

and therefore

$$\mathbb{P}(C_n(n_1, \ldots, n_r)) = \frac{n!}{n_1! \ldots n_r!} \ p^{n_1} \ldots p^{n_r}\,.$$

Problem 1 (Poisson's Theorem)

Prove $$\lim_{n\to\infty} np(n) = \lambda > 0 \implies \lim_{n\to\infty} C_n^k \; p(n)^k \; (1 - p(n))^{n-k} = \frac{\lambda^k}{k!} \; e^{-\lambda}$$

Solution 1

$$p(n) = \frac{\lambda}{n} + o(\frac{1}{n})$$
$$\Downarrow$$

$$C_n^k \; p(n)^k \; (1-p(n))^{n-k}$$

$$= \frac{1}{k!} \; (n(n-1)\dots(n-k+1)) \; \left(\frac{\lambda}{n} + o(\frac{1}{n})\right)^k \; \left(1 - \frac{\lambda}{n} + o(\frac{1}{n})\right)^{n-k}$$

Note

$$(n(n-1)\dots(n-k+1)) \; \left(\frac{\lambda}{n} + o(\frac{1}{n})\right)^k$$

$$= \frac{(n(n-1)\dots(n-k+1))}{n^k} \; (\lambda + o(1))^k \longrightarrow \lambda^k$$

and

$$\left(1 - \frac{\lambda}{n} + o(\frac{1}{n})\right)^{n-k} \longrightarrow \exp -\lambda$$

and the end of the proof is clear. ∎

1.1.5 Random variables

$$\boxed{X : (\Omega, \mathcal{A}) \longrightarrow (E, \mathcal{E}) \text{ state space } (E = \mathbb{N},\ \mathbb{R},\ \mathbb{R}^d, \ldots)}$$

Example 1.12 (2-tosses) $\Omega = \{HH, HT, TH, TT\}$

- $X(\omega) =$*number of heads in* ω

$$X(HH) = 2 \qquad X(HT) = X(TH) = 1 \qquad X(TT) = 0$$

- $X = 1_A$ *with* $A \in \mathcal{A}$.

$$\begin{array}{rcl} X : (\Omega, \mathcal{A}) & \longrightarrow & (E, \mathcal{E}) \\ X^{-1}(A) = \{\omega \in \Omega \ : \ X(\omega) \in A\} & \longleftarrow & A \in \mathcal{E} \qquad (\textit{measurable}) \end{array}$$

Image probability measure on $(E, \mathcal{E})$*:*

$$\boxed{\mathbb{P}^X(A) = \mathbb{P}(X^{-1}(A))}$$

Independence: $(X, Y) : \Omega \to (E \times F)$ *independent r.v.*

$$\boxed{\begin{array}{rcl} \mathbb{P}(X^{-1}(A) \cap Y^{-1}(B)) & = & \mathbb{P}(X^{-1}(A)) \times \mathbb{P}(Y^{-1}(B)) \\ \mathbb{P}^{(X,Y)} & = & \mathbb{P}^X \otimes \mathbb{P}^Y \end{array}}$$

Bernoulli r.v. $X : \Omega \to E \qquad (p+q=1)$

- $E = \{0,1\}$

$$\mathbb{P}^X(\{0\}) = \mathbb{P}(X^{-1}(\{0\})) = p \quad \text{and} \quad \mathbb{P}^X(\{1\}) = \mathbb{P}(X^{-1}(\{1\})) = q.$$

- $E = \mathbb{R}$, $\forall I = [a,b] \in \mathcal{E} = \mathbf{B}(\mathbb{R})$ Borel subsets of $\mathbb{R}$.

$$\mathbb{P}^X(I) = \mathbb{P}(X^{-1}(I)) = \int_\Omega 1_I(X(\omega))\ \mathbb{P}(d\omega) = \int_I \mathbb{P}^X(dx)$$

with

$$\boxed{\mathbb{P}^X(dx) = p\ \delta_0(dx) + \delta_1(dx)}$$

Binomial r.v. $X : \Omega \to E \qquad (p+q=1)$

- $E = \{0,1,\ldots,n\}$

$$\mathbb{P}^X(\{k\}) = \mathbb{P}(X^{-1}(\{k\})) = C_n^k\ p^k\ q^{n-k}\ , \qquad k = 0,\ldots,n.$$

- $E = \mathbb{R}$, $\forall I = [a,b] \in \mathcal{E} = \mathbf{B}(\mathbb{R})$ Borel subsets of $\mathbb{R}$.

$$\mathbb{P}^X(I) = \mathbb{P}(X^{-1}(I)) = \int_\Omega 1_I(X(\omega))\ \mathbb{P}(d\omega) = \int_I \mathbb{P}^X(dx)$$

with

$$\boxed{\mathbb{P}^X(dx) = \sum_{k=0}^n C_n^k\ p^k\ q^{n-k}\ \delta_k(dx)}$$

- $\Omega = \{\omega\ :\ \omega = (a_1,\ldots,a_n)\ ,\ a_i \in \{0,1\}\}$ [failure/success]

$$\omega = (a_1,\ldots,a_n) \in \Omega \ \longrightarrow\ p(\omega) = p^{\sum_{i=1}^n a_i}\ q^{n-\sum_{i=1}^n a_i} \in [0,1]\ , \qquad p+q=1$$

and

$$X(\omega) = \sum_{i=1}^n a_i \in E = \{0,\ldots,n\},\ \mathbb{N},\ \mathbb{R},\ldots$$

- $(Y_i)_{i=1,\ldots,n}$ i.i.d Bernoulli with $\mathbb{P}(Y_1 = 1) = 1 - \mathbb{P}(Y_1 = 0) = p$

$$X = Y_1 + \ldots + Y_n$$

1.2 Elementary estimation methods

1.2.1 Expectation/mean value

Discrete r.v.:

$$X : \Omega \to E = \{x_1, \ldots, x_n\} \subset \mathbb{R} \quad \text{with} \quad \mathbb{P}(X = x_i) = \mathbb{P}^X(x_i) = p_i$$

$$\Longrightarrow \qquad \mathbb{E}(X) = \sum_{i=1}^{n} x_i \ \mathbb{P}(X = x_i)$$

Taking $E = \mathbb{R}$ we write alternatively

$$\mathbb{E}(X) = \int_\Omega X(\omega) \ \mathbb{P}(d\omega) = \int_\mathbb{R} u \ \mathbb{P}^X(du)$$

with

$$\mathbb{P}^X(du) = \sum_{i=1}^{n} p_i \ \delta_{x_i}(du)$$

Real valued r.v.:

$$\mathbb{E}(X) = \int_\Omega X(\omega) \ \mathbb{P}(d\omega) = \int_\mathbb{R} u \ \mathbb{P}^X(du)$$

E-valued r.v.: E abstract measurable space $(E = \mathbb{N}^d \ , \ \mathbb{R}^d, \ldots)$

$$X : (\Omega, \mathcal{A}) \to (E, \mathcal{E})$$

For any bounded Borel-test function $f : E \to \mathbb{R}$

$$\begin{array}{rcccl} f(X) = f \circ X & : & (\Omega, \mathcal{A}) & \longrightarrow & (\mathbb{R}, \mathbf{B}(\mathbb{R})) \\ & & \omega & \mapsto & f(X(\omega)) \end{array}$$

and the change of variable formula (in Lebesgue integrals)

$$\boxed{\mathbb{E}(f(X)) \ = \ \int_\Omega f(X(\omega)) \ \mathbb{P}(d\omega) = \int_E f(x) \ \mathbb{P}^X(dx) = \langle f, \mathbb{P}^X \rangle_{C_b(E), M_1(E)}}$$

with the duality bracket between the set of all probability measures $M_1(E)$ on E and the Banach space $C_b(E)$ of bounded continuous test functions on E.

1.2.2 Variance/dispersion and Covariance/correlation

$\longrightarrow$ X, Y real valued r.v

$$\begin{aligned} Var(X) &= \mathbb{E}((X-\mathbb{E}(X))^2) \\ Cov(X,Y) &= \mathbb{E}((X-\mathbb{E}(X)\ )(Y-\mathbb{E}(Y))) = \mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y) \end{aligned}$$

$$\inf_{a\in\mathbb{R}} \mathbb{E}((X-a)^2) \Longrightarrow a = \mathbb{E}(X)$$

and

$\inf_{(a,b)\in\mathbb{R}^2} \mathbb{E}((X-(a+bY))^2)$

$\Longrightarrow \mathbb{E}(X) = a + b\ \mathbb{E}(Y)$ and $\mathbb{E}((X-(a+bY))Y) = 0$

$\Longrightarrow b = \frac{Cov(X,Y)}{Var(Y)}$ and $a = \mathbb{E}(X) - \frac{Cov(X,Y)}{Var(Y)}\ \mathbb{E}(Y)$

$$\Downarrow$$

Mean square linear estimator

$$\boxed{\widehat{X} = \mathbb{E}(X) + \frac{Cov(X,Y)}{Var(Y)}\ (Y-\mathbb{E}(Y))}$$

1.2.3 Law of large numbers

$$\longrightarrow X^1, \ldots, X^n, \ldots \textbf{i.i.d. Bernoulli} \quad \mathbb{P}(X^i = 1) = p = 1 - \mathbb{P}(X^i = 0)$$

Pb: Using $(X^n)_{n \geq 0}$ find parameter p ?

Sol: [Monte Carlo estimate]

$$\boxed{S_N = \frac{1}{N} \sum_{i=1}^{N} X^i}$$

$$\mathbb{E}(S_N) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}(X^i) = p \qquad \text{(unbiased)}$$

Numerical errors:

Do we have $\forall \epsilon > 0 \quad \forall \omega \in \Omega \qquad |S_N(\omega) - p| \leq \epsilon \quad ?? \longrightarrow$ NO !!

$$\downarrow$$

$$\begin{aligned} \mathbb{P}(S_N = 0) &= \mathbb{P}(X^1 = \ldots = X^N = 0) = q^N \to 0 \\ \mathbb{P}(S_N = 1) &= \mathbb{P}(X^1 = \ldots = X^N = 1) = p^N \to 0 \end{aligned}$$

$\mathbb{L}_2$-mean error estimate:

$$\begin{aligned} \mathbb{E}((S_N - p)^2) &= \frac{1}{N} Var(X^1) = \frac{pq}{N} \\ &\Downarrow \qquad (\mathbb{P}(|X| \geq \epsilon) \leq \frac{1}{\epsilon^2} \mathbb{E}(|X|^2)) \\ \mathbb{P}(|S_N - p| \geq \epsilon) &\leq \frac{pq}{N\epsilon^2} \to 0 \end{aligned}$$

$$\longrightarrow X^1, \ldots, X^N, \ldots \textbf{i.i.d. r.v.} \sim \quad \mathbb{P}(X \in dx) = \mu(dx) \text{ proba. on } (E, \mathcal{E})$$

Pb: Using $(X^n)_{n \geq 0}$ find distribution μ ?

Note:

$$\text{pb} \iff \text{find } \forall f \in \mathcal{C}_b(E) \quad \mu(f) = \int_E f(x)\ \mu(dx)$$

Sol: [Monte Carlo estimate] $\forall f : E \to \mathbb{R}$ bounded we set

$$\boxed{S_N(f) = \frac{1}{N} \sum_{i=1}^{N} f(X^i)}$$

$$\mathbb{E}(S_N(f)) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}(f(X^i)) = \mu(f) \qquad \text{(unbiased)}$$

Numerical errors: [$\mathbb{L}_2$-mean error estimate]

$$\mathbb{E}((S_N(f) - \mu(f))^2) \;=\; \frac{1}{N}\, Var(f(X)) \leq \frac{\|f\|^2}{N}$$

$$\Downarrow \qquad (\mathbb{P}(|X| \geq \epsilon) \leq \frac{1}{\epsilon^2}\ \mathbb{E}(|X|^2))$$

$$\mathbb{P}(|S_N(f) - \mu(f)| \geq \epsilon) \;\leq\; \frac{\|f\|^2}{N\epsilon^2} \to 0$$

1.2.4 Conditional probabilities

$$\longrightarrow \quad \textbf{of event } A \textbf{ assuming event } B\textbf{:} \qquad \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Example 1.13 (family with two children) *(**B**oys or **G**irls?)*

Pb:

$$\mathbb{P}(\text{both children are boys}|\text{the older is a boy}) \quad ??$$
$$\mathbb{P}(\text{both children are boys}|\text{at least one is a boy}) \quad ??$$

Proba. space construction:

$$\Omega = \{BB, BG, GB, GG\} \quad (BG = (\text{Boy older, Girl younger})) \quad \mathbb{P}(XX) = \frac{1}{4}$$

$$\begin{aligned} A &= \text{the older child is a Boy} = \{BB, BG\} \\ B &= \text{the younger child is a Boy} = \{BB, GB\} \end{aligned}$$

$$\Downarrow$$

$$\begin{aligned} A \cup B &= \text{at least one child is a Boy} \\ A \cap B &= \text{both children are Boys} \end{aligned}$$

Sol:

$$\mathbb{P}(\text{both children are boys}|\text{the older is a boy})$$
$$= \mathbb{P}(A \cap B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(BB)}{\mathbb{P}(BB, BG)} = \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{4}} = \frac{1}{2}$$

$$\mathbb{P}(\text{both children are boys}|\text{at least one is a boy})$$
$$= \mathbb{P}(A \cap B|A \cup B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A \cup B)} = \frac{\mathbb{P}(BB)}{\mathbb{P}(BB, BG, GB)} = \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{4} + \frac{1}{4}} = \frac{1}{3}$$

$\longrightarrow$ **finite or countable conditioning** $(\Omega, \mathcal{A}, \mathbb{P})$ proba. space

$$\begin{aligned} \mathcal{D} &= \{D_1, \ldots, D_d\} \text{ decomposition of } \Omega \\ & \quad (\mathbb{P}(D_i) > 0\,, \quad D_i \cap D_j = \emptyset, \quad \cup_i D_i = \Omega) \end{aligned}$$

$$\downarrow \quad (\forall A \in \mathcal{A})$$

$$\boxed{\begin{aligned} \mathbb{P}(A|\mathcal{D}) : (\Omega, \mathcal{D}) &\longrightarrow [0,1] \\ \omega &\mapsto \mathbb{P}(A|\mathcal{D})(\omega) = \sum_{i=1}^{d} \mathbb{P}(A|D_i)\ 1_{D_i}(\omega) \end{aligned}}$$

Example 1.14 (X,Y independent Bernoulli)

$$\mathbb{P}(X=1) = \mathbb{P}(Y=1) = p \qquad \mathbb{P}(X=0) = \mathbb{P}(Y=0) = q$$

Decomposition $\mathcal{D} = \{D_0, D_1\}$ *associated to* Y *that is*

$$D_0 = \{Y=0\} \qquad D_1 = \{Y=1\}$$

$$\begin{aligned} \mathbb{P}(X+Y=k|\mathcal{D}) &= \mathbb{P}(X+Y=k|Y) \\ &= \mathbb{P}(X+Y=k|D_0)\ 1_{D_0} + \mathbb{P}(X+Y=k|D_1)\ 1_{D_1} \\ &= \mathbb{P}(X+Y=k|Y=0)\ 1_0(Y) + \mathbb{P}(X+Y=k|Y=1)\ 1_1(Y) \end{aligned}$$

Then

$$\begin{aligned} \mathbb{P}(X+Y=k|\mathcal{D}) &= \frac{\mathbb{P}(X+0=k, Y=0)}{\mathbb{P}(Y=0)}\ 1_0(Y) + \frac{\mathbb{P}(X+1=k, Y=1)}{\mathbb{P}(Y=1)}\ 1_1(Y) \\ &= \mathbb{P}(X=k)\ 1_0(Y) + \mathbb{P}(X=k-1)\ 1_1(Y) \end{aligned}$$

and therefore

$$\mathbb{P}(X+Y=k|Y) = \begin{cases} q\ 1_0(Y) & \text{if } k=0 \\ p\ 1_0(Y) + q\ 1_1(Y) & \text{if } k=1 \\ p\ 1_1(Y) & \text{if } k=2 \end{cases}$$

Using elementary events:

$$\Omega = \{(0,0),(0,1),(1,0),(1,1)\} \qquad \mathcal{A} = \mathcal{P}(\Omega)$$

$$\mathbb{P}(\{(0,0)\}) = q^2 \quad \mathbb{P}(\{(0,1)\}) = pq = \mathbb{P}(\{(1,0)\}) \qquad \mathbb{P}(\{(1,1)\}) = p^2$$

$$\downarrow$$

independent Bernoulli r.v.

$$X(\omega_1,\omega_2) = \omega_1 \quad \text{and} \quad Y(\omega_1,\omega_2) = \omega_2$$

$$\begin{aligned}
\mathbb{P}(X=1) &= \mathbb{P}(\{(1,0),(1,1)\}) = pq + p^2 = p(p+q) = p \\
\mathbb{P}(X=0) &= \mathbb{P}(\{(0,1),(0,0)\}) = pq + q^2 = q(p+q) = q
\end{aligned}$$

Note

$$D_0 = \{Y=0\} = \{(1,0),(0,0)\} \qquad D_1 = \{Y=1\} = \{(0,1),(1,1)\}$$

and

$$\{X+Y=k\} = \begin{cases} \{(0,0)\} & \text{if } k=0 \\ \{(1,0),(0,1)\} & \text{if } k=1 \\ \{(1,1)\} & \text{if } k=2 \end{cases}$$

This yields for instance

$$\begin{aligned}
\mathbb{P}(X+Y=0|D_0) &= \frac{\mathbb{P}(\{(0,0)\})}{\mathbb{P}(\{(1,0),(0,0)\})} = \frac{q^2}{pq+q^2} = \frac{q}{p+q} = q \\
\mathbb{P}(X+Y=1|D_0) &= \frac{\mathbb{P}(\{(1,0)\})}{\mathbb{P}(\{(1,0),(0,0)\})} = \frac{qp}{pq+q^2} = \frac{p}{p+q} = p \\
\mathbb{P}(X+Y=2|D_0) &= 0
\end{aligned}$$

Note:

$$\begin{aligned}
\sigma(\mathcal{D}) &= \sigma(Y) = \{\Omega,\emptyset,D_0,D_1\} = \{\Omega,\emptyset,\{(1,0),(0,0)\},\{(0,1),(1,1)\}\} \\
&\subset \mathcal{A} \qquad (\text{ex.:} \quad \{(1,0)\} \notin \sigma(\mathcal{D}) \to \textit{missing info on } X)
\end{aligned}$$

$\longrightarrow$ **Abstract conditioning w.r.t** $\mathcal{B} \subset \mathcal{A}$ with$(\Omega, \mathcal{A}, \mathbb{P})$ **proba. space**

$$\downarrow$$

For any $A \in \mathcal{A}$**,** $\mathbb{P}(A|\mathcal{B}) : (\Omega, \mathcal{B}) \to [0,1]$ **such that:**

$$\forall B \in \mathcal{B}\,, \qquad \mathbb{P}(A \cap B) = \int_B \mathbb{P}(A|\mathcal{B})(\omega)\ \mathbb{P}(d\omega) \qquad (1)$$

Note:

$$\begin{aligned}(1) \iff & \forall B \in \mathcal{B} \quad \mathbb{E}(1_A\ 1_B) = \mathbb{E}(\mathbb{E}(1_A|\mathcal{B})\ 1_B) \\ \iff & \forall B \in \mathcal{B} \quad \mathbb{E}([1_A - \mathbb{E}(1_A|\mathcal{B})]\ 1_B) = 0 \iff \mathbb{E}(1_A|\mathcal{B}) = \mathrm{proj}_{L_2(\Omega,\mathcal{B})}(1_A)\end{aligned}$$

$\longrightarrow$ **fund. ex.** $(X,Y) : (\Omega, \mathcal{A}) \to (E \times F, \mathcal{E} \otimes \mathcal{F}) \quad (\mathcal{A} = \sigma(X,Y)\,, \quad \mathcal{B} = \sigma(Y))$

- $|E| \vee |F| < \infty$ **and** $\mathbb{P}(X = x, Y = y)$ **given:**

$$\boxed{\mathbb{P}(X = x|Y) = \sum_{y \in F} \mathbb{P}(X = x|Y = y)\ 1_y(Y)}$$

$$\mathbb{P}(X = x|Y = y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}$$

and

$$\mathbb{P}(Y = y) = \sum_{x \in E} \mathbb{P}(X = x, Y = y)$$

- $(E,\mathcal{E})$, $(F,\mathcal{F})$ **abstract measurable spaces** ($\mathbb{N}^2, \mathbb{N}^p \times \mathbb{R}^q$,...) **and given**

$$\mathbb{P}((X,Y) \in d(x,y)) = p^{X,Y}(x,y)\ m^X(dx)\ m^Y(dy)$$

for some measures m^X, m^Y **and some density** $p^{X,Y}$**:**

$$\mathbb{P}(X \in dx | Y = y) = p^{X|Y}(x|y)\ m^X(dx)$$

$$\Downarrow$$

$$\boxed{p^{X|Y}(x|y) = \frac{p^{X,Y}(x,y)}{p^Y(y)} \quad \text{and} \quad p^Y(y) = \int_E p^{X,Y}(x,y)\ m^X(dx)}$$

Example 1.15

$$E = F = \mathbb{R} \qquad m^X(du) = m^Y(du) = du \quad \text{(Lebesgue meas.)}$$

$$\downarrow$$

$$\mathbb{P}((X,Y) \in d(x,y)) = p^{X,Y}(x,y) \; dx \; dy$$

For instance

$$p^{X,Y}(x,y) = \frac{1}{2\pi} \exp\left[-\frac{1}{2}(y - h(x))^2 - \frac{1}{2}(x-m)^2\right]$$

for some $h : \mathbb{R} \to \mathbb{R}$ *and* $m \in \mathbb{R}$ *leads to*

$$\begin{aligned} p^X(x) &= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(x-m)^2\right] \\ p^{Y|X}(y|x) &= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(y-h(x))^2\right] \end{aligned}$$

and

$$p^{X|Y}(x|y) = \frac{p^{Y|X}(y|x)}{p^Y(y)} \times p^X(x)$$

with

$$p^Y(y) = \int p^{Y|X}(y|x) \; p^X(x) \; dx$$

Problem 2

ξ^i *i.i.d. with distribution* η, $\|f\| \leq 1$,

- $\eta^N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\xi^i} \quad \Longrightarrow \quad \mathbb{E}\left([\eta^N(f) - \eta(f)]^4\right) \leq \frac{c}{N^2}$

- $\mathbb{P}(|\eta^N(f) - \eta(f)| \geq 1/p) \leq \frac{cp^4}{N^2}$

- *If we set*

$$\begin{aligned} A_N(p) &= \{|\eta^N(f) - \eta(f)| \geq 1/p\} \\ A(p) &= \cap_{N\geq 1} \cup_{M\geq N} A_N(p) \end{aligned}$$

  then we have that

$$\{\lim_{N\to\infty} |\eta^N(f) - \eta(f)| > 0\} = \cup_{p\geq 1} A(p)$$

  and

$$\mathbb{P}(A(p)) = \lim_{N\to\infty} \mathbb{P}\left(\cup_{M\geq N} A_N(p)\right) \leq \lim_{N\to\infty} \sum_{M\geq N} \mathbb{P}(A_M(p)) = 0.$$

- *Deduce that* $\lim_{N\to\infty} \eta^N(f) = \eta(f)$.

Solution 2 • *If we set $\tilde{f}(x) = f(x) - \eta(f)$ we have that*

$$
\begin{aligned}
&[\eta^N(f) - \eta(f)]^4 = \eta^N(\tilde{f})^4 \\
&= \frac{1}{N^4} \sum_{i_1,i_2,i_3,i_4=1}^{N} \tilde{f}(\xi^{i_1})\tilde{f}(\xi^{i_2})\tilde{f}(\xi^{i_3})\tilde{f}(\xi^{i_4}) \\
&= \frac{1}{N^4}\sum_{i=1}^{N} \tilde{f}^4(\xi^i) + \frac{1}{N^4}\ \frac{4!}{3!1!}\sum_{i\neq j} \tilde{f}^3(\xi^i)\tilde{f}(\xi^j) + \frac{1}{N^4}\ \frac{4!}{2!2!}\sum_{i<j} \tilde{f}^2(\xi^i)\tilde{f}^2(\xi^j) \\
&+\frac{1}{N^4}\ \frac{4!}{2!1!1!} \sum_{\substack{i\not\in\{j,k\}\\ j<k}} \tilde{f}^2(\xi^i)\tilde{f}(\xi^j)\tilde{f}(\xi^k) + \frac{1}{N^4}\ (4!) \sum_{i<j<k<l} \tilde{f}(\xi^i)\tilde{f}(\xi^j)\tilde{f}(\xi^k)\tilde{f}(\xi^l).
\end{aligned}
$$

Since $\eta(\tilde{f}) = 0$, for distinct indexes i, j, k, l we have that

$$
\mathbb{E}\left(\tilde{f}^3(\xi^i)\tilde{f}(\xi^j)\right) = \mathbb{E}\left(\tilde{f}^2(\xi^i)\tilde{f}(\xi^j)\tilde{f}(\xi^k)\right) = \mathbb{E}\left(\tilde{f}(\xi^i)\tilde{f}(\xi^j)\tilde{f}(\xi^k)\tilde{f}(\xi^l)\right) = 0
$$

Thus,

$$
\begin{aligned}
\mathbb{E}\left([\eta^N(f) - \eta(f)]^4\right) &= \frac{1}{N^4}\ \left(N\ \eta(\tilde{f}^4) + 6N(N-1)\eta(\tilde{f}^2)^2\right) \\
&\leq \frac{c}{N^2}\ \|f\|^4 \leq \frac{c}{N^2}
\end{aligned}
$$

- *By construction we have that*

$$
\mathbb{E}(|\eta^N(f) - \eta(f)|^4) \geq \mathbb{E}(|\eta^N(f) - \eta(f)|^4\ 1_{A_N(p)}) \geq \frac{1}{p^4}\ \mathbb{P}(A_N(p))
$$

- *By monotone convergence theorem we have*

$$
\mathbb{P}(A(p)) = \lim_{N\to\infty} \mathbb{P}\left(\cup_{M\geq N} A_N(p)\right)
$$

and using the above estimates

$$
\lim_{N\to\infty} \sum_{M\geq N} \mathbb{P}(A_M(p)) = 0
$$

- *Since $\{\lim_{N\to\infty} |\eta^N(f) - \eta(f)| > 0\} = \cup_{p\geq 1} A(p)$ and $\mathbb{P}(A(p)) = 0$ the end of the proof is now clear.* ■

1.3 Markov chains

$\longrightarrow$ iterative sequence of r.v. $(X_0 \to X_1 \to \ldots \to X_{n-1} \to X_n \ldots)$

Markov property:

Given the states $X_0, \ldots, X_{n-1}$ the current r.v. X_n only depends on X_{n-1}.

$$\downarrow$$

$$\boxed{\mathbb{P}(X_n \in dx_n | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) = \mathbb{P}(X_n \in dx_n | X_{n-1} = x_{n-1})}$$

1.3.1 Random dynamical systems

$$\longrightarrow \qquad X_n = F_n(X_{n-1}, W_n), \qquad n \geq 1$$

for some initial condition $X_0 = x_0$ and some i.i.d. r.v $(W_n)_{n\geq 0}$ in some state spaces.

$$\mathbb{P}(X_n \in dx_n | X_{n-1} = x_{n-1}) = \mathbb{P}(F_n(x_{n-1}, W_n) \in dx_n)$$

Example 1.16 (linear state space models in $\mathbb{R}^p$)

$$X_n = F_n \ X_{n-1} + G_n \ W_n, \qquad n \geq 1$$

for some i.i.d. r.v X_0, $(W_n)_{n\geq 0}$ in $\mathbb{R}^q$ and matrices F_n and G_n with appropriate dimensions.

$$G_n = Id \, , \ W_n \sim g(w)dw \Longrightarrow \mathbb{P}(X_n \in dx_n | X_{n-1} = x_{n-1}) = g(x_n - F_n x_{n-1}) \ dx_n$$

Example 1.17 (random walk on the real line)

$$X_n = X_{n-1} + W_n, \qquad n \geq 1$$

for some i.i.d. r.v X_0, $(W_n)_{n\geq 0}$ and on the half line

$$X_n = (X_{n-1} + W_n)^+, \qquad n \geq 1$$

with $a^+ = a \vee 0$

Example 1.18 (simple bilinear model)

$$X_n = a\ X_{n-1} + b\ X_{n-1}\ W_n + V_n, \qquad n \geq 1$$

for some i.i.d. r.v X_0, $(W_n, V_n)_{n\geq 0}$.

Example 1.19 (random bits)

$$E = \{(a_1, \ldots, a_p)\ :\ a_i = 0, 1\} \qquad |E| = 2^p$$

Given $X_{n-1} = (a_1, \ldots, a_p)$ *we define the next r.v.* X_n*:*

1. *We choose an index* $I_n = q$ *uniformly in* $\{1, \ldots, p\}$*, that is*

$$\forall q \in E\ , \qquad \mathbb{P}(I_n = q | X_n) = \frac{1}{p}$$

2. *Given the position index* $I_n = q$ *we flip randomly the element* a_q *and we set*

$$X_{n-1} = (a_1, \ldots, a_{q-1}, a_q + \epsilon_n, a_{q+1}, \ldots, a_p)$$

with +-mod(2) and

$$\mathbb{P}(\epsilon_n = 0) = 1/2 = \mathbb{P}(\epsilon_n = 1)$$

1.3.2 Simulated annealing

$\longrightarrow \quad U : E \to \mathbb{R}_+$ potential , $\forall x$, $K(x, dy)$ proba., $T > 0$ fixed temperature.

Given $X_{n-1} = x$ the next random state X_n is defined as follows

1. First we propose a r.v. Y_n with law $K(x, dy)$, that is

$$\mathbb{P}(Y_n \in dy | X_{n-1} = x) = K(x, dy)$$

2. Given $Y_n = y$ we next proceed in two steps

   (a) If $U(y) \le U(x)$ we accept the proposal $Y = y$ and we set $X_n = y$.

   (b) If $U(y) > U(x)$ we sample uniformly in $[0, 1]$ a r.v. U_n

      i. If $U_n \le \exp -([U(y) - U(x)]/T)$ we set $X_n = y$.

      ii. If $U_n \ge \exp -([U(y) - U(x)]/T)$ we reject $Y = y$ and set $X_n = x$.

Nice K **,** $T_n = T_0/log(n + e)$**,** T_0 **large** $\Rightarrow U(X_n) \to \inf_E U$ **in proba.**

Example 1.20 (Commercial traveler)

N cities $\mathcal{C} = \{c_1, \ldots, c_N\}$ with a distance function $d : \mathcal{C} \times \mathcal{C} \to \mathbb{R}_+$.

Obj: *Find $\sigma = \begin{pmatrix} 1 & \ldots & N \\ \sigma(1) & \ldots & \sigma(N) \end{pmatrix} \in E = \mathcal{G}_N$ with minimal distance*

$$U(\sigma) = d(c_{\sigma(1)}, c_{\sigma(2)}) + \ldots + d(c_{\sigma(N-1)}, c_{\sigma(N)}) + d(c_{\sigma(N)}, c_{\sigma(1)})$$

Proposal=index permutation:

$$K(\sigma, \tau) = \frac{1}{|\mathcal{V}(\sigma)|} \; 1_{\mathcal{V}(\sigma)}(\tau)$$

with $\mathcal{V}(\sigma) = \{\tau \in \mathcal{G}_N \; ; \; \tau \sim \sigma\}$. and

$$\sigma \;\sim\; \tau \iff \exists i \leq j \; : \; \sigma = \begin{pmatrix} 1 & \ldots & i & \ldots & j & \ldots N \\ \tau(1) & \ldots & \tau(j) & \ldots & \tau(i) & \ldots \tau(N) \end{pmatrix}$$

1.3.3 Iterated random functions

$\longrightarrow$ **Chaos, fractals and noise, Stochastic aspects of dynamics.**
A. Lasota, M. C. Mackey.
Applied Mathematical Sciences 97, second edition,
Springer-Verlag New-York, 1994.

$\longrightarrow$ **Iterated random functions**
P. Diaconis and D. Freedman.
SIAM Review, Vol. 41, No. 1, pp. 45-76, 1999.

$$\boxed{X_n = S_{\epsilon_n}(X_{n-1})}$$

ϵ_n i.i.d. Bernoulli $\mathbb{P}(\epsilon_n = 1) = 1 - \mathbb{P}(\epsilon_n = 0) = 0.2993$, $S_i(x) = A_i.x + b_i$ with

$$A_0 = \begin{pmatrix} +0.4000 & -0.3733 \\ +0.0600 & +0.6000 \end{pmatrix} \qquad b_0 = \begin{pmatrix} +0.3533 \\ +0.0000 \end{pmatrix}$$

and

$$A_1 = \begin{pmatrix} -0.8000 & -0.1867 \\ +0.1371 & +0.8000 \end{pmatrix} \qquad b_1 = \begin{pmatrix} +1.1000 \\ +0.1000 \end{pmatrix}$$

1.3.4 Robbins Monro algorithm

State space $E = \mathbb{R}^p$

$$X_n = X_{n-1} + \gamma_n \ \left(a - U(X_{n-1}, Y_n)\right), \qquad n \geq 1.$$

for some $U : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^p$, Y_n i.i.d. $\sim \ \mu(dx)$ on $\mathbb{R}^q$, some level $a \in \mathbb{R}^p$ and some parameter $\gamma > 0$.

Reg. conditions and $\sum_{n\geq 0} \gamma_n = \infty$ and $\sum_{n\geq 0} \gamma_n^2 < \infty$

$$\Longrightarrow X_n \to x_a \qquad \text{where} \quad a = \int U(x_a, y) \ \mu(dy)$$

For $p = q = 1$, $a = \frac{1}{2}$ and $U(x, y) = 1_{(x \leq y)}$

$$\mathbb{P}(Y \geq x_{\frac{1}{2}}) = \frac{1}{2} \qquad \text{and} \qquad x_{\frac{1}{2}} = \text{median}(Y)$$

$$\downarrow$$

Equivalent ways to define a Markov chain:

- Random dynamical system with drift function F_n, random perturbations W_n and initial condition X_0.
- Complicated random process/stochastic algorithm.
- Abstract transition distribution

$$K_n(x_{n-1}, dx_n) = \mathbb{P}(X_n \in dx_n | X_{n-1} = x_{n-1})$$

Notations:

$$\mathbb{P}(X_n \in dx_n | X_{n-1} = x_{n-1}) = K_n(x_{n-1}, dx_n)$$

$$\Downarrow$$

$$\mathbb{P}(X_0 \in dx_0, \ldots, X_n \in dx_n)$$

$$= \mathbb{P}(X_n \in dx_n | X_{n-1} = x_{n-1}, \ldots X_0 = x_0)\ \mathbb{P}(X_0 \in dx_0, X_1 \in dx_1, \ldots, X_{n-1} \in dx_{n-1})$$

$$= \mathbb{P}(X_n \in dx_n | X_{n-1} = x_{n-1}) \ldots \mathbb{P}(X_1 \in dx_1 | X_0 = x_0) \mathbb{P}(X_0 \in dx_0)$$

$$\Downarrow$$

$$\boxed{\mathbb{P}((X_0, \ldots, X_n) \in d(x_0, \ldots, x_n)) = \eta_0(dx_0) K_1(x_0, dx_1) \ldots K_n(x_{n-1}, dx_n)}$$

On test functions f:

$$\boxed{\mathbb{E}(f(X_n)|X_{n-1}=x) = \int K_n(x,dy)\ f(y) \stackrel{\text{def.}}{=} K_n(f)(y)}$$

$$\Rightarrow \mathbb{E}(f(X_{n+p})|X_n = x) = K_{n+1}K_{n+2}\ldots K_{n+p}(f)(x)$$

Transition K on measures η:

$$\boxed{\eta(f) = \int \eta(dy)\ f(y) \quad \text{and} \quad (\eta K)(dy) = \int \eta(dx)\ K(x,dy)}$$

With $\eta_n \stackrel{\text{def.}}{=} \text{Law}(X_n)$ this yields

$$\begin{aligned}
\eta_n(f) &= \mathbb{E}(f(X_n)) = \mathbb{E}(\mathbb{E}(f(X_n)|X_{n-1})) \\
&= \int \eta_{n-1}(dx)\ \mathbb{E}(f(X_n)|X_{n-1}=x) \\
&= \int \eta_{n-1}(dx)\ K_n(f)(x) = (\eta_{n-1}K_n)(f)
\end{aligned}$$

$$\Longrightarrow \qquad \eta_n = \eta_{n-1}K_n = \eta_0 K_1 K_2 \ldots K_n$$

Problem 3 *When $E = \{1, \ldots, N\}$ we identify the set of all probability measures on E with the unit simplex*

$$\Delta_N = \{(p(1), \ldots, p(N)) \in \mathbb{R}^N \ ; \ \forall 1 \leq i \leq N, \ \ p(i) \geq 0, \ \ \sum_{i=1}^{N} \ p(i) = 1\}$$

and

$$p(dx) = \sum_{i=1}^{N} \ p(i) \ \delta_i(dx) \longleftrightarrow p = (p(1), \ldots, p(N)) \in \Delta_N$$

Let X_n be a Markov chain on E with transitions K_n.

- *Check that*

$$(\eta_n(1), \ldots, \eta_n(N)) = (\eta_{n-1}(1), \ldots, \eta_{n-1}(N)) \begin{pmatrix} K_n(1,1) & \cdots & K_n(1,N) \\ \vdots & \vdots & \vdots \\ K_n(N,1) & \cdots & K_n(N,N) \end{pmatrix}.$$

- *If we identify*

$$f(x) = \sum_{i=1}^{N} \ f(i) \ 1_i(x) \longleftrightarrow f = \begin{pmatrix} f(1) \\ \vdots \\ f(N) \end{pmatrix}$$

  then show that

$$\begin{pmatrix} K_n(1,1) & \cdots & K_n(1,N) \\ \vdots & \vdots & \vdots \\ K_n(N,1) & \cdots & K_n(N,N) \end{pmatrix} \begin{pmatrix} f(1) \\ \vdots \\ f(N) \end{pmatrix} = \begin{pmatrix} K_n(f)(1) \\ \vdots \\ K_n(f)(N) \end{pmatrix}.$$

Solution 3 *The assertions are simple consequences of the formulae*

$$\begin{aligned} \eta_n &= \eta_{n-1}K_n(=\mathrm{Law}(X_n)) \\ K_n(f)(x) &= \int_E K_n(x,dy)\ f(y)(=\mathbb{E}(f(X_n)|X_{n-1}=x)) \end{aligned}$$

2 Simulation methods

$\longrightarrow$ **Non uniform random variable generation**
L. Devroye , *Springer 1986.*

Simulation/simulate: feign, pretend to be, look like, effect fictitiously,....

$\downarrow$

Simulate a uniform r.v. on $[0,1]$ consists in sampling **a sequence of numbers** U_n between 0 and 1 s.t.:

- The sequence is un-predictable and numbers occur randomly and **independently** of each other.
- The distribution is **uniform** on $[0,1]$ and there is $1/2$ of the numbers in $[0,1/2]$, $(b-a)$ numbers in $[a,b]$,....

2.1 Uniform distributions

Hyp (perfect unif. sampling):

$$\exists (U_n)_{n\geq 0} \qquad \textbf{i.i.d.} \quad \sim \quad \mathbb{P}(U \in du) = 1_{[0,1]}(u)\ du$$

$$\downarrow$$

$\forall n \geq 1\,, \qquad (U_1, \ldots, U_n) \qquad$ **uniform r.v. in the** n**-cube** $[0,1]^n$

Uniform r.v on a Borel set $A \subset \mathbb{R}^d$ **s.t.** $0 < vol(A) = \int_A dx < \infty$**.:**

$$\mathbb{P}(X \in dx) = \frac{1}{vol(A)}\ 1_A(x)\ dx.$$

Example 2.1 ($\rightarrow$ geometric distribution) *Note that for U_n i.i.d. on $[0,1]$ and $p \in [0,1]$ we have that*

$$\begin{aligned} p(1-p)^{n-1} &= \mathbb{P}(U_n \in [0,p[, \text{ and } U_1, \ldots, U_{n-1} \in [p,1]) \\ &= \mathbb{P}(U_n \in [0,p[) \prod_{k=1}^{n-1} \mathbb{P}(U_k \in [p,1]). \end{aligned}$$

Then

$$X = \inf\{n \geq 1\ ;\ U_n \in [0,p[\}$$

is a geometric r.v.

$$\forall n \geq 1, \qquad \mathbb{P}(X = n) = \ p\ (1-p)^{n-1}$$

2.2 Discrete and multinomial

Discrete r.v.: X taking values $\{x_n\ ;\ n \geq 1\} \subset (E, \mathcal{E})$:

$$\mathbb{P}(X_n \in dx) = \sum_{n \geq 1} p_n\ \delta_{x_n}(dx) \qquad (\Leftrightarrow \mathbb{P}(X_n = x_n) = p_n).$$

Decomposition $\quad [0,1[= \cup_{n \geq 1} I_n\,, \quad$ with $\quad dx(I_n) = p_n$

$\downarrow \qquad$ (ex: $\quad I_n = [p_1 + \cdots + p_{n-1}, p_1 + \cdots + p_n[$)

$$\boxed{\sum_{n \geq 1} x_n\ 1_{I_n}(U) = x_m \iff U \in I_m.}$$

$\downarrow$

U unif. r.v. on $[0,1]$

$\Longrightarrow X = \sum_{n \geq 1}\ x_n\ 1_{I_n}(U) \quad$ discrete r.v $\quad \mathbb{P}(X \in dx) = \sum_{n \geq 1}\ p_n\ \delta_{x_n}(dx).$

Example 2.2 (Bernoulli and geometric)

- *With* U unif. r.v. on $[0,1]$

$$X = 1_{[0,p[}(U) \Rightarrow \mathbb{P}(X=1) = p = 1 - \mathbb{P}(X=0)$$

- *With* X_n *i.i.d. Bernoulli*

$$Y = \inf\{n \geq 1\ ;\ X_n = 1\} \Longrightarrow \forall n \geq 1\,, \quad \mathbb{P}(Y=n) = p\ (1-p)^{n-1}.$$

Example 2.3 (Uniform on a finite set)

State space $E = \{x_1, \ldots, x_n\}$ *and* U *unif. r.v. on* $[0,1]$

$$X = \sum_{m=1}^{n} x_m\ 1_{[\frac{m-1}{n}, \frac{m}{n}[}(U) = x_{1+[nU]} \Longrightarrow \mathbb{P}(X = x_i) = \frac{1}{n}$$

with $[a]$ *integer part of* $a \in \mathbb{R}$.

Let $N \geq 1$, $p_1, \ldots, p_n \geq 0$ s.t. $\sum_{i=1}^n p_i = 1$.

$$\begin{array}{rcl} M & = & \text{Multinomial}\,(N; p_1, \ldots, p_n) \\ & \Updownarrow & \\ \mathbb{P}\,(M = m) & = & \dfrac{N!}{m_1! \ldots m_n!}\, p_1^{m_1} \ldots p_n^{m_n} \end{array}$$

for all $m \in \Delta = \{(m_1, \ldots, m_n) \in \mathbb{N}^n \ ; \ \sum_{i=1}^n m_i = N\}$.

Example 2.4 (Binomial) *(n=2, $M_1 = k \Rightarrow M_2 = N - k$)*

$$\forall 0 \leq k \leq N\,, \quad \mathbb{P}\,(M_1 = k) = \frac{N!}{k!\ (N-k)!}\, p_1^k\ (1 - p_1)^{N-k}$$

With U_n i.i.d. uniform on $[0,1]$

$$B = \sum_{i=1}^N\ 1_{[0,p]}(U_i) \Rightarrow \mathbb{P}\,(B = m) = \frac{N!}{m!\ (N-m)!}\ p^m\ (1-p)^{N-m}.$$

Discrete $\to$ multinomial

Let U_i i.i.d. uniform on $[0,1]$, $I_k = [p_1 + \ldots + p_{k-1}, p_1 + \ldots + p_k[$

$$\begin{array}{rcl} X_i & = & \displaystyle\sum_{k=1}^n\ x_k\ 1_{I_k}(U_i) \qquad (\Rightarrow \mathbb{P}(X_i = x_k) = p_k) \\ M_k & = & \text{Card}\{1 \leq i \leq N\ ;\ X_i = x_k\} \end{array}$$

$$\Longrightarrow M = (M_1, \ldots, M_n) = \text{Multinomial}\,(N; p_1, \ldots, p_n)$$

Uniform N-order statistics on $[0,1]$:
Reordering $U_1,\ldots,U_N$ i.i.d. uniform on $[0,1]$ we get an N-statistics

$$V_i = U_{\sigma_u(i)}$$

for some index permutation $\sigma_u \in \mathcal{G}_N$ such that $U_{\sigma_u(1)} \leq \ldots \leq U_{\sigma_u(N)}$.

With $X_i = \sum_{k=1}^{n} \ x_k \ 1_{I_k}(U_i)$ (U_i unif. i.i.d. on $[0,1]$):

$$(X_1,\ldots,X_N)$$

$$= \sum_{m\in\Delta} (\underbrace{x_1,\ldots,x_1}_{m_1-\text{times}},\ldots,\underbrace{x_n,\ldots,x_n}_{m_n-\text{times}}) \ 1_{\underbrace{I_1\times\ldots\times I_1}_{m_1-\text{times}}\times\ldots\times\underbrace{I_n\times\ldots\times I_n}_{m_n-\text{times}}}(V_1,\ldots,V_N).$$

$$\Downarrow$$

Simulation algorithm

$k \longleftarrow 1$
$i \longleftarrow 1$
While $\quad i \leq N$

$$\begin{cases} \text{If} & V_i < p_1 + \cdots + p_k \ \text{ then } \ (X_i = x_k \ \text{ and } \ i \longleftarrow i+1) \\ \text{Otherwise} & (M_k \longleftarrow i-1 \ \text{ and } \ k \longleftarrow k+1). \end{cases}$$

2.3 Change of variables

$$(\mathcal{D},\ \mathcal{D}' \text{ open subsets in} \mathbb{R}^d) \qquad \begin{array}{rcl} \theta :\ \mathcal{D} & \longrightarrow & \mathcal{D}' \subset \mathbb{R}^d \\ x & \mapsto & \theta(x) = \left(\theta^1(x), \ldots, \theta^d(x)\right) \end{array}$$

$$\downarrow \quad (\theta \text{ diffeomorphism (i.e. } \theta, \theta \text{ diff. bijection))}$$

$$\boxed{\int_{\mathcal{D}} f(x)\ dx = \int_{\theta(\mathcal{D})} f(\theta^{-1}(y))\ |\mathrm{Jac}(\theta^{-1})(y)|\ dy}$$

$x = \theta^{-1}(y)$

$$\Rightarrow (dx^1, \ldots, dx^d)' = (d\theta_1^{-1}(y), \ldots, d\theta_d^{-1}(y))' = \begin{pmatrix} \frac{\partial \theta_1^{-1}}{\partial y^1} & \cdots & \frac{\partial \theta_1^{-1}}{\partial y^d} \\ \vdots & \cdots & \vdots \\ \frac{\partial \theta_d^{-1}}{\partial y^1} & \cdots & \frac{\partial \theta_d^{-1}}{\partial y^d} \end{pmatrix} (dy^1, \ldots, dy^d)'.$$

X $\mathcal{D}$-valued r.v., $\theta : \mathcal{D} \to \theta(\mathcal{D})$ diffeo.

$$\mathbb{P}\left(X \in dx\right) \ = \ p_X(x)\ 1_{\mathcal{D}}(x)\ dx$$

$$\Downarrow$$

$$\mathbb{P}\left(\theta(X) \in dy\right) \ = \ p_X(\theta^{-1}(y))\ |\mathrm{Jac}(\theta^{-1})(y)|\ 1_{\theta(\mathcal{D})}(y)\ dy.$$

Example 2.5 *X uniformly distributed in $\mathcal{D} \subset \mathbb{R}^d$*

$$\begin{aligned} \mathbb{P}\left(X \in dx\right) &= \frac{1}{vol(\mathcal{D})}\ 1_{\mathcal{D}}(x)\ dx \\ \mathbb{P}\left(\theta(X) \in dy\right) &= \frac{1}{vol(\mathcal{D})}\ |Jac(\theta^{-1})(y)|\ 1_{\theta(\mathcal{D})}(y)\ dy. \end{aligned}$$

Example 2.6 *U_n i.i.d. unif. on $[0,1]$*

$$(V_1, \ldots, V_n) = (a_1 + (b_1 - a_1)U_1, \ldots, a_n + (b_n - a_n)U_n)$$

is uniformly distributed on $[a_1, b_1] \times \cdots \times [a_n, b_n]$.

$$U, U' \quad \text{i.i.d. unif. on } [0,1] \Rightarrow \begin{cases} X &= \sqrt{U'}\ \cos(2\pi U) \\ Y &= \sqrt{U'}\ \sin(2\pi U) \end{cases}$$

is unif. distributed on the unit disk $\mathcal{D}$.

Proof:

$$\begin{aligned} (x,y) &= \theta(u',u) = (\sqrt{u'}\cos(2\pi u), \sqrt{u'}\sin(2\pi u)) \\ (u',u) &= \theta^{-1}(x,y) = (x^2+y^2, \frac{1}{2\pi}\mathrm{Arctan}(\frac{y}{x})) \end{aligned}$$

$$\Rightarrow |\mathrm{Jac}(\theta^{-1})(x,y)| = |\begin{pmatrix} 2x & 2y \\ \frac{1}{2\pi}\frac{-y}{x^2+y^2} & \frac{1}{2\pi}\frac{x}{x^2+y^2} \end{pmatrix}| = \frac{1}{\pi} = \frac{1}{\mathrm{vol}(\mathcal{D})}$$

$$\downarrow$$

$\frac{1}{2}$-circle distribution:

$$\mathbb{P}(X \in dx) = \mathbb{P}(X \in dx\ ,\ Y \in]-1,1[) = \frac{2}{\pi}\ \sqrt{1-x^2}\ 1_{]-1,1[}(x)\ dx.$$

Box-Muller Gaussian simulation:

If (U,V) are i.i.d. unif. on $[0,1]$ then

$$X = \sqrt{-2\log U}\ \cos(2\pi V) \quad \text{and} \quad Y = \sqrt{-2\log U}\ \sin(2\pi V)$$

are independent and Gaussian r.v.

Proof:

$$\begin{aligned} (x,y) &= \theta(u,v) = \left(\sqrt{-2\log u}\ \cos(2\pi v), \sqrt{-2\log u}\ \sin(2\pi v)\right) \\ (u,v) &= \theta^{-1}(x,y) = \left(e^{-\frac{1}{2}(x^2+y^2)}, \frac{1}{2\pi}\ \arctan(y/x)\right). \end{aligned}$$

$$\Rightarrow \quad |\mathrm{Jac}(\varphi^{-1})(x,y)| = |\begin{pmatrix} -xe^{-\frac{1}{2}(x^2+y^2)} & -ye^{-\frac{1}{2}(x^2+y^2)} \\ \frac{1}{2\pi}\frac{-y}{x^2+y^2} & \frac{1}{2\pi}\frac{x}{x^2+y^2} \end{pmatrix}| = \frac{1}{2\pi}\ e^{-\frac{1}{2}(x^2+y^2)}.$$

2.4 Inversion

Repartition function: (X real valued)

$$F(x) = \mathbb{P}\left(X \leq x\right).$$

If F is strictly increasing then

$$U \quad \text{unif. on } [0,1] \Rightarrow \mathbb{P}\left(F^{-1}(U) \leq x\right) = \mathbb{P}(U \leq F(x)) = \int_0^{F(x)} du = F(x).$$

Otherwise

Generalized inverse $F^{-}(y) = \inf\{x \in \mathbb{R}\ ;\ F(x) \geq y\}$.

$$\begin{aligned} F \text{ continuous} &\implies F^{-}(y) = \inf\{x \in \mathbb{R}\ ;\ F(x) = y\} \\ F \text{ strict. increasing} &\implies F^{-} = F^{-1}. \end{aligned}$$

$$\Downarrow$$

Inversion method:

$$U \quad \text{unif. on } [0,1] \Rightarrow \quad X = F^{-}(U) \quad \text{r.v. with repartition } F$$

Example 2.7 (Exponential(λ))

$$\mathbb{P}(X \in dx) = \lambda\ e^{-\lambda\ x}\ 1_{[0,+\infty[}(x)\ dx \qquad (\lambda > 0)$$

Since

$$F(x) = 1 - e^{-\lambda\ x} \Rightarrow F^{-1}(y) = -\frac{1}{\lambda} \log\left(1-y\right)$$

we have

$$U \quad \text{unif. on } [0,1] \Rightarrow \quad -\frac{1}{\lambda} \log U \quad \sim \quad \text{expo}(\lambda).$$

Example 2.8 (geometric(p))

$$X \sim \exp(\lambda) \quad \text{with} \quad \lambda = -\log(1-p)$$
$$\Downarrow$$
$$\mathbb{P}(n-1 \leq X < n) = \lambda \int_{n-1}^{n} e^{-\lambda\ x}\ dx = p\ (1-p)^{n-1}$$
$$\Downarrow$$
$$Y = 1 + [X] \sim \text{geometric}(p)$$

Example 2.9 (Cauchy(σ))

$$\mathbb{P}(X \in dx) = \frac{\sigma}{\pi\ (x^2+\sigma^2)}\ dx.$$

Since

$$F(x) \;=\; \frac{1}{2} + \frac{1}{\pi}\arctan\frac{x}{\sigma} \quad \text{and} \quad F^{-1}(y) = \sigma\ \tan\left(\pi\ (y-\frac{1}{2})\right)$$

we have

$$U \quad \text{unif. on } [0,1] \Rightarrow \quad \sigma\ \tan\left(\pi\ (U-\frac{1}{2})\right) \sim \text{Cauchy}(\sigma)$$

Example 2.10 (Rayleigh(σ))

$$\mathbb{P}(X \in dx) = \frac{x}{\sigma^2}\ e^{-\frac{x^2}{2\sigma^2}}\ 1_{[0,\infty[}(x)\ dx.$$

Since

$$F(x) \;=\; 1 - \exp\left[-\frac{x^2}{2\sigma^2}\right] \quad \text{and} \quad F^{-1}(y) = \sigma\ \sqrt{-\log(1-y)}.$$

we have

$$U \quad \text{unif. on } [0,1] \Rightarrow \quad \sigma\ \sqrt{-\log U} \sim \text{Rayleigh}(\sigma)$$

Example 2.11 (Pareto(a,b)) *($(a,b) \in \mathbb{R}\times]0,\infty[$)*

$$\mathbb{P}(X \in dx) = \frac{a\ b^a}{x^{a+1}}\ 1_{[b,\infty[}(x)\ dx.$$

Since

$$F(x) \ = \ 1 - \left(\frac{b}{x}\right)^a \quad \text{and} \quad F^{-1}(y) = \frac{b}{(1-y)^{\frac{1}{a}}}$$

we have

$$U \quad \text{unif. on } [0,1] \Rightarrow \quad \frac{b}{U^{\frac{1}{a}}} \sim \text{Pareto}(a,b)$$

Example 2.12 (Triangular)

$$\mathbb{P}(X \in dx) = \frac{2}{a}\ \left(1 - \frac{x}{a}\right)\ 1_{[0,a]}(x)\ dx$$

Since

$$F(x) \ = \ \frac{2}{a}\ \left(x - \frac{x^2}{2a}\right) \quad \text{and} \quad F^{-1}(y) = a\ \left(1 - \sqrt{1-y}\right)$$

we have

$$U \quad \text{unif. on } [0,1] \Rightarrow \quad \left(a\ \left(1 - \sqrt{U}\right)\right) \sim \text{triangular}$$

Example 2.13 (Weibull(a,b)) *$(a, b > 0)$*

$$\mathbb{P}(X \in dx) = \frac{a}{b^a}\ x^{a-1}\ e^{-(x/b)^a}\ 1_{[0,\infty[}(x)\ dx.$$

Since

$$\begin{aligned} F(x) \ &= \ \int_0^x \frac{\partial}{\partial z}(-e^{-(z/b)^a})\ dz = 1 - e^{-(x/b)^a} \\ F^{-1}(y) \ &= \ b\ (-\log(1-y))^{\frac{1}{a}} \end{aligned}$$

we have

$$U \quad \text{unif. on } [0,1] \Rightarrow \quad \left(b\ (-\log U)^{\frac{1}{a}}\right) \sim \text{Weibull}(a,b)$$

Problem 4 *Propose a way to simulate a Bernoulli r.v.*

$$\mathbb{P}(X \in dx) = p \ \delta_1(dx) + (1-p) \ \delta_0(dx).$$

and a discrete r.v

$$\mathbb{P}(X \in dx) = \sum_{i=1}^{n} \ p_i \delta_{x^i}(dx)$$

Solution 4 *The Bernoulli repartition function is clearly given by*

$$F(x) = \begin{cases} 0 & \text{si} \;\; x \in]-\infty, 0[\\ 1-p & \text{si} \;\; x \in [0,1[\\ 1 & \text{si} \;\; x \in [1, \infty[\end{cases}$$

We also prove easily that its generalized inverse is given by

$$F^{-}(y) = 1_{]1-p,1]}(y) = \begin{cases} 0 & \text{si} \;\; y \in [0, 1-p] \\ 1 & \text{si} \;\; y \in]1-p, 1]. \end{cases}$$

The Bernoulli r.v. is simply given by

$$1_{]1-p,1]}(U) \qquad \text{with} \quad \mathbb{P}(U \in du) = 1_{[0,1]}(u)du$$

There is no loss of generality to assume that $x_1 \leq \cdots \leq x_n$. *We first check that*

$$F(x) = \sum_{i=1}^{n} p_i \; 1_{[x_i, \infty[}(x) = \begin{cases} 0 & \text{si} \;\; x \in]-\infty, x_1[\\ p_1 & \text{si} \;\; x \in [x_1, x_2[\\ p_1 + p_2 & \text{si} \;\; x \in [x_2, x_3[\\ \vdots & \vdots \;\; \vdots \\ p_1 + \ldots + p_i & \text{si} \;\; x \in [x_i, x_{i+1}[\\ \vdots & \vdots \;\; \vdots \\ 1 & \text{si} \;\; x \in [x_n, \infty[\end{cases}$$

and

$$F^{-}(y) = \sum_{i=1}^{n} x_i \; 1_{]\sum_{k=1}^{i-1} p_k, \sum_{k=1}^{i} p_k]}(y) = \begin{cases} x_1 & \text{si} \;\; y \in]0, p_1] \\ x_2 & \text{si} \;\; y \in]p_1, p_1 + p_2] \\ \vdots & \vdots \;\; \vdots \\ x_i & \text{si} \;\; y \in]p_1 + \ldots p_{i-1}, p_1 + \ldots p_i] \\ \vdots & \vdots \;\; \vdots \\ x_n & \text{si} \;\; y \in]p_1 + \ldots p_{n-1}, 1] \end{cases}$$

with the convention $\sum_{\emptyset} = 0$. *Then the r.v.*

$$\sum_{i=1}^{n} x_i \; 1_{]\sum_{k=1}^{i-1} p_k, \sum_{k=1}^{i} p_k]}(U) \qquad \text{with} \quad \mathbb{P}(U \in du) = 1_{[0,1]}(u)du$$

is the desired discrete r.v.

2.5 Rejection

Absolute continuity of meas. on $(E, \mathcal{E})$:

$$\mu \prec \nu \iff (\forall A \in \mathcal{A} \quad \nu(A) = 0 \Rightarrow \mu(A) = 0)$$

$$\Downarrow$$

Radon-Nykodim derivative:

$$\exists ! \frac{d\mu}{d\nu}(x) \; ; \quad \nu(A) = \int_A \frac{d\mu}{d\nu}(x) \; \nu(dx)$$

(uniqness mod. ν-null sets)

Example 2.14 (discrete)

$$\begin{aligned}
(q_i = 0 \Rightarrow p_i = 0) \implies & \; \mu(dx) = \sum_{i=1}^{n} p_i \; \delta_{x_i} \prec \nu(dx) = \sum_{i=1}^{n} q_i \; \delta_{x_i} \\
\implies & \; \frac{d\mu}{d\nu}(x) = \sum_{i:\nu(x_i)>0} \frac{p_i}{q_i} \; 1_{x_i}(x)
\end{aligned}$$

In particular $p_i \leq \mathrm{C} \; q_i \Rightarrow \mu \prec \nu$ *and*

$$q_i \vee p_i > 0 \implies \mu \prec \nu \quad \text{and} \quad \frac{d\mu}{d\nu}(x_i) = \frac{p_i}{q_i}$$

Example 2.15 (continuous)

$$\begin{aligned}
f(x) \vee g(x) > 0 \implies & \; \mu(dx) = f(x) \; m(dx) \prec \nu(dx) = g(x) \; m(dx) \\
& \; (\text{ex.:} \quad E = \mathbb{E}^d \quad \text{and} \quad m(dx) = dx) \\
\implies & \; \frac{d\mu}{d\nu}(x) = \frac{f(x)}{g(x)}
\end{aligned}$$

Example 2.16 (Gauss$\prec$Cauchy)

$$\mu(dx) = \frac{1}{\sqrt{2\pi}} \; e^{-\frac{x^2}{2}} \; dx \prec \nu(dx) = \frac{1}{\pi \; (x^2+1)} \; dx$$

$$\Rightarrow \frac{d\mu}{d\nu}(x) = \sqrt{\frac{\pi}{2}} \; (x^2+1) \; e^{-\frac{x^2}{2}}$$

Condition:

$$(\mu \textbf{ and } \nu \textbf{ proba. on } (E,\mathcal{E})) \qquad \mu \prec \nu \qquad \text{and} \qquad \frac{d\mu}{d\nu}(x) \leq m \ (<\infty)$$

$$\Downarrow$$

Rejection method:

Let U_n i.i.d. uniform on $[0,1]$, X_n i.i.d. $\sim\ \nu$ and

$$T_m = \inf\{n \geq 1 \ ; \text{ such that } \ (m\ U_n) < \frac{d\mu}{d\nu}(X_n)\},$$

Then (X_{T_m}, T_m) are independent with

$$\text{Law}(X_{T_m}) = \mu \qquad \text{and} \quad \mathbb{P}(T_m = n) = \frac{1}{m}\ \left(1 - \frac{1}{m}\right)^{n-1}.$$

Note: the mean number of runs to obtain X_{T_m} is $\mathbb{E}(T_m) = m$.

Proof:(sketched)
Using

$$\begin{aligned}
\mathbb{P}(mU_n < \frac{d\mu}{d\nu}(X_n), X_n \in A) &= \mathbb{E}(1_A(X_n)\mathbb{P}(mU_n < \frac{d\mu}{d\nu}(X_n)|X_n)) \\
&= \frac{\mu(A)}{m}
\end{aligned}$$

and

$$\begin{aligned}
\{X_{T_m} \in A\} &= \cup_{n\geq 1} \{X_n \in A\,, T_m = n\} \\
\{X_n \in A\,, T_m = n\} &= \{X_n \in A\,, m\ U_n < \frac{d\mu}{d\nu}(X_n)\} \cap \cap_{p=1}^{n-1}\{m\ U_p \geq \frac{d\mu}{d\nu}(X_p)\}
\end{aligned}$$

we obtain

$$\mathbb{P}(X_{T_m} \in A) \ = \ \frac{1}{m}\ \mu(A) \sum_{n\geq 1} \left(1 - \frac{1}{m}\right)^{n-1} = \frac{1}{m}\ \mu(A)\ m = \mu(A)$$

Example 2.17 (Gauss≺Cauchy)

$$\mu(dx) = \frac{1}{\sqrt{2\pi}}\; e^{-\frac{x^2}{2}}\; dx \prec \nu(dx) = \frac{1}{\pi\;(x^2+1)}\; dx.$$

Since

$$\frac{d\mu}{d\nu}(x) = \sqrt{\frac{\pi}{2}}\;(x^2+1)\; e^{-\frac{x^2}{2}} \quad \text{and} \quad \frac{\partial}{\partial x}\left(\frac{d\mu}{d\nu}(x)\right) = -\sqrt{\frac{\pi}{2}}\;(x^2-1)\; e^{-\frac{x^2}{2}}.$$

the rejection condition holds with $m = \sqrt{\frac{2\pi}{e}}$.

Example 2.18 *Let* $\mathcal{D}(a,b,c)$ *be the set of all proba. densities* $p(x) \leq c$ *with compact support* $[a,b] \subset \mathbb{R}$.

$$p \in \mathcal{D}(a,b,c) \Rightarrow p(x) \leq c\; 1_{[a,b]}(x) = m\; q(x)$$

with $m = (c(b-a))$ *and* $q(x) = \frac{1}{b-a}\; 1_{[a,b]}(x)$.

Example 2.19 (1/2-circle≺ unif.)

$$\mu(dx) = \frac{2}{\pi}\;\sqrt{1-x^2}\; 1_{[-1,1]}(x)\; dx \prec \nu(dx) = \frac{1}{2}\; 1_{[-1,1]}(x)\; dx$$

The rejection condition holds with $m = 4/\pi$.

Simulation of conditional distributions:

For μ proba. on $(E, \mathcal{E})$ and $\mu(A) > 0$ we set

$$\mu_{|A}(dx) = \frac{1}{\mu(A)}\ 1_A(x)\ \mu(dx)$$

Let X_n i.i.d. $\sim \mu$ and $T_A = \inf\{n \geq 1\ ;\ \text{s.t. } X_n \in A\}$.

$$\Rightarrow \text{Law}(X_{T_A}) = \mu_{|A} \qquad \text{and} \qquad \mathbb{P}(T_A = n) = \mu(A)\ \ (1-\mu(A))^{n-1}$$

Proof:

$$\frac{d\mu_{|A}}{d\mu}(x) = \frac{1}{\mu(A)}\ 1_A(x) \leq \frac{1}{\mu(A)}$$

Example 2.20

$$A \subset B \subset \mathbb{R}^d \implies \mu(dx) = \frac{1}{vol(B)}\ 1_B(x)\ dx \prec \mu_{|A}(dx) = \frac{1}{vol(A)}\ 1_A(x)\ dx$$

$$\text{and} \quad \mu(A) = \frac{vol(A)}{vol(B)}$$

Example 2.21 *With* $= [-1,1]^2$ *and* $A = \{(x,y) \in \mathbb{R}^2\ ;\ x^2+y^2 < 1\}$ *we have*

$$\begin{aligned}
\mu(d(x,y)) &= \frac{1}{4}\ 1_B((x,y))\ dx\ dy \\
\nu(d(x,y)) &= \frac{1}{\pi}\ 1_A((x,y))\ dx\ dy = \mu_{|A}(d(x,y))
\end{aligned}$$

2.6 Exponential inter-times/Poisson and n-order statistics

S_n **i.i.d. expo(λ):**

$$\mathbb{P}(S_n \in ds) \quad = \quad \lambda\ e^{-\lambda\ s}\ 1_{[0,\infty[}(s)\ ds$$

$$\downarrow$$

$$T_n = S_1 + \ldots + S_n \ \text{ so that } \ T_n - T_{n-1} = S_n.$$

$$\downarrow$$

$$\boxed{\mathbb{P}(T_{n+1} \in dt) = \lambda\ e^{-\lambda t}\ \frac{(\lambda t)^n}{n!}\ 1_{[0,\infty[}(t)\ dt.}$$

($\lambda = 1 \Longrightarrow$ Law(T_{n+1}) =Gamma($n+1$))

If S_n i.i.d. $\sim$ exp(1) then

$$N = \inf\{n \geq 1\ ;\ S_1 + \ldots + S_{n+1} > \gamma\}$$

is Poisson(γ)

$$\forall n \geq 0\,, \qquad \mathbb{P}(N = n) = \frac{\gamma^n}{n!}\ e^{-\gamma}.$$

and

$$\left(\frac{T_1}{T_{n+1}}, \ldots, \frac{T_n}{T_{n+1}}\right) \sim \text{ n-order statistic on } [0,1]$$

2.7 Markov chain simulation

Simulation of a Markov chain

$$X_0 \longrightarrow \cdots \longrightarrow X_{n-1} \longrightarrow X_n \longrightarrow \cdots$$

with Law$(X_0) = \mu$ **probability transition** $K_n(x_{n-1}, dx_n)$:

- We choose randomly X_0 with distribution μ.
- The n-step transition $X_{n-1} \to X_n$ consists in sampling a random variable with distribution $K_n(X_{n-1}, .)$.

$$\Updownarrow$$

When the Markov chain is given by some recursion

$$X_n = F_n(X_{n-1}, W_n)$$

with X_0, W_n **an i.i.d. sequence:**

- We sample X_0 with distribution μ.
- At time n we sample a random variable W_n and we set

$$X_n = F_n(X_{n-1}, W_n)$$

Problem 5 *Let $(V_1, \ldots, V_N)$ be an uniform N-order statistics.*

- *Prove that*

$$\mathbb{P}\left((V_1, \ldots, V_N) \in d(v_1, \ldots, v_N)\right) = N! \; 1_{C_N(1)}(v_1, \ldots, v_N) \; dv_1 \; \cdots dv_N$$

  with

$$C_N(1) = \{(v_1, \ldots, v_N) \in [0,1]^N \; ; \; 0 < v_1 < \cdots < v_N < 1\}.$$

- *S_n i.i.d. expo(λ)*

$$\mathbb{P}(S_n \in ds) = \lambda \; e^{-\lambda \; s} \; 1_{[0,\infty[}(s) \; ds$$

  If we define

$$T_n = S_1 + \ldots + S_n$$

  then prove that

$$\mathbb{P}\left((T_1, \ldots, T_n) \in d(t_1, \ldots, t_n)\right) = \lambda^n \; e^{-\lambda \; t_n} \; 1_{C_n}(t_1, \ldots, t_n) \; dt_1 \ldots dt_n$$

  where

$$C_n = \{(t_1, \ldots, t_n) \in \mathbb{R}_+^n \; ; \; t_1 < \ldots < t_n\}.$$

- *We next define*

$$\begin{aligned} C_n(t) &= \{(t_1, \ldots, t_n) \in \mathbb{R}_+^n \; ; \; 0 < t_1 < \ldots < t_n < t\} \\ &= \{(t_1, \ldots, t_n) \in C_n \; ; \; t_n < t\} \end{aligned}$$

  and

$$I_n(t) = \int 1_{C_n(t)}(t_1, \ldots, t_n) \; dt_1 \ldots dt_n$$

  - *Prove that $I_n(t) = t^n/n!$*
  - *Deduce that*

$$\mathbb{P}(T_{n+1} \in dt) = \lambda \; e^{-\lambda t} \; \frac{(\lambda t)^n}{n!} \; 1_{[0,\infty[}(t) \; dt$$

- *Check that*

$$\mathbb{P}\left((T_1, \ldots, T_{n+1}) \in d(t_1, \ldots, t_{n+1})\right)$$

$$= \left(n!\ 1_{C_n(t_{n+1})}(t_1, \ldots, t_n)\ t_{n+1}^{-n}\ dt_1 \ldots dt_n\right)$$

$$\times\ \left(\lambda\ e^{-\lambda\ t_{n+1}}\ \tfrac{(\lambda\ t_{n+1})^n}{n!}\ 1_{[0,\infty[}(t_{n+1})\ dt_{n+1}\right)$$

- *Establish that*

$$\mathbb{P}\left((T_1/T_{n+1}, \ldots, T_n/T_{n+1}, T_{n+1}) \in d(v_1, \ldots, v_n, t)\right)$$

$$= \mathbb{P}\left((T_1/T_{n+1}, \ldots, T_n/T_{n+1}) \in d(v_1, \ldots, v_n)\right) \mathbb{P}\left(T_{n+1} \in dt\right).$$

and deduce that

$$\mathbb{P}\left(\left(\frac{T_1}{T_{n+1}}, \ldots, \frac{T_n}{T_{n+1}}\right) \in d(v_1, \ldots, v_n)\right) = n!\ 1_{C_n(1)}(v_1, \ldots, v_n)\ dv_1\ \cdots dv_n$$

Solution 5 • *We first suppose that*

$$(V_1,\dots,V_N) = (U_{\sigma(1)},\dots,U_{\sigma_u(N)})$$

$$\Downarrow$$

$$\mathbb{P}\left((V_1,\dots,V_N)\in d(v_1,\dots,v_N)\,;\ \sigma_u=\sigma\right)$$

$$= \mathbb{P}((U_{\sigma(1)},\dots,U_{\sigma(N)})\in d(v_1,\dots,v_N)\ ;\ U_{\sigma(1)}\le\ldots\le U_{\sigma(N)})$$

$$= 1_{C_N(1)}(v_1,\dots,v_N)\ dv_1\ \cdots dv_N$$

$$\Downarrow \quad (\sum_{\sigma} \to N!)$$

$$\mathbb{P}\left((V_1,\dots,V_N)\in d(v_1,\dots,v_N)\right) = N!\ 1_{C_N(1)}(v_1,\dots,v_N)\ dv_1\ \cdots dv_N$$

- *We apply the change of variable formula with*

$$\theta(S_1,\dots,S_n) = (T_1,\dots,T_n)$$

and the diffeomorphism

$$\begin{array}{ccc} \mathbb{R}_+^n & \longrightarrow & C_n \\ (s_1,\dots,s_n) & \xrightarrow{\theta} & (s_1,s_1+s_2,\dots,s_1+\cdots+s_n) \\ (t_1,t_2-t_1\dots,t_n-t_{n-1}) & \xleftarrow{\theta^{-1}} & (t_1,\dots,t_n) \end{array}$$

with Jacobian

$$\mathrm{Jac}(\theta^{-1})(t_1,\dots,t_n) = \left|\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \vdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & -1 & 1 \end{pmatrix}\right| = 1.$$

  – *By induction, noting that* $I_1(t) = \int_0^t ds = t$ *and*

$$\begin{aligned} \int_0^t I_n(t_{n+1})\ dt_{n+1} &= \int_0^t \left[\int 1_{C_n(t_{n+1})}(t_1,\dots,t_n)\ dt_1\dots dt_n\right] dt_{n+1} \\ &= \int 1_{C_{n+1}(t)}(t_1,\dots,t_n,t_{n+1})\ dt_1\dots dt_n\ dt_{n+1} \\ &= I_{n+1}(t). \end{aligned}$$

– *By definition we have that*

$$\begin{aligned}\mathbb{P}(T_{n+1} \in dt) &= \lambda^{n+1}\ e^{-\lambda t}\ 1_{[0,\infty[}(t)\ \left(\int_{C_n(t)} dt_1 \ldots dt_n\right)\ dt \\ &= \lambda\ e^{-\lambda t}\ \frac{(\lambda t)^n}{n!}\ 1_{[0,\infty[}(t)\ dt.\end{aligned}$$

- *We simply note that*

$$\mathbb{P}\left((T_1, \ldots, T_{n+1}) \in d(t_1, \ldots, t_{n+1})\right)$$

$$= 1_{C_{n+1}}(t_1, \ldots, t_{n+1})\ \lambda^{n+1}\ e^{-\lambda\ t_{n+1}}\quad dt_1 \ldots dt_{n+1}$$

$$= \left(n!\ 1_{C_n(t_{n+1})}(t_1, \ldots, t_n)\ t_{n+1}^{-n}\ dt_1 \ldots dt_n\right)\ \left(\lambda\ e^{-\lambda\ t_{n+1}}\ \tfrac{(\lambda\ t_{n+1})^n}{n!}\ 1_{[0,\infty[}(t_{n+1})\ dt_{n+1}\right)$$

- *We note that*

$$\mathbb{P}\left((T_1/T_{n+1}, \ldots, T_n/T_{n+1}, T_{n+1}) \in d(v_1, \ldots, v_n, t)\right)$$

$$= \left(n!\ 1_{C_n(1)}(v_1, \ldots, v_n)\ dv_1 \ldots dv_n\right)\ \left(\lambda\ e^{-\lambda\ t}\ \tfrac{(\lambda\ t)^n}{n!}\ 1_{[0,\infty[}(t)\ dt\right)$$

$$= \mathbb{P}\left((T_1/T_{n+1}, \ldots, T_n/T_{n+1}) \in d(v_1, \ldots, v_n)\right) \mathbb{P}\left(T_{n+1} \in dt\right).$$

We prove the above assertion applying the change of variable formula to

$$\Psi(T_1, \ldots, T_n, T_{n+1}) = (T_1/T_{n+1}, \ldots, T_n/T_{n+1}, T_{n+1})$$

with the diffeomorphism $\Psi : C_{n+1} \to \Psi(C_{n+1}) = C_n(1) \times]0, \infty[$

$$\begin{aligned}\Psi(t_1, \ldots, t_n, t) &= (t_1/t, \ldots, t_n/t, t) \\ \Psi^{-1}(v_1, \ldots, v_n, t) &= (tv_1, \ldots, tv_n, t)\end{aligned}$$

with Jacobian

$$\mathrm{Jac}\Psi^{-1}(v_1, \ldots, v_n, t) = \left| \begin{pmatrix} t & 0 & 0 & \cdots & 0 & v_1 \\ 0 & t & 0 & \vdots & 0 & v_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \ldots & t & v_n \\ 0 & 0 & \ldots & \ldots & 0 & 1 \end{pmatrix} \right| = t^n.$$

3 Filtering

(signal/observation) $= (X, Y) = \{(X_n, Y_n)\ ;\ n \geq 0\}$ **Markov chain**

$$
\begin{array}{ccccccccc}
X_0 & \longrightarrow & \ldots \longrightarrow & X_{n-1} & \longrightarrow & X_n & \longrightarrow & & \text{(signal)} \\
\downarrow & & \cdots & \downarrow & \cdots & \downarrow & \cdots & & \\
Y_0 & & & Y_{n-1} & & Y_n & & & \text{(observation)}
\end{array}
$$

Objective:

$\longrightarrow$ Law$(X_0, \ldots, X_n | Y_0, \ldots, Y_n)$ and $\longrightarrow$ Law$(X_p | Y_0, \ldots, Y_q)$

$$\downarrow$$

$p > q \rightarrow$ (prediction) , $p = q \rightarrow$ (filtering) , $p < q \rightarrow$ (smoothing)

- $\boxed{\textbf{Signal } X = \{X_n \; ; \; n \geq 0\}\textbf{:}}$

$\sim$ physical system we need to estimate/predict (Markov evolution)

  - temperature evolution
  - chemical reaction
  - target evolution (plane,boat,missile,...)

$\rightarrow$ **signal's randomness represents:**

  - model uncertainties
  - statistics of the unknown parameters (control policies, navigation coefficients,....)

- $\boxed{\textbf{Observation process } Y = \{Y_n \; ; \; n \geq 0\}\textbf{:}}$

$\sim$ partial and noisy informations on the signal X

  - thermic/infrared sensors.
  - radar/sonar
  - global positioning systems

$\rightarrow$ **observation noise represents:**

  - model uncertainties
  - thermic noise $\sim$ electronic components (resistors...)

3.1 Filtering linear/Gaussian signals

$\longrightarrow$ **Lectures on discrete time filtering,**
R.S. Bucy, Signal Processing and Digital Filtering,
Springer Verlag, 1994.

$\longrightarrow$ **Probability**
A.N. Shiryaev, *Second edition, Springer*
Graduate texts in Mathematics.

3.1.1 Gaussian variables

$Z = (Z^1, \ldots, Z^d)$ **Gaussian on** $\mathbb{R}^d$**:**

$$Z \sim \mathcal{N}(\mathbb{E}(Z), Q) \Leftrightarrow \mathbb{E}(e^{i<u,Z>}) = \exp\left[i < u, \mathbb{E}(Z) > -\frac{1}{2} < Qu, u >\right]$$

with covariance matrix

$$Q(i,j) = \mathbb{E}((Z^i - E(Z^i))(Z^j - E(Z^j))).$$

$$\downarrow \ (\exists Q^{-1})$$

$$\mathbb{P}(Z \in dz) = \frac{1}{(2\pi)^{\frac{d}{2}}\sqrt{\det(Q)}} \ \exp\left[-\frac{1}{2}\langle (Q)^{-1}(z - \mathbb{E}(Z)), (z - \mathbb{E}(Z))\rangle\right].$$

Fundamental Properties:

1. $Z \sim$ Gauss $\iff < u, Z > \overset{\forall u}{=} u'Z = \sum_i u^i z^i$ Gauss.
2. Let (Z_1, Z_2) be a Gaussian vector in $\mathbb{R}^{d_1+d_2}$.
   (a) Z_1 and Z_2 are independent iff there are uncorrelated
   $$Q_{z_1,z_2} = \mathbb{E}((Z_1 - \mathbb{E}(Z_1))(Z_2 - \mathbb{E}(Z_2))') = 0.$$
   (b) $\exists Q^{-1}_{z_1,z_2} \Longrightarrow Z_1$ given Z_2 is Gaussian with mean
   $$\mathbb{E}(Z_1|Z_2) = \mathbb{E}(Z_1) + Q_{z_1,z_2} Q^{-1}_{z_2,z_2} (Z_2 - \mathbb{E}(Z_2))$$
   and covariance matrix
   $$\begin{aligned} Q_{z_1|z_2} &= \mathbb{E}((Z_1 - \mathbb{E}(Z_1|Z_2))(Z_1 - \mathbb{E}(Z_1|Z_2))') \\ &= Q_{z_1,z_1} - Q_{z_1,z_2} Q^{-1}_{z_2,z_2} Q_{z_2,z_1}. \end{aligned}$$

3.1.2 Kalman-Bucy filter

$$\left\{ \begin{array}{rcl} X_n & = & A\ X_{n-1} + B\ W_n \\ Y_n & = & C\ X_n + V_n \end{array} \right.$$

((X,V,W) $\mathbb{R}^{d_x+d_v+d_w}$-valued r.v. and A, B, C $\mathbb{R}^{d_x\times d_x}$- $\mathbb{R}^{d_x\times d_w}$, $\mathbb{R}^{d_y\times d_x}$ matrices)

(X_0,V,W) independent and Gaussian with:

$$\begin{array}{rcl} \mathbb{E}(V_n) & = & \mathbb{E}(W_n) = 0\,, \quad Q_n^v = \mathbb{E}\,(V_n\ V_n')\,, \quad Q_n^w = \mathbb{E}\,(W_n\ W_n') \\ \overline{X}_0 & = & \mathbb{E}(X_0), \qquad\quad Q_0^x = \mathbb{E}\left((X_0-\mathbb{E}(X_0))\,(X_0-\mathbb{E}(X_0))'\right) \end{array}$$

$$\downarrow$$

$$\boxed{(X_0,\ldots,X_n,Y_0,\ldots,Y_n) = Span(X_0,W_0,\ldots,W_n,V_0,\ldots,V_0)\ \text{ Gaussian}}$$

$$\Downarrow$$

$$\begin{array}{rcl} \widehat{\eta}_n(dx) & = & \mathbb{P}\left(X_n\in dx\,|Y_0,\ldots,Y_n\right) = \mathcal{N}\left(\widehat{X}_n,P_n\right)(dx) \\ \eta_{n+1}(dx) & = & \mathbb{P}\left(X_{n+1}\in dx\,|Y_0,\ldots,Y_n\right) = \mathcal{N}\left(\widehat{X}_{n+1}^-,P_{n+1}^-\right)(dx) \end{array}$$

with

$$\boxed{\left\{ \begin{array}{rcl} \widehat{X}_n & = & \mathbb{E}(X_n|Y_0,\ldots,Y_n) \\ P_n & = & \mathbb{E}((X_n-\widehat{X}_n)(X_n-\widehat{X}_n)') \end{array} \right.}$$

and

$$\boxed{\left\{ \begin{array}{rcl} \widehat{X}_{n+1}^- & = & \mathbb{E}(X_{n+1}|Y_0,\ldots,Y_n) \\ P_{n+1}^- & = & \mathbb{E}((X_{n+1}-\widehat{X}_{n+1}^-)(X_{n+1}-\widehat{X}_{n+1}^-)'). \end{array} \right.}$$

Kalman-Bucy Filter :

$$\begin{pmatrix} \widehat{X}_n^- \\ P_n^- \end{pmatrix} \xrightarrow{\text{Updating/correction}} \begin{pmatrix} \widehat{X}_n \\ P_n \end{pmatrix} \xrightarrow{\quad\text{Prediction}\quad} \begin{pmatrix} \widehat{X}_{n+1}^- \\ P_{n+1}^- \end{pmatrix}$$

$(\widehat{X}_0^- = \mathbb{E}(X_0),\ P_0^- = Q_0^x)$

Updating/Correction:

$$\begin{array}{c} (X_n, Y_0, \ldots, Y_{n-1}) \quad \text{and} \quad (Y_0, \ldots, Y_n) \quad \text{Gauss} \\ \Downarrow \\ \widehat{X}_n^- = \mathbb{E}(X_n|Y_0, \ldots, Y_{n-1}) \quad \text{and} \quad \widehat{Y}_n^- = \mathbb{E}(Y_n|Y_0, \ldots, Y_{n-1}) \quad \text{Gauss} \\ \Downarrow \end{array}$$

$$(\tilde{X}_n, \tilde{Y}_n) \;=\; (X_n - \widehat{X}_n^-, Y_n - \widehat{Y}_n^-) \qquad \text{Gauss}$$

and

$$\begin{aligned} \mathbb{E}(\tilde{X}_n|\tilde{Y}_n) &= Q_{\tilde{x}_n,\tilde{y}_n} Q_{\tilde{y}_n,\tilde{y}_n}^{-1} \tilde{Y}_n \\ \mathbb{E}((\tilde{X}_n - \mathbb{E}(\tilde{X}_n|\tilde{Y}_n))(\tilde{X}_n - \mathbb{E}(\tilde{X}_n|\tilde{Y}_n))') &= Q_{\tilde{x}_n,\tilde{x}_n} - Q_{\tilde{x}_n,\tilde{y}_n} Q_{\tilde{y}_n,\tilde{y}_n}^{-1} Q_{\tilde{y}_n,\tilde{x}_n}. \end{aligned}$$

Note:

$$\begin{aligned} Q_{\tilde{x}_n,\tilde{x}_n} &= P_n^- \\ \widehat{Y}_n^- &= \mathbb{E}(C\ X_n + V_n|Y_0, \ldots, Y_{n-1}) = C\widehat{X}_n^- \end{aligned}$$

and

$$\begin{aligned} Q_{\tilde{x}_n,\tilde{y}_n} &= \mathbb{E}(\tilde{X}_n \tilde{Y}_n') = \mathbb{E}(\tilde{X}_n (C\tilde{X}_n + V_n)') = P_n^- C' \\ Q_{\tilde{y}_n,\tilde{y}_n} &= \mathbb{E}(\tilde{Y}_n \tilde{Y}_n') = \mathbb{E}((C\tilde{X}_n + V_n)(C\tilde{X}_n + V_n)') = CP_n^- C' + Q_n^v. \end{aligned}$$

$$\tilde{X}_n \quad \text{and} \quad (Y_0, \dots, Y_{n-1}) \quad \text{un-correlated} \Rightarrow \text{independent}$$

$$\Downarrow$$

$$\mathbb{E}(\tilde{X}_n | Y_0, \dots, Y_{n-1}, Y_n) = \mathbb{E}(\tilde{X}_n | Y_0, \dots, Y_{n-1}, \tilde{Y}_n) = \mathbb{E}(\tilde{X}_n | \tilde{Y}_n)$$

Then we have

$$\begin{aligned} \widehat{X}_n &= \widehat{X}_n^- + \mathbb{E}(X_n - \widehat{X}_n^- | Y_0, \dots, Y_n) = \widehat{X}_n^- + \mathbb{E}(\tilde{X}_n | \tilde{Y}_n) \\ &= \widehat{X}_n^- + P_n^- C' [C P_n^- C' + Q_n^v]^{-1} (Y_n - C \widehat{X}_n^-). \end{aligned}$$

Now using the decomposition

$$X_n - \widehat{X}_n = X_n - \widehat{X}_n^- - (\widehat{X}_n - \widehat{X}_n^-) = \tilde{X}_n - \mathbb{E}(\tilde{X}_n | \tilde{Y}_n),$$

we obtain

$$P_n = P_n^- - P_n^- C' [C P_n^- C' + Q_n^v]^{-1} C P_n^-$$

$$\Downarrow$$

Correction/Updating

$$\begin{aligned} \widehat{X}_n &= \widehat{X}_n^- + K_n (Y_n - C \widehat{X}_n^-) \\ P_n &= [I - K_n C] P_n^- \end{aligned}$$

where K_n is the **gain-matrix**

$$K_n = P_n^- C' [C P_n^- C' + Q_n^v]^{-1}.$$

Prediction :

Note

$$\begin{aligned}\widehat{X}^-_{n+1} &= \mathbb{E}(X_{n+1}|Y_0,\ldots,Y_n)\\ &= A\ \mathbb{E}(X_n|Y_0,\ldots,Y_n) = A\widehat{X}_n\end{aligned}$$

Therefore

$$\begin{aligned}P^-_{n+1} &= \mathbb{E}((X_{n+1}-\widehat{X}^-_{n+1})(X_{n+1}-\widehat{X}^-_{n+1})')\\ &= \mathbb{E}((A\ (X_n-\widehat{X}_n)+B\ W_n)(A\ (X_n-\widehat{X}_n)+B\ W_n)')\\ &= A\ P_n\ A' + B\ Q^w_n\ B'.\end{aligned}$$

$$\Downarrow$$

Prediction

$$\begin{aligned}\widehat{X}^-_{n+1} &= A\widehat{X}_n\\ P^-_{n+1} &= A\ P_n\ A' + B\ Q^w_n\ B'\end{aligned}$$

Problem 6 *Consider a signal/observation real-valued model*

$$\left\{ \begin{array}{lll} X_n & = & a\ X_{n-1} + W_n \\ Y_n & = & c\ X_n + V_n \end{array} \right.$$

with X_0, V_n, W_{n+1} *i.i.d.* $\sim \mathcal{N}(0,1)$ *and* $a, c \in \mathbb{R}$.

- *Check that* $(\widehat{X}_0^-, P_0^-) = (0,1)$ *and deduce*

$$K_0 = \frac{c}{1+c^2}, \qquad \text{et} \qquad \widehat{X}_0 = K_0\ Y_0.$$

- *Prove that*

$$K_n = \frac{c\ P_n^-}{c^2\ P_n^- + 1}, \qquad \text{and} \qquad C\ P_n^- = a^2\ K_{n-1} + c.$$

- *Deduce that*

$$\left\{ \begin{array}{lll} \widehat{X}_n & = & a\ \widehat{X}_{n-1} + K_n\ (Y_n - c\ a\ \widehat{X}_{n-1}) \\ K_n & = & \dfrac{a^2\ K_{n-1} + c}{c\ (a^2\ K_{n-1} + c) + 1} \end{array} \right.$$

- *When* $c = 0$ *check that*

$$K_n = 0, \qquad \text{and} \qquad \widehat{X}_n = \mathbb{E}(X_0) = 0.$$

Solution 6 • *We clearly have that*

$$(\widehat{X}_0^-, P_0^-) = (\mathbb{E}(X_0), \mathbb{E}((X_0 - \mathbb{E}(X_0))^2)) = (0, 1)$$

and therefore

$$\begin{aligned} K_0 &= P_0^- \; c \; [c^2 \; P_0^- + 1]^{-1} = \frac{c}{1+c^2} \\ \widehat{X}_0 &= \widehat{X}_0^- \; + \; K_0 \; (Y_0 - C \; \widehat{X}_0^-) = K_0 \; Y_0 \end{aligned}$$

- *By definition of the gain matrix we have that*

$$K_n = \frac{c \; P_n^-}{c^2 \; P_n^- + 1}$$

Using the prediction stage we notice that

$$P_n^- = a^2 \; P_{n-1} + 1 \qquad \Rightarrow \qquad cP_n^- = a^2(cP_{n-1}) + c$$

On the other hand by the updating step we have

$$\begin{aligned} P_n &= (1 - cK_n) \; P_n^- = (1 - \frac{c \; P_n^-}{c^2 \; P_n^- + 1}) \; P_n^- \\ &= \frac{P_n^-}{c^2 \; P_n^- + 1} = K_n / c \end{aligned}$$

From the above we conclude that

$$c \; P_n^- = a^2 \; K_{n-1} + c.$$

- *By definition of the gain matrix we have*

$$K_n \;\; = \;\; cP_n^- / \; [c^2 \; P_n^- + 1] = [a^2 \; K_{n-1} + c] / [c(a^2 \; K_{n-1} + c) + 1].$$

3.2 Non linear filtering

$\longrightarrow$ **Lectures on discrete time filtering,**
R.S. Bucy, Signal Processing and Digital Filtering,
Springer Verlag, 1994.

Model:

$(X_n, Y_n) \rightarrow (E \times F)$-valued Markov chain, initial distribution and Markov transition

$$\begin{array}{rcl} \mu_0(d(x_0, y_0)) & = & \overline{g}_0(x_0, y_0)\ \eta_0(dx_0)\ \gamma_0(dy_0) \\ Q_n((x_1, y_1), d(x_2, y_2)) & = & \overline{g}_n(x_2, y_2)\ K_n(x_1, dx_2)\ \gamma_n(dy_2) \end{array}$$

(γ_n proba. on F, $\overline{g}_n : E \times F \rightarrow \mathbb{R}_+$ and K_n Markov transitions)

Example 3.1 $E = F = \mathbb{R}^d$

$$\left\{ \begin{array}{rcl} X_n & = & F_n(X_{n-1}, W_n) \\ Y_n & = & H_n(X_n, V_n) \end{array} \right.$$

and

$$\mathbb{P}\left(H_n(x, V_n) \in dy\right) = \overline{g}_n(x, y)\ \gamma_n(dy).$$

In this situation

$$K_n(f)(x) = \int\ K_n(x, dx')f(x') = \mathbb{E}\left(f(F(x, W_n))\right).$$

If

$$H_n(x, v) = h_n(x) + v \qquad \text{and} \qquad \mathbb{P}(V_n \in dv) = p_n(v)\ dv$$

then

$$\overline{g}_n(x, y) = \frac{p_n(y - h_n(x))}{p_n(y)} \qquad \text{and} \qquad \gamma_n(dy) = p_n(y)\ dy.$$

In particular

$$V_n \sim \mathcal{N}(0, Q_n^v) \Rightarrow \overline{g}_n(x, y) = \exp\langle (Q_n^v)^{-1}(y - \frac{1}{2}\ h_n(x)), h_n(x)\rangle$$

$$
\begin{aligned}
&\mu_n\left(d(x_0,y_0),\ldots,d(x_n,y_n)\right)\\
&= \mathbb{P}\left((X_0,Y_0)\in d(x_0,y_0),\ldots,(X_n,Y_n)\in d(x_n,y_n)\right)\\
&= \mu_0\left(d(x_0,y_0)\right)\ Q_1((x_0,y_0),d(x_1,y_1))\ldots Q_n((x_{n-1},y_{n-1}),d(x_n,y_n))
\end{aligned}
$$

$$\Downarrow$$

$$\text{With}\qquad X_{[0,n]}=(X_0,\ldots,X_n)\qquad Y_{[0,n]}=(Y_0,\ldots,Y_n)$$

$$\Downarrow$$

$$
\begin{aligned}
&\mathbb{P}\left(X_{[0,n]}\in dx_{[0,n]},Y_{[0,n]}\in dy_{[0,n]}\right)\\
&= \mathbb{P}\left(Y_{[0,n]}\in dy_{[0,n]}|X_{[0,n]}=x_{[0,n]}\right)\ \times\ \mathbb{P}\left(X_{[0,n]}\in dx_{[0,n]}\right)\\
&= [\textstyle\prod_{p=0}^n\ \overline{g}_p(x_p,y_p)\gamma_p(dy_p)]\ \times\ [\eta_0(dx_0)\ K_1(x_0,dx_1)\ldots K_n(x_{n-1},dx_n)]
\end{aligned}
$$

with

$$
\begin{aligned}
\mathbb{P}\left(Y_{[0,n]}\in dy_{[0,n]}|X_{[0,n]}=x_{[0,n]}\right) &= \underbrace{\prod_{p=0}^n\ \overline{g}_p(x_p,y_p)}_{\mathbf{Z}_n\mathbf{(x,y)}}\quad \gamma_0(dy_0)\ldots\gamma_n(dy_n)\\
\text{and}\quad \mathbb{P}\left(X_{[0,n]}\in dx_{[0,n]}\right) &= \eta_0(dx_0)\ K_1(x_0,dx_1)\ldots K_n(x_{n-1},dx_n)
\end{aligned}
$$

Notice

$$\mathbb{P}\left((Y_0,\ldots,Y_n)\in d(y_0,\ldots,y_n)\right)=\overline{g}_{[0,n]}(y_{[0,n]})\ \gamma_0(dy_0)\ldots\gamma_n(dy_n)$$

with

$$\overline{g}_{[0,n]}(y_0,\ldots,y_n)=\int\ Z_n(x,y)\ \ \eta_0(dx_0)\ K_1(x_0,dx_1)\ldots K_n(x_{n-1},dx_n)$$

$$\Downarrow$$

$$\boxed{\begin{aligned}&\mathbb{E}\left(\varphi(X_0,\ldots,X_n)\,|Y_0=y_0,\ldots,Y_n=y_n\right)\\&=\frac{\int\ \varphi(x_0,\ldots,x_n)\ Z_n(x,y)\ \eta_0(dx_0)\ K_1(x_0,dx_1)\ldots K_n(x_{n-1},dx_n)}{\int\ Z_n(x,y)\ \eta_0(dx_0)\ K_1(x_0,dx_1)\ldots K_n(x_{n-1},dx_n)}\end{aligned}}$$

3.2.1 Non linear equations

With $Y = y$ fixed we denote $g_n(x) = \overline{g}_n(x, y_n)$ and we have:

$$\boxed{\mathbb{E}\left(\varphi(X_0,\ldots,X_n)\,|Y_0 = y_0,\ldots,Y_n = y_n\right) = \frac{\mathbb{E}(\varphi(X_0,\ldots,X_n)\ \prod_{p=0}^{n} g_p(X_p))}{\mathbb{E}(\prod_{p=0}^{n} g_p(X_p))}}$$

Thus,

$$\widehat{\eta}_n(f) = \frac{\mathbb{E}(f(X_n)\ \prod_{p=0}^{n} g_p(X_p))}{\mathbb{E}(\prod_{p=0}^{n} g_p(X_p))} \quad \text{and} \quad \eta_n(f) = \frac{\mathbb{E}(f(X_n)\ \prod_{p=0}^{n-1} g_p(X_p))}{\mathbb{E}(\prod_{p=0}^{n-1} g_p(X_p))}$$

$$\Downarrow$$

$$\boxed{\eta_n \xrightarrow{\text{Correction/Updating}} \widehat{\eta}_n \xrightarrow{\text{Prediction}} \eta_{n+1} = \widehat{\eta}_n K_{n+1}}$$

Updating:

$$\Psi_n(\eta_n)(dx) = \frac{1}{\eta(g_n)}\ g_n(x)\ \eta_n(dx) \qquad \text{(non linear)}$$

Prediction:

$$\widehat{\eta}_n K_{n+1}(dx) = \int \widehat{\eta}_n(dx') K_{n+1}(x', dx) \qquad \text{(linear)}$$

Proof:

$$\widehat{\eta}_n(f) = \frac{\mathbb{E}(f(X_n)\ \prod_{p=0}^{n} g_p(X_p))}{\mathbb{E}(\prod_{p=0}^{n} g_p(X_p))} = \frac{\mathbb{E}(f(X_n)g(X_n)\ \prod_{p=0}^{n-1} g_p(X_p))}{\mathbb{E}(g(X_n)\ \prod_{p=0}^{n-1} g_p(X_p))}$$

$$\Downarrow$$

$$\widehat{\eta}_n(f) = \frac{\mathbb{E}(f(X_n)g(X_n)\ \prod_{p=0}^{n-1} g_p(X_p))/\mathbb{E}(\prod_{p=0}^{n-1} g_p(X_p))}{\mathbb{E}(g(X_n)\ \prod_{p=0}^{n-1} g_p(X_p))/\mathbb{E}(\prod_{p=0}^{n-1} g_p(X_p))}$$

$$\Downarrow$$

$$\widehat{\eta}_n(f) = \frac{\eta_n(g_n f)}{\eta_n(g_n)}$$

By Markov property we have

$$\begin{aligned}
\eta_n(f) &= \frac{\mathbb{E}(f(X_n)\ \prod_{p=0}^{n-1} g_p(X_p))}{\mathbb{E}(\prod_{p=0}^{n-1} g_p(X_p))} = \frac{\mathbb{E}(\mathbb{E}(f(X_n)|X_{n-1})\ \prod_{p=0}^{n-1} g_p(X_p))}{\mathbb{E}(\prod_{p=0}^{n-1} g_p(X_p))} \\
&= \frac{\mathbb{E}(K_n(f)(X_{n-1})\ \prod_{p=0}^{n-1} g_p(X_p))}{\mathbb{E}(\prod_{p=0}^{n-1} g_p(X_p))} \\
&= \widehat{\eta}_{n-1}K_n(f)(X_{n-1})
\end{aligned}$$

$$\Downarrow$$

$$\eta_n = \widehat{\eta}_{n-1}K_n$$

■

Non linear equation (1-step Predictors):

$$\eta_{n+1} = \Phi_{n+1}(\eta_n) \qquad \text{with} \qquad \Phi_{n+1}(\eta_n) = \Psi_n(\eta_n)K_{n+1}.$$

Non linear equation (optimal filter):

$$\widehat{\eta}_{n+1} = \widehat{\Phi}_{n+1}(\widehat{\eta}_n) \qquad \text{with} \qquad \widehat{\Phi}_{n+1}(\widehat{\eta}_n) = \widehat{\Psi}_n(\widehat{\eta}_n)\widehat{K}_{n+1}.$$

and

$$\begin{aligned}
\widehat{\Psi}_n(\widehat{\eta}_n) &= \frac{1}{\widehat{\eta}_n(\widehat{g}_n)}\, \widehat{g}_n(x)\ \widehat{\eta}_n(dx) \qquad \text{with} \quad \widehat{g}_n(x) = K_{n+1}(g_n)(x) \\
\widehat{K}_{n+1}(x, dx') &= \frac{1}{K_{n+1}(g_n)(x)}\ K_{n+1}(x, dx')\ g_n(x')
\end{aligned}$$

Proof:

$$\begin{aligned}
\widehat{\eta}_{n+1}(f) &= \Psi_{n+1}(\eta_{n+1})(f) = \Psi_{n+1}(\widehat{\eta}_n K_{n+1})(f) \\
&= \frac{\widehat{\eta}_n K_{n+1}(g_n f)}{\widehat{\eta}_n(K_{n+1}(g_n))}
\end{aligned}$$

$$\Downarrow$$

$$\begin{aligned}
\widehat{\eta}_{n+1}(f) &= \frac{\widehat{\eta}_n(K_{n+1}(g_n)\ [K_{n+1}(g_n f)/K_{n+1}(g_n)])}{\widehat{\eta}_n K_{n+1}(g_n)} \\
&= \frac{\widehat{\eta}_n(\widehat{g}_n\ [K_{n+1}(g_n f)/K_{n+1}(g_n)])}{\widehat{\eta}_n(\widehat{g}_n)}
\end{aligned}$$

$$\Downarrow$$

$$\widehat{\eta}_{n+1}(f) = \frac{\widehat{\eta}_n(\widehat{g}_n\ \widehat{K}_{n+1}(f))}{\widehat{\eta}_n(\widehat{g}_n)} = \widehat{\Psi}_n(\widehat{\eta}_n)\widehat{K}_{n+1}(f)$$

■

3.2.2 Measure valued processes and interacting particle systems

$\longrightarrow$ **Measure valued processes and interacting particle systems. Applications to non linear filtering problems**
P. DM, *The Annals of Applied Probability,* 8(2): 438-495, 1998.

$\longrightarrow$ **On a class of discrete generation interacting particle systems.**
M. Kouritzin, L. Miclo and P. DM.,
Electronic Journal of Probability
Vol. 6 (2001) Paper no. 16, pages 1-26.

Measure valued dynamical system:

$$\eta_n = \Phi_n(\eta_{n-1})$$

$\longrightarrow$ **Pb.:**

- Non Linear dynamics Φ_{n+1}.
- $\in$ Proba. meas. on $E \rightharpoonup \infty$-dimensional.

$\longrightarrow$ **Objective/Solution:**

- E^N-valued Markov model $\xi_n = (\xi_n^1, \ldots, \xi_n^N)$ s.t.

$$\eta_n^N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\xi_n^i} \xrightarrow[N \to \infty]{} \eta_n \qquad \text{(in some sense)}$$

- $\rightleftharpoons$ N-Approximating model for sampling "i.i.d. $\sim\ \eta_n$"

$\longrightarrow$ **Notations/conventions:**

$$\xi_n = \underbrace{(\xi_n^1, \ldots, \xi_n^i, \ldots, \xi_n^N)}_{N \text{ particles}} \longleftarrow \text{configuration/population/system}$$

$$\begin{aligned} i &= \text{particle label} \\ n &= \text{time parameter} \\ N &= \text{size of the system/precision parameter} \end{aligned}$$

Initially ($n=0$): $\xi_0 = (\xi_0^1, \ldots, \xi_0^N)$ i.i.d. $\sim \eta_0$

$$\text{Law large numbers} \qquad \Longrightarrow \qquad \eta_0^N = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_0^i} \xrightarrow[N\to\infty]{} \eta_0$$

Ex. (L_2-mean error): $\mathbb{E}((\eta_0^N(f) - \eta_0(f))^2) = \eta_0((f-\eta_0(f))^2)/N \leq (\|f\|^2)/N$

Step $n \to (n+1)$:

$$\eta_n^N = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i} \;\to\; \eta_n \xRightarrow{\Phi_{n+1}\ \text{cont.}} \Phi_{n+1}\left(\eta_n^N\right) \to \Phi_{n+1}(\eta_n) = \eta_{n+1}$$

$$\Downarrow \qquad (\text{given } \xi_n)$$

$$\begin{aligned} \xi_{n+1} &= (\xi_{n+1}^1, \ldots, \xi_{n+1}^N) \qquad \textbf{i.i.d.} \quad \sim \Phi_{n+1}\left(\frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i}\right) \\ \eta_{n+1}^N &= \frac{1}{N}\sum_{i=1}^N \delta_{\xi_{n+1}^i} \xrightarrow[N\to\infty]{} \eta_{n+1} \end{aligned}$$

$$\Downarrow$$

$$\boxed{\begin{aligned} \mathbb{P}\left(\xi_0 \in d(x^1,\ldots,x^N)\right) &= \prod_{p=1}^n \eta_0(dx^p) \\ \mathbb{P}\left(\xi_n \in d(x^1,\ldots,x^N) \,\middle|\, \xi_{n-1} = (z^1,\ldots,z^N)\right) &= \prod_{p=1}^n \Phi_n\left(\frac{1}{N}\sum_{i=1}^N \delta_{z^i}\right)(dx^p) \end{aligned}}$$

- The initial configuration $\xi_0 = \left(\xi_0^1, \ldots, \xi_0^N\right)$ consists of N i.i.d. with law η_0.
- Given the empirical measure $\eta_{n-1}^N = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_{n-1}^i}$ the next configuration
$$\xi_n = \left(\xi_n^1, \ldots, \xi_n^N\right)$$
consists of N i.i.d. with law $\Phi_n\left(\eta_{n-1}^N\right)$.

Limiting/Exact dynamical system:

$$\begin{array}{ccccccccc}
\eta_0 & & \Phi_1(\eta_0) & & \Phi_2(\eta_1) & & & \Phi_n(\eta_{n-1}) & \\
\| & \nearrow & \| & \nearrow & \| & \cdots & \nearrow & \| & \nearrow \cdots \\
\eta_0 & & \eta_1 & & \eta_2 & & & \eta_n &
\end{array}$$

Approximating N-particle model:

$$\begin{array}{ccccccccc}
\eta_0 & & \Phi_1(\eta_0^N) & & \Phi_2(\eta_1^N) & & & \Phi_n(\eta_{n-1}^N) & \\
\downarrow \text{ (N-i.i.d.)} & \nearrow & \downarrow \text{ (N-i.i.d.)} & \nearrow & \downarrow \text{ (N-i.i.d.)} & \cdots & \nearrow & \downarrow \text{ (N-i.i.d.)} & \nearrow \cdots \\
\eta_0^N & & \eta_1^N & & \eta_2^N & & & \eta_n^N &
\end{array}$$

Measure valued dynamical system:

$$\eta_n = \eta_{n-1}\mathcal{K}_{n,\eta_{n-1}}$$

For some Markov transitions $\mathcal{K}_{n,\eta}(x,dy)$ and

$$\eta_{n-1}\mathcal{K}_{n,\eta_{n-1}}(dy) = \int \eta_{n-1}(dx)\mathcal{K}_{n,\eta_{n-1}}(x,dy)$$

$$\Downarrow$$

Interacting Particle systems:

$$\begin{aligned}
\mathbb{P}\left(\xi_0 \in d(x^1,\ldots,x^N)\right) &= \prod_{p=1}^{n} \eta_0(dx^p) \\
\mathbb{P}\left(\xi_n \in d(x^1,\ldots,x^N)\,\middle|\,\xi_{n-1} = (z^1,\ldots,z^N)\right) &= \prod_{p=1}^{n} \mathcal{K}_{n,\frac{1}{N}\sum_{i=1}^N \delta_{z^i}}(z^p,dx^p)
\end{aligned}$$

Description of the algorithm:

- The initial configuration $\xi_0 = \left(\xi_0^1,\ldots,\xi_0^N\right)$ consists of N i.i.d. with law η_0.
- Given the empirical measure $\eta_{n-1}^N = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_{n-1}^i}$ the next configuration

  $$\xi_n = \left(\xi_n^1,\ldots,\xi_n^N\right)$$

  consists of N independent particles with respective distributions

  $$\mathcal{K}_{n,\frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i}}(\xi_n^1,.),\ldots,\mathcal{K}_{n,\frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i}}(\xi_n^N,.)$$

Example 3.2 (Mean field model) *($E = \mathbb{R}^d$)*

$$\mathcal{K}_{n,\eta}(x,dy) = \frac{1}{(2\pi)^{\frac{d}{2}}\sqrt{\det(Q_n)}}\ \exp -\frac{1}{2}\left[(y-a_n(x,\eta))'\,Q_n^{-1}\,(y-a_n(x,\eta))\right]\ \ dy$$

with $Q_n > 0$ sym.$(d\times d)$-matrix, $a_n : \mathbb{R}^d \times Proba(\mathbb{R}^d) \to \mathbb{R}^d$.

$$\Downarrow$$

$$\eta_n = \mathrm{Law}(\overline{X}_n) \qquad \text{with} \qquad \overline{X}_n = a_n(\overline{X}_{n-1}, \eta_{n-1}) + W_n$$

with W_n i.i.d. with distribution

$$\gamma_n(dy) = \frac{1}{(2\pi)^{\frac{d}{2}}\sqrt{\det(Q_n)}}\ \exp -\frac{1}{2}\left[y'Q_n^{-1}y\right]\ dy$$

In particular $a_n = (a_n^1, \ldots, a_n^d)$ with

$$a_n^i(x,\eta) = \sum_{j\in J}\ \alpha_{n,j}^i(x)\ \eta(h_{n,j}^i)$$

Example 3.3 (McKean gas model) *($E = \{-1,+1\}$)*

$$\mathcal{K}_{n,\eta}(x,dy) = \eta(+1)\ \delta_x(dy) + \eta(-1)\ \delta_{-x}(dy)$$

Since

$$\mathcal{K}_{n,\eta}(x,dy) = \int_E \eta(dx')\ K_n((x,x'),dy) \qquad \text{with} \quad K_n((x,x'),dy) = \delta_{xx'}(dy)$$

we also have

$$\eta_n(f) = \int_{E\times E} \eta_{n-1}(dx)\ \eta_{n-1}\ (dx')\ K_n((x,x'),dy)\ f(y)$$

Theorem: ($p \geq 1$, $\mathcal{K}_{n,\eta}$ or Φ_n suff. regular)

$$\mathbb{E}(|\eta_n^N(f) - \eta_n(f)|^p)^{\frac{1}{p}} \leq \frac{1}{\sqrt{N}}\, C_n(f)$$

If Φ_n suff. stable then for some $\alpha \in (0,1]$

$$\sup_{n\geq 0} \mathbb{E}(|\eta_n^N(f) - \eta_n(f)|^p)^{\frac{1}{p}} \leq \frac{1}{N^{\alpha/2}}\, C(f)$$

(Filtering) If K_n suff. mixing ($\Rightarrow \Phi_n$ expo. stable) then

$$\sup_{n\geq 0} \mathbb{E}(|\eta_n^N(f) - \eta_n(f)|^p)^{\frac{1}{p}} \leq \frac{1}{\sqrt{N}}\, C(f)$$

Proof:

$\longrightarrow$ **Nonlinear filtering: interacting particle solution.**
P. DM, *Markov Process. Related Fields*, 2 (1996), no. 4, 555–579.

$\longrightarrow$ **Branching and Interacting Particle Systems Approximations of Feynman-Kac Formulae with Applications to Non-Linear Filtering,**
L. Miclo and P. DM, *Séminaire de Probabilités XXXIV.*, *Lecture Notes in Mathematics No. 1729*,
J. Azéma, M. Émery, M. Ledoux, M. Yor (Eds.), pp.1-145, 2000.

$\longrightarrow$ **On a class of discrete generation interacting particle systems.**
M. Kouritzin, L. Miclo and P. DM.,
Electronic Journal of Probability
Vol. 6 (2001) Paper no. 16, pages 1-26.

3.2.3 Genetic particle filters

Reminder: ($\mu(dx)$ proba., $K(x,dy)$ Markov transition)

$$\mathbb{P}((X,Y) \in d(x,y)) = \mu(dx)\ K(x,dy)$$
$$\Rightarrow \mathbb{P}(Y \in dy) = \mu K(dy) = \int \mu(dx) K(x,dy)$$

$$\Downarrow$$

$$\boxed{\text{Simulation}\, Y\ \sim\ \mu K \Longleftrightarrow \begin{cases} -\text{Simulation } X\ \sim\ \mu \\ \\ -(\text{Given } X) \text{ Simulation}\, Y\ \sim\ K(X,.) \end{cases}}$$

1-step predictor equation: $(\forall x = (x^1, \ldots, x^N) \in E^N)$

$$\Phi_n\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x^i}\right) \;=\; \Psi_{n-1}\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x^i}\right) K_n$$

$$\Psi_{n-1}\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x^i}\right) \;=\; \sum_{i=1}^{N}\frac{g_{n-1}(x^i)}{\sum_{j=1}^{N} g_{n-1}(x^j)}\,\delta_{x^i}$$

$$\downarrow$$

Simulate a r.v. $z^1 \sim \Phi_n\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x^i}\right)$
$\Longleftrightarrow$

- **Select randomly a site** $\widehat{x}^1$ **according to the discrete distribution:**

$$\Psi_{n-1}\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x^i}\right) = \sum_{i=1}^{N}\frac{g_{n-1}(x^i)}{\sum_{j=1}^{N} g_{n-1}(x^j)}\,\delta_{x^i}$$

- **Given the selected site** $\widehat{x}^1$ **evolve and perform a mutation according to** K_n**, that is sample a r.v.** z^1 **with distribution**

$$K_n(\widehat{x}^1, .)$$

Simulate N i.i.d. $z^1,\ldots,z^N \sim \Phi_n\left(\frac{1}{N}\sum_{i=1}^N \delta_{x^i}\right)$
$\Longleftrightarrow$

- **Select** randomly N sites $\widehat{x}^1,\ldots,\widehat{x}^N$ according to

$$\Psi_{n-1}\left(\frac{1}{N}\sum_{i=1}^N \delta_{x^i}\right) = \sum_{i=1}^N \frac{g_{n-1}(x^i)}{\sum_{j=1}^N g_{n-1}(x^j)}\,\delta_{x^i}$$

- Given the selected sites $(\widehat{x}^1,\ldots,\widehat{x}^N)$ evolve and perform N **mutations** according to K_n.
  That is, sample N independent variable $z^1,\ldots,z^N$ with respective distributions

$$K_n(\widehat{x}^1,.),\ldots,K_n(\widehat{x}^N,.)$$

$$\Downarrow$$

$\boxed{\textbf{Genetic Particle scheme}}$

$$\xi_{n-1} = (\xi_{n-1}^1,\ldots,\xi_{n-1}^N) \xrightarrow{\text{Selection}} \widehat{\xi}_{n-1} = (\widehat{\xi}_{n-1}^1,\ldots,\widehat{\xi}_{n-1}^N) \xrightarrow{\text{Mutation}} \xi_n$$

- **Selection:** (given ξ_{n-1})

$$\widehat{\xi}_{n-1} = (\widehat{\xi}_{n-1}^1,\ldots,\widehat{\xi}_{n-1}^N) \text{ i.i.d } \sim \sum_{i=1}^N \frac{g_{n-1}(\xi_{n-1}^i)}{\sum_{j=1}^N g_{n-1}(\xi_{n-1}^j)}\,\delta_{\xi_{n-1}^i}$$

- **Mutation:** For each $i=1,\ldots,N$, $\quad \xi_n^i \sim K_n(\widehat{\xi}_{n-1}^i,.)$.

$$\Downarrow$$

With $dy = dy^1\times\ldots\times dy^N$, $dx = dx^1\times\ldots\times dx^N$ infinitesimal neighborhoods of $y=(y^1,\ldots,y^N)$ and $x=(x^1,\ldots,x^N)\in E^N$ we have

$$\begin{aligned}
\mathbb{P}(\widehat{\xi}_{n-1}\in dy|\xi_{n-1}=x) &= \prod_{p=1}^N \sum_{i=1}^N \frac{g_{n-1}(x^i)}{\sum_{j=1}^N g_{n-1}(x^j)}\,\delta_{x^i}(dy^p) \\
\mathbb{P}(\xi_n = dx|\widehat{\xi}_n = y) &= \prod_{p=1}^N K_n(y^p,dx^p)
\end{aligned}$$

3.2.4 Interacting particle filters

$$\boxed{(G_{min}) \qquad \exists \epsilon_n > 0 \quad : \qquad g_n(x) \geq \epsilon_n.}$$

$$\Downarrow$$

$$\Psi_n(\eta) = \eta S^{\epsilon}_{n,\eta}$$

where

$$S^{\epsilon}_{n,\eta}(u, dv) = \frac{\epsilon_n}{\eta(g_n)} \; \delta_u(dv) \; + \left(1 - \frac{\epsilon_n}{\eta(g_n)}\right) \; \Psi^{\epsilon}_n(\eta)(dv)$$

and

$$\Psi^{\epsilon}_n(\eta)(f) = \frac{\eta(g^{\epsilon}_n f)}{\eta(g^{\epsilon}_n)} \qquad \text{with} \qquad g^{\epsilon}_n = g_n - \epsilon_n.$$

$$\Downarrow$$

$$\boxed{\eta_{n+1} = \eta_n \mathcal{K}_{n+1,\eta_n} \quad \text{with} \quad \mathcal{K}_{n+1,\eta} = S^{\epsilon}_{n,\eta} K_{n+1}}$$

Example 3.4 $E = F = \mathbb{R}^d$

$$\begin{cases} X_n &= F_n(X_{n-1}, W_n) \\ Y_n &= h_n(X_n) + V_n \end{cases}$$

h_n bounded, X_0, V_n, W_n i.i.d. and $\mathbb{P}(V_n \in dv) = p_n(v) \; dv$.

$$\Rightarrow \quad (G_{min}) \quad \text{with} \quad g_n(x) \geq \epsilon_n(y_n) = \inf_x \frac{p_n(y_n - h_n(x))}{p_n(y_n)}$$

In particular

$$\|h_n\|_\infty = \sup_x \|h_n(x)\| < \infty \quad \text{and} \quad V_n \sim \mathcal{N}(0, Q^v_n)$$

$$\Downarrow$$

$$-\log \epsilon_n(y_n) \leq |||(Q^v_n)^{-1}||| \; \|h_n\|_\infty \left(\frac{1}{2}\|h_n\|_\infty + \|y\|\right) < \infty$$

Reminder: ($K(x,dy)$, $L(y,dz)$ Markov transitions)

$$\begin{aligned} & \mathbb{P}((Y,Z)\in d(y,z)|X=x) = K(x,dy)L(y,dz) \\ \Rightarrow \qquad & \mathbb{P}(Z\in dz|X=x) = KL(x,dz) = \int K(x,dy)L(y,dz) \end{aligned}$$

$$\Downarrow$$

Given X

$$\text{Simulation } Z \ \sim \ KL(X,.) \Longleftrightarrow \begin{cases} -\text{Simulation } Y \ \sim \ K(X,.) \\ \\ -(\text{Given } Y) \text{ Simulation } Z \ \sim \ L(Y,.) \end{cases}$$

$$\downarrow$$

Simulate a r.v. (given $x = (x^1, \ldots, x^N)$ with $m(x) = \frac{1}{N}\sum_{i=1}^N \delta_{x^i}$)

$$z^1 \sim S^{\epsilon}_{n,m(x)} K_{n+1}(x^1, .)$$

$\Longleftrightarrow$

- **Select randomly a site $\widehat{x}^1$ according to the discrete distribution:**

$$S^{\epsilon}_{n,m(x)}(x^1, .) = \frac{\epsilon_n}{m(x)(g_n)}\ \delta_{x^1}\ + \left(1 - \frac{\epsilon_n}{m(x)(g_n)}\right)\ \Psi^{\epsilon}_n(m(x))$$

- **Given the selected site $\widehat{x}^1$ evolve and perform a mutation according to K_{n+1}, that is sample a r.v. z^1 with distribution**

$$K_{n+1}(\widehat{x}^1, .)$$

$$\Downarrow$$

$$\xi_n \xrightarrow{\text{Selection}} \widehat{\xi}_n \xrightarrow{\text{Mutation}} \xi_{n+1}$$

Selection: (given $m(\xi_n) = \frac{1}{N}\sum_{i=1}^N \delta_{\xi^i_n}$)

$\widehat{\xi}^i_n$, $i = 1, \ldots, N$, independent particles with respective distributions

$$\begin{cases} S^{\epsilon}_{n,m(\xi_n)}(\xi^i_n, .) = \dfrac{\epsilon_n}{m(\xi_n)(g_n)}\ \delta_{\xi^i_n}\ + \left(1 - \dfrac{\epsilon_n}{m(\xi_n)(g_n)}\right)\ \Psi^{\epsilon}_n(m(\xi_n)) \\ 1 \leq i \leq N. \end{cases}$$

Mutation: For each $i = 1, \ldots, N$, $\quad \xi^i_n \sim K_n(\widehat{\xi}^i_{n-1}, .)$.

Elementary selection $\quad \xi_n^i \longrightarrow \widehat{\xi}_n^i$

$$\Updownarrow$$

Sampling a Bernoulli b_n^i r.v.

$$\mathbb{P}(b_n^i = 1|\xi_n) = 1 - \mathbb{P}(b_n^i = 0|\xi_n) = \frac{\epsilon_n}{m(\xi_n)(g_n)} = \frac{N\epsilon_n}{\sum_{i=1}^N g_n(\xi_n^i)}$$

- $b_n^i = 1 \Longrightarrow \widehat{\xi}_n^i = \xi_n^i$.
- If $b_n^i = 0$ we sample a r.v. $\widehat{\xi}_n^i$ with distribution

$$\Psi_n^\epsilon(m(\xi_n)) = \sum_{i=1}^N \frac{(g_n(\xi_n^i) - \epsilon_n)}{\sum_{j=1}^N (g_n(\xi_n^j) - \epsilon_n)}\ \delta_{\xi_n^i}.$$

3.2.5 Branching particle filters

$$\widehat{x}^1, \ldots, \widehat{x}^N \quad \textbf{i.i.d.} \sim \Psi(m(x)) = \sum_{i=1}^N W^i \, \delta_{x^i} \quad \text{with} \quad W^i = \frac{g(x^i)}{\sum_{j=1}^N g(x^j)}$$

$$\Downarrow \quad (N \to \infty)$$

$$m(\widehat{x}) = \frac{1}{N} \sum_{i=1}^N \delta_{\widehat{x}^i} \quad \longleftrightarrow \quad \Psi(m(x))$$

Example 3.5

$$\mathbb{E}((m(\widehat{x})(f) - \Psi(m(x))(f))^2) = \frac{1}{N} \Psi(m(x))((f - \Psi(m(x))(f))^2) \leq \frac{\|f\|^2}{N}$$

Note:

$$m(\widehat{x}) = \frac{1}{N} \sum_{i=1}^N M_i \, \delta_{x^i} \qquad \text{with} \quad M^i = \text{Card}\left\{1 \leq j \leq N \; ; \; \widehat{x}^j = x^i\right\}$$

$$\Rightarrow \qquad M = (M^1, \ldots, M^N) = \text{Multinomial}\left(N; W^1, \ldots, W^N\right).$$

Branching variants: ($[a]$=integer part, $\{a\} = a - [a]$)

$$\Psi_n(m(x)) \;=\; \frac{1}{N}\sum_{i=1}^{N} [NW_i]\; \delta_{x^i} + \frac{1}{N}\sum_{i=1}^{N} \{NW_i\}\; \delta_{x^i}$$

$\boxed{\textbf{Baker selection:}}$ [$\longrightarrow$Remainder stochastic sampling]

$$M^i = \overline{M}^i + \tilde{M}^i \qquad \text{with} \quad \overline{M}^i = [NW_i]$$

- If $\tilde{N} = N - \overline{N} = \sum_{i=1}^{N} \{NW_i\} > 0$ with $\overline{N} = \sum_{i=1}^{N} [NW_i]$ then

$$(\tilde{M}^1, \ldots, \tilde{M}^N) = \text{Multinomial}\left(\tilde{N}; \frac{\{NW_1\}}{\sum_{j=1}^{N}\{NW_j\}}, \ldots, \frac{\{NW_N\}}{\sum_{j=1}^{N}\{NW_j\}}\right)$$

- If $N = \overline{N}$ then $\tilde{M}^i = 0$.

$\boxed{\textbf{Bernoulli selection:}}$ [$\longrightarrow$ Minimal variance]

$$M^i = \overline{M}^i + \tilde{M}^i \qquad \text{with} \quad \overline{M}^i = [NW_i]$$

and

$$\mathbb{P}(\tilde{M}^i = 1) = 1 - \mathbb{P}(\tilde{M}^i = 0) = \{NW_i\}$$

$\boxed{\textbf{Poisson selection:}}$

$$\forall m \geq 0 \qquad \mathbb{P}\left(M^i = m\right) = \frac{(NW_i)^m}{m!} \quad \exp\left[-(NW_i)\right]$$

$\boxed{\textbf{Binomial selection:}}$

$$\forall 0 \leq m \leq N \qquad \mathbb{P}\left(M^i = m\right) = \frac{N!}{m!(N-m)!}\, W_i^m \;(1 - W_i)^{N-m}$$

3.3 Examples and Applications

$$\left\{ \begin{array}{lcl} X_n &=& F_n(X_{n-1}, W_n) \\ Y_n &=& h_n(X_n) + V_n \end{array} \right.$$

X_0, V_n, W_n i.i.d. $\mathbb{R}^d$-valued random variables with

$$\mathbb{P}(V_n \in dv) = p_n(v)\ dv\,, \quad \text{with} \quad \forall v \in \mathbb{R}^d\,,\ \ p_n(v) > 0.$$

$\longrightarrow$ $\boxed{\textbf{Genetic algorithm:}}$

$$\xi_n \xrightarrow{\text{Selection}} \widehat{\xi}_n \xrightarrow{\text{Mutation}} \xi_{n+1}$$

- **The initial system** $\xi_0 = (\xi_0^1, \ldots, \xi_0^N)$ consists in N i.i.d. particle $\sim\ X_0$.
- **Selection transition** $\xi_n \longrightarrow \widehat{\xi}_n$**:**
  ($\rightarrow$ depends on the current observation Y_n).

$$\widehat{\xi}_n = \left(\widehat{\xi}_n^1, \ldots, \widehat{\xi}_n^N\right) \quad \text{i.i.d.} \quad \sim \sum_{i=1}^N \frac{W_n^i}{\sum_{j=1}^N\ W_n^j}\ \delta_{\xi_n^i}$$

with

$$W_n^i = \frac{p_n(Y_n - h_n(\xi_n^i))}{p_n(Y_n)}.$$

$$\text{Law}(V_n) = \mathcal{N}(0, Q_n^v) \Longrightarrow W_n^i = \exp\langle (Q_n^v)^{-1}(Y_n - \frac{1}{2}\ h_n(\xi_n^i)), h_n(\xi_n^i)\rangle.$$

- **Mutation transition** $\widehat{\xi}_n \longrightarrow \xi_{n+1}$**:**

$$\xi_{n+1}^i = F_{n+1}(\widehat{\xi}_n^i, W_{n+1}^i)$$

  with $(W_{n+1}^1, \ldots, W_{n+1}^N)$ i.i.d. $\sim\ W_{n+1}$.

3.3.1 linear non Gaussian

$$\begin{cases} X_n &= a\ X_{n-1} \ + \ b\ \epsilon_n\ W_n \\ Y_n &= c\ X_n \ + \ V_n \end{cases}$$

with

- $a, b, c \in \mathbb{R}$, $\text{Law}(X_0) = \mathcal{N}(m_0, \sigma_0)$.
- ϵ_n i.i.d. Bernoulli $\mathbb{P}(\epsilon_n \in dx) = p\ \delta_1(dx) + (1-p)\ \delta_0(dx)$.
- W_n i.i.d. Gauss $\text{Law}(W_n) = \mathcal{N}(0, \sigma_w)$.
- V_n i.i.d. Gauss $\text{Law}(V_n) = \mathcal{N}(0, \sigma_v)$.
- X_0, V_n, W_n and ϵ_n independent.

Related linear/Gaussian model ($\Longrightarrow$ Kalman-Bucy filter)

$$\begin{cases} \tilde{X}_n &= a\ \tilde{X}_{n-1} \ + \ b\ \tilde{W}_n \\ \tilde{Y}_n &= c\ \tilde{X}_n \ + \ V_n \end{cases}$$

with $\tilde{W}_n$ i.i.d. Gauss $\text{Law}(\tilde{W}_n) = \mathcal{N}(0, \sigma_{\tilde{w}})$ and

$$\sigma_{\tilde{w}}^2 = \mathbb{E}(\tilde{W}_n^2) = \mathbb{E}((\epsilon_n W_n)^2) = p\ \sigma_w^2.$$

Comparison Kalman-Bucy/Genetic particle filters

$$\begin{array}{lll} m_0 = 0 & \sigma_w = 5 & p = 0.1 \\ \sigma_0 = 1 & \sigma_v = 1 & N = 100 \\ a = 1/2 & b = 1 & c = 1 \end{array}$$

3.3.2 Cauchy signal noise

$$\begin{cases} X_n &= X_{n-1} + W_n \\ Y_n &= X_n + V_n \end{cases}$$

- X_0 Gauss Law$(X_0) = \mathcal{N}(0,1)$.
- W_n i.i.d. Cauchy with parameter $\sigma_w = 10^{-1}$

$$\mathbb{P}(W_n \in dw) = \frac{\sigma_w}{\pi \ (\sigma_w^2 + w^2)} \ dw.$$

- V_n i.i.d. Gauss Law$(V_0) = \mathcal{N}(0,1)$.
- X_0, V_n and W_n independent.

Related linear/Gauss model

$$\begin{cases} \tilde{X}_n &= \tilde{X}_{n-1} + \tilde{W}_n \\ \tilde{Y}_n &= \tilde{X}_n + V_n \end{cases}$$

with $\tilde{W}_n$ i.i.d. Gauss Law$(\tilde{W}_n) = \mathcal{N}(0, \sigma_{\tilde{w}})$ and

$$\sigma_{\tilde{w}} = \sigma_w \ \sqrt{\frac{\pi}{2}}.$$

Comparison Kalman-Bucy/Genetic particle filters

$$\sigma_w = 1 \qquad \sigma_0 = 1 \qquad \sigma_v = 1$$

3.4 Smoothing and path estimation

Signal/observation

$$
\begin{array}{cccccccc}
\longrightarrow & X'_{n-1} & \longrightarrow & X'_n & \longrightarrow & & \text{(signal)} \\
\cdots & \downarrow & \cdots & \downarrow & \cdots & & \\
 & Y_{n-1} & & Y_n & & & \text{(observation)}
\end{array}
$$

$$
\boxed{\textbf{Objective: } \longrightarrow \mathrm{Law}(X'_0, \ldots, X'_n | Y_0, \ldots, Y_n)}
$$

Key idea: (X'_n E'-valued sequence)

$$
X_n = (X'_0, \ldots, X'_n) \qquad E_n = \underbrace{E' \times \ldots \times E'}_{(n+1)-times} - \text{valued } \underline{\text{ Markov chain}}
$$

$$
\Downarrow
$$

$$
\mathrm{Law}(X'_0, \ldots, X'_n | Y_0, \ldots, Y_n) = \mathrm{Law}(X_n | Y_0, \ldots, Y_n)
$$

Alternative model:

$$
\mathcal{X}_n = (\mathcal{X}_n^p)_{p \geq 0} \qquad (E')^{\mathbb{N}} - \text{valued} \quad \text{with} \quad X'_p = \begin{cases} X'_p & \text{if } \ p \leq n \\ X'_n & \text{if } \ p \geq n \end{cases}
$$

Model:

$(X_n, Y_n) \rightarrow (E_n \times F_n)$-valued Markov chain, initial distribution and Markov transition

$$\begin{aligned} \mu_0(d(x_0,y_0)) &= \overline{g}_0(x_0,y_0)\ \eta_0(dx_0)\ \gamma_0(dy_0) \\ Q_n((x_{n-1},y_{n-1}),d(x_n,y_n)) &= \overline{g}_n(x_n,y_n)\ K_n(x_{n-1},dx_n)\ \gamma_n(dy_n) \end{aligned}$$

- $\gamma_n(d_n)$ proba. on F_n
- $\overline{g}_n : E_n \times F_n \rightarrow \mathbb{R}_+$
- $K_n(x_{n-1}, dx_n)$ Markov transitions from E_{n-1} into E_n

X'_n E'-valued Markov chain with transitions K'

$$X_n = (X'_0, \ldots, X'_n) \qquad E_n = \underbrace{(E' \times \ldots \times E')}_{\text{(n+1)-times}} \text{valued Markov chain}$$

Elementary K'-transitions:

$$X_n = (X'_0, \ldots, X'_n) \longrightarrow X_{n+1} = ((X'_0, \ldots, X'_n), X'_{n+1}) = (X_n, X'_{n+1})$$

with $X'_n \rightarrow X'_{n+1}$.

$$\Updownarrow$$

$\forall x_n = (x'_0, \ldots, x'_n)$, $z_{n+1} = (z'_0, \ldots, z'_{n+1})$ and $dz_{n+1} = dz'_0 \times \ldots \times dz'_{n+1}$

$$K_{n+1}(x_n, dz_{n+1}) = \delta_{(x'_0,\ldots,x'_n)}(d(z'_0, \ldots, z'_n))\ K'_{n+1}(z'_n, dz'_{n+1}).$$

$\longrightarrow Y = y,\ g_n(.) = \overline{g}_n(., y_n) : E_n \to \mathbb{R}_+,\ f : E_n \to \mathbb{R}$

One-step predictor

$$\begin{aligned} \eta_n(f) &= \mathbb{E}(f(X'_0, \ldots, X'_n) | Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}) \\ &= \mathbb{E}(f(X_n) | Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}) \end{aligned}$$

$$\Downarrow$$

$$\eta_n(f) = \frac{\mathbb{E}(f(X'_0, \ldots, X'_n)\ \prod_{p=0}^{n-1} g_p(X'_0, \ldots, X'_p))}{\mathbb{E}(\prod_{p=0}^{n-1} g_p(X'_0, \ldots, X'_p))}$$

$$= \frac{\mathbb{E}(f(X_n)\ \prod_{p=0}^{n-1} g_p(X_p))}{\mathbb{E}(\prod_{p=0}^{n-1} g_p(X_p))}$$

Optimal filter

$$\begin{aligned} \widehat{\eta}_n(f) &= \mathbb{E}(f(X'_0, \ldots, X'_n) | Y_0 = y_0, \ldots, Y_n = y_n) \\ &= \mathbb{E}(f(X_n) | Y_0 = y_0, \ldots, Y_n = y_n) \end{aligned}$$

$$\Downarrow$$

$$\widehat{\eta}_n(f) = \frac{\mathbb{E}(f(X'_0, \ldots, X'_n)\ \prod_{p=0}^{n} g_p(X'_0, \ldots, X'_p))}{\mathbb{E}(\prod_{p=0}^{n} g_p(X'_0, \ldots, X'_p))}$$

$$= \frac{\mathbb{E}(f(X_n)\ \prod_{p=0}^{n} g_p(X_p))}{\mathbb{E}(\prod_{p=0}^{n} g_p(X_p))}$$

Example 3.6 (Path observations)

$$\left\{\begin{array}{lll} X'_n & & E'-\text{valued Markov chain} \\ Y_n & = & H_n(X'_0,\ldots,X'_n)+V_n \end{array}\right.$$

with $H_n : (E')^{n+1} \to \mathbb{R}^d$, V_n *i.i.d., independent of* X'_n *and*

$$\mathbb{P}(V_n \in dv) = p_n(v)\ dv$$

$$\Downarrow (X_n = (X'_0,\ldots,X'_n) \in E_n = (E')^{n+1})$$

$$\left\{\begin{array}{lll} X_n & & E_n-\text{valued Markov chain} \\ Y_n & = & H_n(X_n)+V_n \end{array}\right.$$

For any fixed $Y = y$ *and* $x_n = (x'_0,\ldots,x'_n) \in E_n$ *we set*

$$\begin{array}{rcl} g_n(x_n) & = & g_n(x'_0,\ldots,x'_n) \\ & = & \dfrac{p_n(y_n - H_n(x'_0,\ldots,x'_n))}{p_n(y_n)} \\ & = & \dfrac{p_n(y_n - H_n(x_n))}{p_n(y_n)} \end{array}$$

In particular but <u>more usually</u> if we have $H_n(x'_0,\ldots,x'_n) = h_n(x'_n)$ *and*

$$\left\{\begin{array}{lll} X'_n & & E'-\text{valued Markov chain} \\ Y_n & = & h_n(X'_n)+V_n \end{array}\right.$$

then

$$g_n(x'_0,\ldots,x'_n) = \frac{p_n(y_n - h_n(x'_n))}{p_n(y_n)}$$

In both situations for any test function $f : E_n = (E')^{n+1} \to \mathbb{R}$

$$\begin{array}{rcl} \widehat{\eta}_n(f) & = & \mathbb{E}(f(X'_0,\ldots,X'_n)|Y_0 = y_0,\ldots,Y_n = y_n) \\ & = & \mathbb{E}(f(X_n)|Y_0 = y_0,\ldots,Y_n = y_n) \end{array}$$

Taking $f(x'_0,\ldots,x'_n) = f_n(x'_n)$ *we obtain the* n*-marginals*

$$\widehat{\eta}_n(f) = \mathbb{E}(f_n(X'_n)|Y_0 = y_0,\ldots,Y_n = y_n)$$

3.5 Historical processes and genealogical trees

Path valued signal:

$$X_n = (X'_0, \ldots, X'_n) \in E_n = \underbrace{E' \times \ldots \times E'}_{\text{(n+1)-times}}$$

$$\Downarrow$$

Path particles:

$$\xi^i_n = \left(\xi^i_{0,n}, \xi^i_{1,n}, \ldots, \xi^i_{n,n}\right) \in E_n = \underbrace{E' \times \ldots \times E'}_{\text{(n+1)-times}}, \qquad 1 \leq i \leq N.$$

(E'-valued signal path from the origin up to time n)

$$\eta_n \in \text{Proba}(E_n) \xrightarrow{\text{Updating}} \widehat{\eta}_n \in \text{Proba}(E_n) \xrightarrow{\text{Prediction}} \eta_{n+1} \in \text{Proba}(E_{n+1})$$

$$\Downarrow$$

$$\xi_n \in E_n^N \xrightarrow{\text{Selection}} \widehat{\xi}_n \in E_n^N \xrightarrow{\text{Mutation}} \xi_{n+1} \in E_{n+1}^N$$

- **Initial configuration:**

$$\xi_0^i = \xi_0'^{\,i} \in E_0 = E' \qquad \text{i.i.d.} \quad \eta_0 = \mathrm{Law}(X_0) = \mathrm{Law}(X_0')$$

- **Selection:** $(Y_n = y_n)$

$$\xi_n = \left(\xi'^{\,i}_{0,n}, \ldots, \xi'^{\,i}_{n,n}\right)_{1\leq i\leq N} \in E_n^N \to \widehat{\xi}_n = \left(\widehat{\xi}'^{\,i}_{0,n}, \ldots, \widehat{\xi}'^{\,i}_{n,n}\right)_{1\leq i\leq N} \in E_n^N$$

$$\Updownarrow$$

Selection of N i.i.d. path particles $\widehat{\xi}_n^i = \left(\widehat{\xi}'^{\,i}_{0,n}, \ldots, \widehat{\xi}'^{\,i}_{n,n}\right)$ with discrete law

$$\begin{aligned}
\Psi_n\left(\frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i}\right) &= \sum_{i=1}^N \frac{g_n(\xi_n^i)}{\sum_{j=1}^N g_n(\xi_n^j)}\,\delta_{\xi_n^i} \\
&= \sum_{i=1}^N \frac{g_n(\xi'^{\,i}_{0,n}, \ldots, \xi'^{\,i}_{n,n})}{\sum_{j=1}^N g_n(\xi'^{\,j}_{0,n}, \ldots, \xi'^{\,j}_{n,n})}\,\delta_{(\xi'^{\,i}_{0,n}, \ldots, \xi'^{\,i}_{n,n})}
\end{aligned}$$

- **Mutation:** Each selected path-particle

$$\widehat{\xi}_n^i = \left(\widehat{\xi}'^{\,i}_{0,n}, \ldots, \widehat{\xi}'^{\,i}_{n,n}\right)$$

is extended with an elementary K'_{n+1}-transition

$$\begin{aligned}
\xi_{n+1}^i &= \left(\xi'^{\,i}_{0,n+1}, \xi'^{\,i}_{1,n+1}, \ldots, \xi'^{\,i}_{n,n+1},\ \xi'^{\,i}_{n+1,n+1}\right) \\
&= \Big(\underbrace{(\widehat{\xi}'^{\,i}_{0,n},\ \widehat{\xi}'^{\,i}_{1,n},\ \ldots\ ,\ \widehat{\xi}'^{\,i}_{n,n})}_{=\widehat{\xi}_n^i},\ \xi'^{\,i}_{n+1,n+1}\Big) \in E_{n+1} = E_n \times E'
\end{aligned}$$

with

$$\xi'^{\,i}_{n+1,n+1} \sim K'_{n+1}(\widehat{\xi}'^{\,i}_{n,n},\ .\ )$$

Example 3.7

$$Y_n = H_n(X'_0, \ldots, X'_n) + V_n \qquad V_n \text{ i.i.d. } \sim p_n(v)\ dv$$

$$\Downarrow$$

$$\sum_{i=1}^{N} \frac{g_n(\xi_n^i)}{\sum_{j=1}^{N} g_n(\xi_n^j)}\ \delta_{\xi_n^i} = \sum_{i=1}^{N} \frac{p_n(y_n - H_n(\xi'^{\,i}_{0,n}, \ldots, \xi'^{\,i}_{n,n}))}{\sum_{j=1}^{N} p_n(y_n - H_n(\xi'^{\,j}_{0,n}, \ldots, \xi'^{\,j}_{n,n}))}\ \delta_{(\xi'^{\,i}_{0,n}, \ldots, \xi'^{\,i}_{n,n})}$$

If $H_n(x'_0, \ldots, x'_n) = h_n(x'_n)$ *then*

$$\sum_{i=1}^{N} \frac{g_n(\xi_n^i)}{\sum_{j=1}^{N} g_n(\xi_n^j)}\ \delta_{\xi_n^i} = \sum_{i=1}^{N} \frac{p_n(y_n - h_n(\xi'^{\,i}_{n,n}))}{\sum_{j=1}^{N} p_n(y_n - h_n(\xi'^{\,j}_{n,n}))}\ \delta_{(\xi'^{\,i}_{0,n}, \ldots, \xi'^{\,i}_{n,n})}$$

$$\Downarrow \quad (g_n(x'_0, \ldots, x'_n) = g'_n(x'_n))$$

Historical process/Genealogical tree

$\xi'^{\,i}_n = \xi'^{\,i}_{n,n} \quad \rightarrow \quad$ **genetic algo with selection/mutation** (g'_n, K'_n)

$(\xi'^{\,i}_{0,n}, \ldots, \xi'^{\,i}_{n,n}) =$ **complete ancestral line of** $\xi'^{\,i}_n$

$\boxed{N\textbf{-approximating measures}}$ $(N \to \infty)$

The empirical measures of the genealogical trees

$$\eta_n^N = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i} = \frac{1}{N}\sum_{i=1}^N \delta_{(\xi'^{\,i}_{0,n},\ldots,\xi'^{\,i}_{n,n})} \longrightarrow \eta_n$$

and the weighted measures

$$\widehat{\eta}_n^N = \sum_{i=1}^N \frac{g_n(\xi_n^i)}{\sum_{j=1}^N g_n(\xi_n^j)}\delta_{\xi_n^i} = \sum_{i=1}^N \frac{g_n(\xi'^{\,i}_{0,n},\ldots,\xi'^{\,i}_{n,n})}{\sum_{j=1}^N g_n(\xi'^{\,j}_{0,n},\ldots,\xi'^{\,j}_{n,n})}\delta_{(\xi'^{\,i}_{0,n},\ldots,\xi'^{\,i}_{n,n})} \longrightarrow \widehat{\eta}_n$$

or

$$\widehat{\eta}_n^N = \frac{1}{N}\sum_{i=1}^N \delta_{\widehat{\xi}_n^i} = \frac{1}{N}\sum_{i=1}^N \delta_{(\widehat{\xi}'^{\,i}_{0,n},\ldots,\widehat{\xi}'^{\,i}_{n,n})} \longrightarrow \widehat{\eta}_n$$

with

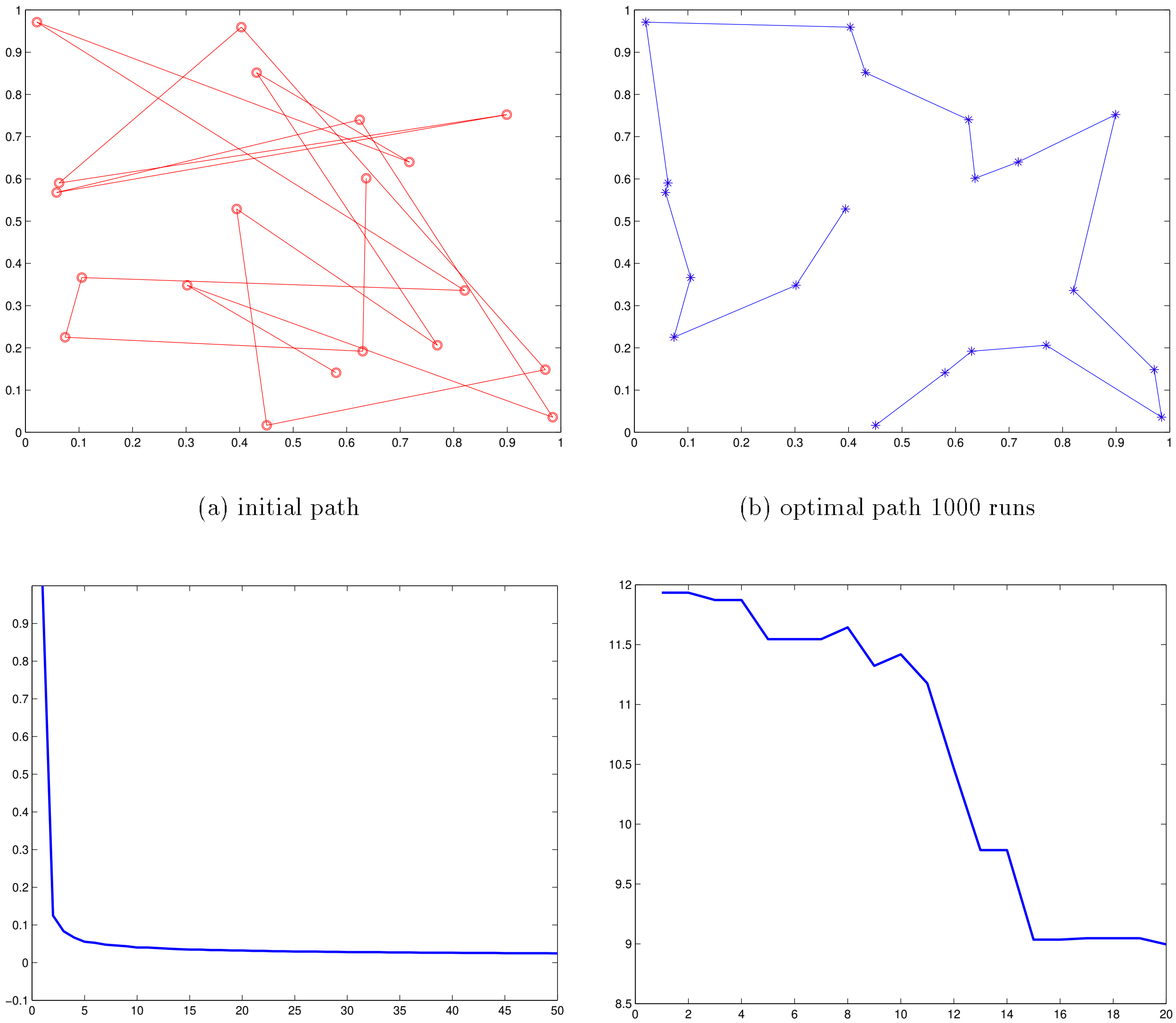
$$\eta_n(f) = \mathbb{E}(f(X'_0,\ldots,X'_n)\,|Y_0 = y_0,\ldots,Y_{n-1} = y_{n-1})$$

and

$$\widehat{\eta}_n(f) = \mathbb{E}(f(X'_0,\ldots,X'_n)\,|Y_0 = y_0,\ldots,Y_{n-1} = y_{n-1}, Y_n = y_n)\,.$$

More details/Proofs

$\longrightarrow$ **Genealogies and increasing propagation of chaos for Feynman-Kac and genetic models**
L.Miclo, P. DM, *The Annals of Applied Probability,*, to appear (2001).



(a) initial path

(b) optimal path 1000 runs

(c) temperature (50 runs)

(d) potential (20 runs)

Figure 1: Optimal path 20 cities (1000 runs).

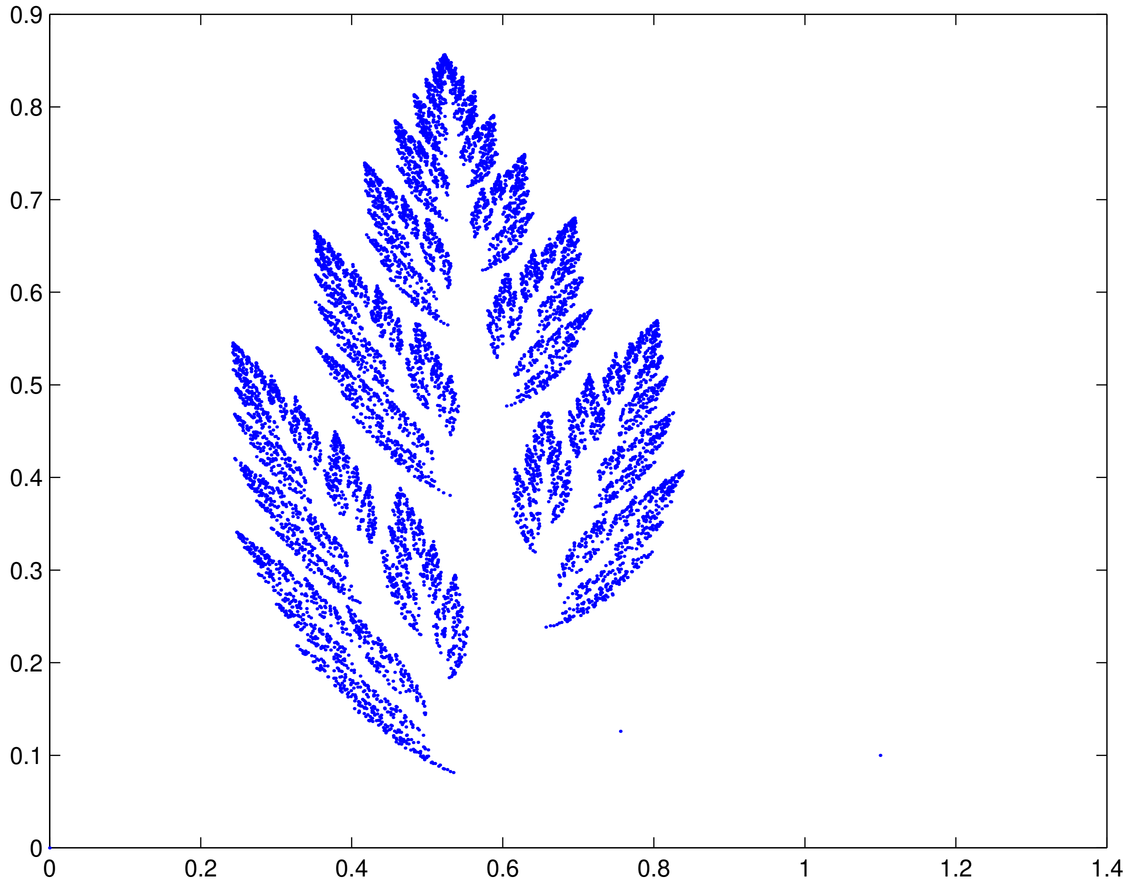


Figure 2: 10000 runs

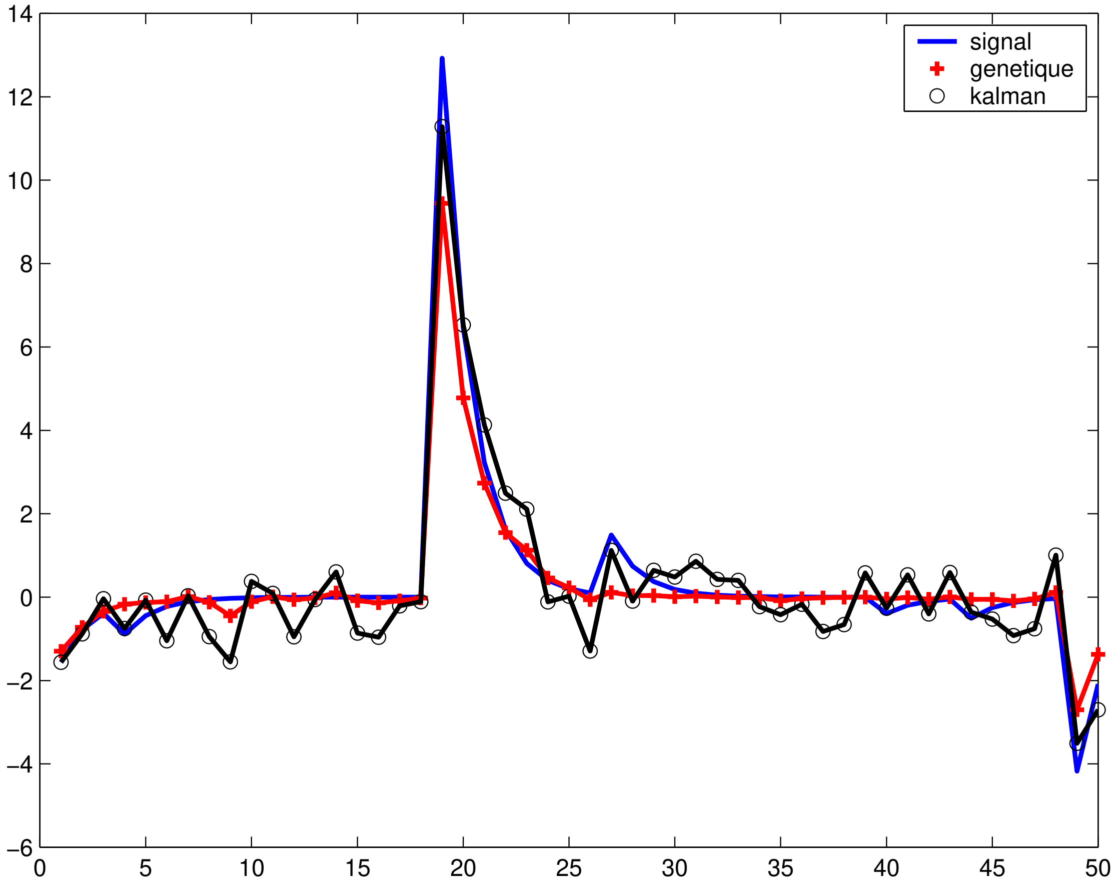
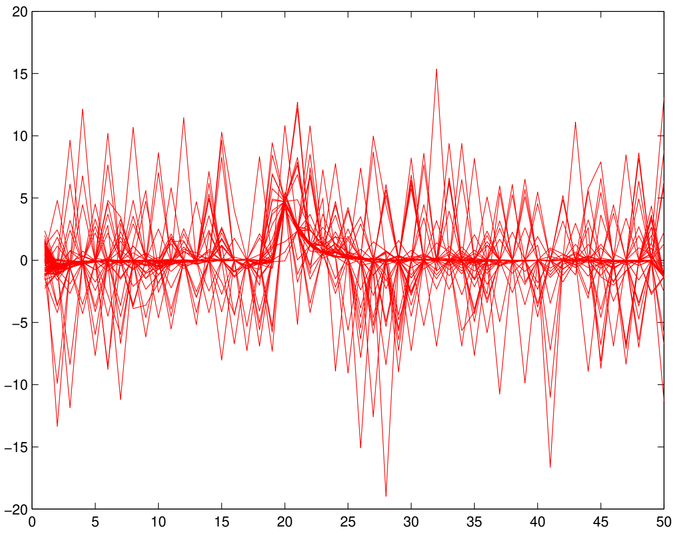
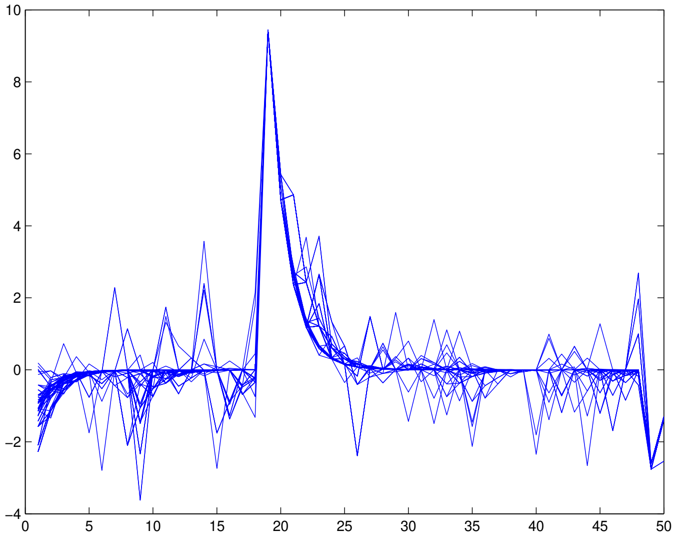


Figure 3: [linear non Gaussian] Kalman-Bucy/Genetic particle filters (100 particles) $a = 0.5$, $b = 1$ et $c = 1$.



(a) Mutation



(b) Selection

Figure 4: [linear non Gaussian] Mutation/Selection with N=100 particles

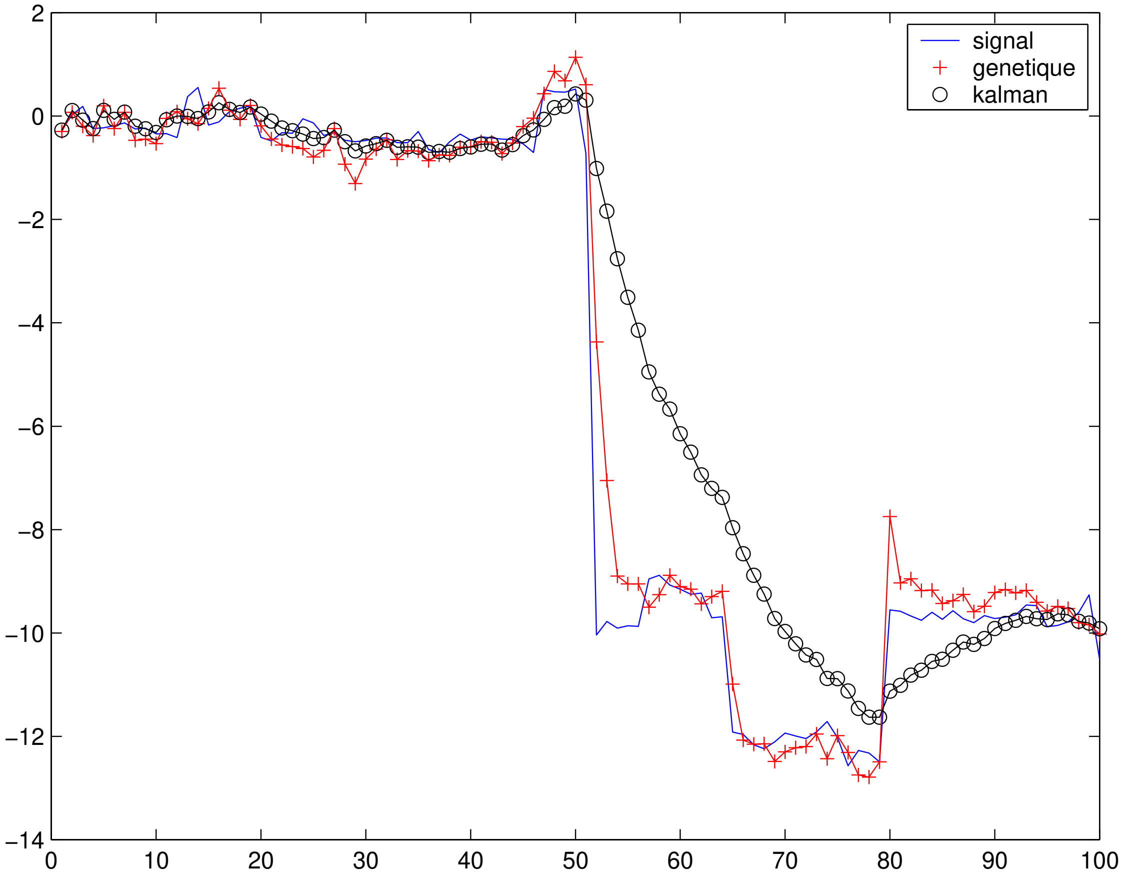


Figure 5: [linear/Cauchy signal noise] Kalman-Bucy/Genetic particle filters (300 particles).