Aula 6

Aula passada

- Limitante da união
- Método do primeiro momento
- Lei dos grandes números (fraca e forte)
- Erro e confiança

Aula de hoje

- Método de Monte Carlo
- Estimando somatórios
- Calculando erro
- Estimando π
- Erro de π
- Integração de Monte Carlo
- Monte Carlo Ray Tracing

Método de Monte Carlo

- Classe de algoritmos baseado em amostragem aleatória repetida
  - obter solução aproximada para problemas determinísticos
- Ideia central: grande número de amostras repetidas acabam revelando a solução
  - no limite, amostras dão a solução

Lei dos grandes números!

$$M_n = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad P[\lim_{n \to \infty} M_n = \mu] = 1$$

$\mu = E[X_i], \forall i$

- Arcabouço teórico para método de Monte Carlo

Calculando Soma

- Calcular o valor de um somatório (problema da aula 1)
  - *N* (número de parcelas) é muito grande
  - calcular valor de cada parcela é fácil

$$G_N = \sum_{i=1}^{N} \underbrace{g(i)}_{\text{parcela}}$$

- Como usar aleatoriedade para resolver (aproximar) este somatório?
- Usando valor esperado!

- Seja X uma v.a. uniforme em [1, N]
  - $P[X = i] = 1/N$ para i=1, …, N

Faço com que não dependa de i → uniforme

$$E[g(X)] = \sum_{i=1}^{N} P[X=i]\, g(i) = \frac{1}{N} \sum_{i=1}^{N} g(i) = \frac{G_N}{N}$$

Figueiredo 2018

Calculando Soma

- Logo, temos que $G_N = N\, E[g(X)]$ → Equação determinística pois N é um parâmetro
- Podemos agora estimar $E[\, g(X)\,]$. Como?

Gerando amostras, fazendo a média!

- Seja X_i sequência iid de v.a. uniforme [1, N]
  - escolher um valor para n (número de amostras)

1º passo de Monte Carlo: escrever o que se quer calcular por um E.

$$M_n = \frac{1}{n} \sum_{i=1}^{n} g(X_i)$$ ← Média amostral com n amostras

Estimador para E

- Temos que $E[M_n] = E[g(X)]$, para todo n → O estimador não possui viés
- $M_n \rightarrow E[g(X)]$ quando $n \rightarrow \infty$ (pela lei dos grds números)
- M_n é estimador de E[g(X)], logo $N*M_n$ é estimador para G_N

Arestas no Facebook

- Quantas arestas tem a rede de amizade do Facebook?
  - função indicadora de aresta, $g(i, j) = 1$ se existe aresta entre perfil i e j
  - número de perfis $n_p = 2*10^9$

Todos os pares

$$T = \sum_{i=1}^{n_p} \sum_{j=i+1}^{n_p} g(i,j)$$

T = número de arestas

Problema: somatório tem ~ 10^{18} termos!

- Como construir um método de Monte Carlo para obter um valor aproximado para T?

Enumerar os pares

Arestas no Facebook

- Seja Z = (*i*, *j*) uma v.a. uniforme no conjunto de pares com n_p perfis

P → conjunto de pares de perfis

$\mathbb{E}(g(i,j)) = \sum_{(i,j)\in E} P[x=i \wedge y=j]\, g(x,y)$

  - $N = n_p(n_p-1)/2$ → número total de pares

$N = \binom{n_p}{2}$

$= \sum_{i=1}^{n_p} \sum_{j=i+1}^{n_p} \frac{1}{N} g(i,j)$

$= \frac{1}{N} T \rightarrow$

$$E[g(Z)] = \sum_{i=1}^{n_p} \sum_{j=i+1}^{n_p} \frac{1}{N} g(i,j) = \frac{1}{N} T$$

↳ Sempre começar com um $\mathbb{E}$

$T = N \cdot \mathbb{E}[g(Z)]$

- Seja Z_i sequência iid de v.a. uniforme sobre pares

↳ i-ésimo par escolhido de maneira uniforme do conjunto de pares

NÃO HÁ NADA ALEATÓRIO

$$M_n = \frac{1}{n} \sum_{i=1}^{n} g(Z_i)$$

← Média amostral com *n* amostras, estima E[g(Z)]

- Logo *T* é aproximado por $M_n * N$

Jogando Paciência

- Qual é a fração de vezes que um algoritmo determinístico *A* vence o jogo de paciência?
  - vencer depende apenas da permutação das cartas do baralho (não há aleatoriedade)
- Seja *s* uma permutação das cartas
  - $f_A(s) = 1$ se algoritmo *A* vence com permutação s, 0 c.c.
- *N* = número total de permutações das cartas, *N* = 52!

Uma maneira óbvia é passar por todas as permutações e verificar a fração de vezes em que o algoritmo venceu

$$F_A = \frac{1}{N} \sum_{s=1}^{N} f_A(s)$$

Fração de vezes que algoritmo *A* vence

Queremos estimar esta fração.

Problema: somatório tem 52! termos!

- *Monte Carlo to the rescue!*

Jogando Paciência

- Seja S uma v.a. uniforme em [1, N] que representa uma permutação
  - uniforme entre todas possíveis permutações das cartas

$P[S=s]=1/N$

$$E[f_A(S)]=\sum_{s=1}^{N}\frac{1}{N}f_A(s)=\frac{1}{N}\sum_{s=1}^{N}f_A(s)=F_A$$

Etapa 1: Trocar um $\sum$ por um $\mathbb{E}$

- Valor esperado já é a fração que queremos!
- Seja S_i sequência iid de v.a. uniforme em [1, N]
  - S_i é uma permutação das cartas → permutação aleatória

$$M_n=\frac{1}{n}\sum_{i=1}^{n}f_A(S_i)$$

← Média amostral da fração de vezes que algoritmo vence

- Logo, F_A é a aproximado por M_n

Etapa 2: Estimar $\mathbb{E}$

Vantagens e Desvantagens

- Trocar $G_N = \sum_{i=1}^{N} g(i)$ por $M_n = \frac{1}{n} \sum_{i=1}^{n} g(X_i)$
- **Quando esta ideia é boa?**

n possui 2 papéis:
- calcular o estimador
- controlar a qualidade do estimador

- Comparação entre *N* e *n*
- Se *N* for pequeno, ideia não é boa
  - podemos calcular G_N diretamente
- Se calcular g(.) é muito caro, boa ideia mesmo quando *N* pequeno

complicada

- Se *g*(*i*) for muito "errática", ideia não é boa

outlier

  - ex. um valor de *g*(*i*) é maior que todo o resto da soma
  - estimador pode ser muito ruim se *n* não for muito grande
- Qualidade da aproximação depende de *N, n, g*()

Figueiredo 2018

Calculando o Erro

$G_N = \sum_{i=1}^{N} g(i)$

- Podemos usar Chebyshev para calcular *n*
  - para precisão ε e confiança β, temos

$$P[M_n \in [\mu - \epsilon, \mu + \epsilon]] > 1 - \frac{\sigma^2}{\epsilon^2 n} = \beta$$

precisão ↲ ↳ confiança

- onde μ e σ² são o valor esperado e variância da v.a. que será aproximada pelo estimador

$$\mu = E[g(X)] \qquad \sigma^2 = Var[g(X)]$$

μ não é um problema para calcular n, mas σ² é algo que precisa ser estimado

- onde *g*( ) é geralmente uma função indicadora (para contar coisas) e *X* é uma v.a. uniforme nos valores que *g* pode assumir

1. Estimar σ²
2. Estimar n

- **Problema**: muitas vezes não sabemos σ²
  - temos que estimar com o estimador!

Figueiredo 2018

Erro da Paciência

- Seja F_A = 0.1 fração de vezes que algoritmo *A* vence
  - *S* é v.a. uniforme nas permutações

    variância da
    indicadora $p(1-p)$
  - $\mu = E[\, f_A(S)\,] = 0.1$, $\sigma^2 = Var[f_A(S)] = 0.1*0.9 = 0.09$

    indicadora (Bernoulli)
- Supor $\varepsilon = 10^{-4}$ e $\beta = 0.99$

$$P[M_n \in [\mu - \epsilon, \mu + \epsilon]] > 1 - \frac{\sigma^2}{\epsilon^2 n} = \beta$$

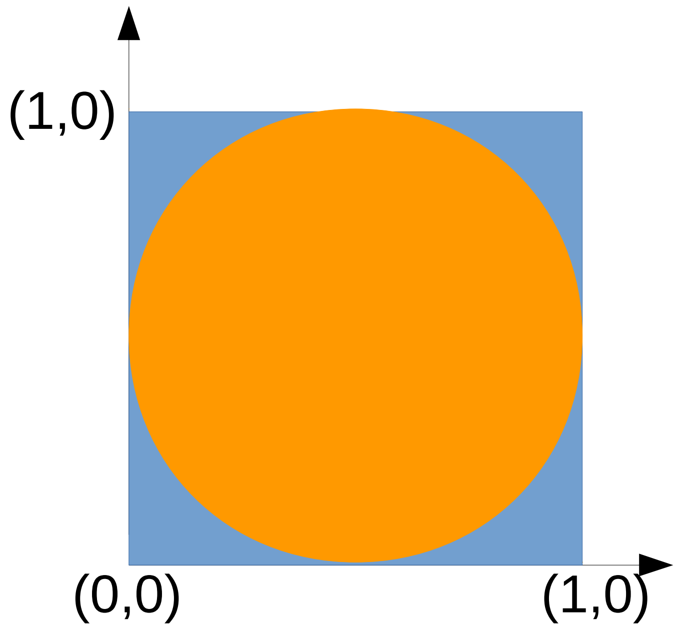
$$1 - \frac{0.09}{(10^{-4})^2 n} = 0.99 \rightarrow n = 9 * 10^8$$

Trocamos um Σ de 10^{68} parcelas por um de 10^8 parcelas e com uma confiança muito alta

- Muito, muito menor do que $52! \sim 10^{68}$

Calculando π

- Como estimar o valor de π?
  - ou qualquer outro valor que tenha relação com geométrica
- **Ideia**
  - Escrever π como relação entre áreas
  - Usar Monte Carlo para estimar relação

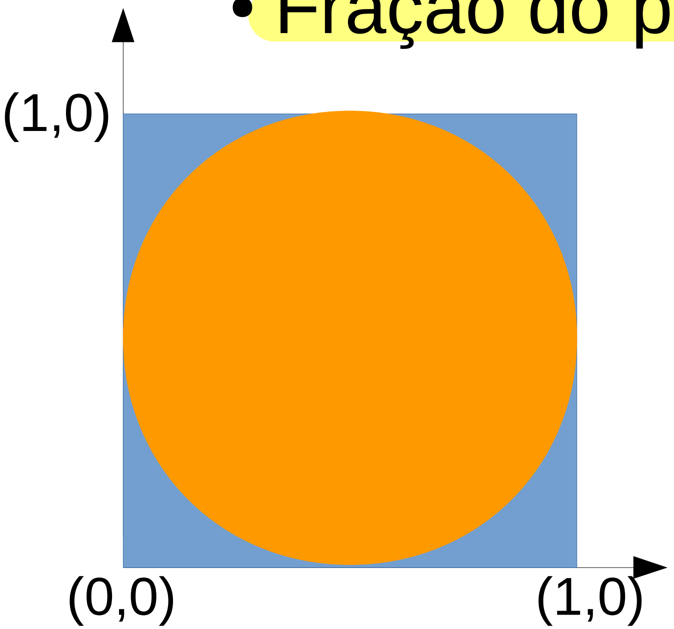


- A_q = Área do quadrado = 1
- A_c = Área do círculo = $\pi r^2 = \pi/4$
- $\pi = 4 * A_c / A_q$
- Estimar A_c / A_q

Figueiredo 2018

Calculando π

- Ideias para estimar A_c / A_q ?
- Gerar *n* pontos uniforme no quadrado
- Fração do pontos que estão dentro do círculo!



$$E[g(x,y)] = \frac{A_c}{A_q}$$

Uniforme no quadrado

- Seja *X* e *Y* duas v.a. uniformes contínuas em [0, 1]
- Seja $g(x, y)$ indicadora do ponto (x, y) estar dentro do círculo
  - $g(x,y) = 1$ se $(x\text{-}0.5)^2 + (y\text{-}0.5)^2 <= 1$
- Temos $E[g(x, y)] = A_c / A_q = \pi/4$
- Como estimar $E[g(x, y)]$?

Calculando π

- Seja X_i, Y_i sequência iid uniforme em [0,1]
  - v.a. contínua

$$M_n = \frac{1}{n}\sum_{i=1}^{n} g(X_i, Y_i) \longleftarrow$$

Fração de pontos que estão dentro do círculo

- M_n converge para $\pi/4$ (pela lei dos grds números)
- Logo π pode ser estimado por $4 * M_n$

O Erro de π

- Podemos usar Chebyshev para calcular *n*
  - para precisão ε e confiança β, temos

$$P[M_n \in [\mu - \epsilon, \mu + \epsilon]] > 1 - \frac{\sigma^2}{\epsilon^2 n} = \beta$$

confiança

* note que não precisamos de μ para fazer este cálculo, mas precisamos de σ^2

$\pi/4$

precisão

$E[g(X,Y)] = \frac{A_c}{A_q} = \pi/4 = p$

- μ= E[g(X,Y)] = P[g(X,Y) = 1] = π/4

$p(1-p)$

- σ^2 = Var[g(X,Y)] = π/4 (1-π/4)

Variância da indicadora: $p \cdot (1-p)$

- Não sabemos π mas σ^2 de uma v.a. indicadora tem valor máximo quando $p = ½ \rightarrow \sigma^2 = ¼$

O n calculado com $p = 1/2$ certamente satisfaz a confiança mínima, pois σ^2 é máximo para $p = 1/2$.

- Supor $\varepsilon = 10^{-6}$ e $\beta = 0.99$
  - então temos que $n = 2.5 * 10^{16}$

MSE e SEM

- Mean Squared Error (MSE)
  - medida clássica para erro de preditores ou estimadores

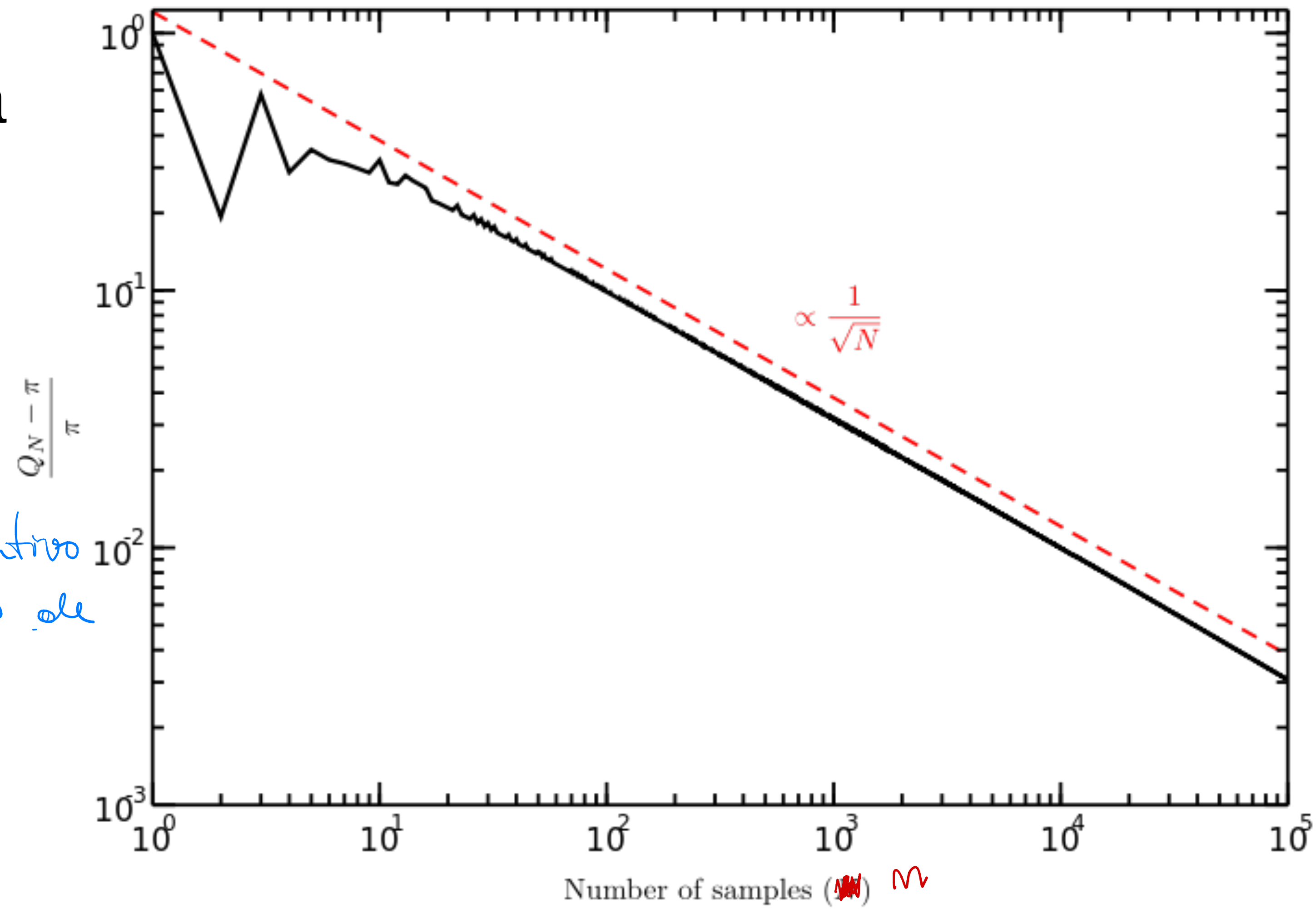
    estimador não - aleatório
  - $\text{MSE}(\phi') = E_{\phi'}[(\phi' - \phi)^2]$, onde ϕ' é o estimador e ϕ o valor a ser estimado aleatório média amostral
- Seja M_n (média amostral) o estimador para o valor μ
  - $\text{MSE}(M_n) = \text{Var}(M_n) = \sigma^2/n$
- Standard Error of the Mean (SEM)
  - medida de erro relativo $\sqrt{\frac{\sigma^2}{n}} = \sqrt{\text{MSE}(M_n)}$
  - $\text{SEM}(M_n) = \sigma/\text{sqrt}(n) = \text{sqrt}(\text{MSE}(M_n))$
- Para qualquer M_n, sempre decresce como sqrt(n)
  - não é muito rápido! SEM decresce muito devagar

Figueiredo 2018

SEM de π

- SEM(M_n) = σ/sqrt(n)
- Erro relativo = $| 4M_n - \pi | / \pi$ = y
- Teoria e prática estão bem de acordo!
- Repare 1/sqrt(n) é devagar !

Para um erro relativo de 10^{-2} precisamos de 10^4 amostras



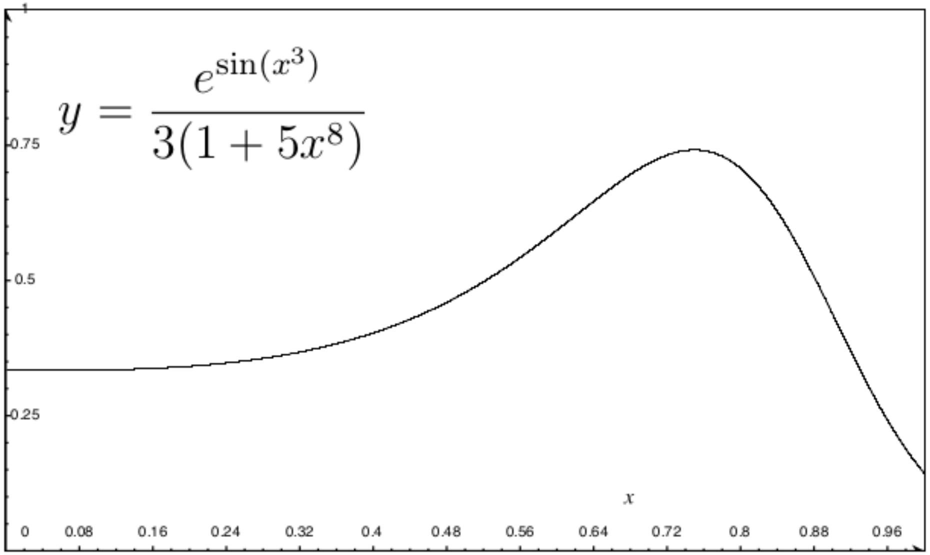
Figueiredo 2018

Integração Numérica

- Encontrar a integral definida de uma função

$$\int_{x=a}^{x=b} f(x)\,dx$$

- **Problema:** função pode não ser integrável (ou muito difícil de integrar)!



- **Solução:** integração numérica
- Conjunto de algoritmos para calcular valor da integral
  - muito usado na física, química, engenharia, etc

- Abordagem via Monte Carlo é uma classe de algoritmos
  - *Monte Carlo integration*

Figueiredo 2018

Integração de Monte Carlo

- Generalização da ideia de computar o valor de π
- Calcular razão entre área de "baixo da curva" (integral) e de um "quadrado" (limites de integração)
- Estimar este valor usando amostras uniformes
- Supor 0 <= f(x) <= 1, para x em [0, 1]

$$I = \int_{x=0}^{x=1} f(x)\,dx$$

- Definir função indicadora para ponto de baixo da curva
  - $g(x, y) = 1$ se $f(x) <= y$, e zero c.c.
- Seja X, Y v.a. contínuas uniformes em [0, 1]

Integração de Monte Carlo

- Valor esperado

$$E[g(X,Y)] = \int_{x=0}^{x=1} \int_{x=0}^{x=1} f_{XY}(x,y)\, g(x,y)\, dx\, dy$$

- $f_{XY}(x,y) = f_X(x)\, f_Y(y) = 1$ (densidade conjunta de duas v.a. contínuas uniformes e independentes)
- $E[\,g(X,Y)\,]$ = área embaixo da curva / área do quadrado [0,1]x[0,1] = I

X e Y tem domínio uniforme

$g(x,y) = 1$ se $f(x) \leq y$

- Como estimar $E[\,g(X,Y)\,]$?
- Seja X_i, Y_i sequência iid de uniformes em [0, 1]

indicadora do ponto (x_i, y_i) abaixo da curva

$$M_n = \frac{1}{n} \sum_{i=1}^{n} g(X_i, Y_i)$$

Fração de pontos que estão baixo da curva

Figueiredo 2018

Integração de Monte Carlo

- Método alternativo: relacionar valor esperado diretamente com valor da integral

função que quero integrar

$$I=\int_{x=0}^{x=1} f(x)\,dx$$

- Seja X v.a. contínua uniforme em [0, 1]

$$E[f(X)]=\int_{x=0}^{x=1} f_X(x) f(x)\,dx=\int_{x=0}^{x=1} f(x)\,dx=I$$

- Pois $f_X(x) = 1$ (densidade da v.a. uniforme em [0,1])
- Como estimar $E[\,f(X)\,]$?
- Seja X_i sequência iid uniforme em [0, 1]

Converge para valor esperado em n (lei dos grds números) ←

$$M_n=\frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

Figueiredo 2018

Generalização

- Funciona para qualquer limite de integração, e qualquer número de dimensões
  - grande vantagem da abordagem Monte Carlo

→ volume dentro dos limites de integração ω

$$I = \int_{\omega} f(\vec{x})\, d\vec{x} \qquad\qquad V = \int_{\omega} d\vec{x}$$

↳ integral múltipla com k dimensões

- onde x é um ponto em um espaço de k dimensões, ω um "pedaço" fechado deste espaço, e V é o volume de ω
- Seja X_i uma sequência iid de v.a. uniforme no espaço ω
  - X_i tem k dimensões
- Temos

INTEGRAL

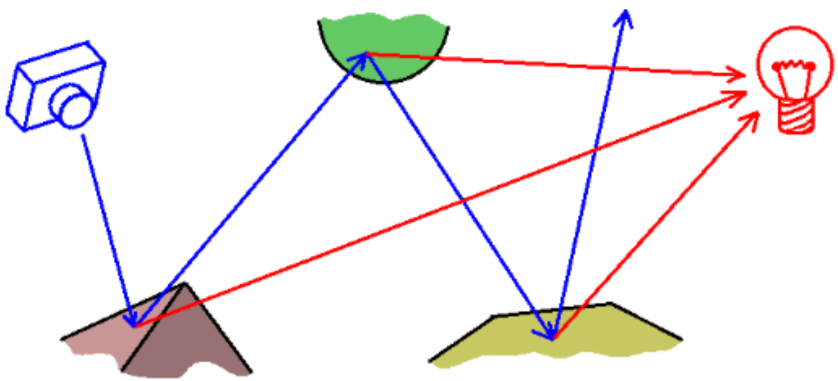
VOLUME

$$M_n = \frac{1}{n} \sum_{i=1}^{n} f(\vec{X}_i)$$ converge para I/V quando n → infinito

Monte Carlo *Ray Tracing*

- Integração de Monte Carlo em Computação Gráfica
- **Problema:** determinar cor (intensidade de luz) em um pixel em uma cena construída – Renderização

Integral formulation ("rendering equation")



$$R = \int_{\Omega} L(y)dy$$ ← Resolvida numericamente usando método de Monte Carlo

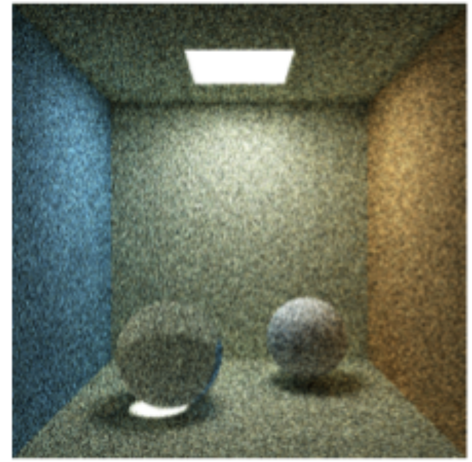
↳ caminho

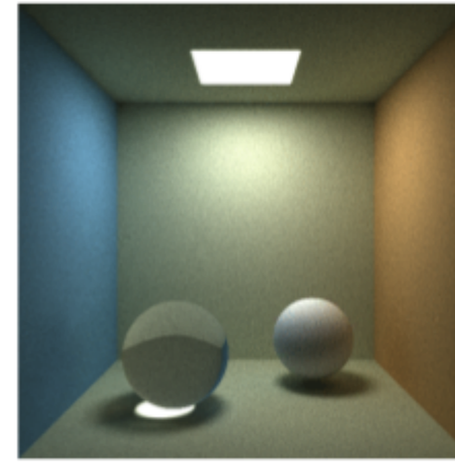
- Integrate over all possible light paths y ← Muitos, muitos caminhos!

Monte Carlo *Ray Tracing*

uniforme

$$\int_\Omega L(y)dy = E\left[\frac{L(Y)}{\underbrace{\text{pdf}[Y]}_{\text{constante}}}\right]$$





Se diferenciam no número de amostras utilizadas na renderização

Advantage of Monte-Carlo

- Good compromise between cost and precision: $\epsilon = O\left(1/\sqrt{N}\right)$

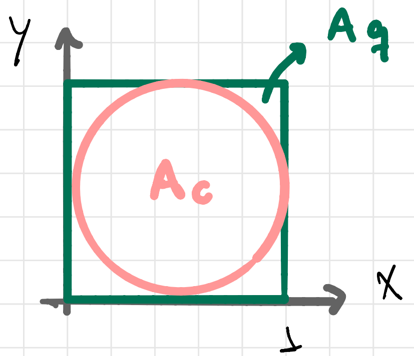
  computacional

  - in contrast, integration by e.g. trapezoid rule has an error of $\epsilon = O\left(1/N^{2/d}\right)$, for d dimensions

Disadvantage of Monte-Carlo

- Output is random → A renderização depende da realização

Estimando π



Seja $X \sim \mathcal{U}[0,1]$ e $Y \sim \mathcal{U}[0,1]$ e $g(x,y)$ uma variável aleatória indicadora tal que

$$g(x,y) = \begin{cases} 1, & \text{se } x^2+y^2 \leq 1 \\ 0, & \text{caso contrário} \end{cases}$$

Desta forma, $g(X,Y)$ pode ser modelada como uma Bernoulli cujo parâmetro p é a fração A_c / A_q. Logo,

$$\mathbb{E}[g(X,Y)] = p = \frac{A_c}{A_q} = \frac{\pi r^2}{(2r)^2} = \frac{\pi}{4}$$

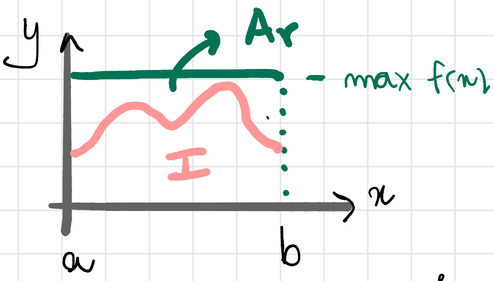
Por Monte Carlo, podemos estimar $\mathbb{E}[g(X,Y)]$ por amostragem

$$\mathbb{E}[g(X,Y)] \approx \frac{1}{n} \sum_{i=1}^{n} g(x,y) = M_n$$

Portanto,

$$\boxed{\pi = 4M_n}$$

Integração de Monte Carlo - Método da Fração



Queremos calcular $I = \int_a^b f(x)\,dx$

$f_X(x) = \frac{1}{b-a}$

Seja $X \sim \mathcal{U}[a,b]$, $Y \sim \max f(x) \cdot \mathcal{U}[0,1]$

e $g(X,Y) = \begin{cases} 1, & \text{se } f(x) \leq y \\ 0, & \text{caso contrário} \end{cases}$

Desta forma, $g(x,y)$ pode ser modelada como uma Bernoulli cujo parâmetro p é a fração I/A_r. Assim,

$$\mathbb{E}[g(X,Y)] = p = \frac{I}{A_r} = \frac{I}{(b-a)\cdot \max f(x)}$$

Por Monte Carlo, podemos estimar $\mathbb{E}[g(X,Y)] \approx \frac{1}{n}\sum_{i=1}^{n} g(x,y) = M_n$

Logo,

$$\boxed{I = (b-a)\cdot \max f(x) \cdot M_n}$$

Integração de Monte Carlo - Método Direto

Queremos calcular $I = \int_a^b f(x)\, dx$.

Seja $X \sim \mathcal{U}[a,b] \rightarrow f_X(x) = \frac{1}{b-a}$. Tomando-se o valor esperado de $f(x)$ em relação a f_X temos

$$\mathbb{E}[f(X)] = \int_a^b f_X(x) \cdot f(x)\, dx = \int_a^b \frac{1}{b-a} \cdot f(x)\, dx = \frac{I}{b-a}$$

Por outro lado, $\mathbb{E}[f(x)]$ pode ser estimado por Monte Carlo:

$$M_n = \frac{1}{n} \sum_{i=1}^{n} f(x_i) \approx \mathbb{E}[f(x)]$$

Logo,

$$\boxed{I \simeq (b-a) M_n}$$