

Introdução à Física Computacional II (4300318)

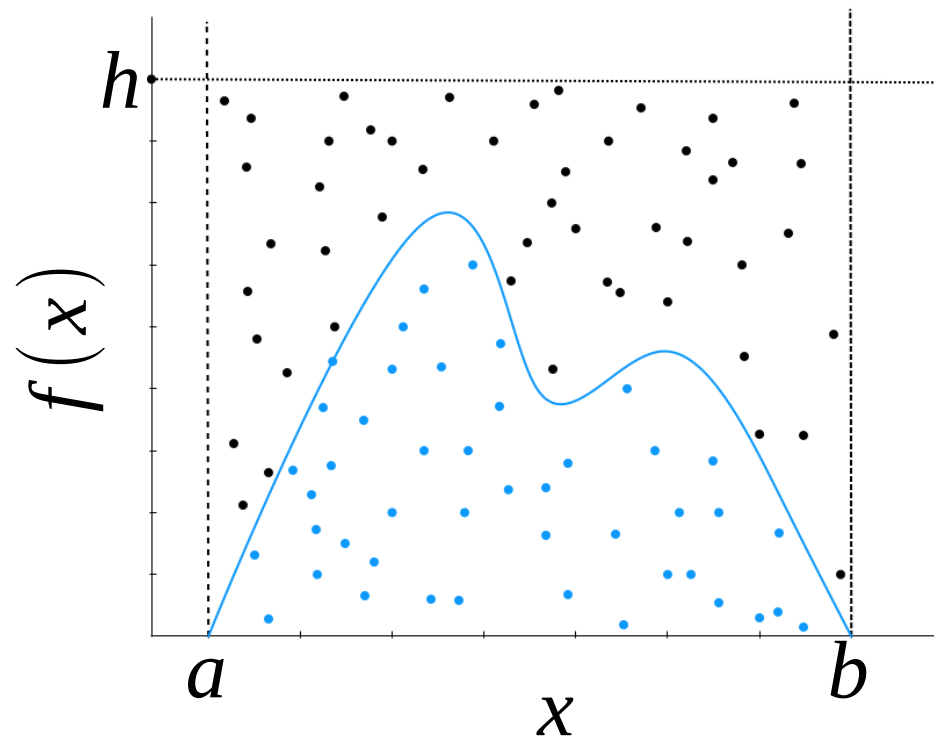
Prof. André Vieira
apvieira@if.usp.br
Sala 3120 – Edifício Principal

Aula 10

Integração por Monte Carlo

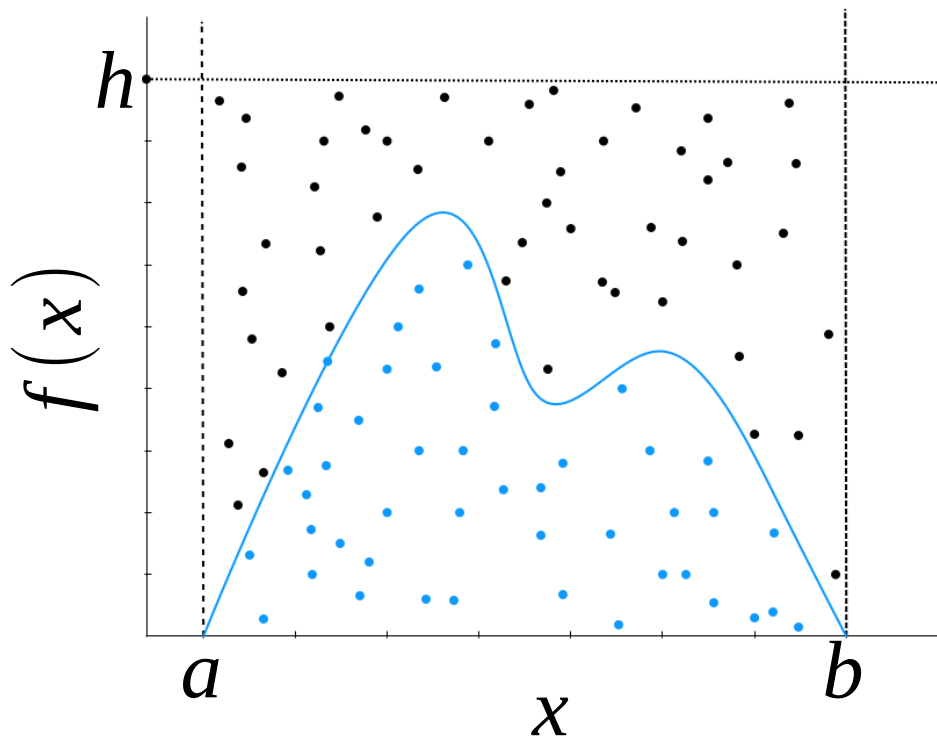
Integrar jogando pedras

- Aplicações para a integração de funções estão na origem dos métodos de Monte Carlo. (Para uma perspectiva histórica, veja a seção 1.4 do livro de [Newman e Barkema.](#))



Integrar jogando pedras

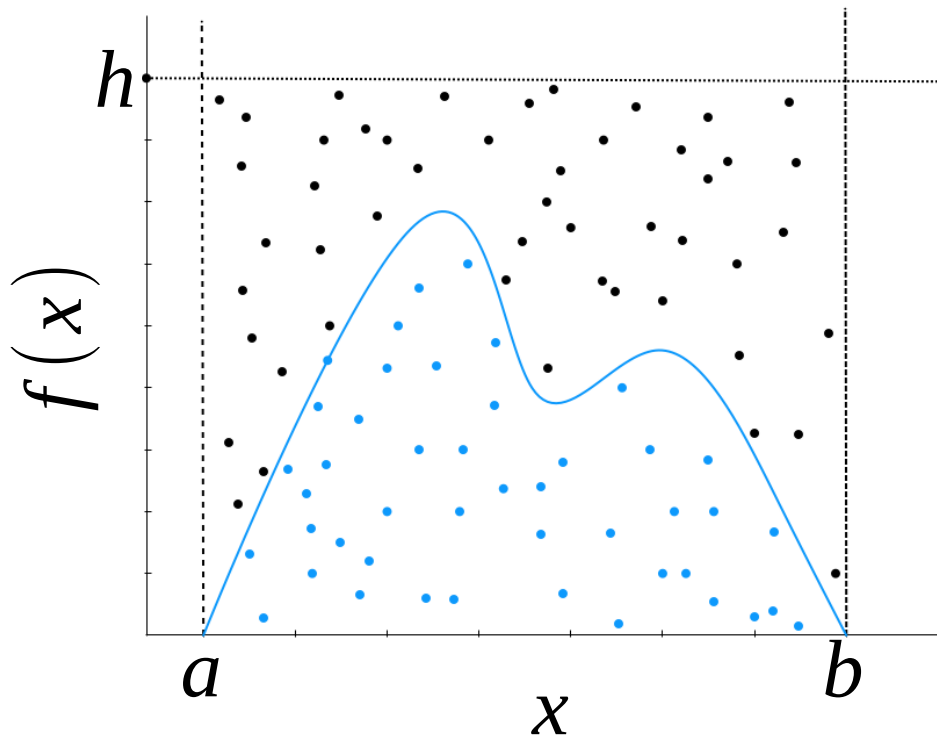
- A ideia é que, para integrar uma função $f(x) \geq 0$ entre $x=a$ e $x=b$, podemos “jogar N pedras” ao acaso em um retângulo que delimita a função e o intervalo de integração e contar o número k de “pedras” que caem abaixo da função.



$$\int_a^b f(x) dx \simeq \frac{k}{N} (b-a) h$$

Integrar jogando pedras

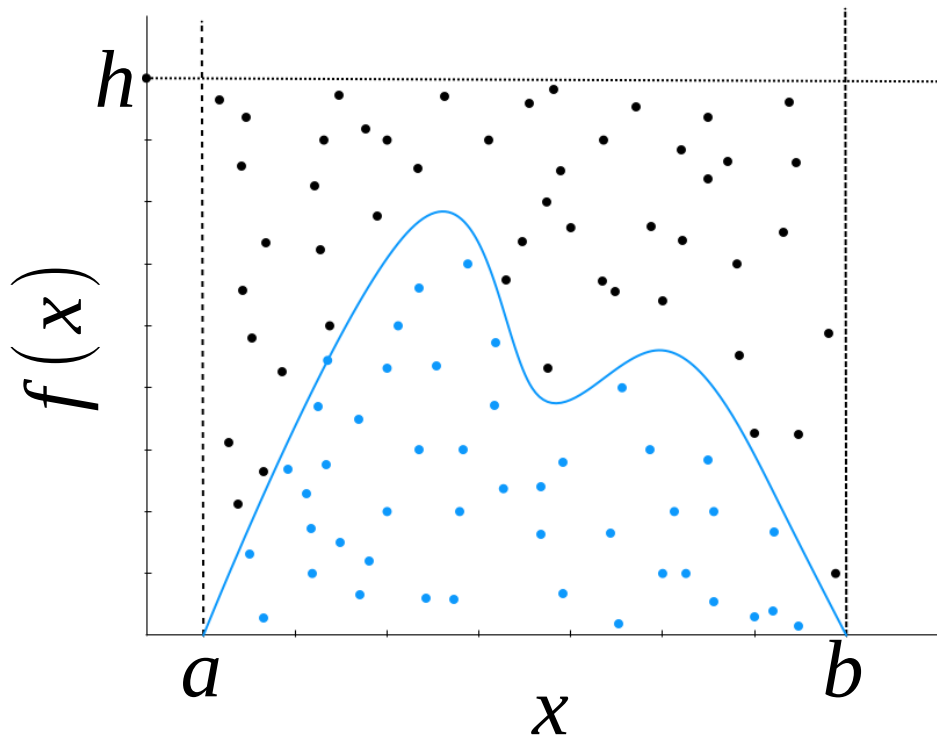
- Em outras palavras, a probabilidade de que uma pedra caia abaixo da função é $p = I/A$, sendo I o valor da integral e $A = (b-a)h$ a área do retângulo delimitante. A fração k/N é uma aproximação para p , e portanto para I .



$$I \simeq \frac{k}{N} A$$

Integrar jogando pedras

- Em outras palavras, a probabilidade de que uma pedra caia abaixo da função é $p = I/A$, sendo I o valor da integral e $A = (b-a)h$ a área do retângulo delimitante. A fração k/N é uma aproximação para p , e portanto para I .



$$I \simeq \frac{k}{N} A$$

O que fazer se a função assumir valores negativos no intervalo de integração?

Integrar jogando pedras

- Se a função $f(x)$ assumir valores negativos no intervalo de integração, podemos somar a ela uma constante positiva c cujo valor seja maior do que o módulo do valor mínimo de $f(x)$ no intervalo, definindo assim uma outra função $g(x)=f(x)+c \geq 0$ tal que

$$\int_a^b f(x) dx = \int_a^b g(x) dx - c(b-a),$$

com a integral de $g(x)$ podendo ser feita sem problemas pelo método de jogar pedras.

Exemplo 1

- Vamos utilizar esse método para estimar o valor de π . Como a área de um círculo de raio r é πr^2 , metade da área de um círculo unitário é $\pi/2$. Da equação que define os pontos da circunferência de raio unitário centrada na origem, obtemos a função a integrar:

$$x^2 + y^2 = 1 \Rightarrow y = f(x) = \sqrt{1 - x^2}.$$

- Queremos então estimar a integral

$$I = \int_{-1}^1 \sqrt{1 - x^2} dx.$$

Exemplo 1a

```
# Parâmetros de integração
a = -1.0      # Limite inferior de integração
b = 1.0      # Limite superior de integração
h = 1.0      # Altura do retângulo de integração
N = 10**4    # Número de pontos a utilizar para a integração

# Definição da função a ser integrada
def f(x):
    return sqrt(1-x*x)

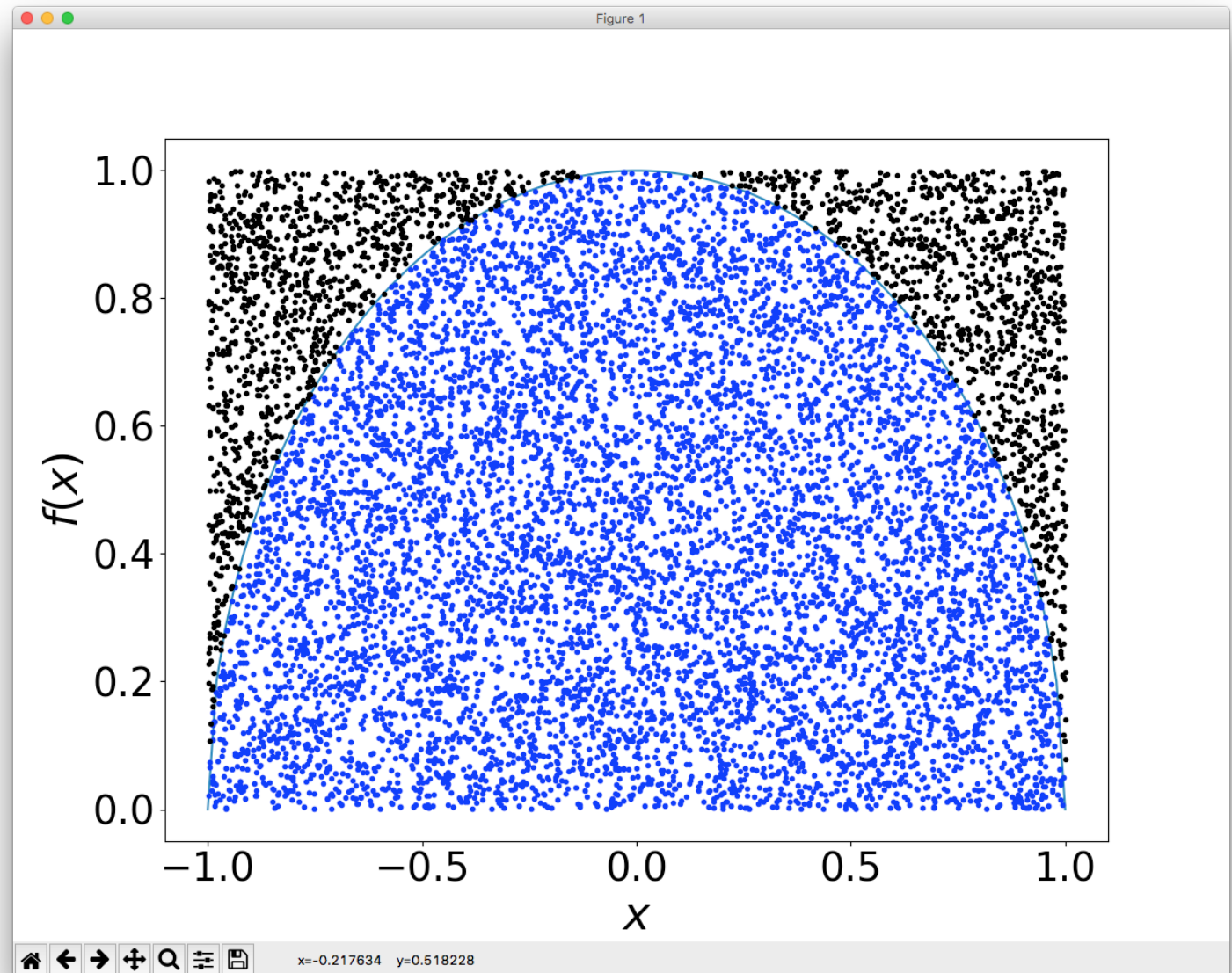
# Laço de integração. Note que contamos os pontos sorteados e aqueles
# abaixo da curva da função, produzindo listas dos valores correspondentes.
k = 0
x_acima, x_abaixo, y_acima, y_abaixo = [], [], [], []
for n in range(N):
    x = a+(b-a)*random() # Sorteamos a coordenada x do ponto
    y = h*random()       # Sorteamos a coordenada y do ponto
    if y < f(x):          # O ponto sorteado está abaixo da curva da função?
        k += 1           # Atualizamos o contador desses pontos
        x_abaixo.append(x)
        y_abaixo.append(y)
    else:                 # O ponto sorteado está acima da curva da função?
        x_acima.append(x)
        y_acima.append(y)

integral = k/N*(b-a)*h   # Estimativa para a integral
print("A estimativa para pi é ",2*integral)
```

No programa acima são criadas listas dos pontos acima e abaixo da função apenas para efeito de exibição em gráficos.

Exemplo 1a

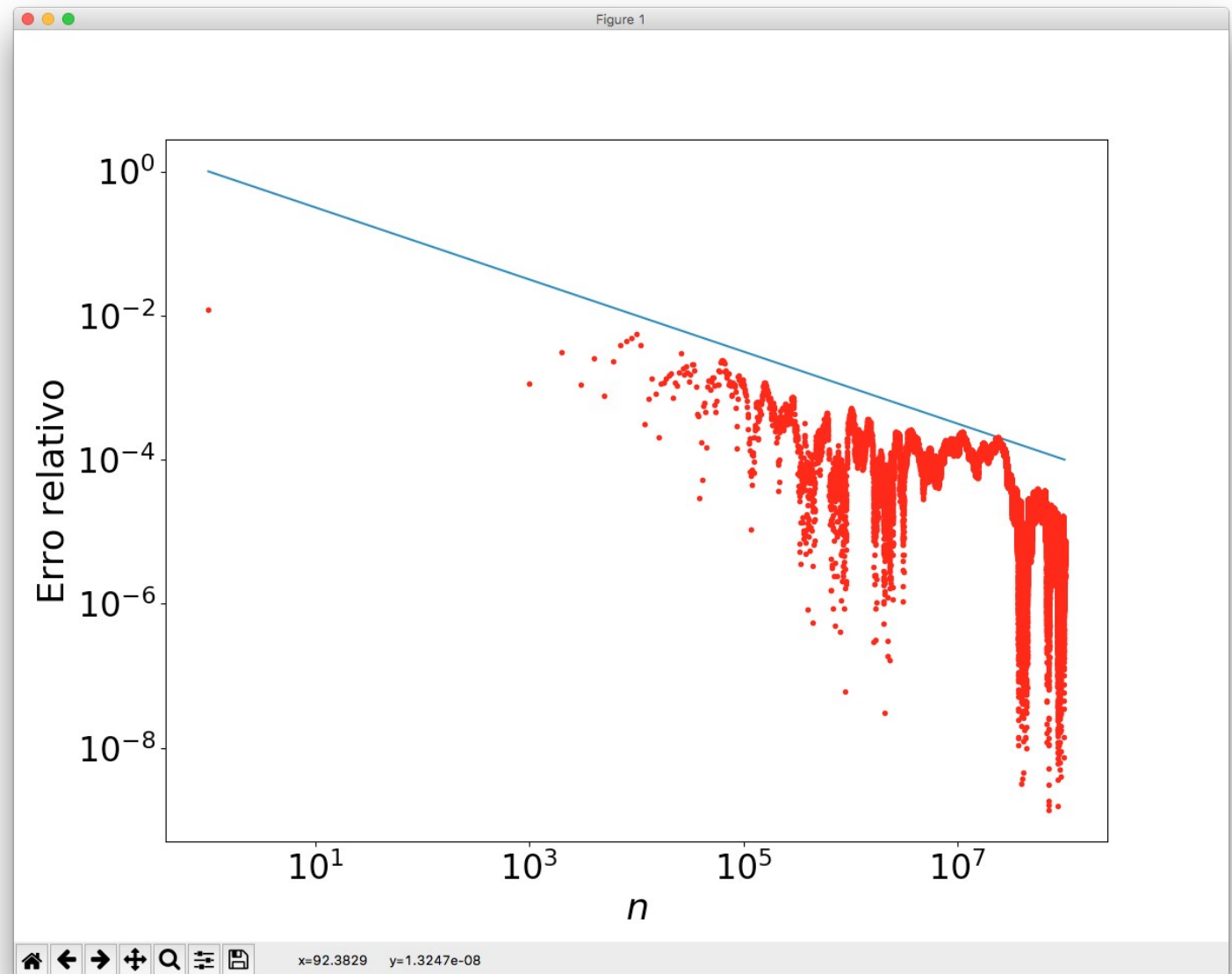
$$I_N \approx \frac{k}{N} A$$



Os pontos em azul (preto) foram sorteados abaixo (acima) da curva da função.

Exemplo 1b

$$\text{erro}(\pi_n) = |2I_n - \pi|$$



Os pontos em **vermelho** vêm da integração por Monte Carlo, e a curva em **azul** é o inverso da raiz quadrada de n .

Estimando o erro

- A probabilidade de que um ponto sorteado caia abaixo da curva da função é $p=I/A$, enquanto a probabilidade de que exatamente k entre N pontos caiam abaixo da função é

$$P_N(k) = \binom{N}{k} p^k (1-p)^{N-k}.$$

- O valor esperado de k segundo essa distribuição é Np , enquanto a variância é

$$\text{var } k = \overline{(k - \bar{k})^2} = Np(1-p).$$

Interlúdio: deduções do slide anterior

- Substituindo $1-p$ por q , temos formalmente

$$\begin{aligned}\bar{k} &= \sum_{k=0}^N k P_N(k) = \lim_{q \rightarrow 1-p} \left\{ \sum_{k=0}^N k \binom{N}{k} p^k q^{N-k} \right\} \\ &= \lim_{q \rightarrow 1-p} \left\{ p \frac{\partial}{\partial p} \sum_{k=0}^N \binom{N}{k} p^k q^{N-k} \right\} = \lim_{q \rightarrow 1-p} \left\{ p \frac{\partial}{\partial p} (p+q)^N \right\} \\ &= \lim_{q \rightarrow 1-p} \left\{ Np (p+q)^{N-1} \right\} = Np\end{aligned}$$

- Para calcular a variância, notemos primeiro que

$$\text{var } k = \overline{(k - \bar{k})^2} = \overline{k^2 - 2k\bar{k} + \bar{k}^2} = \overline{k^2} - \bar{k}^2.$$

Interlúdio: deduções do slide anterior

- Substituindo $1-p$ por q , temos formalmente

$$\begin{aligned}\overline{k^2} &= \sum_{k=0}^N k^2 P_N(k) = \lim_{q \rightarrow 1-p} \left\{ \sum_{k=0}^N k^2 \binom{N}{k} p^k q^{N-k} \right\} \\ &= \lim_{q \rightarrow 1-p} \left\{ p \frac{\partial}{\partial p} \left[p \frac{\partial}{\partial p} \sum_{k=0}^N \binom{N}{k} p^k q^{N-k} \right] \right\} \\ &= \lim_{q \rightarrow 1-p} \left\{ p \frac{\partial}{\partial p} [Np(p+q)^{N-1}] \right\} \\ &= \lim_{q \rightarrow 1-p} \left\{ N(N-1)p^2(p+q)^{N-2} + Np(p+q)^{N-1} \right\} \\ &= N(N-1)p^2 + Np = N^2 p^2 + Np(1-p)\end{aligned}$$

Interlúdio: deduções do slide anterior

- Finalmente,

$$\begin{aligned}\text{var } k &= \overline{k^2} - \bar{k}^2 = N^2 p^2 + Np(1-p) - N^2 p^2 \\ &= Np(1-p)\end{aligned}$$

Estimando o erro

- A raiz quadrada da variância de k fornece uma estimativa para o erro em k , de modo que uma estimativa para o erro na integral é

$$\sigma \equiv \text{erro}(I_N) = \text{erro}\left(k \frac{A}{N}\right) = \sqrt{\text{var } k} \frac{A}{N} = \sqrt{Np(1-p)} \frac{A}{N}.$$

- Lembrando que $p=I/A$, obtemos por fim

$$\sigma = \frac{\sqrt{I(A-I)}}{\sqrt{N}}.$$

Note que a dependência em N concorda com o que observamos no slide do exemplo 1b.

O método do valor médio

- Um outro método de integração por Monte Carlo tira proveito da definição de valor médio de uma função. Entre $x=a$ e $x=b$ o valor médio da função $f(x)$ é definido como

$$\langle f \rangle = \frac{1}{(b-a)} \int_a^b f(x) dx = \frac{I}{(b-a)},$$

de modo que podemos estimar o valor da integral estimando o valor médio da função:

$$\langle f \rangle \simeq f_N = \frac{1}{N} \sum_{i=1}^N f(x_i) \Rightarrow I \simeq (b-a) f_N,$$

sendo x_i sorteados uniformemente entre a e b .

O método do valor médio

- O erro associado a esse método é proporcional ao erro em f_N . A variância de f_N é

$$\text{var } f_N = \text{var} \left\{ \frac{1}{N} \sum_{i=1}^N f(x_i) \right\}.$$

Agora, se x é uma variável aleatória e c é uma constante,

$$\text{var} \{c x\} = \overline{(c x)^2} - \overline{c x}^2 = c^2 \text{var } x$$

enquanto

$$\text{var} \left\{ \sum_{i=1}^N x_i \right\} = \overline{\left[\left(\sum_{i=1}^N x_i \right) - N \bar{x} \right]^2} = \overline{\left[\sum_{i=1}^N (x_i - \bar{x}) \right]^2}.$$

O método do valor médio

- Mas

$$\begin{aligned} \overline{\left[\sum_{i=1}^N (x_i - \bar{x}) \right]^2} &= \overline{\left[\sum_{i=1}^N (x_i - \bar{x}) \right]} \overline{\left[\sum_{j=1}^N (x_j - \bar{x}) \right]} \\ &= \overline{\sum_{i=1}^N (x_i - \bar{x})^2} + \overline{\sum_{i=1}^N \sum_{j \neq i}^N (x_i - \bar{x})(x_j - \bar{x})} \end{aligned}$$

- Agora, como os x_i são independentes entre si,

$$\overline{\sum_{i=1}^N \sum_{j \neq i}^N (x_i - \bar{x})(x_j - \bar{x})} = \sum_{i=1}^N \sum_{j \neq i}^N \overline{(x_i - \bar{x})(x_j - \bar{x})} = 0$$

O método do valor médio

- Portanto,

$$\overline{\left[\sum_{i=1}^N (x_i - \bar{x}) \right]^2} = \overline{\sum_{i=1}^N (x_i - \bar{x})^2},$$

de modo que

$$\text{var} \left\{ \sum_{i=1}^N x_i \right\} = \overline{\left[\sum_{i=1}^N (x_i - \bar{x}) \right]^2} = \overline{\sum_{i=1}^N (x_i - \bar{x})^2} = N \text{ var } x.$$

O método do valor médio

- Combinando os resultados dos dois últimos slides, concluimos que

$$\text{var } f_N = \text{var} \left\{ \frac{1}{N} \sum_{i=1}^N f(x_i) \right\} = \frac{1}{N^2} N \text{var } f = \frac{1}{N} \text{var } f.$$

- Como uma estimativa do erro é a raiz quadrada da variância, temos para a integral um erro

$$\sigma \equiv \text{erro}(I_N) = \text{erro}((b-a)f_N) = (b-a) \frac{\sqrt{\text{var } f}}{\sqrt{N}},$$

novamente proporcional ao inverso da raiz quadrada de N , embora com um prefator menor do que com o método anterior. (Ver Newman.)

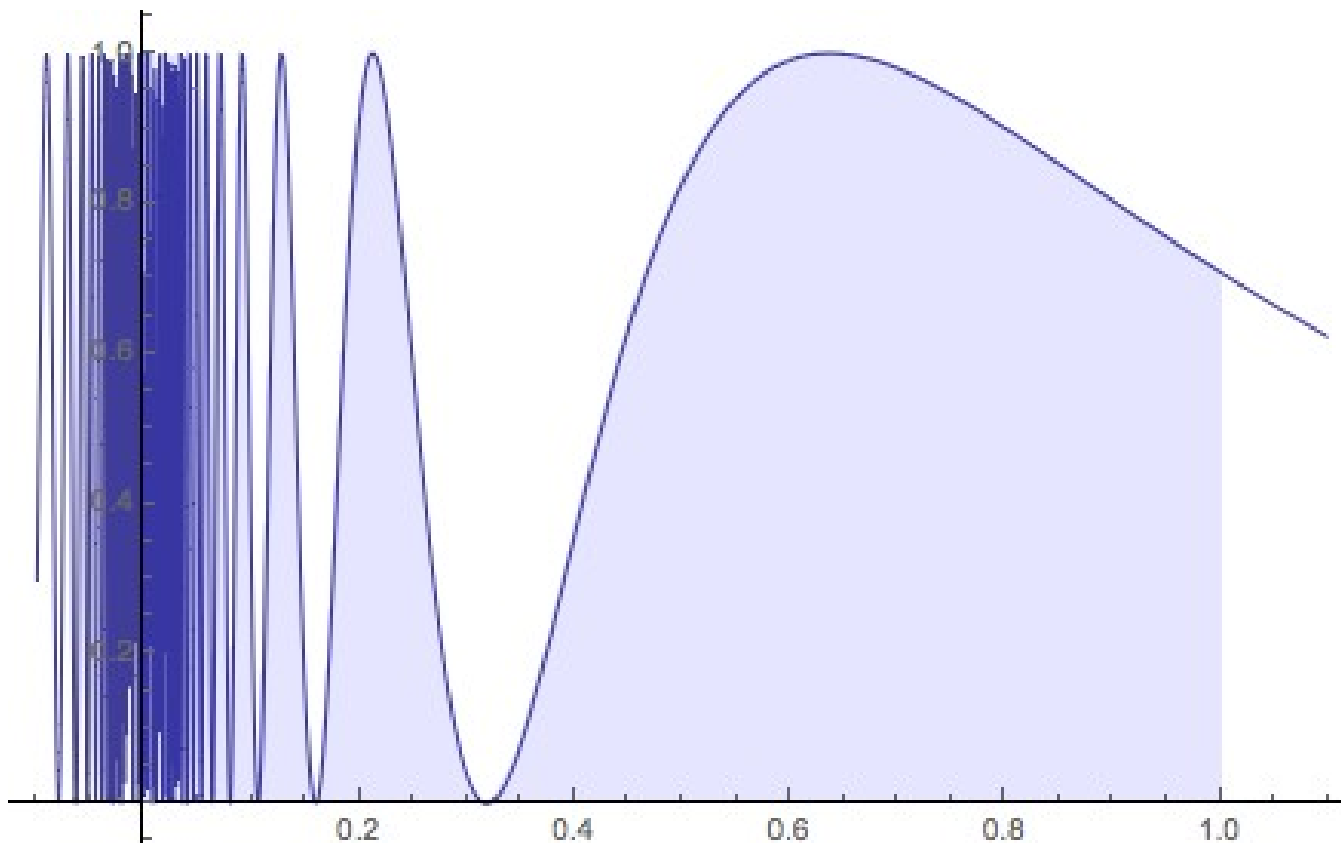
Por que integrar por Monte Carlo?

- Um método com um erro que decai apenas com o inverso da raiz quadrada do número de operações é bastante ineficiente em comparação com as técnicas aprendidas em Introdução à Física Computacional I. Por que utilizá-lo?
 - 1) Há funções que variam muito rapidamente, de tal forma que os métodos usuais requerem um grande número de passos.
 - 2) Para integrais múltiplas de funções bem comportadas, os métodos mais eficientes com integrais simples tornam-se proibitivos.

Exemplo 2

- Considere a integral

$$I = \int_0^1 \sin^2\left(\frac{1}{x}\right) dx.$$



Exemplo 2

- Calculando a integral utilizando a regra do trapézio, o método de “jogar pedras” e o método do valor médio, sempre com $N=10^6$ pontos, os resultados são os mostrados abaixo.

O valor exato da integral entre $x=0$ e $x=1$ é 0.673457

A estimativa para a integral pela regra do trapézio é 0.6734672687672397
O tempo de cálculo em segundos foi 0.8062820434570312

A estimativa para a integral pelo método de jogar pedras é 0.6738209999326179
O tempo de cálculo em segundos foi 1.0446670055389404

A estimativa para a integral pelo método do valor médio é 0.6734217886854258
O tempo de cálculo em segundos foi 0.8824360370635986

- Note que os tempos de execução são comparáveis e que o método do ponto médio é um pouco mais preciso. O método do trapézio ainda é, contudo, o mais preciso.

O método do valor médio em muitas dimensões

- Para integrais múltiplas, o método do valor médio é generalizado para

$$\langle f \rangle \simeq f_N = \frac{1}{N} \sum_{i=1}^N f(\vec{r}_i) \Rightarrow I \simeq V f_N,$$

sendo V o “volume” de integração, a generalização multidimensional do intervalo de integração em integrais simples.

Exemplo 3

- O volume de uma hiperesfera de raio unitário em 10 dimensões é aquele delimitado pela hipersuperfície

$$\sum_{i=1}^{10} x_i^2 = 1,$$

em que supomos que a hiperesfera é centrada na origem e um ponto de sua hipersuperfície é descrito pelo vetor

$$\vec{r} = \sum_{i=1}^{10} \vec{e}_i x_i.$$

Exemplo 3

- Tratando x_{10} como função das demais 9 coordenadas, e lembrando que um dos 2^{10} setores do espaço de 10 dimensões é aquele em que as coordenadas são todas positivas, o volume desejado é

$$\text{Volume} = 2^{10} \int_0^1 dx_1 \int_0^{\sqrt{1-x_1^2}} dx_2 \cdots \int_0^{\sqrt{1-\sum_{i=1}^8 x_i^2}} dx_9 \sqrt{1-\sum_{i=1}^9 x_i^2}.$$

- O cálculo usual exigiria dividir o hiperespaço em pequenos hipervolumes e calcular o integrando em cada um. Se em cada dimensão utilizarmos 100 divisões, teríamos 10^{18} hipervolumes.

Exemplo 3

- Tratando x_{10} como função das demais 9 coordenadas, e lembrando que um dos 2^{10} setores do espaço de 10 dimensões é aquele em que as coordenadas são todas positivas, o volume desejado é

$$\text{Volume} = 2^{10} \int_0^1 dx_1 \int_0^{\sqrt{1-x_1^2}} dx_2 \cdots \int_0^{\sqrt{1-\sum_{i=1}^8 x_i^2}} dx_9 \sqrt{1-\sum_{i=1}^9 x_i^2}.$$

- Se a cada segundo pudéssemos realizar 1 bilhão de operações, demoraríamos ainda 10^9 segundos para chegar a uma estimativa do volume da hiperesfera.

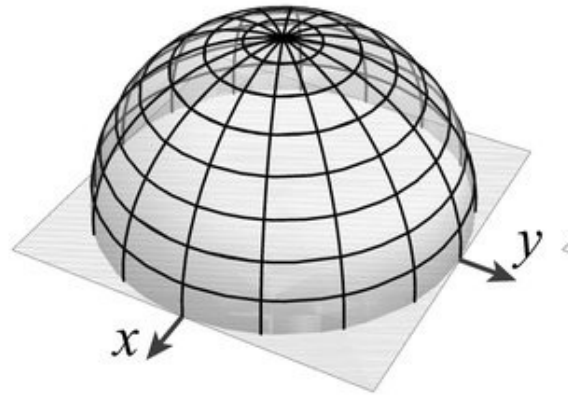
Exemplo 3

- Tratando x_{10} como função das demais 9 coordenadas, e lembrando que um dos 2^{10} setores do espaço de 10 dimensões é aquele em que as coordenadas são todas positivas, o volume desejado é

$$\text{Volume} = 2^{10} \int_0^1 dx_1 \int_0^{\sqrt{1-x_1^2}} dx_2 \cdots \int_0^{\sqrt{1-\sum_{i=1}^8 x_i^2}} dx_9 \sqrt{1-\sum_{i=1}^9 x_i^2}.$$

- Em vez de aguardar um século pelo resultado, vamos utilizar o método do valor médio, escolhendo 10^7 valores das 9 coordenadas.

Exemplo 3



Se estivéssemos trabalhando em 3D, o intervalo de integração corresponderia aos pontos do plano xy no primeiro quadrante e no interior da calota. Podemos contudo utilizar todo o primeiro quadrante como intervalo de integração se definirmos o integrando como nulo fora da calota. Procedemos de forma análoga em dimensões superiores.

Exemplo 3

```
from math import sqrt
from random import random
from numpy import empty
import time

tempo_inicial = time.time()

# Parâmetros de integração
ndim = 10      # Número de dimensões da hiperesfera
N = 10**7      # Número de pontos a utilizar para a integração

# Definição da função a ser integrada
def f(r):
    soma = 0.0
    for i in range(ndim-1):
        soma += r[i]**2
    if soma >= 1.0:
        return 0.0      # A função é nula fora da "calota"
    else:
        return sqrt(1-soma)

# Laço de integração pelo método do valor médio.
# Note que a função a ser integrada retorna zero se o ponto sorteado
# em ndim-1 dimensões está fora da "calota", por isso podemos utilizar
# como intervalo de integração todo o "hiperquadrante".
r = empty(ndim-1)
f_medio = 0.0
for n in range(1,N+1):
    for i in range(ndim-1):
        r[i] = random()    # Sorteamos cada coordenada do ponto
    f_medio += f(r)
f_medio /= N
integral = f_medio # Estimativa para a integral
hipervolume = (2**ndim)*integral # Multiplicamos pelo número de "hiperquadrantes"
print("A estimativa para o volume pelo método do valor médio é ",hipervolume)
```

O resultado exato é 2.5502, contra uma estimativa de 2.5479 obtida em 117 s.

Amostragem por importância

- Se o integrando possui divergências, o método do valor médio é problemático, uma vez que pode haver enormes variações de resultado entre diferentes tentativas de estimar a integral.
- Considere por exemplo a integral

$$I = \int_0^1 \frac{x^{-1/2}}{e^x + 1} dx,$$

que é de interesse na teoria de gases de Fermi. Estimá-la pelo método do valor médio utilizando 100 pontos gera resultados como 0.8137, 0.8249 e 1.0479, contra um resultado exato de 0.8389.

Amostragem por importância

- Para evitar grandes flutuações entre uma estimativa e outra, utiliza-se o método de amostragem por importância. A ideia é escolher os pontos x_i não uniformemente, mas segundo uma distribuição de probabilidades que elimine a divergência no integrando.
- Vamos definir a média ponderada por $w(x)$ de uma função $g(x)$:

$$\langle g \rangle_w = \frac{\int_a^b w(x) g(x) dx}{\int_a^b w(x) dx}.$$

Amostragem por importância

- Fazendo $g(x) = f(x)/w(x)$, temos

$$\left\langle \frac{f(x)}{w(x)} \right\rangle_w = \frac{\int_a^b w(x) f(x)/w(x) dx}{\int_a^b w(x) dx} = \frac{\int_a^b f(x) dx}{\int_a^b w(x) dx},$$

de modo que a integral de $f(x)$ é dada por

$$\int_a^b f(x) dx = \left\langle \frac{f(x)}{w(x)} \right\rangle_w \int_a^b w(x) dx,$$

e podemos estimar a integral se estimarmos a média ponderada de $f(x)/w(x)$.

Amostragem por importância

- Para estimar a média ponderada, supomos $w(x) \geq 0$ e definimos a distribuição de probabilidades

$$p(x) = \frac{w(x)}{\int_a^b w(x) dx}, \quad \int_a^b p(x) dx = 1,$$

e amostramos N pontos X_i a partir dessa distribuição. O número médio aproximado de amostras no intervalo entre $x_j = a + j\Delta x$ e $x_j + \Delta x$ é $Np(x_j)\Delta x$, com $\Delta x = (b-a)/M$, de modo que

$$\sum_{i=1}^N g(X_i) \simeq \sum_{j=1}^M g(x_j) Np(x_j) \Delta x \simeq N \int_a^b p(x) g(x) dx.$$

Amostragem por importância

- A média ponderada será

$$\langle g \rangle_w = \frac{\int_a^b w(x) g(x) dx}{\int_a^b w(x) dx} = \int_a^b p(x) g(x) dx \simeq \frac{1}{N} \sum_{i=1}^N g(X_i).$$

e portanto temos

$$\int_a^b f(x) dx = \left\langle \frac{f(x)}{w(x)} \right\rangle_w \int_a^b w(x) dx \simeq \left[\frac{1}{N} \sum_{i=1}^N \frac{f(X_i)}{w(X_i)} \right] \int_a^b w(x) dx,$$

com os X_i amostrados da distribuição $p(x)$.

Exemplo 4

- Vamos utilizar a amostragem por importância para calcular a integral

$$I = \int_0^1 \frac{x^{-1/2}}{e^x + 1} dx$$

eliminando a divergência na origem. Para isso, vamos escolher $w(x) = x^{-1/2}$, de modo que

$$I = \left\langle \frac{1}{e^x + 1} \right\rangle_w \int_0^1 x^{-1/2} dx \simeq \left[\frac{1}{N} \sum_{i=1}^N \frac{1}{e^{x_i} + 1} \right] \times 2,$$

com x_i amostrado a partir de

$$p(x) = \frac{1}{2} x^{-1/2} \quad (0 \leq x < 1).$$

Exemplo 4

- Para realizar a amostragem, vamos utilizar o método da transformação. Para isso, amostramos z uniformemente entre 0 e 1 e resolvemos a equação

$$\int_0^x p(x') dx' = x^{1/2} = z,$$

ou seja, fazemos

$$x = z^2.$$

Exemplo 4

```
from math import exp,sqrt
from random import random

# Parâmetros de integração
N = 10**2      # Número de pontos a utilizar para a integração

# Definição da função cujo valor médio será calculado.
# Note que aqui se trata da função de integração dividida por  $w(x)=x^{(-1/2)}$ 
def f(x):
    return 1/(exp(x)+1)

print("O valor exato da integral entre x=0 e x=1 é",0.83893296)
print("")

# Laço de integração pelo método do valor médio
# com amostragem por importância
fsw_medio = 0.0
intw = 2.0 # Integral de  $w(x)=x^{(-1/2)}$  entre x=0 e x=1
for n in range(1,N+1):
    x = random()*2 # Sorteamos a coordenada x do ponto não uniformemente
    fsw_medio += f(x)
fsw_medio /= N
integral = fsw_medio*intw # Estimativa para a integral
print("A estimativa para a integral pelo método do valor médio é ",integral)
print("")
```

Agora obtemos, na maioria das execuções, valores entre 0.83 e 0.85 para a estimativa da integral, cujo valor exato é 0.838933.

Amostragem por importância

- Voltaremos a discutir amostragem por importância no contexto das simulações de Monte Carlo em mecânica estatística, a partir da próxima aula.

Exercício no moodle

Há um único exercício, explorando o conteúdo da aula de hoje, e que pode ser feito com base nos programas dos exemplos desta aula e da aula anterior, disponíveis no moodle. O exercício deve ser entregue até o dia **27 de maio**.

O EP4 está liberado, e deve ser entregue até o início da aula do dia **3 de junho**.