Přednáška 7 - Monte Carlo metody

option pricing pro 1 komoditu (zlato):

$$V_T(K) = \int_{\log(K)}^{+\infty} (e^x - e^{\log(K)})\, q_T(x)\, dx$$

$$\approx \sum_0^n w_i \cdot (e^{x_i} - K)\, q_T(x_i)$$ ... x_i kvadraturní body ... w_i kvadraturní váhy

option pricing pro 50 komodit

$$V_T(\vec{K}) \equiv V_T(K_1, .., K_{50}) = \int_{k_1}^{+\infty} \cdots \int_{k_{50}}^{+\infty} f(x_1, .., x_{50}, \vec{K}, ..)\, dx_{50} \cdots dx_1$$

$\rightsquigarrow$ numerická integrace na základě aproximace $\underline{f}$ trvá příliš dlouho $\rightsquigarrow$ potřebovali bychom „n^{50} vyhodnocení f".

V čem je problém?

pokud chci zadávat konstantní počet kvadraturních bodů v každé dimenzi, pak mi exponenciálně roste počet bodů ve kterých musím f vyhodnotit v závislosti na dimenzi $f: \mathbb{R}^d \to \mathbb{R}$

to chceme, protože to nám určuje chybu naší kvadratury

to je problém, protože v praxi se běžně dostaneme k $d \approx 10^2$ a i hrubé odhady ukazují, že tohle nejde upočítat

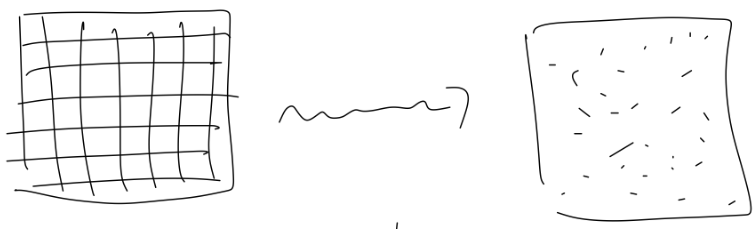
Řešení: musíme se vzdát přístupu „jedna dimenze po druhé"

→ možnost 1:

⇝ grid bodů změníme v mesh $\triangle$, ..., tzv. simplexy

⇝ trochu lepší, ale neřeší jádro problému

→ možnost 2:



⇝ grid bodů vyměníme za náhodně generované body

možnost 2 ≡ Monte-Carlo integrační metody

$$\int_\Omega f(x)\,dx \approx \sum_{1}^{N} w_i f(x_i) \quad \text{- - - -}$$

stejné jako pro kvadratury
aprox. $\int$ lineární komb. $f(x_i)$

- pro rovnoměrně rozdělené náh. body chceme integrovat přesně konstantní funkce

  ⇝ $f(x) = C \in \mathbb{R} \Rightarrow \int_\Omega f\,dx = \mathrm{vol}(\Omega)\cdot C$ & $\sum_{i=1}^{N} w_i f(x_i) = C \cdot \sum_{i=1}^{N} w_i$

- $x_1, \dots, x_N$ jsou rovnoměrné & náhodné → dává smysl aby měli „stejnou důležitost"

  ⇝ $w_1 = w_2 = \dots = w_N \Rightarrow \sum_{i=1}^{N} w_i f(x_i) = C \cdot N \cdot w$ & $w_i = w \; \forall i$.

- abychom integrovali přesně konstanty musí platit $\mathrm{vol}(\Omega)\cdot C = C \cdot N \cdot w$

Dohromady tedy dostaneme
$$\int_\Omega f(x)\,dx \approx \frac{\mathrm{vol}(\Omega)}{N} \cdot \sum_{i=1}^{N} f(x_i)$$

Opáčko teorie pravděpodobnosti

Mějme X náhodnou veličinu na Ω spojenou s pravděpodobností mírou $\mathbb{P}$:

$$\mathbb{P}(|z-z_0|\leq\delta) = \int_\Omega \chi_{\{|z-z_0|\leq\delta\}}(z)\, d\mathbb{P}_z \equiv \int_\Omega \chi_{\{|z-z_0|\leq\delta\}}(z)\cdot w(z)\, dz$$

$$E(Z) = \int_\Omega z\, d\mathbb{P}_z = \int_\Omega z\cdot w(z)\, dz \qquad var(Z) = \mathbb{E}\left((Z-\mathbb{E}(Z))^2\right) = \mathbb{E}(Z^2) - \mathbb{E}(Z)^2$$

$w(\cdot)$ je tzv. probability density function (pdf) naší pravděpodobnostní míry $\mathbb{P}$.... např. pro rovnoměrné dělení máme $w(z) = 1/vol(\Omega)$

Platí tzv. Chebyshevovo lemma:

$$\mathbb{P}(|Z-\mathbb{E}(Z)|\geq\delta) \leq \frac{var(Z)}{\delta^2}.$$

Naše použití:

X .... náhodná veličina s hodnotami v Ω s rovnoměrným rozděl.

$X_1, \dots, X_N$ .... N náhodných, nezávislých veličin (také rovnoměrným rozdel.)

$Y_N := \frac{vol(\Omega)}{N}\sum_{i=1}^{N} f(X_i)$ .... také náhodná veličina & víme:

- $\mathbb{E}(Y_N) = vol(\Omega)\cdot\mathbb{E}(f(X)) = \int_\Omega f(x)\,dx \qquad \left(\text{„}\Leftrightarrow\text{"} \ Y_N \approx \int_\Omega f(x)\,dx\right)$
- $var(Y_N) = \frac{vol(\Omega)^2}{N^2}\cdot N\cdot var(f(X)) = vol(\Omega)^2\cdot\frac{var(f(X))}{N}$

Chebyshevovo lemma:

$$\underbrace{\mathbb{P}\left(\left|Y_N-\mathbb{E}(Y_N)\right|\geq\delta\right)}_{\text{pravděpodobnost, že integrace metodou Monte-Carlo má chybu větší než } \delta} \leq \underbrace{vol(\Omega)^2\cdot\frac{1}{\delta^2}\cdot\frac{var(f(X))}{N}}_{\delta^{-2}\times\frac{\text{„velikost } \Omega\text{"}^2\times\text{ rozptyl } f(X) \text{ v } \Omega}{\text{počet samplovaných bodů}}}$$

Přepíšeme úhledněji: zafixujeme $\varepsilon > 0$ a zadefinujeme $\delta > 0$ jako

$$\delta := \sqrt{\mathrm{vol}(\Omega)^2 \cdot \frac{\mathrm{var}(f(X))}{\varepsilon \cdot N}}.$$ Pak platí

$$\mathbb{P}\left(\left| Y_N - \int_\Omega f(x)\,dx \right| \geq \delta \right) \leq \varepsilon$$ a odtud dostáváme

$$\mathbb{P}\left(\left| Y_N - \int_\Omega f(x)\,dx \right| \leq \sqrt{\mathrm{vol}(\Omega)^2 \cdot \frac{\mathrm{var}(f(X))}{\varepsilon \cdot N}} \right) > 1 - \varepsilon$$

samplujeme iid rovnoměrně
&
$\varepsilon > 0$ je velmi malé
&
$\sqrt{\mathrm{vol}(\Omega)^2 \cdot \frac{\mathrm{var}(f(X))}{\varepsilon \cdot N}}$ je velmi malé

$\Longrightarrow$ metoda Monte-Carlo mám dá s velmi vysokou pravděpodobností velmi dobrou aproximaci integrálu

- pro iid $\{X_i\}_{i=1}^N$ s rovnoměrným rozdělením & pro fixní ε chyba klesá (s pravděp. $1-\varepsilon$) asymptoticky jako $\mathcal{O}(1/\sqrt{N})$.
  $\rightsquigarrow$ velmi pomalu (porovnání s kvadraturami, viz cvičko).

Jak zlepšit?
- $\nearrow$ snížit rozptyl $\mathrm{var}(f(X))$
- $\searrow$ zvětšit počet bodů bez zvýšení rozptylu

- co když mám Ω neomezenou $\to$ jak samplovat rovnoměrné rozdělení? A jak spravit „$\mathrm{vol}(\Omega)$"?
- lze i jiná rozdělení než rovnoměrná?