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

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

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

$\mathbb{P}(X_n \in A) \to \mathbb{P}(X \in A)$

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

$\underbrace{\mathbb{P}(X_n \leq x)} \underset{n\to\infty}{\longrightarrow} \mathbb{P}(X \leq x)$

cdf (cumulative distribution function).

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

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

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

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

Chap 3.

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 $\tilde{Z} \sim \mathcal{N}(0, Var(X_1))$.

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

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

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

because (X_i) are independent

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

An important tool: Slutsky's Lemma:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$Z_n U_n + a = X_n \xrightarrow{\mathcal{L}} Z \times 0 + a = a.$$

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

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

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

$$\underbrace{\sqrt{N}\, \frac{(S_N - \mathbb{E}(f(X_1)))}{\sqrt{\sigma^2}}}_{\xrightarrow{\mathcal{L}} G} \qquad \underbrace{\sqrt{\frac{\sigma^2}{V_N}}}_{\xrightarrow{\mathbb{P}\text{-prob}} 1}$$

where $G \sim \mathcal{N}(0,1)$. (CLT). (by Lemma 3.6)

Slutsky's lemma.

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

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

According to: $(*)$. ,

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

$(a = 1.96)$.

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

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

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

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

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

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

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

Chapter 4:

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

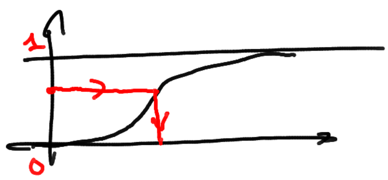
- F_Y is in $[0,1]$. , $\begin{cases} \lim\limits_{t\to-\infty} F_Y(t) = 0. \\ \lim\limits_{t\to+\infty} F_Y(t) = 1. \end{cases}$

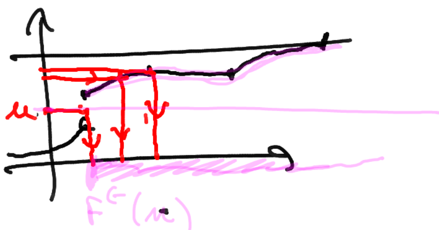
  $F_Y \nearrow$. (not always strictly increasing).

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

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

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





Mathematically speaking,

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

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

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

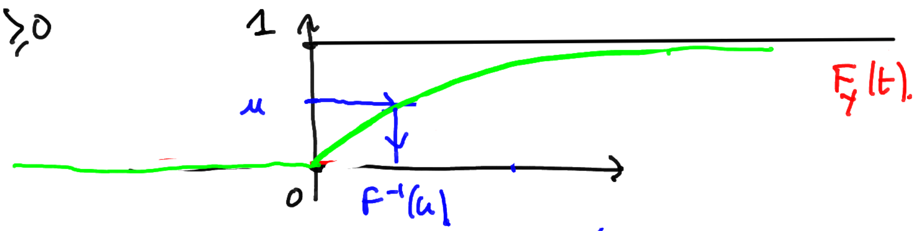
<u>Example:</u> $Y \sim \exp(\lambda)$ ie: Y has the density $f(y) = \lambda e^{-\lambda y} \mathbb{1}(y \geq 0)$.

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

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

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

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



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

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

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

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

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

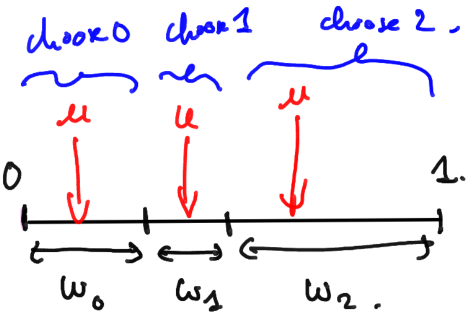
if $U \sim \mathrm{Unif}\,[0,1]$.

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

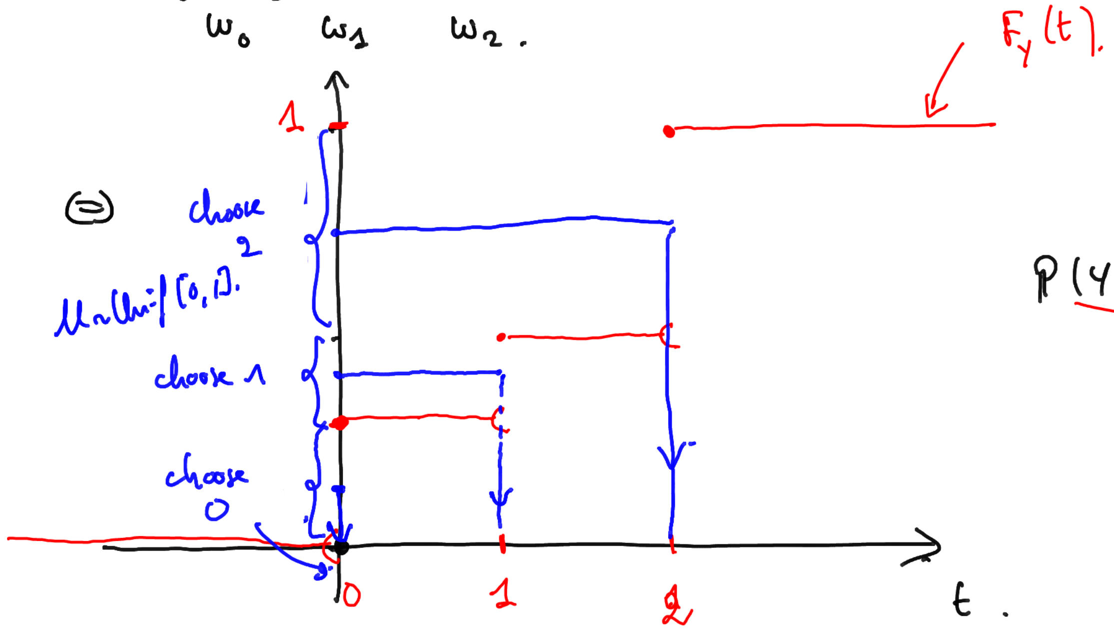
(where $Y \sim \exp(\lambda)$)

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

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



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



$\mathbb{P}(Y \leq t)$.

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 $U \sim \mathrm{Unif}[0,1]$.
- While $U > \frac{f(X)}{M g(X)}$, do: sample $X \sim g$, $U \sim \mathrm{Unif}[0,1]$.
- output $Y = X$.

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

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

then: $Y = X_T$.

We have that $Y \sim f$.

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

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

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

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

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

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

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

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

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

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

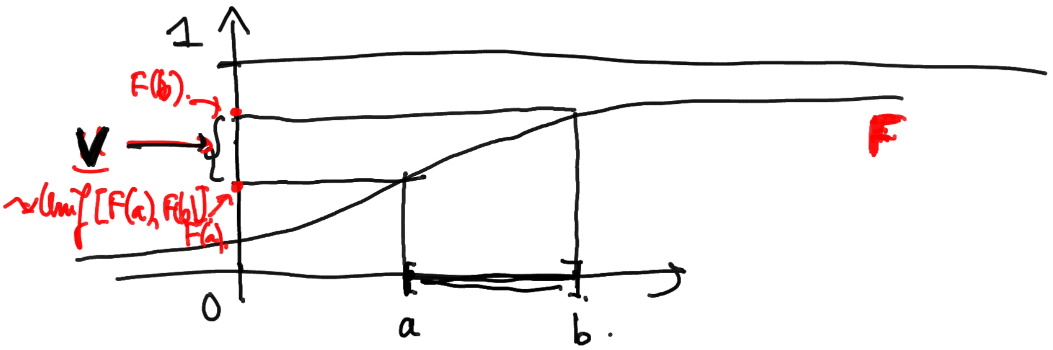
$Y \sim f$, $T \sim \text{Geom}\left(\frac{1}{M}\right)$.

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

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

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

<u>Importance sampling methods</u>:

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

$\underbrace{}_{\mathbb{E}_f[h(X)]}$ (ie $\mathbb{E}(h(X))$ when $X\sim f$).

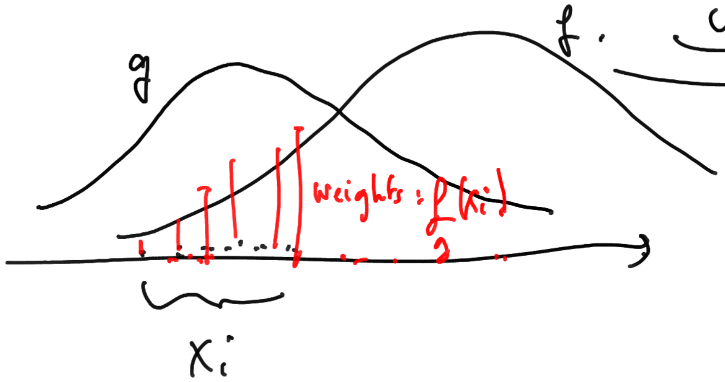
$$I=\int\underbrace{g(x)}\;\overbrace{\frac{f(x)}{g(x)}}^{\text{weight}}\;\underbrace{h(x)}\,dx.$$

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

where $X\sim g$.

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

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



<u>Variant</u>: if $f(x)$ is known up to a multiplicative constant:

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

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

$$\longrightarrow \frac{\int f(x) h(x)\,dx}{\left(\int f(x)\,dx\right) = 1} = \underbrace{\int f(x) h(x)\,dx}_{I}.$$

- $\displaystyle\frac{1}{N}\sum_{i=1}^{N} \frac{f(X_i)}{g(X_i)} h(X_i) = \hat{I}_N(f)$ 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.

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

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

(because (X_i) indep),

(they have the same law),

Y.

$$= \frac{1}{N} \underbrace{\mathrm{Var}\left(\frac{f(X_1)}{g(X_1)} h(X_1) \right)}_{\underbrace{\int \left(\frac{f(x)}{g(x)} h(x)\right)^2 g(x)\,dx}_{E(Y^2)} - \underbrace{\left(\int \frac{f(x)}{g(x)} h(x) \cdot g(x)\,dx \right)^2}_{E(Y)^2}}.$$

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

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

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

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

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

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

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

Computersession 2:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 $Law(X_n) \xrightarrow[n\to+\infty]{} \pi$.

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

- $X_{n+1} = \underbrace{f}(X_n, \varepsilon_n)$ 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. $\varepsilon_1 \sim \mathcal{N}(0,1)$.

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

//

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

↓

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 : $\int p(x_n, x_{n+1})\, dx_n$

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

$$= \pi_{n+1}(x_{n+1})$$

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

$\rightarrow$ kernel density.

Letting $n \to +\infty$.

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

Drawindep.

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

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

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

<u>Normalized importance sampling.</u>

- 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 $I = \mathbb{E}_\pi(h(X)) = \int h(x)\,\pi(x)\,dx$ by:

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

where $\begin{vmatrix} X_i \text{ iid.} \\ X_1 \sim g. \end{vmatrix}$

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

(2) if $\tilde{\pi}(x) = c\,\pi(x)$. Then:

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

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

Indeed:

$$\hat{I}_n = \frac{\frac{1}{n}\sum_{i=1}^{n} \frac{\pi(X_i)}{g(X_i)} h(X_i)}{\frac{1}{n}\sum_{j=1}^{n} \frac{\pi(X_j)}{g(X_j)}} \xrightarrow[n\to+\infty]{} \frac{I}{1} = I.$$

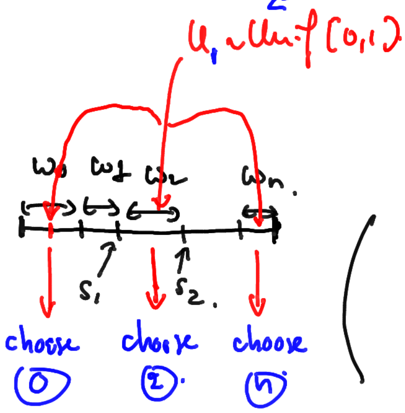
Numerator $\xrightarrow{LLN}$ $\mathbb{E}_g\left[\frac{\pi(X)h(X)}{g(X)}\right] = \int \cancel{g(x)} \frac{\pi(x)h(x)}{\cancel{g(x)}} dx = \int \pi(x)h(x)dx = I.$

Denominator $\xrightarrow{LLN}$ $\mathbb{E}_g\left[\frac{\pi(X)}{g(X)}\right] = \int \cancel{g(x)} \frac{\pi(x)}{\cancel{g(x)}} dx = \int \pi(x)\,dx = 1$

Computer session 4: 1) Implement the normalized importance sampling.

2) Resample among the $(X_1,..,X_n)$ according to the weights $(w_1,..,w_n)$

(this is possible because: $w_1 + \dots + w_n = 1$)



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

Where: $\forall i \in \{1,..,m\}, \forall j \in \{1,..,n\}$,

$$\mathbb{P}(\tilde{X}_i = X_j \mid X_1,..,X_n) = w_j.$$

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

1) $X_1, .., X_n$ iid according to g.

$w_1, \dots, w_n$.

2) $\tilde{X}_1 \to X_1$ wp w_1, $\to X_2$ — w_2, $\to X_n$ wp w_n. $\Longleftrightarrow$ Draw $U_1 \sim Unif[0,1]$, choose index k, where $S_{k-1} = w_0 + \dots + w_{k-1} \le U \le w_0 + \dots + w_k = S_k$. set $\tilde{X}_1 = X_k$.

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

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

$$\sum_{i=1}^n \frac{1}{n} h(\hat{X}_i)$$

$\mathbb{E}(h(\tilde{X}_1))$

$$K = \begin{cases} 1 & \text{wp } w_1 \\ \vdots & \\ n & \text{—} \ w_n \end{cases}$$

$\tilde{X}_1 = X_K$.

$$\| $$

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

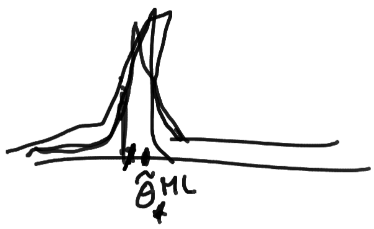
$$= \mathbb{E}(h(X_1) \mathbb{1}_{K=1}) + \dots + \mathbb{E}(h(X_n) \mathbb{1}_{K=n}).$$

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

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

- Usually: (in Bayesian statistics).

$$\begin{cases} \theta \sim \pi_0 \\ (Y_i) \text{ iid} \quad Y_1 \sim f_\theta \end{cases}$$



You observe: $Y_1, \dots, Y_n$.

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

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

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

Example: (1) $\theta \sim \mathcal{U}nif[0, \lambda]$ ← prior $\quad \lambda = 1$?

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

from Bayesian,

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

Law $\theta \mid Y_{1:n}$ 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:

$Y_t \sim q(X_{t-1}, \cdot)$ (density which depends on X_{t-1})

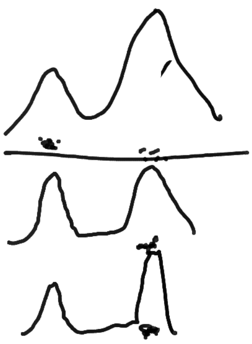
Draw $U_t \sim \mathrm{Unif}\,[0,1]$, s.t U_t and Y_t indep.

if $U_t \leq \alpha(X_{t-1}, Y_t)$, $X_t = Y_t$.

otherwise $X_t = X_{t-1}$

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

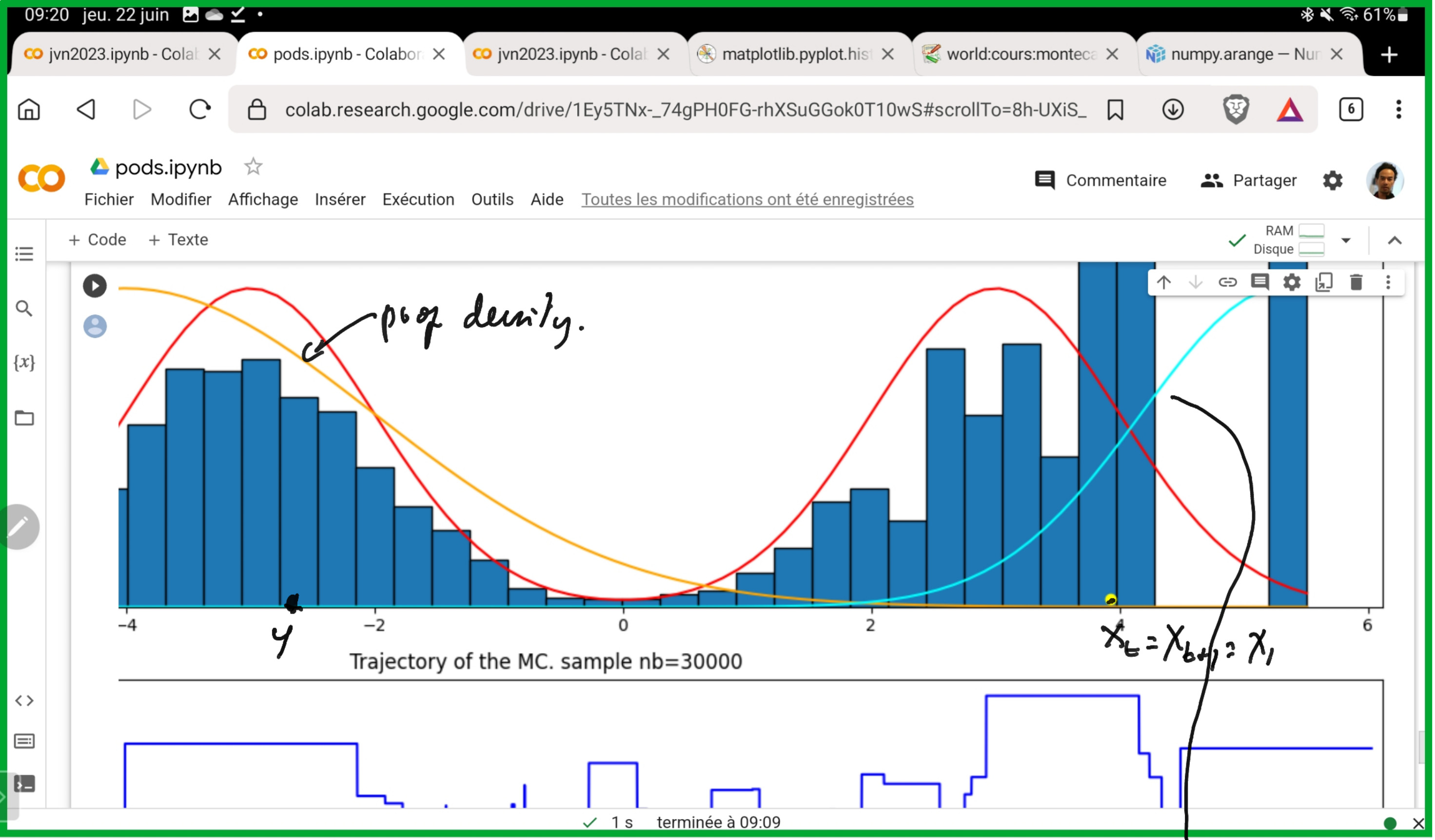
Independance sampler : we use : $q(x,y) = g(y)$.



$\boxed{\text{MCMC}}$: error : $\rho^n \downarrow 0$ $\rho \in]0,1[$.

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

SMC : (Sequential Monte Carlo) for target tracking.



π/q

if Y proposed and $X_t = X$

you accept Y wp $\left(\frac{\pi/q(Y)}{\pi/q(X)} \wedge 1 \right)$.

Chap 6 Antithetic variates.

$I = \mathbb{E}[h(U)]$ is approximated by $\underbrace{S_n = \frac{1}{n} \sum_{i=1}^{n} h(U_i)}_{\mathrm{Var}(S_n) = \frac{\mathrm{Var}(h(U_1))}{n}}$.

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

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

I may be approximated by: $\underbrace{S_{2n} = \frac{1}{2n} \sum_{i=1}^{2n} h(U_i)}$ — $\mathrm{Var}(S_{2n}) = \frac{\mathrm{Var}(h(U_1))}{2n}$

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

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

$\sim U_i / [0,1]$

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

Both are unbiased estimators of I.

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

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

$$= \left(\frac{1}{2n}\right)^2 \underbrace{\mathrm{Var}\left(\sum_{i=1}^{n} (h(U_i) + h(1-U_i))\right)}_{\sum_{i=1}^{n} \underbrace{\mathrm{Var}(h(U_i) + h(1-U_i))}_{\mathrm{Var}(h(U_1)+h(1-U_1))}}.$$

(because (U_i) are indep.).

(because (U_i) have the same Law).

$$= \frac{1}{4n^2} \times n \times \mathrm{Var}(h(U_1) + h(1-U_1))$$

$$= \frac{1}{4n}\left[\underbrace{\mathrm{Var}(h(U_1) + h(1-U_1))}_{\mathrm{Var}(h(U_1)) + \mathrm{Var}(h(\underbrace{1-U_1}_{\cong U_1})) + 2\,\mathrm{cov}(h(U_1),\, h(1-U_1))}\right].$$

$$= \frac{1}{4n}\left(2\,\mathrm{Var}(h(U_1))\right) + \frac{2}{4n}\mathrm{cov}(h(U_1), h(1-U_1))$$

$$= \underbrace{\frac{1}{2n}\mathrm{Var}(h(U_1))}_{\mathrm{Var}(S_{2n})} + \frac{1}{2n}\mathrm{cov}(h(U_1), h(1-U_1)),$$

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

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

Lemma: if $\left|\begin{array}{l} X \text{ random variable.} \\ h \text{ and } g \text{ are monotone functions then:} \end{array}\right.$

$$\mathrm{Cov}(h(X), g(X)) \leq 0.$$

Let

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

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

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

$$\begin{cases} X \overset{\mathcal{L}}{=} Y \\ X \overset{\mathcal{L}}{=} Z \\ (X, Y, Z) \text{ indep.} \end{cases}$$

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

- $A = \mathbb{E}\big[\underbrace{(h(X) - \mathbb{E}(h(Y)))(g(X) - \mathbb{E}(g(Z)))}\big]$.

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

$\shortparallel$

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

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

$$= \mathbb{E}\big[\underbrace{(h(X) - h(Y))}_{''}(g(X) - g(Z))\big]. = A.$$

$$A = \mathbb{E}\big[\underbrace{(-h(Y) + h(X))}(-g(Y) + g(Z))\big]$$

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

$$= \mathbb{E}\big[(\underbrace{h(X) - h(Y)})(\underbrace{g(X) - g(Y)})\big].$$

of $\neq$ sign.

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

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

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

Control variates.

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

A control variate approximation I is:

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

usually: Y_i and X_i are correlated.

$$\| $$

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n} Y_i}_{\substack{\xrightarrow{n\to\infty} \mathbb{E}(Y_1). \\ \text{LLN}}} - \Big(\underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}(X_1)}_{\substack{\xrightarrow{n\to\infty} 0. \\ \text{LLN}}}\Big).$$

X_i

Compute $\mathbb{E}(X^{2.2})$ where $X \sim \mathcal{N}(0,1)$.

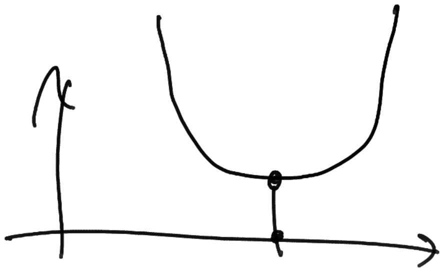
$$\frac{1}{n}\sum_{i=1}^{n} X_i^{2.2} - \Big(X_i^2 - \underbrace{\mathbb{E}(X_1^2)}_{1}\Big).$$

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

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

<u>indeed:</u> $\underset{\lambda}{\text{Minimize}}\ \mathrm{Var}\left(Y_1 - \lambda(X_1 - \mathbb{E}(X_1))\right)$

$$= \mathrm{Var}(Y_1 - \lambda X_1)$$
$$= \mathrm{Var}(Y_1) - 2\,\mathrm{Cov}(Y_1, X_1)\,\lambda + \lambda^2\,\mathrm{Var}(X_1).$$
$$= \varphi(\lambda).$$



$$\varphi'(\lambda) = 0 \iff 2\lambda\,\mathrm{Var}(X_1) - 2\,\mathrm{cov}(Y_1, X_1) = 0$$
$$\iff \lambda = \mathrm{cov}(Y_1, X_1) / \mathrm{Var}(X_1).$$

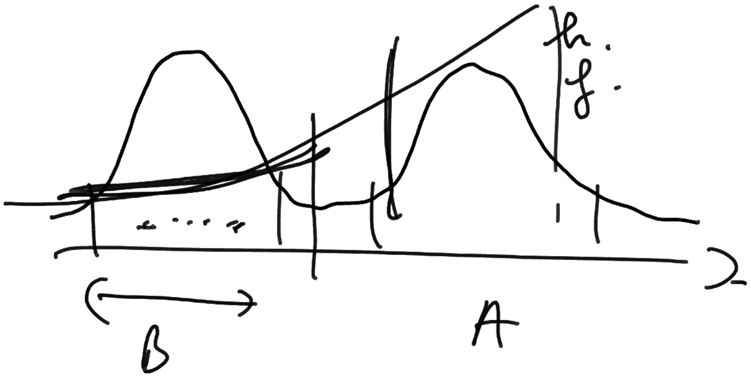
Stratified Sampling:

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

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

$$= \sum_{i=1}^{p} \underbrace{\mathbb{P}(Y \in S_i)}_{\substack{\alpha_i \\ \shortparallel \\ \left(\int_{S_i} f(y)\,dy\right)}} \cdot \underbrace{\frac{\int_{S_i} h(y) f(y)\,dy.}{\mathbb{P}(Y \in S_i).}}_{\int_{S_i} h(y) \left[\frac{f(y) \mathbb{1}_{S_i}(y)}{\int_{S_i} f(y)\,dy}\right] dy.}$$

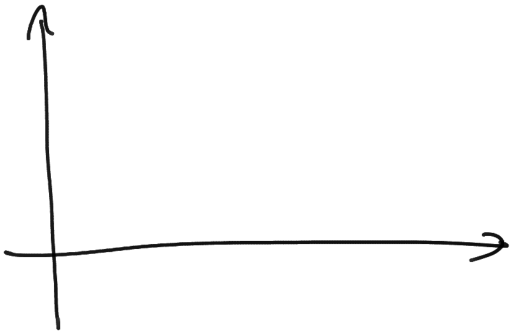
$\left[\frac{f(y) \mathbb{1}_{S_i}(y)}{\int_{S_i} f(y)\,dy}\right]$: Law of $Y|_{Y \in S_i}$



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

2) The approximation is:

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



if $X \sim exp(\lambda)$.

the cdf, $F(x) = \int_0^x \lambda e^{-\lambda t}\, dt$

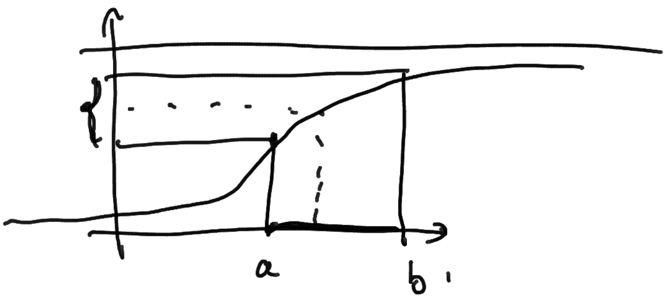
$= 1 - e^{-\lambda x}$.

$\mathbb{P}(X \le x)$

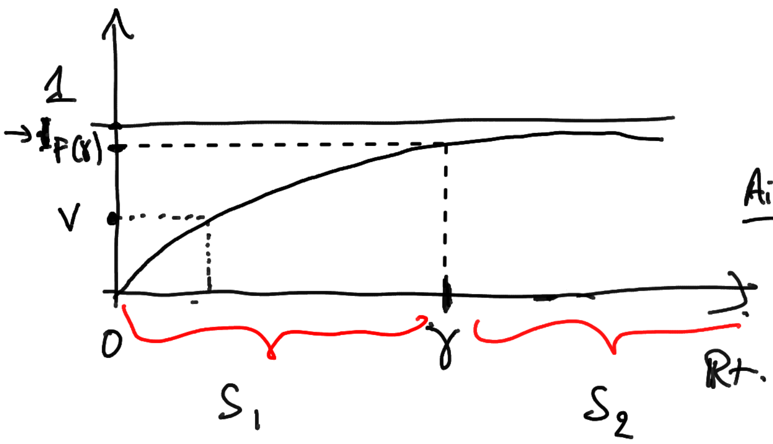
$$\mathbb{E}\left[X^{2.2}\right] = \int_0^{+\infty} x^{2.2} \underbrace{\lambda e^{-\lambda x}}_{f(x)} dx.$$

$S_1 = [0, x]$. $\qquad \mathbb{P}(X \in S_1) = 1 - e^{-\lambda x} = \mathbb{P}(X \le x)$.

$S_2 = [x, +\infty[$. $\qquad \mathbb{P}(X \in S_2) = \mathbb{P}(X > x) = e^{-\lambda x}$.



Computer Session 6



$F(x) = 1 - e^{-\lambda x}$ $\qquad f(x) = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x)$.

$F^{-1}(u) = -\frac{\log(1-u)}{\lambda}$

<u>Aim</u>: Approximate $I = \mathbb{E}(X^{\alpha})$ where $X \sim exp(\lambda)$.

using stratification with regions: $S_1 = [0, \gamma]$.

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

I is approximated by: $\begin{cases} \underbrace{\mathbb{P}(X\in S_1)} \frac{1}{n_1}\sum_{i=1}^{n_1} h(X_{1,i}) + \underbrace{\mathbb{P}(X\in S_2)} \frac{1}{n_2}\sum_{i=1}^{n_2} h(X_{2,i}). \\ \frac{n_1}{n} = \mathbb{P}(X\in S_1) \quad , \quad \frac{n_2}{n} = \mathbb{P}(X\in S_2). \end{cases}$

and: $\begin{cases} (X_{1,i}) \text{ are iid } X_{1,i} \overset{\mathcal{L}}{=} Y \mid Y\in S_1. \\ (X_{2,i}) \text{ are iid } X_{2,i} \overset{\mathcal{L}}{=} Y \mid Y\in S_2. \end{cases}$

- $\mathbb{P}(X\in S_1) = \mathbb{P}(X\in [0,\gamma]) = \mathbb{P}(X\leq \gamma) = F(\gamma) = 1-e^{-\lambda\gamma}$
- $\mathbb{P}(X\in S_2) = \mathbb{P}(X\geq \gamma) = e^{-\lambda\gamma}.$

- to draw $(X_{1,i})_{1\leq i\leq n_1}$, we draw $U_1,\dots,U_{n_1}$ iid with $\mathrm{Unif}[0,1]$.

$.\ V_i = U_i\, F(\gamma) = U_i\,(1-e^{-\lambda\gamma}) \sim \mathrm{Unif}\,[0, F(\gamma)].$

For $i=1,\dots,n_1$, $\left[\ .\ Y_{1,i} = -\frac{\log(1-V_i)}{\lambda}.\right.$

- Similarly, we draw: $\tilde{U}_1,\dots,\tilde{U}_{n_2}$ iid with $\mathrm{Unif}[0,1]$.

$$\tilde{V}_i = F(\gamma) + \tilde{U}_i\,(1-F(\gamma)) \qquad \sim \mathrm{Unif}\,[F(\gamma), 1].$$
$$\tilde{V}_i = 1-e^{-\lambda\gamma} + \tilde{U}_i\,(e^{-\lambda\gamma})$$

For $i=1,\dots,n_2$, set: $Y_{2,i} = -\frac{\log(1-\tilde{V}_i)}{\lambda}$