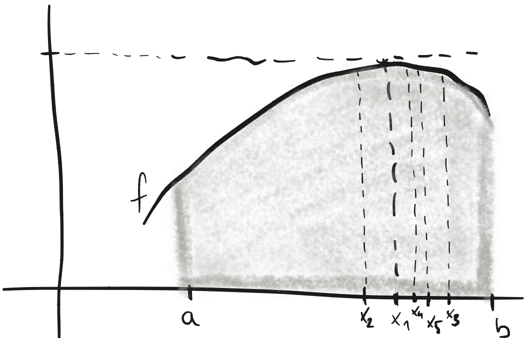
Importance sampling MC

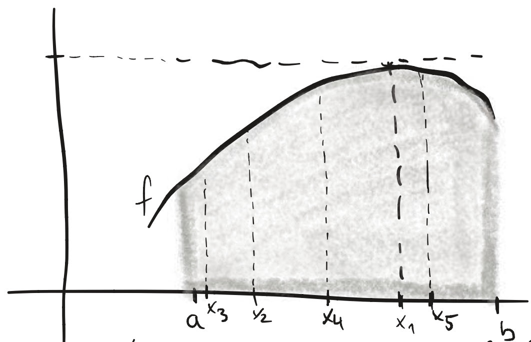
- co když je $\Omega \subseteq \mathbb{R}^d$ neomezená? Pak neexistuje rovnoměrné rozdělení
- co když umím samplovat body z Ω odpovídající jen skrze nějaké dané rozdělení $\mathcal{W}$ s distrib. funkcí $W(\vec{x})$ a hustotou $w(\vec{x})$?

$\Rightarrow$ potřebujeme upravit $I_{MC}(N)$ pro $\{X_i\}_{i=1}^N$ odpovídající danému rozdělení $\mathcal{W}$ s distrib. funkcí $W(\vec{x})$ a hustotou $w(\vec{x})$.

Nebo-li, potřebujeme zvládnout odhadnout $\int_a^b f(x)\,dx$ v „obou případech níže"



$X_i \sim$ nerovnoměrné rozdělení



$X_i \sim$ rovnoměrné rozdělení

Řešení: náš odhad $\int_a^b f(x)\,dx$ bude vážený průměr – každá funkční hodnota $f_i \equiv f(x_i)$ bude škálovaná „pravděpodobností výskytu x_i"

Zapsáno formálně:

$w(\vec{x}) \neq 0$ skoro všude, protože je to hustota

definice $W(\vec{x})$

$$\int_\Omega f(\vec{x})\,d\vec{x} = \int_\Omega \frac{f(\vec{x})}{w(\vec{x})} \cdot w(\vec{x})\,d\vec{x} = \int_\Omega \frac{f(\vec{x})}{w(\vec{x})}\,dW(\vec{x})$$

a tedy $$\int_\Omega f(\vec{x})\,d\vec{x} = \mathbb{E}_W\left(\frac{f}{w}\right)$$

„X_i má velkou pravděpodobnost v $\mathcal{W}$ $\Leftrightarrow$ $w(X_i)$ je vysoká" a tedy hodnoty $f_i = f(X_i)$ jsou škálovány nepřímo úměrně jejich šanci na samplování.

střední hodnota funkce f/w vzhledem k míře dW → při výpočtu této střední hodnoty „už nejsou všechny jevy stejně pravděpodobné" jako u klasické definice → naopak zohledňují jejich pravděpodobnost, právě skrze pravděpodob. míru dW

Pokud máme $\{X_i\}_1^N$ náhodné veličiny, i.i.d, odpovídající rozdělení W (hust. $w(\vec{x})$, distr. $W(\vec{x})$)

$$I_{MCIS}(N) := \frac{1}{N} \sum_{i=1}^{N} \frac{f(X_i)}{w(X_i)}$$

Hodnoty $f_i = f(X_i)$ jsou škálovány „pravděpodob.", že se při samplování z W objeví $\Rightarrow$ $\Rightarrow$ „vzácné" hodnoty ovlivní odhad více než ty „běžné"

pak platí:

- $I_{MCIS}(N)$ je náhodná veličina
- $\mathbb{E}_W(I_{MCIS}(N)) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}_W\left(\frac{f(X_i)}{w(X_i)}\right) = \mathbb{E}_W\left(\frac{f}{w}\right) = \int_\Omega f(\vec{x})\, d\vec{x}$
- $\mathrm{var}_W(I_{MCIS}(N)) = \frac{1}{N^2} \sum_{i=1}^{N} \mathrm{var}_W\left(\frac{f(X_i)}{w(X_i)}\right) = \frac{1}{N} \mathrm{var}_W\left(\frac{f}{w}\right)$

a tedy za použití Chebyshevova lemma dostáváme

$$\mathbb{P}_W\left(\left| I_{MCIS}(N) - \int_\Omega f(\vec{x})\, d\vec{x} \right| < \sqrt{\frac{1}{N}} \cdot \frac{\mathrm{var}_W(f(X)/w(X))}{\sqrt{\varepsilon}} \right) \geq 1 - \varepsilon$$

$\Rightarrow$ všimněme si, že se nám nutně mění i význam toho odhadu ve smyslu „jak měříme, že je něco ne/pravděpodobné".

To by nemuselo být tak překvapivé $\rightarrow$ Ω může být neomezená a my samplujeme v I_{MCIS} z celé Ω. Tudíž větto nelze „klasicky, rovnoměrně" a tedy ani nedává a-priori smysl měřit pravděpodobnost výsledku ve stylu „klasicky, rovnoměrně", ale naopak použít pravděpodobnost indukovanou tím procesem generování $\{X_i\}_{i=1}^N$.

$\rightarrow$ konvergence opět řádu $(\varepsilon \cdot N)^{-1/2}$ s pravděpodobností alespoň $1-\varepsilon$

$\rightarrow$ k rozmyšlení: lze získat I_{MC} jako specifický případ I_{MCIS}?

Další adaptace MC metod

- ## adaptivní MC / MCIS

Dejme tomu, že chceme celkem použít N bodů pro aproximaci $\int_{\mathcal{R}} f(x)\,dx$.
Nejdřív nasampluju $N/2$ bodů z daného rozdělení w (klidně rovnoměrné) a spočtu si odhad $I_{MC/IS}(N/2) \approx \int_{\mathcal{R}} f(x)\,dx$.
Zbývá mi ještě dalších $N/2$ bodů → nasampluju je tam z oblasti $\tilde{\mathcal{R}} \subseteq \mathcal{R}$, kde na základě hodnot $\{f(x_i)/w(x_i)\}_{i=1}^{N/2}$ je největší rozptyl $\mathrm{var}_w(f(x_i)/w(x_i))$

- ## stratifikované MC / MCIS

Dejme tomu, že chceme celkem použít N bodů pro aproximaci $\int_{\mathcal{R}} f(x)\,dx$.
Rozdělím si $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2 \cup \ldots \cup \mathcal{R}_m$ a použiju MC/IS s $N/2m$ body na každou z oblastí $\mathcal{R}_i$ → stále mi zbývá $N/2$ bodů. Ty využiju k tomu abych „resamploval" $\mathcal{R}_i$ s nejvyšší hodnotou $\mathrm{var}_w(f(x_i)/w(x_i))$.

→ souhrnná idea: přidat body tam, kde je vysoký rozptyl a tím ho zmenšit → zlepšujeme odhad

- ## Quazi Monte-Carlo

Idea je založená na pozorování, že pro klasické algoritmy generující náhodná čísla z rovnoměrného rozdělení pozorujeme jejich „shlukování" → demo v pythonu.
Jak jsme tyto čísla dostali? Klasicky číslici po číslici jako zbytky po dělení
Lze i lépe! Lze generovat „také náhodně vypadající", ale bez shlukování
→ tzv. quazi random numbers.

Dobrou volbou lze dosáhnout konvergence $\mathcal{O}\left(\frac{\log^d(N)}{N}\right)$ místo $\mathcal{O}(N^{-1/2})$