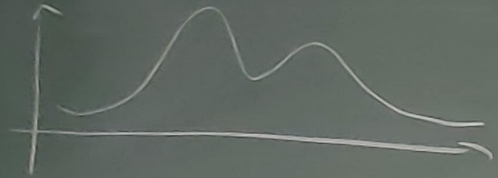
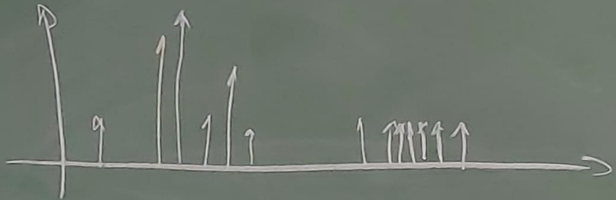
II Approximation (ponctuelle) de MC d'une pdf

$$* \; p(x) \approx \hat{p}(x) = \sum_{i=1}^{N} w^{(i)} \delta\left(x - x^{(i)}\right)$$



$\approx$



$\Downarrow$

$$\mathbb{E}_{p(x)}\left[\phi(x)\right] \approx \mathbb{E}_{\hat{p}(x)}\left[\phi(x)\right] = \int \phi(x)\, \hat{p}(x)\, dx = \sum_{i=1}^{N} w^{(i)} \, \phi\left(x^{(i)}\right)$$

ainsi $\mathbb{E}_{p(x)}[x] \approx \sum_{i=1}^{N} w^{(i)} x^{(i)} =: \bar{x}$

$\mathrm{Cov}_{p(x)}[x] = \mathbb{E}_{p(x)}\left[(x - \mathbb{E}_{p(x)}[x])(x - \mathbb{E}_{p(x)}[x])^T\right] \approx \sum_{i=1}^{N} w^{(i)} (x^{(i)} - \bar{x})(x^{(i)} - \bar{x})^T =: \bar{P}$

en particulier, si

$p(x_k | z_{0:k})$ est approximé par $\boxed{\sum_{i=1}^{N} w_k^{(i)} \delta(x_k - x_k^{(i)})}$

alors

$\boxed{\hat{x}_{k|k} = \mathbb{E}_{p(x_k|z_{0:k})}[x_k]} \approx \boxed{\sum_{i=1}^{N} w_k^{(i)} x_k^{(i)}} =: \overline{x_{k|k}}$

$\boxed{P_{k|k} = \mathrm{Cov}_{p(x_k|z_{0:k})}[x_k]} \approx \boxed{\sum_{i=1}^{N} w_k^{(i)} (x_k^{(i)} - \bar{x}_{k|k})(x_k^{(i)} - \bar{x}_{k|k})^T =: \overline{P_{k|k}}}$

* Désormais, l'ensemble $\underbrace{\{x^{(i)}\}_{i=1..N}}_{\substack{\text{particules} \\ (\text{"particles"})}}$ est <u>aléatoire</u>

$\Longrightarrow$ $\boxed{\text{méthodes de Monte Carlo}}$

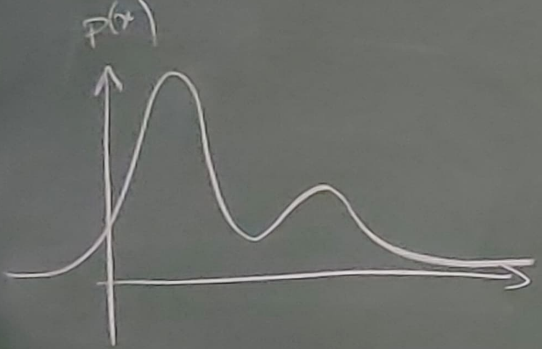
- $x^{(1)} \dots x^{(N)} \underset{\sim}{iid}$ $q(x)$ loi d'importance

et les $w^{(1)} \dots w^{(N)}$ seront déterminés de telle sorte que $\hat{p}(x)$ constitue un <u>estimé</u> "cohérent" de $p(x)$

- En fait, $\mathbb{E}_{\hat{p}(x)}[\phi(x)]$ constitue un <u>estimé</u> de $\mathbb{E}_{p(x)}[\phi(x)]$, sur lequel on raisonne en termes statistiques.

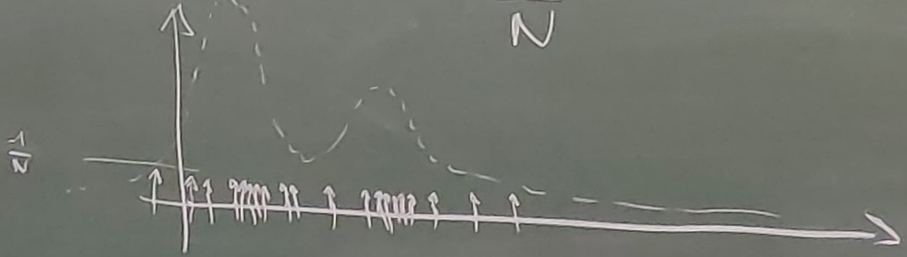
* Échantillonnage idéal

$$\cdot\ p(x) \approx \hat{p}_N(x) = \sum_{i=1}^{N} w^{(i)} \delta(x - x^{(i)})$$



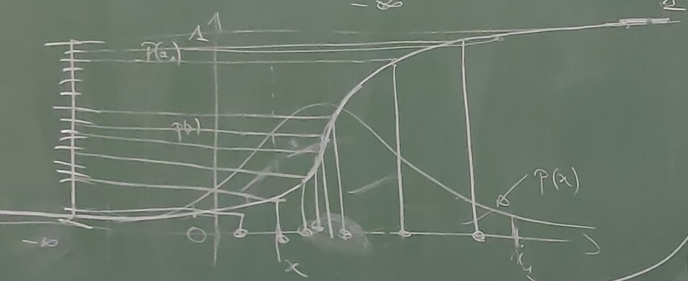
où $x^{(1)} \dots x^{(N)} \overset{iid}{\sim} p(x)$

$w^{(1)} \dots w^{(N)} = \frac{1}{N}$



• Nota : (on doit disposer d'un générateur aléatoire de nombres aléatoires selon $p(x)$!

$$\text{cdf} = P(x) = \int_{-\infty}^{x} p(\xi)\, d\xi = \mathbb{P}(X \leq x)$$



• Dans ce cas, on montre que l'estimateur

$$\hat{I}_N(\phi) := \mathbb{E}_{\hat{p}_N(x)}[\phi(x)] = \sum_{i=1}^{N} \frac{1}{N} \phi(x^{(i)}) \quad \text{de} \quad I(\phi) := \mathbb{E}_{p(x)}[\phi(x)]$$

est
- non biaisé $\forall N$
- "converge" vers $I(\phi)$ lorsque $N \to \infty$
- admet une variance $\propto \frac{1}{N}$
- pour $N \to \infty$, est tq l'erreur est Gaussienne et

"beat the curse of dimensionality"?

* Échantillonnage d'importance

- On suppose qu'on ne peut / veut pas échantillonner selon $p(x)$

- On introduit une pdf $q(x)$
  - qu'on sait échantillonner
  - ______ évaluer
  - dont le support $\{x : q(x) > 0\}$ couvre le support $\{x : p(x) > 0\}$ de $p(x)$

$$\forall x, \ (p(x) \neq 0) \Rightarrow (q(x) \neq 0)$$

(c.à.d. que $p(x) > 0$ et $q(x) = 0$ est impossible)

$$\cdot\ p(x) \approx \mathring{p}_N(x) = \sum_{i=1}^{N} w^{(i)}\,\delta(x - x^{(i)})$$

où $x^{(1)} \ldots x^{(N)} \overset{iid}{\sim} q(x)$

$$w^{(i)} \propto \boxed{w^{*(i)} = \frac{p(x^{(i)})}{q(x^{(i)})}}$$

① poids non normalisé associé à $x^{(i)}$

② puis, lorsque tous les $w^{*(i)}$ ont été calculés, on normalise

$$\forall i,\ w^{(i)} = w^{*(i)} \Big/ \left(\sum_{j=1}^{N} w^{*(j)}\right)$$

• Dans ce cas, on montre que l'estimateur

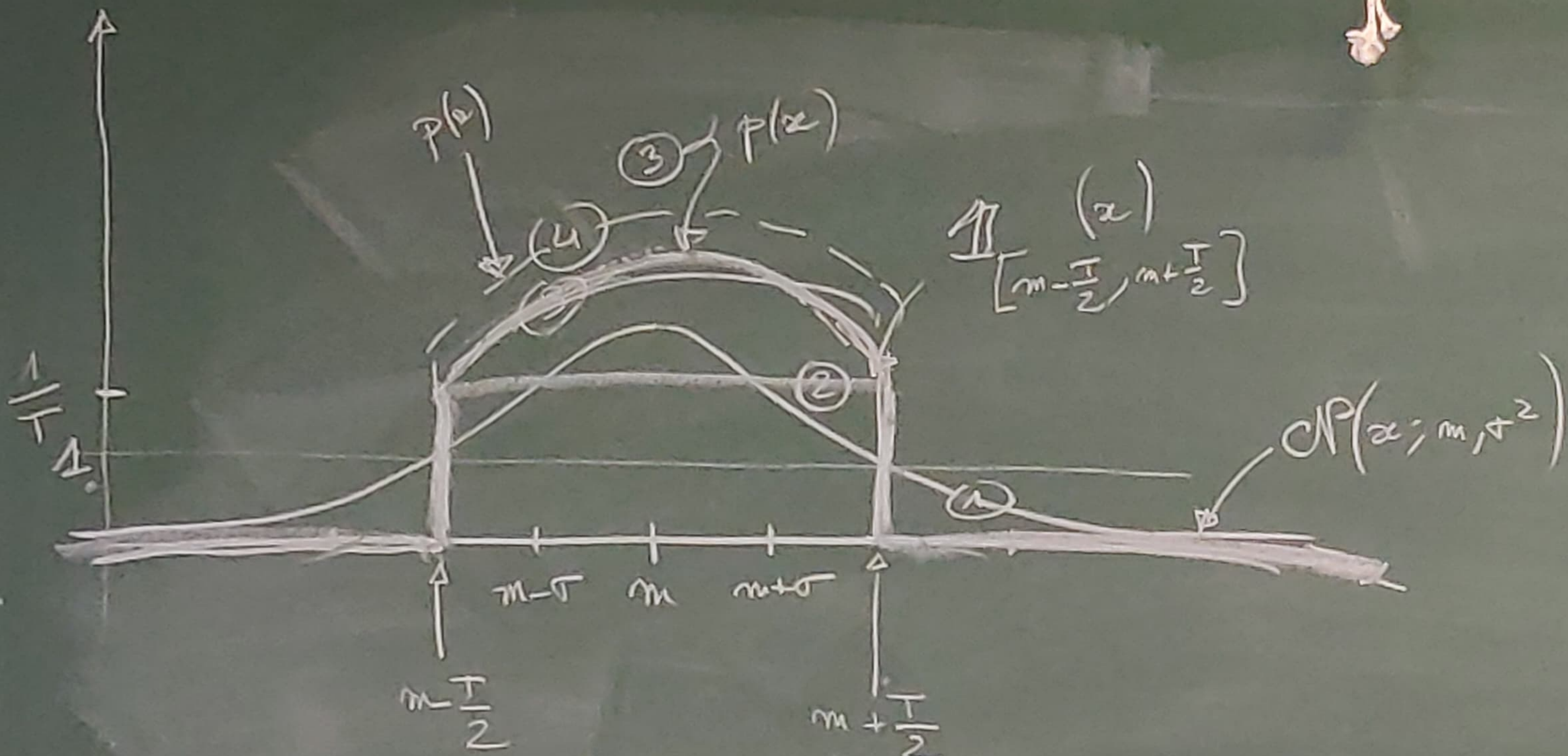
$$\mathring{I}_N(\phi) = \mathbb{E}_{\mathring{P}_N(x)}\left[\phi(x)\right] \left(= \sum_{i=1}^{N} w^{(i)} \phi(x^{(i)})\right)$$ de $I(\phi) := \mathbb{E}_{p(x)}\left[\phi(x)\right]$

est
- ASYMPT° non biaisé (pour $N \to \infty$)
- "converge" vers $I(\phi)$ lorsque $N \to \infty$, dès lors que le support de $q(x)$ couvre celui de $\phi(x)$
- admet une variance finie décroissante avec N

- <u>Nota</u> : on est capables d'estimer $p(x)$ au moyen de $\hat{p}_N(x)$ même si $p(x)$ n'est connu qu'à sa constante de normalisation près.

- <u>Nota 2</u> : le calcul des poids peut être coûteux, instable numériquement, etc.

- **Exercice :** $p(x) \propto \mathcal{N}(x; m, \sigma^2) \cdot \mathbb{1}_{[m-\frac{T}{2}, m+\frac{T}{2}]}(x)$



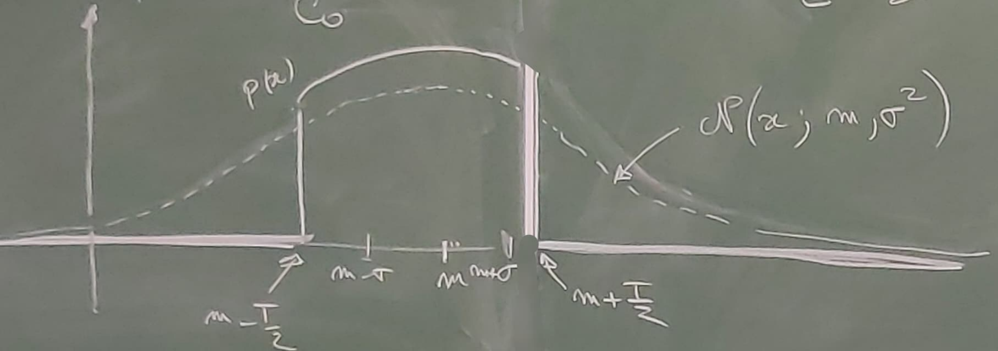
$$p(x) = \frac{\mathcal{N}(x; m, \sigma^2) \cdot \mathbb{1}_{[m-\frac{T}{2},\, m+\frac{T}{2}]}(x)}{\left[\int_{\mathbb{R}} \mathcal{N}(\xi; m, \sigma^2)\, \mathbb{1}_{[\cdot]}(\xi)\, d\xi \right]}$$

notée C_0

Définissons trois approx MC de

$$p(x) = \frac{1}{C_0} \mathcal{N}(x ; m, \sigma^2) \, \mathbb{1}_{[m-\frac{T}{2} ; m+\frac{T}{2}]}(x)$$

(Nota : on ne sait pas écrire C_0 autrement que par une éval numérique)



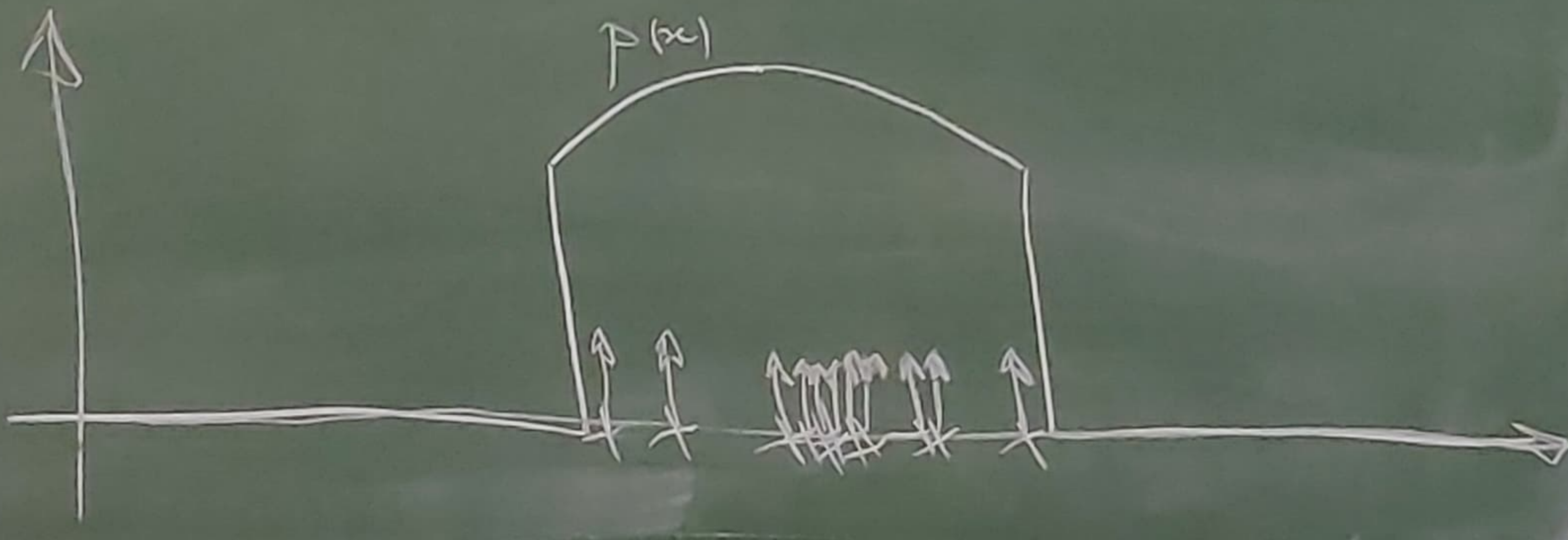
A. Supposons qu'on dispose d'un moyen d'échantillonner selon $p(x)$

- $x^{(1)} \dots x^{(N)} \overset{iid}{\sim} p(x)$

$$w^{(1)} \dots w^{(N)} \propto \frac{p(x^{(i)})}{p(x^{(i)})} = 1$$

c.à d.

$$w^{(1)} = \dots = w^{(N)} = \frac{1}{N}.$$



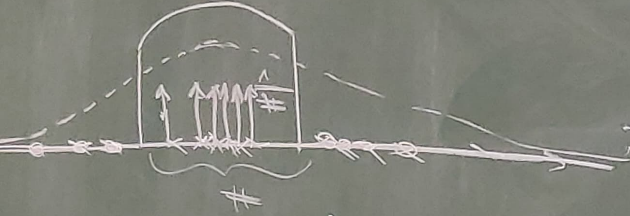
B. Éch d'importance

$$q(x) = \mathcal{N}(x ; m, \sigma^2)$$

• $x^{(1)} \dots x^{(N)} \overset{iid}{\sim} \mathcal{N}(x, m, \sigma^2)$

on évalue en $x^{(i)}$

puis $\forall i$, $w^{(i)} \propto w^{*(i)} := \dfrac{p(x^{(i)})}{q(x^{(i)})} = \dfrac{\frac{1}{C_0} \cancel{\mathcal{N}(x^{(i)} ; m, \sigma^2)} \, \mathbb{1}_{[m-\frac{T}{2}, m+\frac{T}{2}]}(x^{(i)})}{\cancel{\mathcal{N}(x^{(i)}, m, \sigma^2)}}$



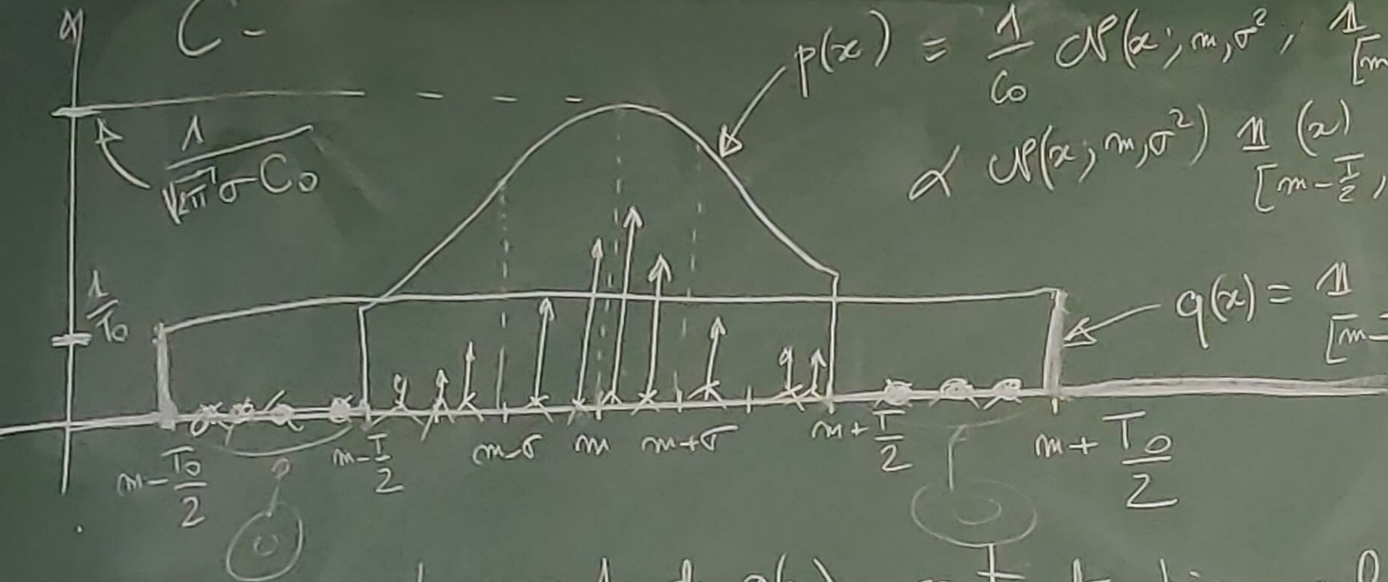
$$= \begin{cases} \dfrac{1}{C_0} \dfrac{1}{T} & \text{si } x^{(i)} \in [m-\frac{T}{2}, m+\frac{T}{2}] \\ 0 & \text{sinon} \end{cases}$$

de sorte que $w^{(i)} = 0$ si $x^{(i)} \notin [m-\frac{T}{2} ; m+\frac{T}{2}]$

$= \dfrac{1}{\#\text{particules} \in [m-T/2, m+T/2]}$ sinon

$\equiv$ "acceptation - reject"

C -



$$p(x) = \frac{1}{C_0} \mathcal{N}(x; m, \sigma^2) \, \mathbb{1}_{[m-\frac{T}{2}, m+\frac{T}{2}]}(x)$$

$$\propto \mathcal{N}(x; m, \sigma^2) \, \mathbb{1}_{[m-\frac{T}{2}, m+\frac{T}{2}]}(x)$$

$$q(x) = \mathbb{1}_{[m-\frac{T_0}{2}, m+\frac{T_0}{2}]}(x)$$

- Le support de $q(x)$ contient bien celui de $p(x)$ car $T_0 \geq T$
- $x^{(1)} \ldots x^{(N)}$ iid $\sim q(x) = \mathbb{1}_{[m-\frac{T_0}{2}; m+\frac{T_0}{2}]}(x)$

$$\cdot\ \forall i,\quad w^{(i)} \propto w^{*(i)} = \frac{p(x^{(i)}) \leftarrow \frac{1}{C_0}\mathcal{N}(x^{(i)}; m, \sigma^2)\,\mathbb{1}_{[m-\frac{T}{2}, m+\frac{T}{2}]}(x^{(i)})}{q(x^{(i)}) \leftarrow \frac{1}{T_0} \propto 1}$$

(Nota : on ne sait pas écrire C_0 autrement que par une éval numérique)

$$\Leftrightarrow w^{(i)} \propto \begin{cases} \mathcal{N}(x^{(i)}, m, \sigma^2) & \text{si } x^{(i)} \in [m-\frac{T}{2}, m+\frac{T}{2}] \\ 0 & \text{sinon} \end{cases}$$

III Application au filtrage Bayésien

III.1 Les équations récursives exactes

* On montre que

$$\underbrace{p(x_{0:k-1} | z_{1:k-1})}_{\text{loi JOINTE de filtrage à } k-1} \text{ et } \underbrace{p(x_{0:k} | z_{1:k-1})}_{\text{loi JOINTE de prédiction à l'instant } k}$$

sont liées par

$$p(x_{0:k} | z_{1:k-1}) = p(x_k | x_{k-1}) \, p(x_{0:k-1} | z_{1:k-1})$$

* D'autre part, $p(x_{0:k} \mid z_{1:k-1})$, z_k et $p(x_{0:k} \mid z_{1:k})$
loi de filtrage à k
sont unies par

$$p(x_{0:k} \mid z_{1:k}) = \frac{p(z_k \mid x_k)\, p(x_{0:k} \mid z_{1:k-1})}{p(z_k \mid z_{1:k-1})}$$

$\equiv \int$ numérateur $dx_{0:k}$

$\equiv$ constante de normalisation.