Introduction

- <u>What is MC?</u>
  Set of methods for
  - sampling distributions $p(x)$
  - generate samples/realization/data from prob. models $X \sim p(x)$
  - estimate distributions / learn prob. models $\hat{p}(x)$
  - estimate expectations/integrals $E[g(x)]$

- Can also be used to solve optimization problems
- Use noise/randomness for solving complex problems

- <u>Why MC in ML?</u>

  - ML $\equiv$ learning distributions
  - Learn prob. models
  - Generate data from prob. models (generative)
  - Learn parameters / Bayesian inference
  $$P(\theta | D) = \frac{P(D|\theta) \, P(\theta)}{P(D)}$$
  $$\propto P(D|\theta) \, P(\theta)$$
  posterior   model likelihood   Prior

- <u>Key ideas/concepts:</u>
  - Distribution over high dim space
  - Exploration vs reinforcement
    $\uparrow$
    noise/randomness
  - Estimation vs generation: Data $\rightarrow \hat{p}(x) \rightarrow$ new data
  - learning $\equiv$ optimization