$\boxed{X_n \overset{W}{\Rightarrow} X} \iff$ (a) or (b) or (c) or (d).

(a) : $\forall h$ bounded and continuous, $\mathbb{E}(h(X_n)) \xrightarrow[n\to\infty]{} \mathbb{E}(h(X))$.

(b) : $\forall$ set A s.t $\mathbb{P}(X \in \partial A) = 0$, $\underbrace{\mathbb{E}(\mathbb{1}_A(X_n))}_{\mathbb{P}(X_n \in A)} \xrightarrow[n\to+\infty]{} \underbrace{\mathbb{E}(\mathbb{1}_A(X))}_{\mathbb{P}(X \in A)}$

(c) $\forall x \in \mathbb{R}$ s.t $\mathbb{P}(X = x) = 0$, $\underbrace{\mathbb{E}(\mathbb{1}_{]-\infty,x]}(X_n))}_{\mathbb{P}(X_n \leq x)} \longrightarrow \mathbb{E}(\mathbb{1}_{]-\infty,x]}(X))$.

$\mathbb{P}(X_n \leq x) \xrightarrow[n\to\infty]{} \mathbb{P}(X \leq x)$

cdf (cumulative distribution function).

(d) $\forall u \in \mathbb{R}$, $\mathbb{E}(\underbrace{e^{iuX_n}}_{h_u(X_n)}) \xrightarrow[n\to+\infty]{} \mathbb{E}(\underbrace{e^{iuX}}_{h_u(X)})$. (Characteristic function).

$\boxed{X_n \xrightarrow{\mathbb{P}\text{-prob}} X} \iff \forall \varepsilon > 0, \quad \mathbb{P}(|X_n - X| \geq \varepsilon) \xrightarrow[n\to+\infty]{} 0.$

$\boxed{X_n \xrightarrow{\mathbb{P}\text{-a.s}} X} \iff \mathbb{P}\left(\lim_{n\to+\infty} X_n = X\right) = 1.$

Prop:

If : $\begin{cases} X_n \xrightarrow{W} X. \\ X_n \xrightarrow{\mathbb{P}\text{-prob}} X. \\ X_n \xrightarrow{\mathbb{P}\text{-a.s}} X \end{cases}$ then: $\forall f$ continuous, $\begin{cases} f(X_n) \xrightarrow{W} f(X). \\ f(X_n) \xrightarrow{\mathbb{P}\text{-prob}} f(X) \\ f(X_n) \xrightarrow{\mathbb{P}\text{-a.s}} f(X). \end{cases}$

We have: $\boxed{X_n \xrightarrow{\mathbb{P}\text{-a.s}} X \Rightarrow X_n \xrightarrow{\mathbb{P}\text{-prob}} X \Rightarrow X_n \overset{W}{\Rightarrow} X.}$

Strong Law of Large Numbers:

If : $\begin{cases} (1)\ (X_i)_{i\geq 1} & \text{are i.i.d (independent and identically distributed)} \\ (2)\ \mathbb{E}(|X_1|) < \infty. \end{cases}$

then: $\lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mathbb{E}(X_1)$. $\mathbb{P}$-a.s.

Central limit theorem.

If : $\begin{cases} (1)\ (X_i)_{i\geq 1}. & \text{are i.i.d.} \\ (2)\ \mathbb{E}(X_1^2) < \infty. \end{cases}$

then:

$$Z_n = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}(X_1)}{\sqrt{\frac{Var(X_1)}{n}}} \xrightarrow{\mathcal{L}} Z \quad \text{where } Z \sim \mathcal{N}(0,1).$$

Equivalently : $\tilde{Z}_n = \sqrt{n}\left(\underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i}_{\overline{X}_n.} - \mathbb{E}(X_1)\right) \xrightarrow{\mathcal{L}} \tilde{Z}$ where $\tilde{Z} \sim \mathcal{N}(0, Var(X_1))$.

$$\mathbb{E}(Z_n.) = \frac{1}{\sqrt{\frac{Var(X_1)}{n}}} \underbrace{\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}(X_1)\right)}_{\underbrace{\frac{1}{n}\sum_{i=1}^{n} \underbrace{\mathbb{E}(X_i)}_{\mathbb{E}(X_1).} - \mathbb{E}(X_1)}_{\underbrace{\mathbb{E}(X_1) - \mathbb{E}(X_1).}_{=0.}}} = 0$$

$$Var(Z_n) = \frac{1}{\frac{Var(X_1)}{n}} \underbrace{Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}(X_1).\right)}_{\boxed{\underbrace{Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)}_{\frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right)}}}.$$

because (X_i) are independent

$$= \frac{1}{n^2}\sum_{i=1}^{n} \overbrace{Var(X_i)}^{Var(X_1).} = \frac{1}{n^2}\, n.Var(X_1)$$

$$= \frac{1}{n}\operatorname{Var}(X_1).$$

$$= \frac{1}{\cancel{\frac{\operatorname{Var}(X_1)}{n}}} \cdot \cancel{\frac{\operatorname{Var}(X_1)}{n}} = 1.$$

<u>Remark:</u> $Z_n \xrightarrow{\mathcal{L}} Z \Rightarrow \tilde{Z}_n \xrightarrow{\mathcal{L}} \tilde{Z}$.

<u>Indeed:</u> $\left.\begin{cases} Z_n \xrightarrow{\mathcal{L}} Z \\ f(z) = z\sqrt{\operatorname{Var}(X_1)} \text{ continuous} \end{cases}\right\} \Rightarrow \begin{cases} \tilde{Z}_n = f(Z_n) \\ \quad = Z_n \times \sqrt{\operatorname{Var}(X_1)} \end{cases} \xrightarrow{\mathcal{L}} \underbrace{f(Z)}_{\tilde{Z}} = Z\sqrt{\operatorname{Var}(X_1)}$

Since: $Z \sim \mathcal{N}(0,1)$, $\tilde{Z} = Z\sqrt{\operatorname{Var}(X_1)} \sim \mathcal{N}(\underset{\mathbb{E}(\tilde{Z})}{\underset{\uparrow}{0}}, \underbrace{\operatorname{Var}(Z)}_{1}\cdot\operatorname{Var}(X_1))$

$$\tilde{\sigma}_N^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mathbb{E}(X))^2, \quad \hat{\sigma}_N^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X}_N)^2.$$

Let us show: $\mathbb{E}[\tilde{\sigma}_N^2] = \mathbb{E}(\hat{\sigma}_N^2) = \sigma^2$ (where: $\sigma^2 = \operatorname{Var}(X_1)$).

$$\tilde{\sigma}_N^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mathbb{E}(X_1))^2. \qquad \text{Set: } \tilde{X}_i = X_i - \mathbb{E}(X_1).$$

$$\tilde{\sigma}_N^2 = \frac{1}{N}\sum_{i=1}^{N}\tilde{X}_i^2$$

$$\mathbb{E}[\tilde{\sigma}_N^2] = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\tilde{X}_i^2\right] = \frac{1}{N}\mathbb{E}\left[\sum_{i=1}^{N}\tilde{X}_i^2\right].$$

$$= \frac{1}{N}\sum_{i=1}^{N}\underbrace{\mathbb{E}[\tilde{X}_i^2]}_{\mathbb{E}[\tilde{X}_1^2]}.$$

$$= \frac{\cancel{N}}{\cancel{N}}\mathbb{E}[\tilde{X}_1^2].$$

$$= \mathbb{E}[\tilde{X}_1^2] = \mathbb{E}[(X_1 - \mathbb{E}(X_1))^2] = \operatorname{Var}(X_1) = \sigma^2$$

$$\boxed{\hat{\sigma}_N^2} = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \overline{X}_N\right)^2 \quad . \text{ Setting } \tilde{X}_i = X_i - E(X_1) \qquad \overline{X}_N = \frac{1}{N}\sum_{j=1}^{N} X_j .$$

$$= \frac{1}{N-1}\sum_{i=1}^{N}\left(\tilde{X}_i + \cancel{E(X_1)} - \left(\frac{1}{N}\sum_{j=1}^{N}\tilde{X}_j + \cancel{E(X_1)}\right)\right)^2 \qquad \overline{\tilde{X}}_N = \underbrace{\frac{1}{N}\sum_{i=1}^{N}\tilde{X}_i}$$

$$= \frac{1}{N-1}\sum_{i=1}^{N}\left(\tilde{X}_i - \overline{\tilde{X}}_N\right)^2$$

$$= \frac{1}{N-1}\sum_{i=1}^{N}\left(\tilde{X}_i^2 + \overline{\tilde{X}}_N^2 - 2\tilde{X}_i . \overline{\tilde{X}}_N\right)$$

$$= \frac{1}{N-1}\sum_{i=1}^{N}\tilde{X}_i^2 + \frac{N}{N-1}\overline{\tilde{X}}_N^2 - \frac{2}{N-1}\underbrace{\sum_{i=1}^{N}\tilde{X}_i}_{N.\overline{\tilde{X}}_N} . \overline{\tilde{X}}_N$$

$$\hat{\sigma}_N^2 = \frac{1}{N-1}\sum_{i=1}^{N}\tilde{X}_i^2 - \frac{N}{N-1}\overline{\tilde{X}}_N^2$$

$$\boxed{\mathbb{E}\left(\hat{\sigma}_N^2\right)} = \mathbb{E}\left[\frac{1}{N-1}\sum_{i=1}^{N}\tilde{X}_i^2 - \frac{N}{N-1}\overline{\tilde{X}}_N^2\right]$$

$$= \frac{1}{N-1}\sum_{i=1}^{N}\underbrace{\mathbb{E}(\tilde{X}_i^2)}_{\underbrace{\mathbb{E}(\tilde{X}_1^2)}_{\underbrace{\mathbb{E}\left[\left(X_1 - \mathbb{E}(X_1)\right)^2\right]}_{\sigma^2}}} - \frac{N}{N-1}\underbrace{\mathbb{E}\left(\overline{\tilde{X}}^2\right)}_{\underbrace{\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^{N}(X_i - \mathbb{E}(X_1))\right)^2\right]}_{\underbrace{Var\left(\frac{1}{N}\sum_{i=1}^{N}X_i\right)}_{=\frac{Var(X_1)}{N} = \frac{\sigma^2}{N}.}}}$$

$$= \frac{N}{N-1}\sigma^2 - \frac{\cancel{N}}{N-1} . \frac{\sigma^2}{\cancel{N}}.$$

$$= \frac{N-1}{N-1}\sigma^2 = \boxed{\sigma^2.}$$

Recall that $\tilde{\sigma}_N^2 = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \mathbb{E}(X_1)\right)^2$.

By the Law of Large Numbers, $\frac{1}{N}\sum_{i=1}^{N}\left(X_i - \mathbb{E}(X_1)\right)^2 \xrightarrow[N\to+\infty]{a.s.} \underbrace{\mathbb{E}\left(\left(X_1 - \mathbb{E}(X_1)\right)^2\right)}_{\sigma^2.}$

$$\hat{\sigma}_N^2 = \frac{1}{N-1}\underbrace{\sum_{i=1}^N \tilde{X}_i^2}_{,} - \frac{N}{N-1}\underbrace{\tilde{\bar{X}}_N^2}_{\left(\frac{1}{N}\sum_{i=1}^N X_i - \mathbb{E}(X_1)\right)^2}$$

$$= \underbrace{\frac{N}{N-1}}_{\substack{\longrightarrow 1 \\ N\to+\infty}}\left(\underbrace{\frac{1}{N}\sum_{i=1}^N (X_i - \mathbb{E}(X_1))^2}_{\substack{\xrightarrow{a.s} \sigma^2 \\ (LLN)}}\right) - \underbrace{\frac{N}{N-1}}_{\substack{\longrightarrow 1 \\ N\to+\infty}}\left(\underbrace{\frac{1}{N}\sum_{i=1}^N X_i - \mathbb{E}(X_1)}_{\substack{\longrightarrow 0 \\ N\to+\infty}}\right)^2$$

$$\xrightarrow[N\to+\infty]{a.s} 1\times\sigma^2 - 1\times 0 = \sigma^2.$$

An important tool: Slutsky's Lemma:

$$\left[\ \text{if } \begin{cases} X_n \xrightarrow{\mathcal{L}} X \\ \left(Y_n \xrightarrow{\mathbb{P}\text{-prob}} a\right) \Leftrightarrow \left(Y_n \xrightarrow{\mathcal{L}} a\right) \end{cases} \quad \text{then } \forall f \text{ continuous, } f(X_n, Y_n) \xrightarrow{\mathcal{L}} f(X, a).\right.$$

Remark: $Y_n \xrightarrow{\mathbb{P}\text{-prob}} Y \Rightarrow Y_n \xrightarrow{\mathcal{L}} Y$.

If Y is a constant called "a" then: $Y_n \xrightarrow{\mathcal{L}} a \Rightarrow Y_n \xrightarrow{\mathbb{P}\text{-prob.}} a$.

Let us show: $\dfrac{\bar{X}_n - \mathbb{E}(X_1)}{\sqrt{\dfrac{\tilde{\sigma}_N^2}{N}}} \xrightarrow{\mathcal{L}} Z$ where $Z \sim \mathcal{N}(0,1)$

By the Central Limit Theorem, $Z_n = \dfrac{\bar{X}_n - \mathbb{E}(X_1)}{\sqrt{\dfrac{\sigma^2}{n}}} \xrightarrow{\mathcal{L}} Z$ where $Z \sim \mathcal{N}(0,1)$

Moreover: $U_n = \sqrt{\dfrac{\sigma^2}{\tilde{\sigma}_N^2}} \xrightarrow[n\to+\infty]{a.s} 1$ because: $\tilde{\sigma}_N^2 \to \sigma^2$ ($\tilde{\sigma}_N^2$ strongly convergent).

So: $U_n \xrightarrow{\mathbb{P}\text{-prob}} 1 = U$.

Set $f(z,u) = z \times u$, f is continuous

By Slutsky's Lemma,

$$\frac{\bar{X}_n - \mathbb{E}(X_1)}{\sqrt{\frac{\tilde{\sigma}_N^2}{N}}} = \overbrace{\frac{\bar{X}_n - \mathbb{E}(X_1)}{\sqrt{\frac{\sigma^2}{n}}}}^{Z_n} \times \overbrace{\sqrt{\frac{\sigma^2}{\tilde{\sigma}_N^2}}}^{U_n.} = f(Z_n, U_n) \xRightarrow{\mathcal{L}} \underbrace{f(Z, \underset{=1}{U})}_{Z \times 1 = Z}.$$

. δ-method

If: $\left\{\begin{array}{l} \overbrace{\sqrt{n}(X_n - a)}^{Z_n} \xrightarrow{\mathcal{L}} Z \\ g \text{ differentiable at } a \end{array}\right\} \Rightarrow \underbrace{\sqrt{n}\,(g(X_n) - g(a))}_{\tilde{Z}_n.} \xrightarrow{\mathcal{L}} g'(a) Z.$

$$\tilde{Z}_n = \sqrt{n}\,(g(X_n) - g(a)) = \underbrace{\sqrt{n}\,(X_n - a)}_{Z_n}. \times \underbrace{\frac{g(X_n) - g(a)}{X_n - a.}}_{U_n.} = f(Z_n, U_n)$$

. **Now:** if $\boxed{U_n \xrightarrow{\mathbb{P}\text{-prob}} g'(a)}$ (*)

then, since: $Z_n \xrightarrow{\mathcal{L}} Z$, we have by Slutsky's Lemma:

$$f(Z_n, U_n) = Z_n \times U_n = \tilde{Z}_n \xrightarrow{\mathcal{L}} f(Z, g'(a)) = g'(a) Z.$$

To conclude, it remains to prove: $U_n \xrightarrow{\mathbb{P}\text{-prob}} g'(a).$

if $\left\{\begin{array}{l} \boxed{X_n \xrightarrow{\mathbb{P}\text{-prob}} a.} \quad (**) \\ x \overset{\varphi}{\longmapsto} \left\{\begin{array}{ll} \frac{g(x) - g(a)}{x - a} & x \neq a. \\ g'(a) & x = a. \end{array}\right. \quad \Bigg| \ \varphi \text{ continuous} \end{array}\right] \Rightarrow \begin{array}{l} \varphi(X_n) \xrightarrow{\mathbb{P}\text{-prob}} \varphi(a). \\ \underbrace{\frac{g(X_n) - g(a)}{X_n - a}}_{U_n} \xrightarrow{\mathbb{P}\text{-prob}} g'(a) \end{array}$

Let us show (**), $\left(\begin{array}{l} Z_n = \sqrt{n}\,(X_n - a) \xrightarrow{\mathcal{L}} Z. \\ U_n = \frac{1}{\sqrt{n}} \xrightarrow{\mathbb{P}\text{-prob}} 0. \end{array}\right.$

$$X_n = Z_n \times \frac{1}{\sqrt{n}} + a = Z_n \times U_n + a := \Psi(Z_n, U_n)$$

where : $\varphi(z,u) = z \times u + a$. is continuous.

By Slutsky's lemma: $\underbrace{\varphi(Z_n, U_n)}_{Z_n U_n + a = X_n} \xrightarrow{\mathcal{L}} \underbrace{\varphi(Z, 0)}_{Z \times 0 + a = a}$.

Then $X_n \xrightarrow{\mathcal{L}} a$ which is equivalent to: $X_n \xrightarrow{\mathbb{P}\text{-prob}} a$.

Day 2:

Confidence intervals. $S_N = \frac{1}{N} \sum_{i=1}^{N} f(X_i)$, $V_N = \frac{1}{N-1} \sum_{i=1}^{N} (f(X_i) - S_N)^2$.

We have seen that: $V_n \xrightarrow{\mathbb{P}\text{-a.s}} \sigma^2 = \text{Var}(f(X_1))$. (Lemma 3.6)

$$A_N = \sqrt{N}\, \frac{S_N - \mathbb{E}(f(X))}{\sqrt{V_N}} = \overbrace{\frac{\sqrt{N}\,(S_N - \mathbb{E}(f(X_1)))}{\sqrt{\sigma^2}}}^{Z_n} \times \overbrace{\sqrt{\frac{\sigma^2}{V_N}}}^{U_n}$$

$\frac{\sqrt{N}\,(S_N - \mathbb{E}(f(X_1)))}{\sqrt{\sigma^2}} \xrightarrow{\mathcal{L}} G$ where $G \sim \mathcal{N}(0,1)$ (CLT).

$\sqrt{\frac{\sigma^2}{V_N}} \xrightarrow{\mathbb{P}\text{-prob}} 1$ (by Lemma 3.6)

Slutsky's lemma.

$$\left(\begin{array}{l} Z_n \xrightarrow{\mathcal{L}} G. \\ U_n \xrightarrow{\mathbb{P}\text{-prob}} 1. \end{array} \right) \Rightarrow \underbrace{f(Z_n, U_n)}_{= Z_n U_n = A_N} \xrightarrow{\mathcal{L}} \underbrace{f(G,1)}_{= G \times 1 = G}.$$

where $f(z,u) = z \times u$. continuous.

(*)

According to: (*).,

$$\underbrace{\mathbb{P}\left(-a \leq A_N = \sqrt{N}\left|\frac{S_N - \mathbb{E}(f(X))}{\sqrt{V_N}}\right| \leq a\right)}_{\mathbb{E}(f(X)) \in I_n} \xrightarrow[n \to +\infty]{} \underbrace{\mathbb{P}(-a \leq G \leq a)}_{0.95}.$$

$(a = 1.96)$.

(where $G \sim \mathcal{N}(0,1)$)

$$-a \leq \sqrt{N} \left(\frac{S_N - \mathbb{E}(f(X))}{\sqrt{V_N}} \right) \leq a$$

$$\iff -a\sqrt{V_N} \leq \sqrt{N}\left(S_N - \mathbb{E}(f(X))\right) \leq a\sqrt{V_N}.$$

$$\iff \sqrt{N}\, S_N - a\sqrt{V_N} \leq \sqrt{N}\, \mathbb{E}(f(X)) \leq \sqrt{N}(S_N) + a\sqrt{V_N}$$

$$\iff S_N - a\sqrt{\frac{V_N}{N}} \leq \mathbb{E}(f(X)) \leq S_N + a\sqrt{\frac{V_N}{N}}.$$

$$\iff \mathbb{E}(f(X)) \in \underbrace{\left[S_N - a\sqrt{\frac{V_N}{N}}, S_N + a\sqrt{\frac{V_N}{N}}\right]}_{I_N}.$$

Length of $I_N \sim 2a\sqrt{\frac{\sigma^2}{N}} \underset{N\to+\infty}{\longrightarrow} 0.$

Chapter 4:

. $t \overset{F_Y}{\mapsto} \mathbb{P}(Y \leq t)$ is the cdf (cumulative distribution function) of a random variable Y.

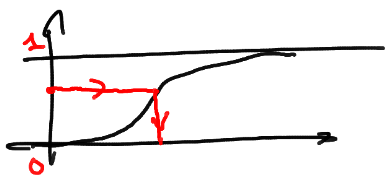
. F_Y is in $[0,1]$. , $\begin{cases} \lim\limits_{t\to-\infty} F_Y(t) = 0. \\ \lim\limits_{t\to+\infty} F_Y(t) = 1. \end{cases}$

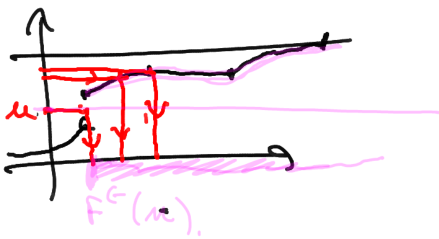
$F_Y \nearrow$. (not always strictly increasing).

. if $a<b$ and $\underbrace{F_Y(a)}_{\mathbb{P}(Y\leq a)} = F_Y(b)$ $\iff$ $\mathbb{P}(a<Y\leq b) = 0$

$\mathbb{P}(Y\leq a) = \mathbb{P}(Y\leq b) = \mathbb{P}(Y\leq a) + \mathbb{P}(a<Y\leq b)$

. if $\lim\limits_{\substack{t\to y \\ t<y}} \underbrace{F_Y(t)}_{\mathbb{P}(Y\leq t)} < \underbrace{F_Y(y)}_{\mathbb{P}(Y\leq y)}$. then: $\mathbb{P}(Y=y) = \underbrace{\mathbb{P}(Y\leq y)}_{F_Y(y)} - \lim\limits_{\substack{t\to y \\ t<y}} \underbrace{\mathbb{P}(Y\leq t)}_{F_Y(t)}$





Mathematically speaking,

$F^{\leftarrow}$ $\quad F_Y^{-1}(u) = \inf\{y \,;\, F_Y(y) \geq u\}$

such that.

<u>Proposition 4.2</u>: if $U \sim \mathrm{Unif}[0,1]$,

$$F_Y^{-1}(U) \overset{\mathcal{L}}{=} Y.$$

<u>Example</u>: $Y \sim \exp(\lambda)$ ie: Y has the density $f(y) = \lambda e^{-\lambda y} \mathbb{1}(y \geq 0)$.

$$\forall t \in \mathbb{R}, \quad \underbrace{\mathbb{P}(Y \leq t)}_{F_Y(t)} = \int_{-\infty}^{t} f(y)\, dy .$$

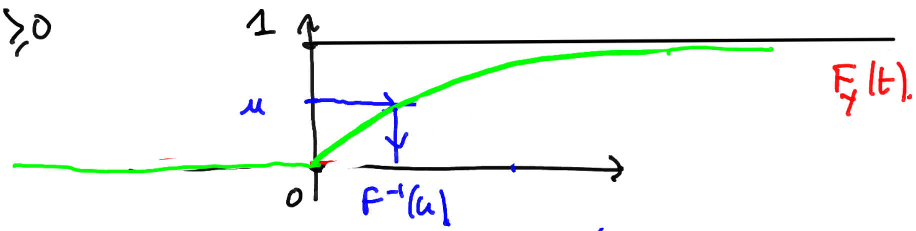
$$= \begin{cases} 0 & \text{if } t \leq 0 \\ \int_0^t f(y)dy & \text{if } t \geq 0 \end{cases}$$

$$= \begin{cases} 0 & \text{if } t \leq 0. \\ \int_0^t \lambda e^{-\lambda y} dy = \left[-e^{-\lambda y}\right]_0^t = -e^{-\lambda t} + 1 & \text{if } t \geq 0 \end{cases}$$

$$F_Y(t) = \begin{cases} 0 & \text{if } t \leq 0 \\ 1 - e^{-\lambda t} & \text{if } t \geq 0 \end{cases}$$

$f_\lambda(t) = \lambda e^{-\lambda t}$.

$f_\lambda'(t) = -\lambda^2 e^{-\lambda t}$.



$F_Y(t) = u = 1 - e^{-\lambda t}$ <u>ie</u>: $\quad 1 - u = e^{-\lambda t}$

<u>ie</u>: $\ln(1-u) = -\lambda t$.

<u>ie</u>: $t = -\dfrac{\ln(1-u)}{\lambda}$.

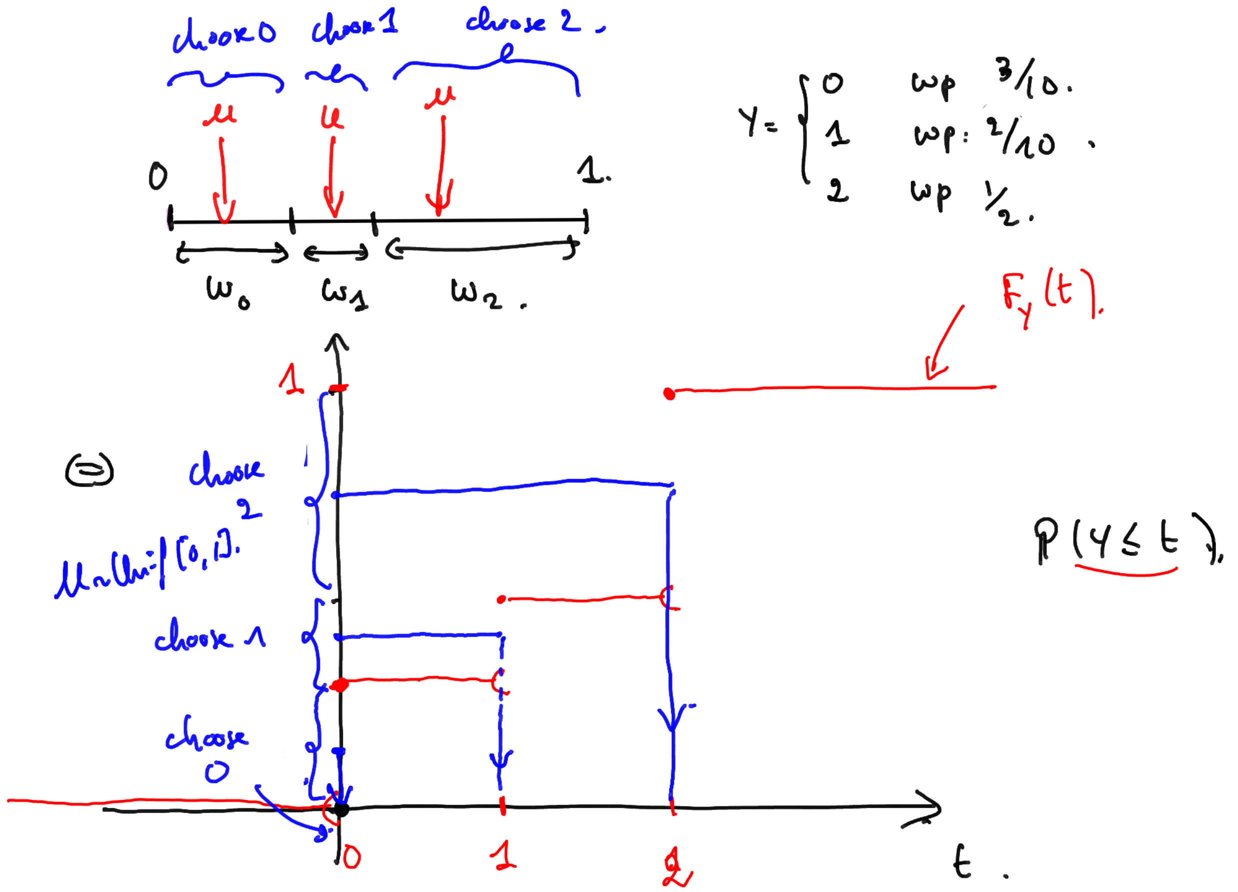
if $U \sim \mathrm{Unif}[0,1]$.

then: $-\dfrac{\ln(1-U)}{\lambda} \overset{\mathcal{L}}{=} Y$

(where $Y \sim \exp(\lambda)$)

if $V \sim \mathrm{Unif}[0,1]$.

$$-\frac{\ln V}{\lambda} \sim \exp(\lambda).$$



Rejection sampling.

f is the target distribution.

Assume that | $\cdot\ \forall x \in \mathbb{R},\quad f(x) \leq M g(x)$, i.e. $\dfrac{f(x)}{Mg(x)} \leq 1$.
| $\cdot$ You can sample from g.

Algorithm:

- Sample $X \sim g$. and $U \sim \mathrm{Unif}[0,1]$.
- While $U > \dfrac{f(X)}{Mg(X)}$, do: sample $\left| \begin{array}{l} X \sim g \\ U \sim \mathrm{Unif}[0,1]. \end{array} \right.$
- output $Y = X$.

i.e: (X_i, U_i) iid, X_i indep. of U_i, $\left| \begin{array}{l} X_i \sim g. \\ U_i \sim \mathrm{Unif}[0,1]. \end{array} \right.$

$$T = \inf \left\{ t \in \mathbb{N}_* \text{ s.t.: } U_i \leq \frac{f(X_i)}{Mg(X_i)} \right\}.$$

Then: $Y = X_T$.

We have that $Y \sim f$.

<u>Proof:</u> $\forall A,\ \forall k \in \mathbb{N}_* = \{1, 2, 3, \dots\}$.

$$\mathbb{P}(Y \in A,\ T = k)$$

$$= \mathbb{P}(X_T \in A,\ T = k)$$

$$= \mathbb{P}(X_k \in A,\ T = k) = \mathbb{P}\left(X_k \in A,\ U_k \leq \frac{f(X_k)}{Mg(X_k)},\ U_{k-1} > \frac{f(X_{k-1})}{Mg(X_{k-1})}, \dots, U_1 > \frac{f(X_1)}{Mg(X_1)}\right).$$

$$= \underbrace{\mathbb{P}\left(X_k \in A,\ U_k \leq \frac{f(X_k)}{Mg(X_k)}\right)}_{\iint \mathbb{1}_A(x)\, \mathbb{1}\left(u \leq \frac{f(x)}{Mg(x)}\right) \cdot \underbrace{g(x)\, \mathbb{1}_{[0,1]}(u)}_{\text{joint density of } (X,U)}\, dx\, du} \times \underbrace{\mathbb{P}\left(U_{k-1} > \frac{f(X_{k-1})}{Mg(X_{k-1})}\right)}_{} \times \dots \times \underbrace{\mathbb{P}\left(U_1 > \frac{f(X_1)}{Mg(X_1)}\right)}_{}.$$

$$\times \left(\iint \mathbb{1}\left(u > \frac{f(x)}{Mg(x)}\right) g(x)\, \mathbb{1}_{[0,1]}(u)\, dx\, du\right)^{k-1}$$

$$= \int_{-\infty}^{+\infty} \mathbb{1}_A(x) \underbrace{\left(\int_0^{\frac{f(x)}{Mg(x)}} du\right)}_{\frac{f(x)}{M\cancel{g(x)}}} \cancel{g(x)}\, dx \times \left[\int_{-\infty}^{+\infty} \underbrace{\left[\int_{\frac{f(x)}{Mg(x)}}^{1} du\right]}_{\left(1 - \frac{f(x)}{Mg(x)}\right)} \underline{g(x)}\, dx\right]^{k-1}$$

$$= \int_{-\infty}^{+\infty} \mathbb{1}_A(x) \frac{f(x)}{M}\, dx \times \underbrace{\left[\int_{-\infty}^{+\infty} \left(g(x) - \frac{f(x)}{M}\right) dx\right]^{k-1}}_{\left(1 - \frac{1}{M}\right)^{k-1}}$$

$$\left(\text{since: } \int_{-\infty}^{+\infty} g(x)\, dx = 1, \quad \int_{-\infty}^{+\infty} f(x)\, dx = 1\right)$$

$$= \underbrace{\int_A f(x)\, dx}_{\mathbb{P}(Y \in A)} \times \underbrace{\frac{1}{M} \times \left(1 - \frac{1}{M}\right)^{k-1}}_{\mathbb{P}(T = k)} = \mathbb{P}(Y \in A,\ T = k).$$

<u>ie:</u> $Y \perp\!\!\!\perp T$ (Y is indep. of T).

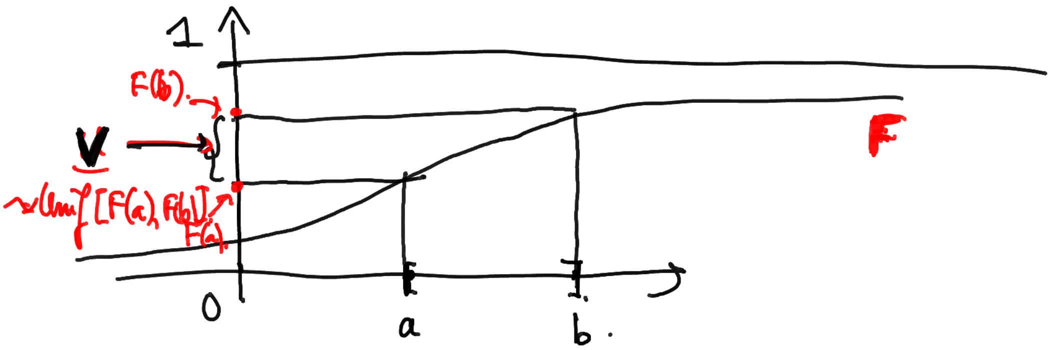
$\underline{Y \sim f}$, $\quad \underline{T \sim \text{Geom}\left(\frac{1}{M}\right)}$.

Sampling a conditional distribution. <u>Aim</u>: sample $X|_{X\in[a,b]}$, where $X \sim g$.

1) <u>Method 1</u>.

. Repeat $X \sim g$. until $X \in [a,b]$. set $Y = X$.

2) <u>Method 2</u>: Sampling by the quantile function.



if $U \sim \text{Unif}[0,1]$.

$$V = F(a) + U(F(b) - F(a)) \sim \text{Unif}([F(a), F(b)]).$$

$$F^{-1}(V) = F^{-1}(F(a) + U(F(b) - F(a))) \sim \frac{g(x)\,\mathbb{1}_{[a,b]}(x)}{\int_a^b g(u)\,du}.$$

<u>Importance sampling methods</u>:

Approximate: $\overbrace{\int f(x)\, h(x)\, dx}^{I}$. where f is a density.

$$\underbrace{\int f(x)\, h(x)\, dx}_{\mathbb{E}_f[h(X)]}$$

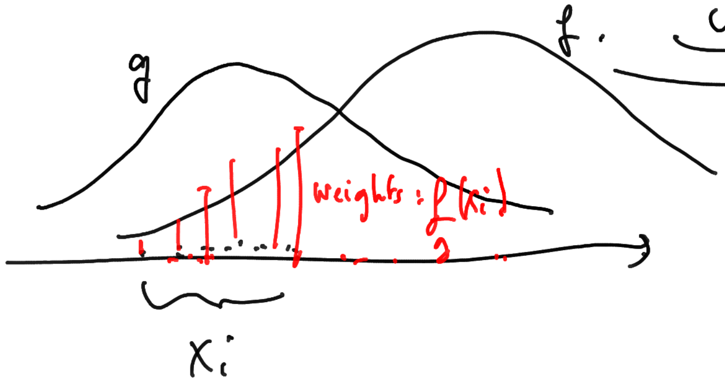
(ie $\mathbb{E}(h(X))$ when $X \sim f$).

$$I = \underbrace{\int g(x) \overbrace{\frac{f(x)}{g(x)}}^{\text{weight}} h(x)\, dx}_{= \mathbb{E}_g\left[\frac{f(X)}{g(X)} h(X)\right]}.$$

where $X \sim g$.

Sample: X_i iid, $X_1 \sim g$.

Approximate I by: $\left.\frac{1}{N}\sum_{i=1}^{N} \underbrace{\overbrace{\frac{f(X_i)}{g(X_i)}}^{\text{weights}}\, h(X_i)}_{f}\right) \xrightarrow[N\to\infty]{LLN} I$



Variant: if $f(x)$ is known up to a multiplicative constant:

$f(x) = C\,\tilde{f}(x)$ where C is unknown.

$$\underbrace{\frac{\sum_{i=1}^{N} \frac{\tilde{f}(X_i)}{g(X_i)}\, h(X_i)}{\sum_{i=1}^{N} \frac{\tilde{f}(X_i)}{g(X_i)}}} = \frac{\left.\frac{1}{N}\sum_{i=1}^{N} \frac{f(X_i)}{g(X_i)}\, h(X_i)\right] \to \int \cancel{g(x)}\frac{f(x)h(x)}{\cancel{g(x)}}\,dx}{\left.\frac{1}{N}\sum_{i=1}^{N} \frac{f(X_i)}{g(X_i)}.\right) \to \int \cancel{g(x)}\frac{f(x)}{\cancel{g(x)}}\,dx.}$$

$$\to \frac{\int f(x)\,h(x)\,dx}{\int f(x)\,dx} = \underbrace{\int f(x)\,h(x)\,dx}_{I}.$$

with $\int f(x)\,dx = 1$

- $\frac{1}{N}\sum_{i=1}^{N} \frac{f(X_i)}{g(X_i)}\, h(X_i) = \hat{I}_N(f)$ approximates $I(f) = \int f(x)\,h(x)\,dx$.

$$\mathbb{E}\left[\hat{I}_N(f)\right] = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\left[\frac{f(X_i)}{g(X_i)}\, h(X_i)\right]. \qquad X_i \sim g.$$

$$= \frac{1}{N}\sum_{i=1}^{N} \int \cancel{g(x)}\, \frac{f(x)}{\cancel{g(x)}}\, h(x)\,dx. = \int f(x)\,h(x)\,dx = I(f).$$

unbiased estimator.

$$\operatorname{Var}\left(\hat{I}_N(f)\right) = \operatorname{Var}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{f(X_i)}{g(X_i)}h(X_i)\right).$$

$$= \frac{1}{N^2}\underbrace{\operatorname{Var}\left(\sum_{i=1}^{N}\frac{f(X_i)}{g(X_i)}h(X_i)\right)}_{\displaystyle\underbrace{\sum_{i=1}^{N}\operatorname{Var}\left(\frac{f(X_i)}{g(X_i)}h(X_i)\right)}_{N\operatorname{Var}\left(\frac{f(X_1)}{g(X_1)}h(X_1)\right)}}$$

(because (X_i) indep).

Y. (they have the same Law).

$$= \frac{1}{N}\operatorname{Var}\left(\frac{f(X_1)}{g(X_1)}h(X_1)\right).$$

$$\underbrace{\int\left(\frac{f(x)}{g(x)}h(x)\right)^2 g(x)\,dx}_{E(Y^2)} - \underbrace{\left(\int\frac{f(x)}{g(x)}h(x).\ g(x)\,dx\right)^2}_{E(Y)^2}.$$

$$= \frac{1}{N}\left[\underbrace{\int\frac{f^2(x)h^2(x)}{g(x)}\,dx} - \left(\int f(x)h(x)\right)^2\right].$$

<u>Aim</u>: search for density g such that it mimizes: $\displaystyle\int\frac{f^2(x)h^2(x)}{g(x)}\,dx$.
$\left(\int g(x)\,dx = 1\right)$

<u>But</u>: $$\operatorname{Var}(|Y|) = \operatorname{Var}\left(\frac{f(x)}{g(x)}|h(x)|\right) = \int\frac{f^2(x)h^2(x)}{g(x)}\,dx - \left(\int f(x)|h(x)|\,dx\right)^2 \geq 0.$$

This implies: $\displaystyle \int \underbrace{\frac{f^2(x)\, h(x)^2}{g(x)}}\, dx \geq \left(\int f(x)\, |h(x)|\, dx\right)^2$

We attain the lower bound for: $|Y|$ constant

That is: $\displaystyle \frac{f(x)\,|h(x)|}{g^*(x)} = c \quad \forall x.$

That is: $\displaystyle \boxed{g^*(x) = \frac{f(x)\,|h(x)|}{\int f(y)\,|h(y)|\, dy}}$ (then: $\int g(x)\,dx = 1$).

Computersession 2:

$f(x) = \dfrac{1}{\pi(1+x^2)}$.

$$F(x) = \int_{-\infty}^{x} \frac{1}{\pi(1+t^2)}\, dt.$$

$$= \left[\frac{1}{\pi} \operatorname{Atan}(t)\right]_{-\infty}^{x}.$$

$$= \frac{1}{\pi}\operatorname{Atan}(x) - \frac{1}{\pi}\left(-\frac{\pi}{2}\right).$$

$$F(x) = \frac{1}{\pi}\operatorname{Atan}(x) + \frac{1}{2} = u$$

$$\frac{1}{\pi}\operatorname{atan}(x) + \frac{1}{2} = u \implies \operatorname{atan} x = \left(u - \frac{1}{2}\right)\pi.$$

$$x = \underbrace{\tan\left(\pi u - \frac{\pi}{2}\right)}_{F^{-1}(u)} \left(= -\operatorname{cotan}(\pi u)\right)$$

Draw: $U \sim \text{Unif}\,[0,1]$

Set: $X = \tan\left(\pi U - \frac{\pi}{2}\right)$.

For exponential distribution of parameter λ:

$X \sim exp(\lambda)$, $\begin{cases} F(x) = (1-e^{-\lambda x}) &, \forall x \geq 0. \\ F^{-1}(u) = -\frac{\ln(1-u)}{\lambda}. \end{cases}$

$$\mathbb{E}(X) = \frac{1}{\lambda}, \quad Var(X) = \frac{1}{\lambda^2}.$$

CLT: $\underbrace{\sqrt{n}\left(\frac{\overline{X}_n - \frac{1}{\lambda}}{\sqrt{1/\lambda^2}}\right)}_{\sqrt{n}\left(\lambda \overline{X}_n - 1\right)} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$.

$\sqrt{n}\left(\lambda \overline{X}_n - 1\right) \xrightarrow{\mathcal{L}} G$ where $G \sim \mathcal{N}(0,1)$.

$$\mathbb{P}(\underbrace{-a \leq \sqrt{n}(\lambda \overline{X}_n - 1) \leq a}_{-\frac{a}{\sqrt{n}} \leq \lambda \overline{X}_n - 1 \leq a.}) \xrightarrow[n \to +\infty]{} \mathbb{P}(-a \leq G \leq \underset{\substack{\| \\ 1.96.}}{a}) = 0.95.$$

$\Leftrightarrow -\frac{a}{\sqrt{n}\,\overline{X}_n} \leq \lambda - \frac{1}{\overline{X}_n} \leq \frac{a}{X_n}$.

$\Leftrightarrow \lambda \in \underbrace{\left[\frac{1}{\overline{X}_n} \pm \frac{a}{\sqrt{n}(\overline{X}_n)}\right]}$.

$\Leftrightarrow \lambda \in \underbrace{\left[\frac{1}{\overline{X}_n} - \frac{a}{\sqrt{n}\,\overline{X}_n}, \frac{1}{\overline{X}_n} + \frac{a}{\sqrt{n}\,\overline{X}_n}.\right]}_{a = 1.96.}$

$Y \sim \mathcal{N}(0,1)$. Y has density: $\tilde{f}(y) = \frac{e^{-y^2/2}}{\sqrt{2\pi}}$.

$Y|_{Y \in \mathbb{R}^+}$ this condition distribution has density: $\frac{\tilde{f}(y)\,\mathbb{1}_{\mathbb{R}^+}(y)}{\int \tilde{f}(z)\,\mathbb{1}_{\mathbb{R}^+}(z)\,dz}$. $= f(y)$

$$f(y) = \frac{\frac{e^{-y^2/2}}{\sqrt{2\pi}} \mathbb{1}_{\mathbb{R}^+}(y)}{\underbrace{\int_{\mathbb{R}^+} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz}_{= \frac{1}{2}}} = \underbrace{2\frac{e^{-y^2/2}}{\sqrt{2\pi}}}\mathbb{1}_{\mathbb{R}^+}(y).$$

$$g(y) = e^{-y}\mathbb{1}_{\mathbb{R}^+}(y).$$

$$f(y) \leq M\, g(y) = Me^{-y} \quad \forall y \geq 0.$$

$$\frac{f(y)}{\underbrace{e^{-y}}} \leq M \qquad \forall y \geq 0.$$

$$\sup_{y\geq 0} \frac{\frac{2}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}}{e^{-y}} = \sqrt{\frac{2}{\pi}} \sup_{y\geq 0} e^{-\frac{y^2}{2}+y} = \sqrt{\frac{2}{\pi}} \sup_{y\geq 0} e^{-\frac{1}{2}\underbrace{(y^2-2y)}_{((y-1)^2-1)}}$$

$$= \sqrt{\frac{2}{\pi}} \sup_{y\geq 0} e^{-\frac{1}{2}(y-1)^2} e^{\frac{1}{2}}$$

$$= \sqrt{\frac{2}{\pi}} e^{1/2}$$

Then: $f(y) \leq M g(y)$ where $M = \sqrt{\frac{2}{\pi}} e^{1/2}$.

Chap. 5. Markov chain Monte Carlo. (MCMC)

Idea: Given a target density π, we construct a Markov chain $X_0, X_1, \dots, X_n, \dots$ such that $\text{Law}\,(X_n) \xrightarrow[n\to+\infty]{} \pi$.

. Recall that a Markov chain satisfies: Law of $X_n /_{X_0,\dots,X_{n-1}}$ = Law of $X_n /_{X_{n-1}}$.

- $X_{n+1} = \underbrace{f}(X_n, \varepsilon_n)$ where ε_n iid. then (X_n) is a Markov chain.

if f does not depend on n, the Markov chain is homogeneous.
f does depend on n, ———— unhomogeneous.

$$\mathbb{E}(h(X_n) \mid X_{n-1}) = \int h(x) \underbrace{Q(X_{n-1}, dx)}$$

Q is the Markov kernel.

Example: $X_n = a X_{n-1} + \underbrace{\varepsilon_n}$ where (ε_n) iid. $\varepsilon_1 \sim \mathcal{N}(0,1)$.

$$X_n \big|_{X_{n-1}} \sim \mathcal{N}(a X_{n-1}, 1).$$

$$\underbrace{X_n \big|_{X_{n-1}}}_{\downarrow} \sim \underbrace{Q(X_{n-1}, \cdot)}_{\text{kernel}}. \qquad y \mapsto \underbrace{q(X_{n-1}, y)}_{\text{kernel density}} = \underbrace{\frac{e^{-(y - a X_{n-1})^2/2}}{\sqrt{2\pi}}}_{\text{density of } \mathcal{N}(aX_{n-1}, 1).}$$

distribution de X_n conditionnally on X_{n-1}.

Q is a kernel if
- $\forall x \in X$, $Q(x, \cdot)$ is a probability measure
- $\forall A \in \mathcal{B}(X)$ $x \mapsto Q(x, A)$. is measurable.

if X_n has density π_n.
X_{n+1} ———— π_{n+1}.

$\underbrace{(X_n, X_{n+1})}$ has density : $\underbrace{\pi_n(X_n)}\ \underbrace{p(X_{n+1} \mid X_n)}$.

to obtain π_{n+1} (density of X_{n+1}),
we marginalize : $\int p(x_n, x_{n+1})\, dx_n$

$$= \int \pi_n(x_n)\, p(x_{n+1} \mid x_n)\, dx_n.$$
$$= \pi_{n+1}(x_{n+1})$$

$$\underbrace{\pi_{n+1}}(\underbrace{x_{n+1}}) = \int \underbrace{\pi_n(x_n)}\ \underbrace{p(\widehat{x_{n+1}} \mid x_n)}_{q(x_n, x_{n+1})}\, dx_n.$$

$q(x_n, x_{n+1}) \to$ kernel density.

Letting $n \to +\infty$.

$$\pi(y) = \int \pi(x)\, q(x,y)\, dx. \quad \Big| \to . \ \pi \text{ is invariant.}$$

Draw indep.

$$\begin{cases} Z \begin{cases} \nearrow 0 & \text{wp: } \alpha. \\ \searrow 1 & \text{wp } 1-\alpha. \end{cases} \\ \\ \begin{vmatrix} Y_1 \sim \mathcal{N}(a,1). \\ Y_2 \sim \mathcal{N}(b,1). \end{vmatrix} \end{cases} \qquad \begin{cases} \text{if } Z=0 \text{ then } Y = Y_1. \\ \text{if } Z=1 \text{ then } Y = Y_2 \end{cases}$$

$$Y = Z Y_2 + (1-Z) Y_1.$$

$$\begin{aligned} \mathbb{E}(Y^2) &= \mathbb{E}(Y^2 \mathbb{1}(Z=0)) + \mathbb{E}(Y^2 \mathbb{1}(Z=1)). \\ &= \mathbb{E}(Y_1^2 \mathbb{1}(Z=0)) + \mathbb{E}(Y_2^2 \mathbb{1}(Z=1)). \\ &= \alpha \underbrace{\mathbb{E}(Y_1^2)}_{\underbrace{Var(Y_1)}_{1} + \underbrace{\mathbb{E}(Y_1)^2}_{a^2}} + (1-\alpha) \underbrace{\mathbb{E}(Y_2^2)}_{\underbrace{Var(Y_2)}_{1.} + \underbrace{\mathbb{E}(Y_2)^2}_{b^2.}} \\ &= \alpha(1+a^2) + (1-\alpha)(1+b^2) \end{aligned}$$