

$$X_n \xrightarrow{w} X \Leftrightarrow (a) \text{ or } (b) \text{ or } (c) \text{ or } (d).$$

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

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

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

(d) $\forall u \in \mathbb{R}$, $\mathbb{E}[e^{iuX_n}] \xrightarrow{n \rightarrow \infty} \mathbb{E}[e^{iuX}]$. (Characteristic function)
 \downarrow
 cdf (cumulative distribution function).

$$X_n \xrightarrow{P\text{-prob}} X \Leftrightarrow \forall \varepsilon > 0, P(|X_n - X| \geq \varepsilon) \xrightarrow{n \rightarrow \infty} 0.$$

$$X_n \xrightarrow{P\text{-a.s.}} X \Leftrightarrow P\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1.$$

Prop:

If: $\left\{ \begin{array}{l} X_n \xrightarrow{w} X. \\ X_n \xrightarrow{P\text{-prob}} X. \\ X_n \xrightarrow{P\text{-a.s.}} X. \end{array} \right.$ Then: $\forall f$ continuous, $\left\{ \begin{array}{l} f(X_n) \xrightarrow{w} f(X). \\ f(X_n) \xrightarrow{P\text{-prob}} f(X). \\ f(X_n) \xrightarrow{P\text{-a.s.}} f(X). \end{array} \right.$

We have: $X_n \xrightarrow{P\text{-a.s.}} X \Rightarrow X_n \xrightarrow{P\text{-prob}} X \Rightarrow X_n \xrightarrow{w} X.$

Chap 3.

Strong law of Large Numbers:

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

Then: $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = \mathbb{E}(X_1) \quad \text{P-a.s.}$

Central limit theorem.

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

Then:

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

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

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

$$\begin{aligned} \text{Var}(Z_n) &= \frac{1}{\frac{\text{Var}(X_1)}{n}} \text{Var} \left(\frac{1}{n} \sum_{i=1}^n X_i - \mathbb{E}(X_1) \right) \\ &= \frac{1}{\frac{\text{Var}(X_1)}{n}} \underbrace{\text{Var} \left(\frac{1}{n} \sum_{i=1}^n X_i \right)}_{\frac{1}{n^2} \text{Var} \left(\sum_{i=1}^n X_i \right)} \quad \text{because } (X_i) \text{ are independent} \\ &= \frac{1}{\frac{\text{Var}(X_1)}{n}} \left(\frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) \right) = \frac{1}{\frac{\text{Var}(X_1)}{n}} \left(\frac{1}{n^2} n \cdot \text{Var}(X_1) \right) = 1. \end{aligned}$$

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

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

Remark:

$$Z_n \stackrel{\mathcal{L}}{\Rightarrow} Z \Rightarrow \tilde{Z}_n \stackrel{\mathcal{L}}{\Rightarrow} \tilde{Z}.$$

Indeed: $\left\{ \begin{array}{l} Z_n \stackrel{\mathcal{L}}{\Rightarrow} Z \\ f(z) = z \cdot \sqrt{\text{Var}(X_1)} \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \tilde{Z}_n = f(Z_n) \\ = Z_n \times \sqrt{\text{Var}(X_1)} \end{array} \right. \xrightarrow{\mathcal{L}} \underbrace{f(Z)}_{\tilde{Z}} = Z \sqrt{\text{Var}(X_1)}$

continuous.

since: $Z \sim \mathcal{N}(0,1)$, $\tilde{Z} = Z \sqrt{\text{Var}(X_1)} \sim \mathcal{N}\left(0, \underbrace{\text{Var}(Z)}_1 \cdot \text{Var}(X_1)\right)$

\uparrow
 $\mathbb{E}(\tilde{Z})$

$$\sigma_N^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \mathbb{E}(X_1))^2.$$

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

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

$$\sigma_N^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \mathbb{E}(X_1))^2.$$

Set: $\tilde{X}_i = X_i - \mathbb{E}(X_1)$.

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

$$\mathbb{E}[\sigma_N^2]$$

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

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

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

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

$$\begin{aligned}
\hat{\sigma}_N^2 &= \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X}_N)^2 \quad . \text{ Setting } \tilde{X}_i = X_i - E(X_1) \quad \bar{X}_N = \frac{1}{N} \sum_{j=1}^N X_j . \\
&= \frac{1}{N-1} \sum_{i=1}^N \left(\tilde{X}_i + E(X_1) - \left(\frac{1}{N} \sum_{j=1}^N \tilde{X}_j + E(X_1) \right) \right)^2 \quad \bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i \\
&= \frac{1}{N-1} \sum_{i=1}^N \left(\tilde{X}_i - \bar{\tilde{X}}_N \right)^2 \\
&= \frac{1}{N-1} \sum_{i=1}^N \left(\tilde{X}_i^2 + \bar{\tilde{X}}_N^2 - 2 \tilde{X}_i \cdot \bar{\tilde{X}}_N \right) \\
&= \frac{1}{N-1} \sum_{i=1}^N \tilde{X}_i^2 + \frac{N}{N-1} \bar{\tilde{X}}_N^2 - \frac{2}{N-1} \underbrace{\sum_{i=1}^N \tilde{X}_i}_{N \cdot \bar{\tilde{X}}_N} \cdot \bar{\tilde{X}}_N \\
\hat{\sigma}_N^2 &= \frac{1}{N-1} \sum_{i=1}^N \tilde{X}_i^2 - \frac{N}{N-1} \bar{\tilde{X}}_N^2
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}(\hat{\sigma}_N^2) &= \mathbb{E} \left[\frac{1}{N-1} \sum_{i=1}^N \tilde{X}_i^2 - \frac{N}{N-1} \bar{\tilde{X}}_N^2 \right] \\
&= \frac{1}{N-1} \sum_{i=1}^N \underbrace{\mathbb{E}(\tilde{X}_i^2)}_{\mathbb{E}(\tilde{X}_1^2)} - \frac{N}{N-1} \underbrace{\mathbb{E}(\bar{\tilde{X}}_N^2)}_{\mathbb{E}\left[\left(\frac{1}{N} \sum_{i=1}^N (X_i - E(X_1))\right)^2\right]} \\
&\quad \underbrace{\mathbb{E}\left[(X_1 - E(X_1))^2\right]}_{\sigma^2} \\
&= \frac{N}{N-1} \sigma^2 - \frac{N}{N-1} \cdot \frac{\sigma^2}{N} \\
&= \frac{N-1}{N-1} \sigma^2 = \sigma^2 .
\end{aligned}$$

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

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

$$\sigma_n^2 = \frac{1}{N-1} \sum_{i=1}^N \tilde{x}_i^2 - \frac{1}{N-1} \left(\sum_{i=1}^N \tilde{x}_i \right)^2$$

$$= \frac{1}{N-1} \left(\frac{1}{N} \sum_{i=1}^N (X_i - \mathbb{E}(X_1))^2 \right) - \frac{1}{N-1} \left(\frac{1}{N} \sum_{i=1}^N (X_i - \mathbb{E}(X_1)) \right)^2$$

$\xrightarrow[N \rightarrow \infty]{\text{LLN}} 1$ $\xrightarrow[\text{a.s.}]{\text{LLN}} \sigma^2$ $\xrightarrow[N \rightarrow \infty]{\text{LLN}} 1$ $\xrightarrow[N \rightarrow \infty]{} 0$

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

An important tool: Slutsky's Lemma:

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

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

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

Let us show: $\frac{\bar{X}_n - \mathbb{E}(X_1)}{\sqrt{\frac{\sigma_n^2}{N}}} \xrightarrow{\mathcal{L}} z$ where $z \sim \mathcal{N}(0,1)$

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

Moreover: $U_n = \sqrt{\frac{\sigma_n^2}{\sigma^2}} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 1$ because: $\sigma_n^2 \rightarrow \sigma^2$ (σ_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,

$$\frac{\overline{X}_n - E(X_1)}{\sqrt{\frac{\sigma^2}{n}}} = \underbrace{\frac{X_n - E(X_1)}{\sqrt{\sigma^2}}}_{Z_n} \times \underbrace{\sqrt{\frac{\sigma^2}{n}}}_{U_n} = f(Z_n, U_n) \xrightarrow{\mathcal{L}} \underbrace{f(z, u)}_{z \times 1 = z}$$

• δ -method

$$\text{If: } \begin{cases} \sqrt{n}(X_n - a) \xrightarrow{\mathcal{L}} Z \\ g \text{ differentiable at } a \end{cases} \Rightarrow \underbrace{\sqrt{n}(g(X_n) - g(a))}_{\tilde{Z}_n} \xrightarrow{\mathcal{L}} g'(a)Z$$

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

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

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: $U_n \xrightarrow{\text{P-prob}} g'(a)$.

if $X_n \xrightarrow{\text{P-prob}} a$ (**).

$$\left. \begin{array}{l} \left\{ \begin{array}{l} x \mapsto \frac{g(x) - g(a)}{x - a} \quad x \neq a \\ g'(a) \quad x = a \end{array} \right. \\ \varphi \text{ continuous} \end{array} \right\} \Rightarrow \underbrace{\varphi(X_n)}_{\frac{g(X_n) - g(a)}{X_n - a}} \xrightarrow{\text{P-prob}} \underbrace{\varphi(a)}_{g'(a)}$$

U_n

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

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

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

By Slutsky's Lemma: $\underbrace{\Psi(z_n, u_n)}_{z_n u_n + a = X_n} \xrightarrow{\mathcal{L}} \underbrace{\Psi(z, 0)}_{z \times 0 + a = a}$.

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

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: $V_n \xrightarrow{\text{P-prob}} \sigma^2 = \text{Var}(f(X_1))$. (Lemma 3.6).

$$A_N = \frac{\sqrt{N} (S_N - \mathbb{E}(f(X_1)))}{\sqrt{V_N}} = \underbrace{\frac{\sqrt{N} (S_N - \mathbb{E}(f(X_1)))}{\sqrt{\sigma^2}}}_{\xrightarrow{\mathcal{L}} G} \times \underbrace{\frac{\sqrt{\sigma^2}}{\sqrt{V_N}}}_{\xrightarrow{\text{P-prob}} 1} \quad (\text{by Lemma 3.6})$$

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

Slutsky's Lemma. (CLT). $\left(\begin{array}{l} Z_n \xrightarrow{\mathcal{L}} G \\ U_n \xrightarrow{\text{P-prob}} 1 \end{array} \right) \Rightarrow \underbrace{f(z_n, u_n)}_{= z_n u_n = A_N} \xrightarrow{\mathcal{L}} \underbrace{f(G, 1)}_{= G \times 1 = G}$. where $f(z, u) = z \times u$. continuous.

(*)

According to: (*),

$$\underbrace{\mathbb{P}(-\alpha \leq A_N = \frac{\sqrt{N} (S_N - \mathbb{E}(f(X_1)))}{\sqrt{V_N}} \leq \alpha)}_{\mathbb{E}(f(X_1)) \in I_m} \xrightarrow{n \rightarrow \infty} \underbrace{\mathbb{P}(-\alpha \leq G \leq \alpha)}_{0.95} \quad (\alpha = 1.96),$$

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

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

$$\Leftrightarrow -a \sqrt{V_N} \leq \sqrt{N} (S_N - E(f(X))) \leq a \sqrt{V_N}.$$

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

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

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

$$\text{Length of } I_N \sim 2a \sqrt{\frac{\sigma^2}{N}} \xrightarrow{N \rightarrow +\infty} 0.$$

Chapter 4:

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

F_Y is in $[0, 1]$.

$$\begin{cases} \lim_{t \rightarrow -\infty} F_Y(t) = 0. \\ \lim_{t \rightarrow +\infty} F_Y(t) = 1. \end{cases}$$

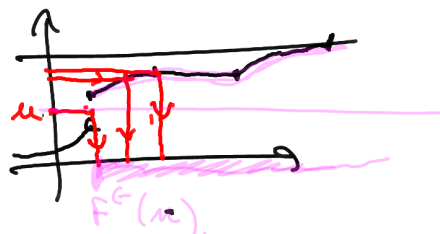
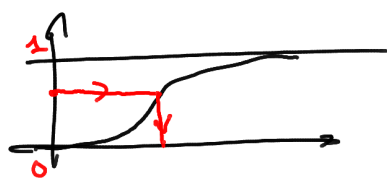
$F_Y \uparrow$. (not always strictly increasing).

if $a < b$ and $F_Y(a) = F_Y(b)$ $\Leftrightarrow P(a < Y \leq b) = 0$

$$P(Y \leq a) \approx P(Y \leq b) = P(Y \leq a) + P(a < Y \leq b)$$

if $\lim_{t \rightarrow y} F_Y(t) < F_Y(y)$. Then: $P(Y=y) = P(Y \leq y) - \lim_{t \rightarrow y} P(Y \leq t)$

$\lim_{t \rightarrow y} P(Y \leq t) = F_Y(y)$



Mathematically speaking,

$$F \leftarrow F_Y^{-1}(u) = \inf \{ y; \underbrace{F_Y(y)}_{\text{such that}} \geq u \}$$

Proposition 4.2: if $U \sim \text{Unif}[0,1]$,
 $F_Y^{-1}(U) \stackrel{\mathcal{L}}{=} Y$.

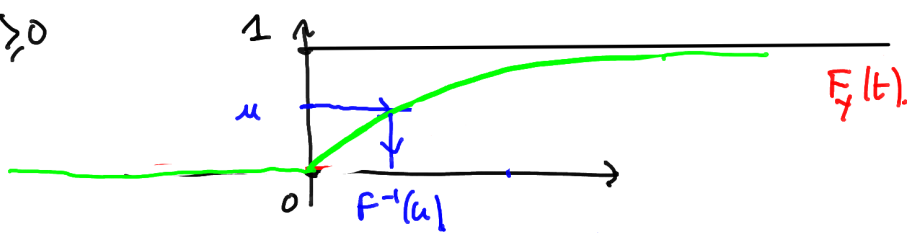
Example: $Y \sim \exp(\lambda)$. ie: Y has the density $f(y) = \lambda e^{-\lambda y} \mathbb{1}_{(y \geq 0)}$.

$$\begin{aligned} \forall t \in \mathbb{R}, \quad \underbrace{P(Y \leq t)}_{F_Y(t)} &= \int_{-\infty}^t f(y) dy, \\ &= \begin{cases} 0 & \text{if } t \leq 0 \\ \int_0^t f(y) dy & \text{if } t \geq 0 \end{cases} \\ &= \begin{cases} 0 & \text{if } t \leq 0. \\ \int_0^t \lambda e^{-\lambda y} dy = [-e^{-\lambda y}]_0^t = -e^{-\lambda t} + 1 & \text{if } t \geq 0 \end{cases} \end{aligned}$$

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

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

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



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

$$\text{ie:} \quad \ln(1-u) = -\lambda t$$

$$\text{ie:} \quad t = \frac{-\ln(1-u)}{\lambda}$$

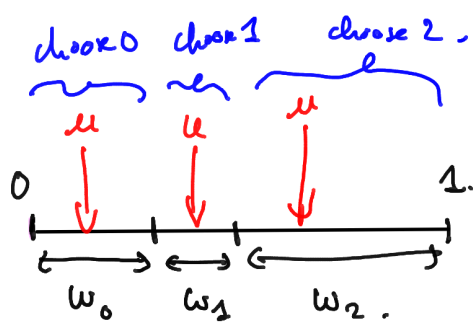
if $U \sim \text{Unif}[0,1]$.

$$\text{then:} \quad -\frac{\ln(1-U)}{\lambda} \stackrel{\mathcal{L}}{=} Y$$

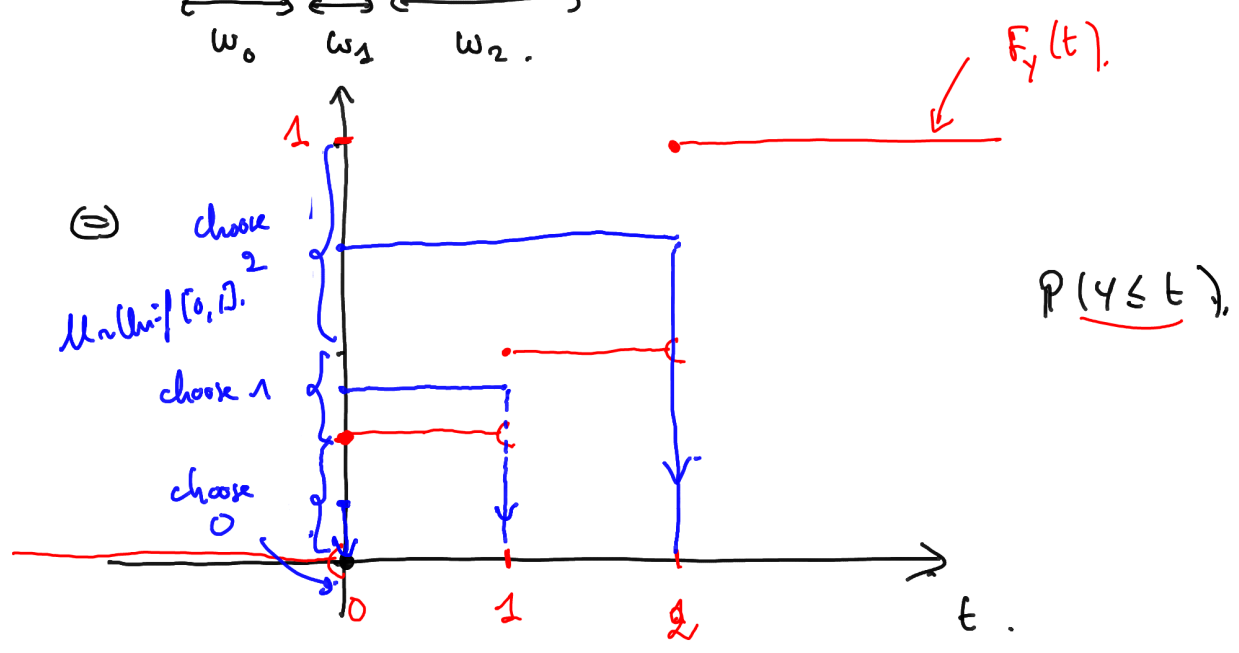
(where $Y \sim \exp(\lambda)$)

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

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



$$Y = \begin{cases} 0 & \text{wp } 3/10. \\ 1 & \text{wp } 2/10. \\ 2 & \text{wp } 1/2. \end{cases}$$



Rejection sampling.

f is the target distribution.

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

Algorithm:

Sample $X \sim g$, and $U \sim \text{Unif}[0,1]$.
 While $U > \frac{f(X)}{M g(X)}$, do: sample $\begin{cases} X \sim g \\ U \sim \text{Unif}[0,1] \end{cases}$.
 output $Y = X$.

i.e.: $(X_i, U_i) \text{ iid}, X_i \text{ indep. of } U_i, \begin{cases} X_i \sim g. \\ U_i \sim \text{Unif}[0,1]. \end{cases}$

$$T = \inf \{ t \in \mathbb{N}_+ \text{ s.t. } U_i \leq \frac{f(X_i)}{M g(X_i)} \}$$

Then: $Y = X_T$.

We have that $Y \sim f$.

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

$$P(Y \in A, T=k).$$

$$= P(X_T \in A, T=k).$$

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

$$= P(X_k \in A, U_k \leq \frac{f(X_k)}{ng(X_k)}) \times P(U_{k-1} > \frac{f(X_{k-1})}{ng(X_{k-1})}) \times \dots \times P(U_1 > \frac{f(X_1)}{ng(X_1)}).$$

$$\int \mathbb{1}_A(x) \mathbb{1}(u \leq \frac{f(x)}{ng(x)}) \cdot \underbrace{g(x) \mathbb{1}_{[0,1]}(u)}_{\text{joint density of } (X,U)} dx du \times \left(\int \mathbb{1}(u > \frac{f(x)}{ng(x)}) g(x) \mathbb{1}_{[0,1]}(u) dx du \right)^{k-1}$$

$$= \int_{-\infty}^{+\infty} \mathbb{1}_A(x) \left(\int_0^{\frac{f(x)}{ng(x)}} du \right) g(x) dx \times \left[\int_{-\infty}^{+\infty} \left[\int_{\frac{f(x)}{ng(x)}}^1 du \right] g(x) dx \right]^{k-1}$$

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

$\left(1 - \frac{1}{n} \right)^{k-1}$

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

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

ie: $Y \perp\!\!\!\perp T$ (Y is indep. of T).

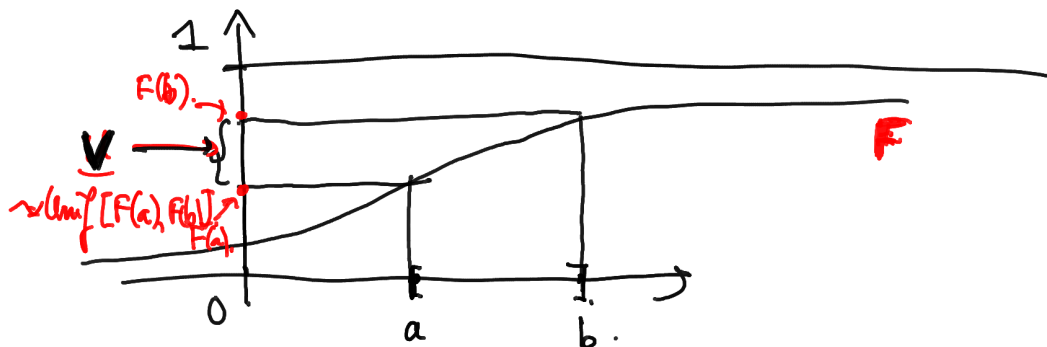
$Y \sim f,$ $T \sim \text{Geom}\left(\frac{1}{n}\right).$

Sampling a conditional distribution. Aim: sample $X \mid X \in [a, b]$, where $X \sim g$.

1) Method 1.

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

2) Method 2: Sampling by the quantile function.



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

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

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

Importance sampling methods:

Approximate: $\int f(x) h(x) dx$ where f is a density.

$\mathbb{E}_f[h(X)]$ (i.e. $\mathbb{E}(h(X))$ when $X \sim f$).

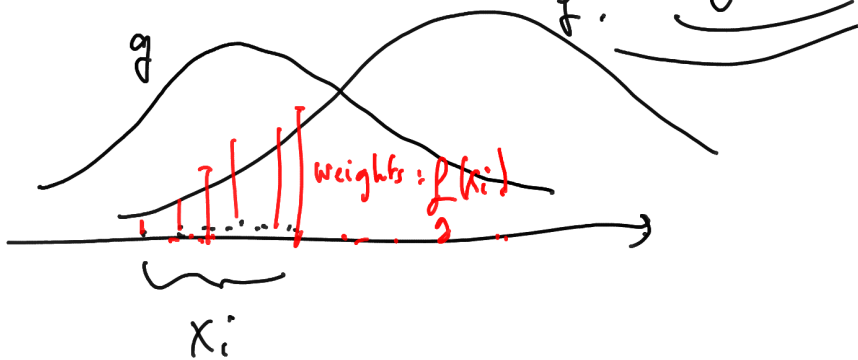
$$I = \int g(x) \underbrace{\frac{f(x)}{g(x)}}_{\text{weight}} h(x) dx.$$

$$= \mathbb{E}_g \left[\frac{f(X) h(X)}{g(X)} \right].$$

where $X \sim g$.

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

Approximate \mathbb{I} by: $\frac{1}{N} \sum_{i=1}^N \frac{f(X_i) h(X_i)}{g(X_i)}$ $\xrightarrow{N \rightarrow \infty} \mathbb{I}$ LLN.



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

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

$$\frac{\frac{1}{N} \sum_{i=1}^N \frac{f(X_i) h(X_i)}{g(X_i)}}{\frac{1}{N} \sum_{i=1}^N \frac{\tilde{f}(X_i)}{g(X_i)}} = \frac{\frac{1}{N} \sum_{i=1}^N \frac{f(X_i) h(X_i)}{g(X_i)}}{\frac{1}{N} \sum_{i=1}^N \frac{f(X_i)}{g(X_i)}} \rightarrow \frac{\int g(x) \frac{f(x) h(x)}{g(x)} dx}{\int g(x) \frac{f(x)}{g(x)} dx}$$

$$\rightarrow \frac{\int f(x) h(x) dx}{\int f(x) dx} = \frac{\int f(x) h(x) dx}{\mathbb{I}}$$

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

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

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

unbiased estimator.

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

$$= \frac{1}{N^2} \text{Var}\left(\sum_{i=1}^N \frac{f(x_i) h(x_i)}{g(x_i)}\right).$$

$$\sum_{i=1}^N \text{Var}\left(\frac{f(x_i) h(x_i)}{g(x_i)}\right) \quad (\text{because } (x_i) \text{ indep.}),$$

$$N \text{Var}\left(\frac{f(x_1) h(x_1)}{g(x_1)}\right) \quad (\text{they have the same Law}),$$

$$= \frac{1}{N} \text{Var}\left(\frac{f(x_1) h(x_1)}{g(x_1)}\right).$$

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

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

Aim: search for density g such that it minimizes: $\int \frac{f^2(x) h^2(x)}{g(x)} dx$.
($\int g(x) dx = 1$)

But: $\text{Var}(Y) = \text{Var}\left(\frac{f(x) |h(x)|}{g(x)}\right) = \int \frac{f^2(x) h^2(x)}{g(x)} dx - \left(\int f(x) |h(x)| dx\right)^2 \geq 0.$

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

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

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

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

Computersession 2:

$$f(x) = \frac{1}{\pi(1+x^2)},$$

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

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

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

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

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

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

Draw: $U \sim \operatorname{Uni}([0, 1])$

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

For exponential distribution of parameter λ :

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

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

CLT:

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

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

$$\underbrace{P(-a \leq \sqrt{n} (\lambda \bar{X}_n - 1) \leq a)}_{-a \leq \lambda \bar{X}_n - 1 \leq a} \xrightarrow{n \rightarrow +\infty} P(-a \leq G \leq a) = 0.95.$$

$\stackrel{||}{=} 1.96.$

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

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

$$\Leftrightarrow \lambda \in \left[\frac{1}{\bar{X}_n} - \frac{a}{\sqrt{n} \bar{X}_n}, \frac{1}{\bar{X}_n} + \frac{a}{\sqrt{n} \bar{X}_n} \right]$$

$a = 1.96.$

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

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

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

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

$$\frac{f(y)}{e^{-y}} \leq \pi \quad \forall y \geq 0.$$

$$\begin{aligned} \sup_{y \geq 0} \frac{\frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}}{e^{-y}} &= \sqrt{\frac{2}{\pi}} \sup_{y \geq 0} e^{-\frac{1}{2}y^2 + y} = \sqrt{\frac{2}{\pi}} \sup_{y \geq 0} e^{-\frac{1}{2}(y^2 - 2y)} \\ &= \sqrt{\frac{2}{\pi}} \sup_{y \geq 0} e^{-\frac{1}{2}(y-1)^2 - 1} = \sqrt{\frac{2}{\pi}} \sup_{y \geq 0} e^{-\frac{1}{2}(y-1)^2} e^{-\frac{1}{2}} \\ &= \sqrt{\frac{2}{\pi}} e^{-\frac{1}{2}} \end{aligned}$$

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

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

Idea: Given a target density π , we construct a Markov chain

$$X_0, X_1, \dots, X_n, \dots \quad \text{such that } \text{Law}(X_n) \xrightarrow{n \rightarrow \infty} \pi.$$

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

$X_{n+1} = f(X_n, \varepsilon_n)$ 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.

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

Q is the Markov kernel.

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

$$X_n | X_{n-1} \sim \mathcal{N}(aX_{n-1}, 1)$$

$$X_n | X_{n-1} \sim \underbrace{Q(X_{n-1}, \cdot)}_{\text{kernel}}$$

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

$$y \mapsto \underbrace{q(X_{n-1}, y)}_{\text{kernel density}} = \frac{e^{-\frac{(y - aX_{n-1})^2}{2}}}{\underbrace{\sqrt{2\pi}}_{\text{density of } \mathcal{N}(aX_{n-1}, 1)}}$$

Q is a kernel if $\left\{ \begin{array}{l} \forall x \in X, Q(x, \cdot) \text{ is a probability measure} \\ \forall A \in \mathcal{B}(X), x \mapsto Q(x, A) \text{ is measurable.} \end{array} \right.$

if X_n has density π_n .

$$X_{n+1} \longrightarrow \pi_{n+1}$$

$$\underbrace{(X_n, X_{n+1})}_{\text{has density}} : \underbrace{\pi_n(X_n)}_{\text{density}} \underbrace{p(X_{n+1} | X_n)}_{\text{kernel density}}$$

to obtain π_{n+1} (density of X_{n+1}),

$$\begin{aligned} \text{we marginalize : } & \int p(X_n, X_{n+1}) dx_n \\ &= \int \pi_n(x_n) p(X_{n+1} | x_n) dx_n \\ &= \pi_{n+1}(X_{n+1}) \end{aligned}$$

$$\underbrace{\pi_{n+1}(x_{n+1})}_{\text{density}} = \int \underbrace{\pi_n(x_n)}_{\text{density}} \underbrace{p(x_{n+1} | x_n)}_{\text{kernel density}} dx_n$$

Letting $n \rightarrow +\infty$.

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

Draw indep.

$$\left. \begin{array}{l} z \rightarrow 0 \text{ wp: } \alpha \\ \quad \searrow 1 \text{ wp: } 1-\alpha \end{array} \right\} \begin{array}{l} \text{if } z=0 \text{ then } Y = Y_1 \\ \text{if } z=1 \text{ then } Y = Y_2 \end{array}$$

$$\left. \begin{array}{l} Y_1 \sim \mathcal{N}(a, 1) \\ Y_2 \sim \mathcal{N}(b, 1) \end{array} \right\}$$

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

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

Normalized importance sampling.

• Let π be a target density.

• Let g be a "instrumental" (or "proposal") density s.t. we can sample from g .

For any function h , we can approximate $\mathbb{I} = \mathbb{E}_\pi(h(X)) = \int h(x) \pi(x) dx$ by:

$$\hat{\mathbb{I}}_n = \frac{\sum_{i=1}^n \frac{\pi(x_i)}{g(x_i)} h(x_i)}{\sum_{j=1}^n \frac{\pi(x_j)}{g(x_j)}} = \sum_{i=1}^n w_i h(x_i) \quad \text{where } w_i = \frac{\frac{\pi(x_i)}{g(x_i)}}{\sum_{j=1}^n \frac{\pi(x_j)}{g(x_j)}}$$

where $\{X_i\}$ iid. $X_i \sim g$.

Remark (1) $\sum_{i=1}^n w_i = 1$ (normalized weights).

(2) if $\tilde{\pi}(x) = c \pi(x)$ then:

$$\hat{I}_n = \frac{\sum_{i=1}^n \frac{\tilde{\pi}(x_i) h(x_i)}{g(x_i)}}{\sum_{i=1}^n \frac{\tilde{\pi}(x_i)}{g(x_i)}}$$

Property: $\hat{I}_n \xrightarrow[n \rightarrow \infty]{a.s.} I$

Indeed:

$$\hat{I}_n = \frac{\frac{1}{n} \sum_{i=1}^n \frac{\pi(x_i) h(x_i)}{g(x_i)}}{\frac{1}{n} \sum_{j=1}^n \frac{\pi(x_j)}{g(x_j)}} \xrightarrow[n \rightarrow \infty]{LLN} \frac{E_g \left[\frac{\pi(x) h(x)}{g(x)} \right]}{E_g \left[\frac{\pi(x)}{g(x)} \right]} = \frac{\int \pi(x) h(x) dx}{\int \pi(x) dx} = I$$

Computer mission 4:

- 1) Implement the normalized importance sampling.
- 2) Resample among the (x_1, \dots, x_n) according to the weight (w_1, \dots, w_n)

(this is possible because:

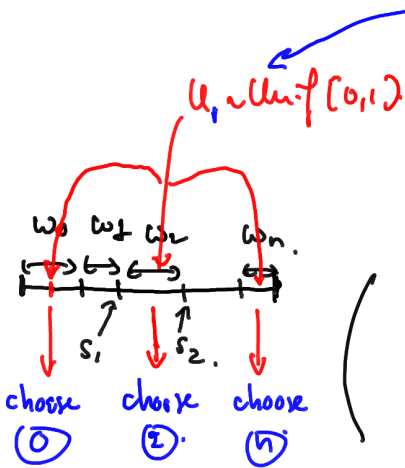
$$w_1 + \dots + w_n = 1)$$

we get: $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_m$ iid s.t.

where: $\forall i \in \{1, \dots, m\}, \forall j \in \{1, \dots, n\}$

$$P(\tilde{X}_i = X_j | X_1, \dots, X_n) = w_j$$

- 3) Check empirically that the law of \tilde{X}_i is close to π .



- 1) X_1, \dots, X_n iid according to g .

w_1, \dots, w_n

- 2) $\tilde{X}_i \rightarrow X_1$ wp w_1
 $\tilde{X}_i \rightarrow X_2$ wp w_2
 $\tilde{X}_i \rightarrow X_n$ wp w_n

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

choose index k where

$$s_{k-1} = w_0 + \dots + w_{k-1} \leq U_i \leq w_0 + \dots + w_k = s_k$$

set $\tilde{X}_i = X_k$

$X_j \stackrel{iid}{\sim} g$.

$$\hat{I}_n = \sum_{j=1}^n w_j h(X_j) \approx \int \pi(x) h(x) dx = I \approx \frac{1}{n} \sum_{i=1}^n h(\hat{X}_i) \quad \text{where } \hat{X}_i \stackrel{iid}{\sim} \pi$$

$$\mathbb{E}(h(\tilde{X}_1)) \quad K = \begin{cases} 1 & \text{w.p. } w_1 \\ \vdots \\ n & \text{--- } w_n \end{cases}$$

$$\tilde{X}_1 = X_K$$

$$= \mathbb{E}(h(\tilde{X}_1) \mathbb{1}_{K=1}) + \dots + \mathbb{E}(h(\tilde{X}_1) \mathbb{1}_{K=n})$$

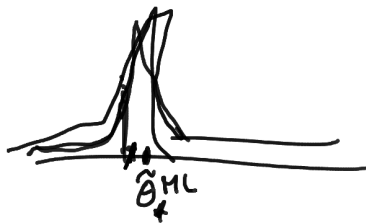
$$= \mathbb{E}(h(X_1) \mathbb{1}_{K=1}) + \dots + \mathbb{E}(h(X_n) \mathbb{1}_{K=n})$$

$$= \mathbb{E}(h(X_1) w_1) + \dots + \mathbb{E}(h(X_n) w_n)$$

$$= \mathbb{E}\left(\underbrace{\sum_{i=1}^n w_i h(X_i)}_{\hat{I}_n}\right) = \mathbb{E}\left(\underbrace{\hat{I}_n}_{\approx I}\right) \approx I = \int \pi(x) h(x) dx$$

Usually: (in Bayesian statistics).

$$\begin{cases} \theta \sim \pi_0 \\ (Y_i)_{i=1}^n \text{ iid} \end{cases} \quad Y_i \sim f_\theta$$



You observe: Y_1, \dots, Y_n .

Law of $\theta | Y_1, \dots, Y_n$ is of density:

$$\pi(\theta | Y_{1:n}) = \frac{p(\theta, Y_{1:n})}{p(Y_{1:n})} = \frac{\pi_0(\theta) \prod_{i=1}^n f_\theta(Y_i)}{\int \pi_0(\theta) \prod_{i=1}^n f_\theta(Y_i) d\theta}$$

$$= \frac{\pi_0(\theta) \prod_{i=1}^n f_\theta(Y_i)}{\int \pi_0(\theta) \prod_{i=1}^n f_\theta(Y_i) d\theta} = C \pi_0(\theta) \prod_{i=1}^n f_\theta(Y_i)$$

Example: (1) $\theta \sim \text{Unif}[0, \lambda]$ ← prior $\lambda = 1$?

(2) $Y_1, \dots, Y_n \text{ iid } \sim \mathcal{N}(\theta, 1)$

from Bayesian

$$\pi(\theta) \propto \prod_{i=1}^n f_\theta(Y_i) = \frac{e^{-\frac{1}{2} \sum_{i=1}^n (Y_i - \theta)^2}}{(2\pi)^{n/2}}$$

Law $\theta | y_{1:n}$ is the a-posteriori Law.
a-posteriori

Metropolis-Hastings algorithm (also called Monte Carlo by Markov Chain, MCMC).
General algorithm.

π is the target density.

x_0 : arbitrary.

For $t=0, \dots, n-1$ do:

$y_t \sim \widehat{q}(x_{t-1}, \cdot)$ (density which depends on x_{t-1}).

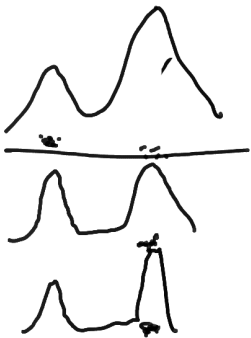
Draw $u_t \sim \text{Unif}[0,1]$, s.t. u_t and y_t indep.

if $u_t \leq \alpha(x_{t-1}, y_t)$, $x_t = y_t$.

otherwise $x_t = x_{t-1}$

where $\alpha(x, y) = \min\left(\frac{\pi(y) \widehat{q}(y, x)}{\pi(x) \widehat{q}(x, y)}, 1\right)$.

Independence sampler: we use: $q(x, y) = q(y)$.

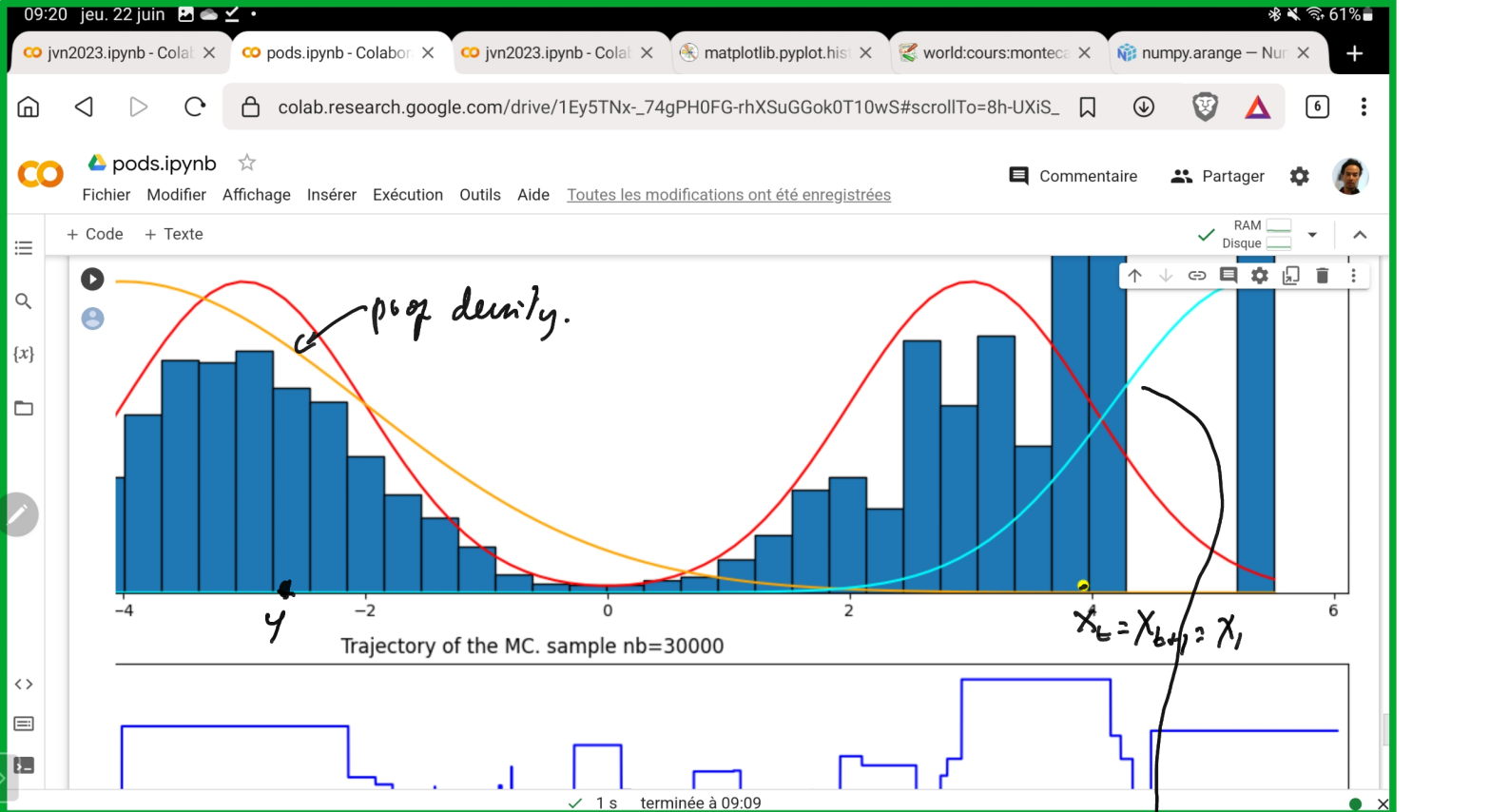


MCMC: error: $p^m \rightarrow 0$ $p \in]0, 1[$.

Resampling / Importance sampling: error: $\frac{1}{n}$.

SMC: (Sequential Monte Carlo) for target tracking.





if Y proposed and $X_t = X$
 you accept Y w.p $\left(\frac{\pi/q(Y)}{\pi/q(X)} \wedge 1 \right)$.

Chap 6 Antithetic variates.

$I = \mathbb{E}[h(U)]$ is approximated by $S_n = \frac{1}{n} \sum_{i=1}^n h(U_i)$.

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

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

$$\text{Var}(S_n) = \frac{\text{Var}(h(U_i))}{n}$$

I may be approximated by: $S_{2n} = \frac{1}{2n} \sum_{i=1}^{2n} h(U_i)$ — $\text{Var}(S_{2n}) = \frac{\text{Var}(h(U_i))}{2n}$

or $\bar{S}_{2n} = \frac{1}{2n} \sum_{i=1}^n h(U_i) + h(1-U_i)$

$$\mathbb{E}(S_{2n}) = \mathbb{E}(\bar{S}_{2n}) = I.$$

$$\frac{1}{2n} \sum_{i=1}^n \underbrace{\mathbb{E}(h(U_i))}_I + \underbrace{\mathbb{E}(h(1-U_i))}_I = \frac{2n}{2n} I = I.$$

$\sim U_i \sim \text{Unif}[0,1]$

Both are unbiased estimators of I .

$$\text{Var}(\bar{S}_{2n}) \stackrel{?}{\leq} \text{Var}(S_{2n}).$$

$$\begin{aligned} \text{Var}(\bar{S}_{2n}) &= \text{Var}\left(\frac{1}{2n} \sum_{i=1}^n h(U_i) + h(1-U_i)\right) \\ &= \left(\frac{1}{2n}\right)^2 \text{Var}\left(\sum_{i=1}^n (h(U_i) + h(1-U_i))\right). \\ &= \frac{1}{4n^2} \times n \times \text{Var}(h(U_1) + h(1-U_1)) \quad \left. \begin{array}{l} \text{(because } (U_i) \\ \text{are indep.)} \\ \text{(because } (U_i) \text{ have} \\ \text{the same Law)} \end{array} \right\} \\ &= \frac{1}{4n} \left[\text{Var}(h(U_1) + h(1-U_1)) \right] \\ &= \frac{1}{4n} \left[\text{Var}(h(U_1)) + \text{Var}(h(1-U_1)) + 2 \text{cov}(h(U_1), h(1-U_1)) \right] \\ &= \frac{1}{4n} \left(2 \text{Var}(h(U_1)) \right) + \frac{2}{4n} \text{cov}(h(U_1), h(1-U_1)) \\ &= \frac{1}{2n} \text{Var}(h(U_1)) + \frac{1}{2n} \text{cov}(h(U_1), h(1-U_1)). \end{aligned}$$

So: $\text{Var}(\bar{S}_n) \leq \text{Var}(S_{2n})$ if and only if: $\boxed{\text{cov}(h(U_1), h(1-U_1)) \leq 0}$

Lemma: if X random variable. h and g are monotone functions then:

$$g(x) = h(1-x)$$

$$\text{cov}(h(X), g(X)) \leq 0.$$

Let

$$A = \text{cov}(h(X), g(X)).$$

$$A = \mathbb{E} \left[(h(X) - \mathbb{E}(h(X))) (g(X) - \mathbb{E}(g(X))) \right].$$

Let Y, Z 2 Random variables s.t.:

$$\begin{cases} X \stackrel{d}{=} Y \\ X \stackrel{d}{=} Z \\ (X, Y, Z) \text{ indep.} \end{cases}$$

$$\mathbb{E}(h(X)) \mathbb{E}(g(X))$$

$$\begin{aligned} \bullet A &= \mathbb{E} \left[\underbrace{(h(X) - \mathbb{E}(h(Y)))}_{\text{red}} \underbrace{(g(X) - \mathbb{E}(g(Z)))}_{\text{red}} \right] \\ &= \mathbb{E} \left[\underbrace{h(X) g(X)}_{\text{red}} - \underbrace{h(X) \mathbb{E}(g(Z))}_{\text{red}} - \underbrace{\mathbb{E}(h(Y)) g(X)}_{\text{red}} + \underbrace{\mathbb{E}(h(Y)) \mathbb{E}(g(Z))}_{\text{red}} \right] \end{aligned}$$

||

$$= \mathbb{E} \left(h(X) g(X) \right) - \underbrace{\mathbb{E}(h(X)) \mathbb{E}(g(Z))}_{\mathbb{E}(h(X) g(Z))} - \underbrace{\mathbb{E}(h(Y)) \mathbb{E}(g(X))}_{\mathbb{E}(h(Y) g(X))} + \underbrace{\mathbb{E}(h(Y)) \mathbb{E}(g(Z))}_{\mathbb{E}(h(Y) g(Z))}$$

$$= \mathbb{E} \left[h(X) g(X) - h(X) g(Z) - h(Y) g(X) + h(Y) g(Z) \right]$$

$$= \mathbb{E} \left[\underbrace{(h(X) - h(Y))}_{\text{red}} \underbrace{(g(X) - g(Z))}_{\text{red}} \right] = A$$

$$A = \mathbb{E} \left[\underbrace{(h(X) - h(Y))}_{\text{red}} \underbrace{(g(X) - g(Z))}_{\text{red}} \right]$$

$$2A = \mathbb{E} \left[(h(X) - h(Y)) (g(X) - \cancel{g(Z)} - g(Y) + \cancel{g(Z)}) \right]$$

$$= \mathbb{E} \left[\underbrace{(h(X) - h(Y))}_{\text{red}} \underbrace{(g(X) - g(Y))}_{\text{red}} \right]$$

of \neq sign.

$$h(X) = X \quad g(X) = -X$$

$$\mathbb{E}(h(X)) \mathbb{E}(g(X)) =$$

$$= \underbrace{\mathbb{E}(X)}_{=0} \underbrace{\mathbb{E}(-X)}_{=0}$$

Control variates.

$I = \mathbb{E}(Y)$ is approximated by $\frac{1}{n} \sum_{i=1}^n Y_i$ where (Y_i) iid.
 $Y_i \stackrel{\text{d}}{=} Y$.

A control variate approximation \tilde{I} is:

$$\frac{1}{n} \sum_{i=1}^n (Y_i - (X_i - \mathbb{E}(X_1))) \quad \text{where } (X_i) \text{ are iid,}$$

usually: Y_i and X_i are correlated.

$$\underbrace{\frac{1}{n} \sum_{i=1}^n Y_i}_{\substack{\xrightarrow{\text{LLN}} \\ \mathbb{E}(Y_1)}} - \left(\underbrace{\frac{1}{n} \sum_{i=1}^n X_i}_{\substack{\xrightarrow{\text{LLN}} \\ \mathbb{E}(X_1)}} - \mathbb{E}(X_1) \right).$$

Compute $\mathbb{E}(X^{2.2})$ where $X \sim \text{d}^r(0,1)$.

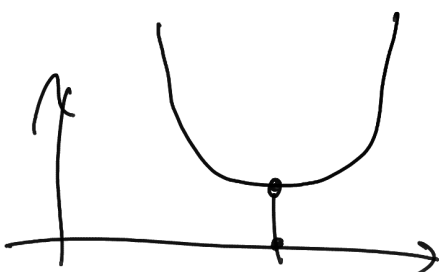
$$\frac{1}{n} \sum_{i=1}^n X_i^{2.2} - (X_i^2 - \underbrace{\mathbb{E}(X_1^2)}_1).$$

The best constant is: $\lambda = \frac{\text{Cov}(Y, X)}{\text{Var}(X)}$.

$$\frac{1}{n} \sum_{i=1}^n \left[Y_i - \lambda (X_i - \mathbb{E}(X_1)) \right]$$

indeed: minimize $\text{Var}(Y_1 - \lambda (X_1 - \mathbb{E}(X_1)))$

$$\begin{aligned} &= \text{Var}(Y_1 - \lambda X_1) \\ &= \text{Var}(Y_1) - 2 \text{Cov}(Y_1, X_1) \lambda + \lambda^2 \text{Var}(X_1) \\ &= \varphi(\lambda). \end{aligned}$$



$$\begin{aligned} \varphi'(\lambda) = 0 &\Leftrightarrow 2 \lambda \text{Var}(X_1) - 2 \text{Cov}(Y_1, X_1) = 0 \\ &\Leftrightarrow \lambda = \text{Cov}(Y_1, X_1) / \text{Var}(X_1). \end{aligned}$$

Stratified Sampling:

$$E(h(Y)) = \int h(y) f(y) dy = \int_{S_1 \cup \dots \cup S_p} h(y) f(y) dy$$

$$= \sum_{i=1}^p \int_{S_i} h(y) f(y) dy.$$

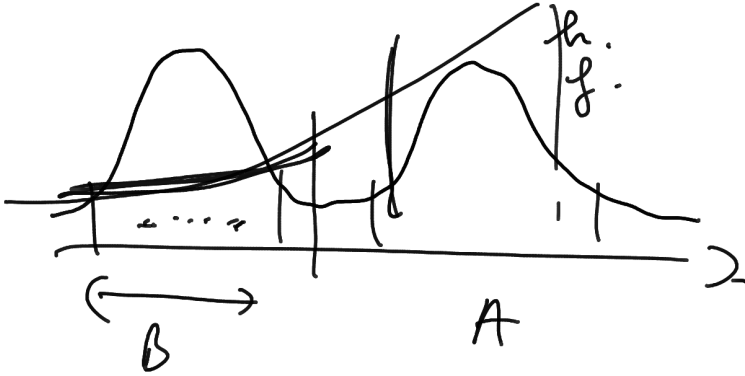
$$= \sum_{i=1}^p \underbrace{P(Y \in S_i)}_{\alpha_i} \cdot \int_{S_i} h(y) f(y) dy.$$

$$\underbrace{\left(\int_{S_i} f(y) dy \right)}_{\alpha_i}$$

$$\underbrace{P(Y \in S_i)}_{\alpha_i}$$

$$\int h(y) \left[\frac{f(y) \mathbb{1}_{S_i}(y)}{\int_{S_i} f(y) dy} \right] dy.$$

Law of $Y|Y \in S_i$

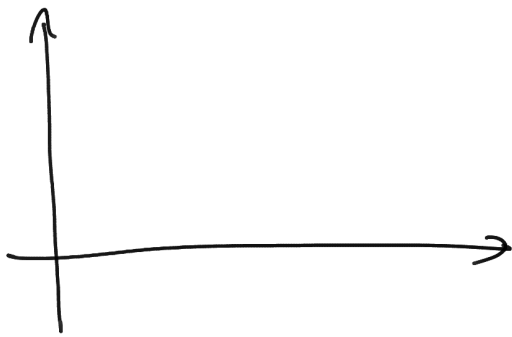


n

Stratification: 1) We sample for each region S_i , $(Y_{i,j} | 1 \leq j \leq n_i)$ such that $Y_{i,j} \stackrel{\mathcal{L}}{=} Y|Y \in S_i$

2) The approximation is:

$$\sum_{i=1}^p \underbrace{P(Y \in S_i)}_{\text{should be explicit}} \times \frac{1}{n_i} \sum_{j=1}^{n_i} h(Y_{i,j}).$$



if $X \sim \text{exp}(\lambda)$.

$$\begin{aligned} \text{the cdf, } F(x) &= \int_0^x \lambda e^{-\lambda t} dt \\ &= 1 - e^{-\lambda x} \\ &P(X \leq x) \end{aligned}$$

$$E[X^{2.2}] = \int_0^{+\infty} x^{2.2} \underbrace{\lambda e^{-\lambda x}}_{f(x)} dx.$$

$$S_1 = [0, \pi].$$

$$P(X \in S_1) = 1 - e^{-\lambda \pi} = P(X \leq \pi).$$

$$S_2 = [\pi, +\infty[.$$

$$P(X \in S_2) = P(X > \pi) = e^{-\lambda \pi}.$$

