

FUNDAÇÃO GETULIO VARGAS  
ESCOLA DE MATEMÁTICA APLICADA  
CURSO DE GRADUAÇÃO EM  
MATEMÁTICA APLICADA

**Métodos de Simulação de Monte Carlo para  
Equações Diferenciais Estocásticas**

por

Arthur da Silva Pereira Carneiro

Rio de Janeiro  
2016

FUNDAÇÃO GETULIO  
VARGAS

FUNDAÇÃO GETULIO VARGAS  
ESCOLA DE MATEMÁTICA APLICADA  
CURSO DE GRADUAÇÃO EM  
MATEMÁTICA APLICADA

Métodos de Simulação de Monte Carlo para  
Equações Diferenciais Estocásticas

**”Declaro ser o único autor do presente projeto de monografia que refere-se ao plano de trabalho a ser executado para continuidade da monografia e ressalto que não recorri a qualquer forma de colaboração ou auxílio de terceiros para realizá-lo a não ser nos casos e para os fins autorizados pelo professor orientador”**

---

**Arthur da Silva Pereira Carneiro**

**Orientador: Hugo A. de la Cruz Cancino**

Rio de Janeiro  
2016

**Arthur da Silva Pereira Carneiro**

**Métodos de Simulação de Monte Carlo para  
Equações Diferenciais Estocásticas**

“Monografia apresentada à Escola de Matemática Aplicada  
como requisito parcial para obtenção do grau de Bacharel  
em Matemática Aplicada”

Aprovado em \_\_\_\_ de \_\_\_\_\_ de \_\_\_\_ .  
Grau atribuído ao Projeto de Monografia: \_\_\_\_ .

---

**Professor Orientador: Hugo A. de la Cruz Cancino**  
**Escola de Matemática Aplicada**  
**Fundação Getulio Vargas**



# Sumário

<b>Resumo</b>	<b>3</b>
<b>1 Introdução</b>	<b>3</b>
1.1 Objetivo do estudo . . . . .	3
<b>2 Preliminares teóricas</b>	<b>4</b>
2.1 Teoria de Probabilidade . . . . .	4
2.1.1 Espaço de Probabilidade . . . . .	4
2.1.2 Variáveis aleatórias . . . . .	5
2.1.3 Independência . . . . .	5
2.1.4 Valor Esperado . . . . .	6
2.1.5 Inequações e convergências - Lei dos Grandes Números . . . . .	7
2.1.6 Processos Estocásticos . . . . .	8
2.2 Cálculo Estocástico . . . . .	9
2.2.1 Integração Estocástica . . . . .	9
2.2.2 Equação Diferencial Estocástica . . . . .	10
2.2.3 Fórmula de Itô . . . . .	10
<b>3 Simulações de Monte Carlo</b>	<b>11</b>
3.1 Apresentação de Monte Carlo . . . . .	11
3.2 Redução de Variância . . . . .	11
3.2.1 Variável Antitética . . . . .	12
3.2.2 Variáveis de Controle . . . . .	13
3.3 Monte Carlo Multinível . . . . .	13
3.4 Resultados numéricos . . . . .	16
3.4.1 Movimento Browniano Geometrico . . . . .	16
3.4.2 Modelo de Heston de volatilidade estocástica . . . . .	17
<b>4 Conclusão</b>	<b>19</b>
<b>5 Apêndice</b>	<b>20</b>
<b>6 Referências Bibliográficas</b>	<b>26</b>

# **1 Introdução**

## **1.1 Objetivo do estudo**

Neste trabalho apresentamos técnicas de simulação de Monte Carlo (MC), para o cálculo de funcionais de soluções de Equações Diferenciais Estocásticas (EDE's). Estudaremos também a nova metodologia Multilevel Monte Carlo (MLMC), recentemente propostas na literatura, que permite uma implementação mais eficiente.

Para cumprir o nosso objetivo, primeiramente vários aspectos avançados da teoria de probabilidade e processos estocásticos serão apresentados. Isto inclui: aspectos relevantes de probabilidade com fundamentos na teoria da medida e integração, espaços medíveis de funções e processos aleatórios, integração estocástica de Itô e fundamentos da teoria de EDE's, incluindo à análise de métodos de integração e simulação computacional de EDE's.

Alguns experimentos numéricos serão realizados comparando as diferentes técnicas a serem analisadas neste projeto.

## 2 Preliminares teóricas

### 2.1 Teoria de Probabilidade

#### 2.1.1 Espaço de Probabilidade

Define-se uma tripla de probabilidade ou espaço medido de probabilidade ou simplesmente espaço de probabilidade a tripla  $(\Omega, \mathcal{F}, \mathcal{P})$ , onde:

- O espaço amostral  $\Omega$  é um conjunto não vazio;
- A  $\sigma$ -álgebra  $\mathcal{F}$  é uma coleção de subconjuntos de  $\Omega$ , contendo o próprio  $\Omega$  e o conjunto vazio  $\emptyset$ , além disso é fechado por complementos -  $F \in \mathcal{F} \rightarrow (\Omega \setminus F) \in \mathcal{F}$  - e por enumeráveis uniões -  $\{F_i\}_{i=1}^{\infty} \in \mathcal{F} \rightarrow (\cup_{i=1}^{\infty} F_i) \in \mathcal{F}$ .
- A medida de probabilidade  $\mathcal{P}$  é um mapeamento de  $\mathcal{F}$  em  $[0, 1]$ , com  $\mathcal{P}(\emptyset) = 0$  e  $\mathcal{P}(\Omega) = 1$ , tal que  $\mathcal{P}$  é aditivamente contável -  $F_1, F_2, \dots$  subconjuntos disjuntos de  $\mathcal{F} \rightarrow \mathcal{P}(\cup_{i=1}^{\infty} F_i) = (\sum_{i=1}^{\infty} \mathcal{P}(F_i))$

A  $\sigma$ -álgebra  $\mathcal{F}$  é a coleção de todos os eventos. Podemos utilizar a  $\sigma$ -álgebra  $\mathcal{F}$  como o conjunto de informações de  $\Omega$  dado que  $\mathcal{F}$  talvez não contenha todos os subconjuntos de  $\Omega$ , ou seja, esperamos que contenha apenas os que representam de forma geral as propriedades do espaço amostral.

Para espaços de probabilidade discretos, isto é, com o espaço amostral finito ou enumerável, sempre conseguimos a tripla de probabilidade pelo teorema a seguir:

**Teorema:** Seja  $\Omega$  um conjunto enumerável não vazio. Seja  $p : \Omega \rightarrow [0, 1]$  uma função satisfazendo  $\sum_{\omega \in \Omega} p(\omega) = 1$ . Então, existe uma tripla de probabilidade  $(\Omega, \mathcal{F}, \mathcal{P})$ , onde  $\mathcal{F}$  é a coleção de todos os subconjuntos de  $\Omega$  e para  $F \in \mathcal{F}$ ,  $\mathcal{P}(F) = \sum_{\omega \in F} p(\omega)$ .

Para o caso em que  $\Omega$  é não vazio e finito, podemos utilizar  $\mathcal{F} = \frac{|F|}{|\Omega|}$ ,  $\forall F \in \mathcal{F}$ , sendo  $\mathcal{F}$  a coleção de todos os subconjuntos de  $\Omega$ . Dessa maneira, obtemos  $(\Omega, \mathcal{F}, \mathcal{P})$  conhecida como distribuição uniforme em  $\Omega$  - Uniforme( $\Omega$ ).

Para expandir a construção para espaços amostrais mais complicados precisamos de algumas outras definições e teoremas.

**Definição:**  $\mathcal{J}$  é uma semi-álgebra de subconjuntos de  $\Omega$ , se  $\Omega \in \mathcal{J}$  e  $\emptyset \in \mathcal{J}$ , além disso  $\mathcal{J}$  é fechado em interseções finitas e o complemento de qualquer elemento de  $\mathcal{J}$  é igual a união finita de elementos de  $\mathcal{J}$ .

**Teorema da Extensão:** Seja  $\mathcal{J}$  uma semi-álgebra de subconjuntos de  $\Omega$ . Seja  $\mathcal{P} : \mathcal{J} \rightarrow [0, 1]$  com  $\mathcal{P}(\emptyset) = 0$  e  $\mathcal{P}(\Omega) = 1$ , satisfazendo a propriedade da superaditividade finita que

$$\mathcal{P}(\cup_{i=1}^k A_i) \geq \sum_{i=1}^k \mathcal{P}(A_i) \text{ sempre que } A_1, \dots, A_k \in \mathcal{J}, \cup_{i=1}^k A_i \in \mathcal{J} \text{ e } A_i \cap A_j = \emptyset, \forall i \neq j \quad (1)$$

e também tem a propriedade de monotocidade enumerável que

$$\mathcal{P}(A) \leq \sum_n \mathcal{P}(A_n) \text{ para } A_1, A_2, \dots \in \mathcal{J} \text{ com } A \subseteq \cup_n A_n. \quad (2)$$

Então existe uma  $\sigma$ -álgebra  $\mathcal{M} \supseteq \mathcal{J}$ , e uma probabilidade aditiva mensurável  $\mathcal{P}^*$  em  $\mathcal{M}$ , tal que  $\mathcal{P}^*(A) = \mathcal{P}(A)$ ,  $\forall A \in \mathcal{J}$ .

Um exemplo de construção é para a distribuição **Uniforme**([0,1]) com  $\mathcal{J} = \{\text{todos os intervalos contidos em } [0, 1]\}$  e o seguinte teorema:

**Teorema:** Existe uma tripla de probabilidade  $(\Omega, \mathcal{M}, \mathcal{P}^*)$ , tal que  $\Omega = [0, 1]$ ,  $\mathcal{M}$  contém todos os intervalos de  $[0, 1]$  e para qualquer intervalo  $I \subseteq [0, 1]$ ,  $\mathcal{P}^*$  é a largura de  $I$ .

Essa tripla de probabilidade é chamada tanto de distribuição uniforme de  $[0, 1]$  quanto medida de Lebesgue em  $[0, 1]$ . Além disso, temos que  $\mathcal{B} = \sigma(\mathcal{J})$  a  $\sigma$ -álgebra gerada por  $\mathcal{J}$ , ou seja, a menor  $\sigma$ -álgebra contendo  $\mathcal{J}$ , onde  $\mathcal{B}$  é chamado de Borel  $\sigma$ -álgebra de subconjuntos de  $[0, 1]$ .

### 2.1.2 Variáveis aleatórias

Intuitivamente, variável aleatória é a relação entre um evento do espaço amostral e um número real. Expressando isso na linguagem mais formal:

**Definição:** Dada uma tripla de probabilidade  $(\Omega, \mathcal{F}, \mathcal{P})$ , uma variável aleatória é uma função  $X$  de  $\Omega$  nos números reais  $R$  tal que

$$\{\omega \in \Omega; X(\omega) \leq x\} \in \mathcal{F}, \quad x \in R \quad (3)$$

Pode-se escrever também  $\{X \leq x\} \in \mathcal{F}$ , ou  $X^{-1}((-\infty, x]) \in \mathcal{F}$ ,  $\forall x \in R$ , ou seja, uma variável aleatória é uma função com a propriedade que a imagem inversa de subintervalos de  $R$  está na  $\sigma$ -álgebra  $\mathcal{F}$ . E isso é equivalente a  $X^{-1}(\mathcal{B}) \in \mathcal{F}$  para todos conjunto de Borel  $\mathcal{B}$ . Além disso, nomea-se  $\mathcal{B}(R)$  como Borelianos da reta.

### 2.1.3 Independência

Eventos são independentes quando eles não afetam a probabilidade de ocorrência um do outro. Dada essa intuição, uma formalização é:

**Definição** Um conjunto infinito  $\{A_\alpha\}_{\alpha \in I}$  de eventos é dito independente se para cada  $i \in \mathbb{N}$  e para cada sequência finita distinta  $\alpha_1, \dots, \alpha_i$ , temos

$$P(A_{\alpha_1} \cap \dots \cap A_{\alpha_j}) = P(A_{\alpha_1}) \cap \dots \cap P(A_{\alpha_j}) \quad (4)$$

Com a definição temos algumas proposições:

**Proposição** Sejam  $X$  e  $Y$  variáveis aleatórias independentes, e  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  funções Borel-mensuráveis. Então, as variáveis aleatórias  $f(X)$  e  $g(Y)$  são independentes.



**Proposição** Sejam  $X$  e  $Y$  variáveis aleatórias definidas em uma tripla de probabilidade  $(\Omega, \mathcal{F}, \mathcal{P})$ . Então,  $X$  e  $Y$  são independentes se, e somente se,  $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$ ,  $\forall x, y \in \mathbb{R}$ .

#### 2.1.4 Valor Esperado

Seja  $X$  uma variável aleatória definida na tripla de probabilidade  $(\Omega, \mathcal{F}, \mathcal{P})$ .

**Definição** A variável aleatória  $X$  é dita simples se  $\{X(\omega); \omega \in \Omega\}$  é finito, ou seja, ela assume um número finito de valores.

Assumindo  $X$  simples, temos os diferentes elementos  $\{x_1, \dots, x_n\}$  e assim,  $X = \sum_{i=1}^n x_i 1_{A_i}$  com  $A_i = \{\omega \in \Omega; X(\omega) = x_i\}$ .

**Definição:** Para uma variável aleatória  $X$  simples, define-se o valor esperado de  $X$

$$\mathbb{E}(X) = \sum_{i=1}^n x_i \mathcal{P}(A_i) \quad (5)$$

**Definição:** Define-se a variância de  $X$

$$\text{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) \quad (6)$$

Para um caso mais geral, seja  $X$  uma variável aleatória não-negativa e não simples:

**Definição:** Para uma variável aleatória  $X$  não-negativas, define-se o valor esperado de  $X$

$$\mathbb{E}(X) = \sup\{\mathbb{E}(Y); Y \text{ simples}, Y \leq X\} \quad (7)$$

Dado essa definição conseguimos observar o seguinte teorema

**Teorema da Convergência Monótona:** Supondo  $X_1, X_2, \dots$  variáveis aleatórias com  $\mathbb{E}(X_1) > -\infty$ , e  $X_1 \leq X_2 \leq \dots$  e  $\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)$ ,  $\forall \omega \in \Omega$  - sequência monotamente convergente. Então,  $X$  é uma variável aleatória e  $\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \mathbb{E}(X)$ .

Para o caso de uma variável aleatória arbitrária  $X$ , podemos reescrevê-la como  $X = X^+ - X^-$ ,  $X^+$ ,  $X^-$  são não negativos e assim podemos aplicar de forma mais geral o conceito de valor esperado:

$$\mathbb{E}(X) = \mathbb{E}(X^+) - \mathbb{E}(X^-) \quad (8)$$

---

<sup>A</sup> $X^+(\omega) = \max\{X(\omega), 0\}$  e  $X^-(\omega) = \max\{-X(\omega), 0\}$

### 2.1.5 Inequações e convergências - Lei dos Grandes Números

Importantes inequações envolvendo os conceitos já apresentados:

**Desigualdade de Markov:** Se  $X$  é uma variável aleatória não-negativa, então para todo  $\alpha > 0$ ,

$$P(X \geq \alpha) \leq \frac{\mathbb{E}(X)}{\alpha} \quad (9)$$

**Desigualdade de Chebyshev:** Seja  $X$  uma variável aleatória, com média finita  $\mu_x$ . Então para todo  $\alpha > 0$ ,

$$P(|X - \mu_x| \geq \alpha) \leq \frac{\text{Var}(X)}{\alpha^2} \quad (10)$$

**Desigualdade de Cauchy-Schwarz:** Seja  $X$  e  $Y$  variáveis aleatórias com  $\mathbb{E}(X) < \infty$  e  $\mathbb{E}(Y) < \infty$ . Então  $\mathbb{E}|XY| \leq \sqrt{\mathbb{E}(Y^2)\mathbb{E}(X^2)}$ .

**Desigualdade de Jensen:** Seja  $X$  uma variável aleatória com média finita e seja  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  uma função convexa<sup>B</sup>. Então  $\mathbb{E}(\phi(X)) \geq \phi(\mathbb{E}(X))$ .

Há alguns tipos de convergência, no caso envolvendo variáveis aleatórias as seguintes definições são usadas:

**Convergência em probabilidade:** Seja  $\{X_1, X_2, \dots\}$  uma sequência de variáveis aleatórias e  $X$  uma variável aleatória definidas na mesma tripla de probabilidade. Dizemos que  $X_n$  converge em probabilidade para  $X$  se para todo  $\varepsilon > 0$ ,  $\lim_{n \rightarrow \infty} \mathcal{P}(|X_n - X| \geq \varepsilon) = 0$

**Convergência quase certa:** Seja  $\{X_1, X_2, \dots\}$  uma sequência de variáveis aleatórias e  $X$  uma variável aleatória definidas na mesma tripla de probabilidade. Dizemos que  $X_n$  converge quase certamente para  $X$ , isto é,  $X_n \xrightarrow{a} X$  se  $\mathcal{P}(\lim_{n \rightarrow \infty} X_n = X) = 1$

Assim temos alguns teoremas:

**Teorema** Sejam  $Z_1, Z_2, \dots$  variáveis aleatórias. Se para cada  $\varepsilon > 0$ , tem-se  $\mathcal{P}(|Z_n - Z| \geq \varepsilon) = 0$ . Então,  $\mathcal{P}(Z_n \rightarrow Z) = 1$ , ou seja,  $\{Z_n\}$  converge quase certamente para  $Z$ .

**Teorema** Sejam  $Z_1, Z_2, \dots$  variáveis aleatórias. Se para cada  $\varepsilon > 0$ , tem-se  $\sum_n \mathcal{P}(|Z_n - Z| \geq \varepsilon) < \infty$ . Então,  $\mathcal{P}(Z_n \rightarrow Z) = 1$ , ou seja,  $\{Z_n\}$  converge quase certamente para  $Z$ .

**Teorema** Sejam  $Z_1, Z_2, \dots$  variáveis aleatórias. Então, se  $Z_n \rightarrow Z$  em probabilidade.

**Teorema - Lei Fraca dos Grandes Números** Seja  $X_1, X_2, \dots$  uma sequência de variáveis aleatórias independentes, cada uma com média  $m$  e com variância  $v < \infty$ . Então, para todo  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \mathcal{P}\left(\left|\frac{1}{n}(X_1 + X_2 + \dots + X_n) - m\right| \geq \varepsilon\right) = 0 \quad (11)$$

---

<sup>B</sup> $\lambda\phi(x) + (1 - \lambda)\phi(y) \geq \phi(\lambda x + (1 - \lambda)y)$  para  $x, y, \lambda \in \mathbb{R}$

Ou seja,  $\frac{1}{n}(X_1 + X_2 + \dots + X_n)$  converge em probabilidade para  $m$ .

**Teorema - Lei Forte dos Grandes Números** Seja  $X_1, X_2, \dots$  uma sequência de variáveis aleatórias independentes, cada uma com média finita  $m$  e com  $\mathbb{E}((X_i - m)^4) \leq a < \infty$ . Então

$$\mathcal{P}\left(\lim_{n \rightarrow \infty} \frac{1}{n}(X_1 + X_2 + \dots + X_n) = m\right) = 1 \quad (12)$$

Ou seja,  $\frac{1}{n}(X_1 + X_2 + \dots + X_n)$  converge quase certamente para  $m$ .

**Lema de Fatou** Se  $X_n \geq C$  para todo  $n$  e para uma constante  $C > -\infty$  então

$$\mathbb{E}[\liminf_{n \rightarrow \infty} X_n] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[X_n] \quad (13)$$

**Teorema da Convergência Dominada** Seja  $X_1, X_2, \dots$  uma sequência de variáveis aleatórias, se  $X_n \rightarrow X$  q.c. e existe  $Y$  variável aleatória com  $|X_n| \leq Y, \forall n$  e se  $\mathbb{E}(Y) < \infty$ , então  $\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \mathbb{E}(X)$ .

### 2.1.6 Processos Estocásticos

Seja uma tripla de probabilidade  $(\Omega, \mathcal{F}, \mathcal{P})$  e o intervalo  $I = [t_0, t_f]$ . Um processo estocástico é uma sequência de variáveis aleatórias nesse espaço de probabilidade. Temos como um importante processo estocástico o seguinte movimento:

**Movimento Browniano:** É uma função  $W : I \times \Omega \rightarrow \mathbb{R}$  que atende:

- $\mathcal{P}(\omega : W(0, \omega) = 1) = 1$
- $\forall 0 \leq s \leq t : W(t) - W(s) \sim N(0, t - s)$
- $\forall 0 \leq t_1 \leq t_2 \leq \dots \leq t_f$  as variáveis aleatórias

$$W(t_1), W(t_2) - W(t_1), \dots, W(t_f) - W(t_{f-1})$$

são independentes

- $\mathcal{P}(\omega : W(\cdot, \omega) \text{ é contínuo}) = 1$ , ou seja, praticamente todos os caminhos simples são funções constantes

O Movimento Browniano  $W(t)$  possui algumas propriedades interessantes:

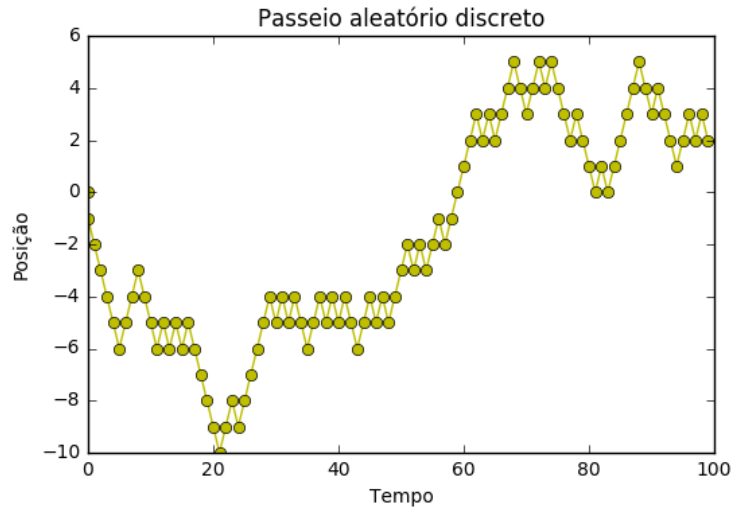
- Caminhos simples não são diferenciáveis em nenhum ponto
- $W(t)$  não possui variação finita nos subintervalos de  $I^C$

Um exemplo comumente utilizado para descrever Movimento Browniano discretamente é o passeio aleatório, dado as variáveis aleatórias  $X_1(t), X_2(t), \dots$  independentes e identicamente distribuídas tais que  $\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = 0.5$ , assim definindo a

---


$$C_{\sup\{\sum_{i=0}^{f-1} |W(t_{i+1}) - W(t_i)| = \infty\}}$$

nova variável aleatória  $S_n = \sum_{i=1}^n X_i$ , ou seja,  $(S_n)_{n \in \mathbb{N}}$  é um processo estocástico, que se pegarmos uma realização específica, teremos o seguinte exemplo<sup>D</sup>:



## 2.2 Cálculo Estocástico

### 2.2.1 Integração Estocástica

Seja a tripla de probabilidade  $(\Omega, \mathcal{F}, \mathcal{P})$  para o Processo Browniano  $W(t)$ ,  $t \in [0, T]$ . E defini-se a Fitração como:

**Fitração:**  $\mathcal{F}_t = \sigma\{W(s), s \leq t\}$ <sup>E</sup>

Tem-se que  $f(t)$  é **adaptado** a  $\mathcal{F}_t$  se  $f(t)$  é  $\mathcal{F}_t$ -mensurável.

Seja  $L_{ad}^2([a, b] \times \Omega) := \{f(t, \omega) | a \leq t \leq b, \omega \in \Omega\}$  tal que

- $f(t)$  é adaptado à  $\mathcal{F}_t$
- $\int_a^b |f(t)| dt < \infty$  quase certamente

Definindo a integral estocástica  $I(f) = \int_a^b |f(t)| dW(s) < \infty$  para  $f \in L_{ad}^2([a, b] \times \Omega)$  em três etapas:

- Para  $f$  um passo processo estocástico em  $L_{ad}^2([a, b] \times \Omega)$ , ou seja,  $f(t, \omega) = \sum_{i=1}^{\infty} \xi_{i-1}(\omega) 1_{[t_{i-1}, t_i]}(t)$  ( $\xi_{i-1} : \mathcal{F}_{t_{i-1}}$ -mensurável).

$$\text{Definindo } I(f) = \sum_{n=1}^{\infty} \xi_{i-1}(W(t_i) - W(t_{i-1}))$$

- Para  $f$  geral em  $L_{ad}^2([a, b] \times \Omega)$ : seja  $\{f_n(t)\}$  um passo do processo tal que

$$\lim_{n \rightarrow \infty} \int_a^b \mathbb{E}(|f(t) - f_n(t)|^2) dt = 0$$

<sup>D</sup>Código no apêndice

<sup>E</sup>a menor  $\sigma$ -álgebra que contém o conjunto  $W(s \leq t)$

- Como  $\{I(f_n)\}$  é uma sequência de Cauchy em  $L^2_\Omega$ , define-se

$$I(f) = \lim_{n \rightarrow \infty} I(f_n)$$

Algumas propriedades interessantes da integral são: valor esperado igual a zero, é contínua com respeito ao extremo direito  $b$ , é Martingal<sup>F</sup>.

### 2.2.2 Equação Diferencial Estocástica

Define-se uma Equação Diferencial Estocástica como uma equação da seguinte forma:

$$x(t) = x_0 + \int_{t_0}^t \mathbf{a}(s, x(s)) ds + \int_{t_0}^t \mathbf{B}(s, x(s)) dW(s), \quad t \in [t_0, T] \quad (14)$$

**Drift:**  $a : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$

**Processo de difusão:**  $B = (b_1, \dots, b_m) : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d+m}$

**Processo Wiener-Browniano:**  $W = (W^1, \dots, W^m)^T : \mathbb{R}^m$

As seguintes proposições garantem a existência e unicidade:

- Condição de Lipschitz:  $\exists K > 0 : \forall x, y \in \mathbb{R}^d, t \in [t_0, T],$

$$|a(t, x) - a(t, y)| + |B(t, x) - B(t, y)| \leq K|x - y| \quad (15)$$

- Condição de Crescimento:  $\exists K > 0:$

$$|a(t, x)|^2 + |B(t, x)|^2 \leq K(1 + |x|^2) \quad (16)$$

### 2.2.3 Fórmula de Itô

Seja  $x(t)$  a solução da seguinte equação:

$$dx(t) = a(t, x(t))dt + b(t, x(t))dW(t) \quad (17)$$

Podemos observar que  $Y_t = U(t, x(t))$  satisfaz:

$$dY_t = \left( \frac{\partial U}{\partial t} + a \frac{\partial U}{\partial x} + \frac{b^2}{2} \frac{\partial^2 U}{\partial x^2} \right) dt + b \frac{\partial U}{\partial x} dW(t) \quad (18)$$

Tendo os operadores diferenciais:

$$L^0 = \left( \frac{\partial}{\partial t} + a \frac{\partial}{\partial x} + \frac{b^2}{2} \frac{\partial^2}{\partial x^2} \right) \text{ e } L^1 = b \frac{\partial}{\partial x}$$

Tendo assim a EDE:

$$U(t, x_t) = U_0 + \int_{t_0}^t L^0 U(s, x(s)) ds + \int_{t_0}^t L^1(s, x(s)) dW(s), \quad t \in [t_0, T] \quad (19)$$

---

<sup>F</sup> $(M_n)_{n \in \mathbb{N}}$  é martingal a filtração  $\mathcal{F}$  com respeito a filtração  $(\mathcal{F}_t)_{t \leq T}$ , se  $M_n \in \mathcal{F}$ ,  $M_n \in L^1$  e  $\mathbb{E}[M_n] = M_{n-1}$

## 3 Simulações de Monte Carlo

### 3.1 Apresentação de Monte Carlo

Um problema relevante ao lidar com variáveis aleatórias é o cálculo do valor esperado  $\mathbb{E}(\phi(X(T)))$ , onde  $X(T)$  é a solução de uma Equação Diferencial Estocástica em  $T$ .

Sabemos a partir da Lei dos Grandes Números que dada uma sequência de variáveis aleatórias identicamente distribuídas e independentes  $X^i(T)$ , a sequência  $\hat{X}_n^i(T)$ ,

$$\hat{X}_n^i = \frac{1}{n} \sum_{i=1}^n X^i(T) \quad (20)$$

converge - em probabilidade e em determinadas condições quase certamente - para  $\mathbb{E}(X(T))$  quando  $n \rightarrow \infty$ . Se gerarmos  $n$  variáveis aleatórias identicamente distribuídas e independentes,  $X^i(T)$ , teremos um estimador não-viesado para  $\mathbb{E}(X(T))$ , denotado  $\hat{X}_n^i(T)$  e chamado de estimador Padrão de Monte Carlo.

Um exemplo interessante de uso de Monte Carlo é a aproximação de integrais, um caso especial que podemos observar isso é no cálculo de  $\pi$ , pois sabemos que um círculo de raio 1 tem área  $\pi$ , se consideramos o primeiro quadrante de uma circunferência com equação  $x^2 + y^2 = 1$  no plano cartesiano, teremos que

$$\frac{\pi}{4} = \int_0^1 \sqrt{1-x^2} dx = \int_{-\infty}^{\infty} \sqrt{1-x^2} 1_{[0,1]} dx = \mathbb{E}[\sqrt{1-x^2}] \approx \frac{1}{n} \sum_{i=1}^n \sqrt{1-X_i^2} = \sum_{i=1}^n \sqrt{1-U[0,1]^2}^G$$

Como usual medida de qualidade de um estimador, definimos a média dos erros ao quadrado:

$$MEQ(\mathbb{E}(\hat{X}_n^i(T))) = \mathbb{E}((\hat{X}_n^i(T) - \mathbb{E}(X(T)))^2) = \mathbb{E}((\hat{X}_n^i(T) - \mathbb{E}(\hat{X}_n^i(T)))^2) + (\mathbb{E}(\hat{X}_n^i(T) - \mathbb{E}(X(T))))^2$$

$$MEQ(\mathbb{E}(\hat{X}_n^i(T))) = Var(\hat{X}_n^i(T)) + Vies(\hat{X}_n^i(T))$$

E raiz da média do error ao quadrado:

$$RMEQ(\mathbb{E}(\hat{X}_n^i(T))) = \sqrt{MEQ(\mathbb{E}(\hat{X}_n^i(T)))} \quad (21)$$

### 3.2 Redução de Variância

Já sabemos que o método de Monte Carlo é uma técnica que tem o propósito de aproximar uma específica medida definida em um domínio dado por meio de uma amostra de acordo com uma determinada distribuição.

Sendo  $\{X_k^i(T)\}_{i=1}^N$  uma amostra aproximada da distribuição  $X(T)$ , temos que o estimador padrão de Monte Carlo é  $\mu = \hat{X}_N^i = \frac{1}{N} \sum_{i=1}^N X_k^i(T)$  (20). Assim, podemos calcular o erro:

$$\mathbb{E}[X(T)] - \mu = \mathbb{E}[X(T) - X_k + X_k] - \mu = \mathbb{E}[X(T) - X_k] + \mathbb{E}[X_k] - \mu$$

---

<sup>G</sup>Implementação no apêndice

O termo  $\mathbb{E}[X(T) - X_k] \sim O(h^p)$  pela convergência fraca do erro<sup>H</sup> e  $(\mathbb{E}[X_k] - \mu) \sim O(\frac{\text{Var}(X)}{\sqrt{N}})$  é o erro inerente ao Monte Carlo pelo Teorema Central do Limite<sup>I</sup>. Assim, temos um erro  $O(h^p) + O(\frac{\text{Var}(X)}{\sqrt{N}})$ .

Técnicas de redução de variância consiste em mudar o clássico método de Monte Carlo para diminuir a magnitude do erro de simulação que depende da variância da variável. A principal ideia nessas técnicas é reescrever a esperança a ser calculada, exemplo  $\mathbb{E}(X)$ , por a esperança de outra variável aleatória,  $Y$ , que possui uma variância menor,  $\text{Var}(Y) \leq \text{Var}(X) | \mathbb{E}(X) = \mathbb{E}(Y) + C$ , sendo  $C$  uma constante computacional.

Os mais usados métodos de redução de variância são variáveis antitéticas e variáveis de controle.

### 3.2.1 Variável Antitética

Variáveis antitéticas são amplamente usadas em simulações de Monte Carlo pela sua generalidade e pela facilidade de implementação. Ou seja, pode ser interpretado como um truque para aumentar a precisão do Monte Carlo.

Para entender como a técnica de variável antitética funciona, considere  $I = \mathbb{E}(f(U))$  com  $U \sim U(0, 1)$  para uma dada função  $f$ . Podemos ter  $U_1, \dots, U_{2n}$  variáveis aleatórias iid's com  $U_i \sim U(0, 1)$  O estimador padrão de Monte Carlo será:

$$I_{2n} = \frac{1}{2n} \sum_{i=1}^{2n} f(U_i)$$

mas uma alternativa é usar a antitéticas:

$$\hat{I}_{2n} = \frac{1}{n} \sum_{i=1}^n \frac{f(U_i) + f(1 - U_i)}{2}$$

As variâncias dos dois estimadores serão:

$$\text{Var}(I_{2n}) = \frac{1}{2n} \text{Var}(f(U_i))$$

$$\text{Var}(\hat{I}_{2n}) = \frac{1}{2n} (\text{Var}(f(U_i)) + \text{cov}(f(U_i), f(1 - U_i)))$$

É fácil observar que  $\text{Var}(\hat{I}_{2n}) \leq \text{Var}(I_{2n})$  se, e somente se,  $\text{cov}(f(U_i), f(1 - U_i)) \leq 0$ . Quando  $f$  é uma função não-decrescente ou não-crescente, tem-se que  $\text{Var}(\hat{I}_{2n}) \leq \text{Var}(I_{2n})$ , ou seja, realizar o método de Monte Carlo com variáveis antitéticas é melhor que usar o método clássico.

Essa ideia pode ser generalizada. Se  $X$  é uma variável aleatória em  $\mathbb{R}^d$  e se  $G$  opera em  $\mathbb{R}^d$  tal que a distribuição de  $X$  é preservada por  $G$ , pode-se construir um método antitético generalizado com a igualdade:

$$\mathbb{E}[f(X)] = \frac{1}{2} (\mathbb{E}[f(X)] + \mathbb{E}[f(G(X))])$$

E assim,

$$\hat{I}_{2n} = \frac{1}{n} \sum_{i=1}^n \frac{f(X_i) + f(G(X_i))}{2}$$

<sup>H</sup>  $\sup_{0 \leq nh \leq T} (\mathbb{E}[X(nh)] - \mathbb{E}[X_n]) = O(h)$ ,  $p \in \mathbb{N}$  sendo  $p$  dependente do grau da aproximação de Euler

<sup>I</sup> Assumindo 95% de confiança

$$I_{2n} = \frac{1}{2n} \sum_{i=1}^{2n} f(X_i)$$

E como anteriormente,  $\hat{I}_{2n}$  é melhor que  $I_{2n}$  se, e somente se,  $\text{cov}(f(X), f(G(X))) \leq 0$ . Um exemplo é  $X \sim N(0, 1)$  e  $G(x) = -x$ , tem-se que  $X$  e  $-X$  são idênticas.

### 3.2.2 Variáveis de Controle

A técnica de variável de controle é menos genérica que do uso de variável antitética, pois necessita de um conhecimento sobre as variáveis aleatórias na simulação. Seja  $X$  uma variável aleatória que deseja-se calcular  $\mathbb{E}[X]$  e tem-se  $Y$  outra variável aleatória parecida com  $X$ , com  $\mathbb{E}[Y]$ . Assim,  $Z = X + [Y] - Y$  uma variável aleatória e

$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y] - \mathbb{E}[Y] = \mathbb{E}[X] = \mathbb{E}[X - Y] + \mathbb{E}[Y]$$

Nesse contexto,  $Y$  é a variável de controle e isso essa troca de variáveis só serão úteis se  $\text{Var}(Z) = \text{Var}(X - Y) \leq \text{Var}(X)$ .

Observando que  $Z_\theta = X + \theta(\mathbb{E}[Y] - Y)$ ,  $\theta \in \mathbb{R}^+$ , implica  $\mathbb{E}[Z_\theta] = \mathbb{E}[X]$  e  $\text{Var}(Z_\theta) = \text{Var}(X - \theta Y) = \text{Var}(X) + \theta^2 \text{Var}(Y) - 2\theta \text{cov}(X, Y)$ , tendo  $\theta$  que minimiza  $\text{Var}(Z_\theta)$  :

$$\theta_{\min} = \frac{\text{cov}(X, Y)}{\text{Var}(Y)}$$

Assim,  $\text{Var}(Z) < \text{Var}(X)$  se e somente se,  $\theta \in [0, 2\theta_{\min}]$ .

### 3.3 Monte Carlo Multinível

Tem-se para o método padrão de Monte Carlo para que a raiz da média ao quadrado  $RMEQ$  (21) seja proporcional a  $\epsilon$ , temos que ter  $MEQ = O(\epsilon^2)^J$  e portanto  $\frac{1}{N} = O(\epsilon^2)$  e  $h^2 = O(\epsilon^2)$ , logo  $N = O(\epsilon^{-2})$  e  $h = O(\epsilon)$ .

O método de Monte Carlo Multinível (MLMC - multilevel Monte Carlo) busca diminuir essa complexidade computacional. De maneira geral, o MLMC usa níveis de resolução,  $l = 0, 1, 2, \dots, L$ . Em simulações de Equações Diferenciais Estocásticas (EDE), o nível 0 tem um intervalo de tempo, ou seja,  $[0, T]$ , enquanto o nível  $L$  terá  $M^L$  intervalos com  $h_L = M^{-L}T$ , sendo  $M$  uma constante de refino dos intervalos.

Se  $P$  denota uma função a ser analisada, um payoff, por exemplo e seja  $P_l$  sua aproximação no nível  $l$ , então  $\mathbb{E}[P_L]$  no nível final é igual ao valor esperado  $\mathbb{E}[P_0]$  somado à correções dadas pelas diferenças de esperanças nas simulações subsequentes,

$$\mathbb{E}(P_L) = \mathbb{E}[P_0] + \sum_{l=1}^L (\mathbb{E}[P_l - P_{l-1}])$$

A ideia do MLMC é estimar independentemente cada esperança do lado direito da equação acima buscando minimizar a variância total. Seja  $Y_l$  uma estimação para  $\mathbb{E}[P_l - P_{l-1}]$  usando  $N_l$  amostras, assumindo  $P_{-1} = 0$ . O estimador para média de  $N_l$  amostras será:

$$Y_l = \frac{1}{N_l} \sum_{i=1}^{N_l} (P_l^i - P_{l-1}^i)$$

---

<sup>J</sup> $MEQ = O(\frac{1}{N}) + O(h^2)$ , o custo computacional é proporcional ao número de caminhos  $N$  multiplicado pelo custo de gerar o caminho, ou seja, o número de pedaços  $h$  em cada caminho.



$P_l^i - P_{l-1}^i$  poderá ser obtida pela discretização para as mesmas amostras estocásticas, assim para níveis finos a diferença é pequena e portanto a variância é pequena também. O estimador  $\hat{Y}$  MLMC

$$\hat{Y} = \sum_{l=0}^L Y_l$$

Pode-se observar

$$\mathbb{E}[Y_l] = \frac{1}{N_l} \sum_{i=1}^{N_l} \mathbb{E}[P_l^i - P_{l-1}^i] = \mathbb{E}[P_l^i - P_{l-1}^i]$$

$$\mathbb{E}[\hat{Y}] = \sum_{l=0}^L \mathbb{E}[Y_l] = \mathbb{E}[P_0] + \sum_{l=1}^L \mathbb{E}[P_l - P_{l-1}] = \mathbb{E}[P_L]$$

Observa-se que mesmo utilizando diferentes discretizações de nível para estimar  $\mathbb{E}[P]$ , a aproximação final depende da aproximação do nível  $L$ . Mas o principal ponto se dá por  $P_l^i - P_{l-1}^i$ , pois depende de duas aproximações em diferentes níveis, ou seja, com diferentes intervalos, porém do mesmo caminho Browniano.

Este caminho pode ser obtido através da aproximação de equações diferenciais estocásticas da seguinte forma

$$dx(t) = f(x(t))dt + g(t)d\omega(t) \quad (22)$$

Assumindo  $\omega(t)$  Browniano multidimensional definido em  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , onde  $x(t) \in \mathbb{R}^d, \forall t \geq 0, f \in C^2(\mathbb{R}^d, \mathbb{R}^d), g \in C^2(\mathbb{R}^d, \mathbb{R}^{d \times m})$  e que haja  $x_0 \in \mathbb{R}^d$ . A aproximação mais simples da EDE (22) é o utilizando Euler-Maruyana.

**Euler-Maruyana** dado um passo  $\Delta t_l$ , definimos a partição  $P_{\Delta t_l} := \{n\Delta t_l : n = 0, 1, M, \dots, M^l\}$  de um intervalo  $[0, T]$ ,  $M^L \Delta t = T > 0$ . A aproximação  $X_n^l \approx (n\Delta t_l)$  será:

$$X_{n+1}^l = X_n^l + f(X_n^l)\Delta t_l + g(X_n^l)\Delta \omega_{n+1}^l \quad (23)$$

Onde  $\Delta \omega_{n+1}^l = \omega((n+1)\Delta t_l) - \omega(n\Delta t_l)$ . E a equação acima na forma de vetor terá sua  $i$ -ésima componente escrita por:

$$X_{i,n+1}^l = X_{i,n}^l + f_i(X_n^l)\Delta t_l + \sum_{j=1}^m g_{ij}(X_n^l)\Delta \omega_{j,n+1}^l$$

Analisando a variância temos que  $\text{Var}(\hat{Y}) = \sum_{l=0}^L \text{Var}(Y_l) = \sum_{l=0}^L \frac{\text{Var}(P_l^i - P_{l-1}^i)}{N_l}$ . Minimizando a variância de  $\hat{Y}$  para um custo computacional fixo  $C = \sum_{l=0}^L N_l \Delta t_l^{-1}$ , tratando  $N_l$  como variável contínua e usando a lagrange para achar o mínimo de

$$L = \sum_{l=0}^L \frac{\text{Var}(P_l^i - P_{l-1}^i)}{N_l} + \lambda \left( \sum_{l=0}^L N_l \Delta t_l^{-1} - C \right)$$

Pela condição de primeira ordem temos  $N_l = \lambda^{-0.5} \sqrt{\text{Var}(P_l^i - P_{l-1}^i) \Delta t_l}$  e como queremos  $\text{Var}(\hat{Y}) \leq \frac{\epsilon^2}{2}$ , e assim  $\lambda^{0.5} \geq 2\epsilon^{-2} \sum_{l=0}^L \sqrt{\frac{\text{Var}(P_l^i - P_{l-1}^i)}{\Delta t_l}}$ . Assim o número ótimo

de amostras para o nível  $l$  é

$$N_l = \lceil 2\varepsilon^{-2} \sqrt{\text{Var}(P_l^i - P_{l-1}^i) \Delta t_l} \left( \sum_{i=0}^L \sqrt{\frac{\text{Var}(P_l^i - P_{l-1}^i)}{\Delta t_l}} \right) \rceil \quad (24)$$

No método MLMC, a estimação da correção  $\mathbb{E}[\hat{P}_l - \hat{P}_{l-1}]$  pode ser usado para calcular a variância. Em particular, para a discretização de Euler com uma função Lipschitz, temos que  $l \rightarrow \infty$ ,  $\mathbb{E}[P - \hat{P}_l] \approx c_1 \Delta t_l$ , sendo  $c_1$  uma constante e portanto,

$$\mathbb{E}[\hat{P}_l - \hat{P}_{l-1}] \approx (M-1)c_1 h_1 \approx \mathbb{E}[P - \hat{P}_l] \quad (25)$$

Assim a variância que desejamos seja menor que  $\frac{\varepsilon}{\sqrt{2}}$  e assim temos que  $|\hat{Y}_L| < \frac{(M-1)\varepsilon}{\sqrt{2}}$ . Podemos utilizar assim

$$\max\{M^{-1} \|\hat{Y}_{L-1}\|, |\hat{Y}_L|\} < \frac{1}{\sqrt{2}}(M-1)\varepsilon \quad (26)$$

Outra forma de observamos a variância, assumindo  $O(\Delta t_l)$  converge fracamente, o viés será igual a  $c\Delta t_l = cTM^{-L}$ , sendo  $c$  uma constante, se quisermos proporcional a  $\frac{\varepsilon}{\sqrt{2}}$ . Teremos assim

$$L_{\max} = \frac{\log(\varepsilon^{-1}cT)}{\log(M)} \quad (27)$$

Desse modo conseguimos apresentar um algoritmo do MLMC:

```

L = 0
while L < Lmax ou não é válido (26) do
    Estime Vl usando Nl = 104;
    Defina ótimo Nl, usando (24);
    Gere adicionais caminhos de acordo com o novo Nl;
end

```

### Algorithm 1: Método Monte Carlo Multinível

Um teorema amplamente usando para ver a complexidade computacional é:

**Teorema** Seja  $P$  o funcional da solução de uma equação diferencial estocástica e  $P_l$  a aproximação numérica no nível  $l$  utilizando a discretização com passo  $h_l = M^{-l}T$ . Se existem estimadores independentes  $Y_l$  baseados em  $N_l$  amostras de Monte Carlo e constantes positivas  $\alpha, \beta, \gamma, c_1, c_2, c_3$  tais que  $\alpha \geq \frac{1}{2} \min(\beta, \gamma)$  e

- $|\mathbb{E}[P_l - P]| \leq c_1 M^{-\alpha l}$
- $\mathbb{E}[Y_l] = \begin{cases} \mathbb{E}[P_0], l = 0 \\ \mathbb{E}[P_l - P_{l-1}], l > 0 \end{cases}$
- $\text{Var}[Y_l] \leq \frac{c_2 M^{-\beta l}}{N_l}$
- $C_l \leq c_3 N_l M^{\gamma l}$ , onde  $C_l$  é a complexidade computacional de  $Y_l$

Então existe uma constante positiva  $c_4$  tal que para qualquer  $\varepsilon < e^{-1}$  existem valores de  $L$  e  $N_l$  para cada estimador de nível

$$Y = \sum_{l=0}^L Y_l$$

E tem MEQ limitado

$$MEQ = \mathbb{E}[(Y - \mathbb{E}[P])^2] < \varepsilon^2$$

com complexidade computacional  $C$  limitada

$$C \leq \begin{cases} c_4 \varepsilon^{-2}, \beta > \gamma \\ c_4 \varepsilon^{-2} (\log \varepsilon)^2, \beta = \gamma \\ c_4 \varepsilon^{-2 - (\gamma - \beta)/\alpha} (\log \varepsilon)^2, 0 < \beta < \gamma \end{cases}$$

### 3.4 Resultados numéricos

Podemos definir o custo computacional do MLMC como sendo:

$$C = N_0 + \sum_{l=1}^L (M^l + M^{l-1})$$

O termo  $M^l + M^{l-1}$  mostra o efeito que cada amostra nonível  $l > 0$  requer a computação de um caminho com dois passos diferentes. Já o Método de Monte Carlo clássico, pode ser calculado como:

$$C = \sum_{l=1}^L 2\varepsilon^{-2} \text{Var}(P_l) M^l$$

Sendo  $\frac{\varepsilon^2}{2}$  a variância do estimador assim como MLMC.

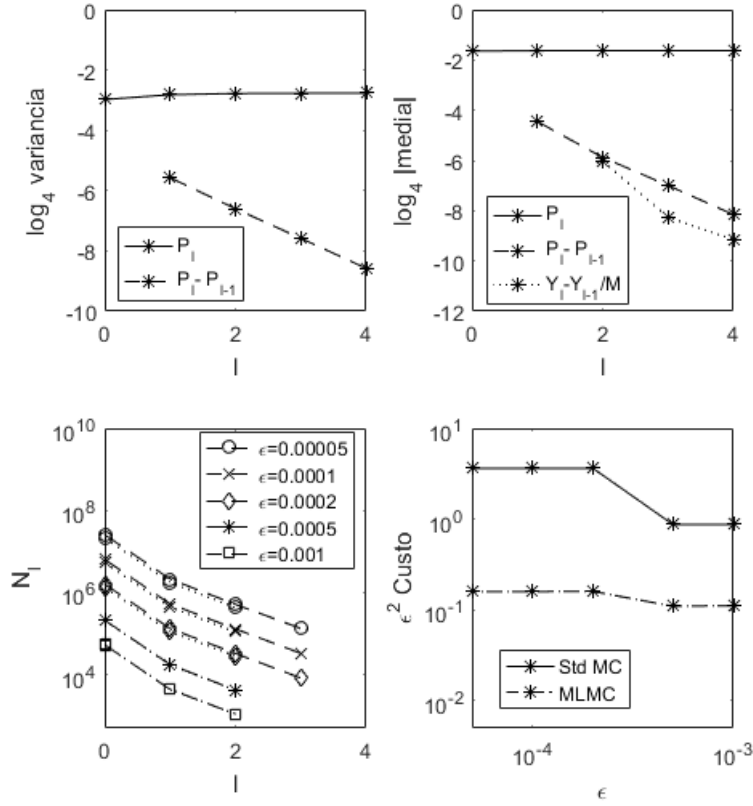
#### 3.4.1 Movimento Browniano Geometrico

Em muitos casos tomamos assumimos que a série de preços de um ativo listado na bolsa segue uma equação diferencial estocástica da seguinte forma:

$$dS = rSdt + \sigma SdW$$

Podemos utilizar essa equação para precificar certos tipos de derivativos, nesse caso apresentaremos a solução numérica, apesar de ter uma solução exata ao considerarmos a nova variável  $X = \log(S)$ , para uma Call Europeia que é um contrato que dá o direito de compra no futuro, ou maturidade  $T$ , do comprador e a obrigação de venda do vendedor da opção. Podemos escrever matematicamente o Pay-off desse contrato como  $P = e^{-r} \max(0, S(T) - S(0))$ , sendo  $r$  a taxa de juros  $S(0)$  o preço do ativo no momento atual e  $S(T)$  o preço do ativo na maturidade  $T$ , na equação diferencial acima  $\sigma$  representa a volatilidade do preço do ativo considerado, neste caso será uma constante.

Neste experimento assumiremos  $S(0) = 1$ ,  $r = 0.05$  e  $\sigma = 0.2$  e utilizaremos  $M = 4$



Esses quatro gráficos mostram características do método de Monte carlo Multinível além de comparações com o modelo padrão. O gráfico de acima a esquerda mostra como a variância cai de acordo com  $l$ , mostrando assim que a  $\text{Var}(P_l - P_{l-1}) = O(h)$  sendo proporcional a  $M^{-l}$ . Além disso,  $\text{Var}(P_l - P_{l-1})$  é bem menor que  $\text{Var}(P_l)$  do modelo clássico de Monte Carlo. O segundo gráfico acima mostra o valor esperado e seus ajustes em cada nível  $l$ , onde a inclinação -1 mostra novamente que  $\mathbb{E}[\hat{P}_l - \hat{P}_{l-1}] \rightarrow O(h)$ .

Os gráficos na parte inferior mostram o comportamento com diferentes valores de precisão, o da esquerda mostra que  $N_l$  decresce com  $l$  devido ao decréscimo de  $\text{Var}(P_l - P_{l-1})$  e  $h_l$ . Além disso, podemos observar que o valor para o máximo nível  $L$  aumenta de acordo com o decréscimo de  $\epsilon$ . Já o da direita colocamos  $\epsilon$  contra  $\epsilon * \text{Custo}$  para mostrar que o  $\epsilon * \text{Custo}$  é muito pouco dependente do valor de  $\epsilon$  para o modelo Multinível.

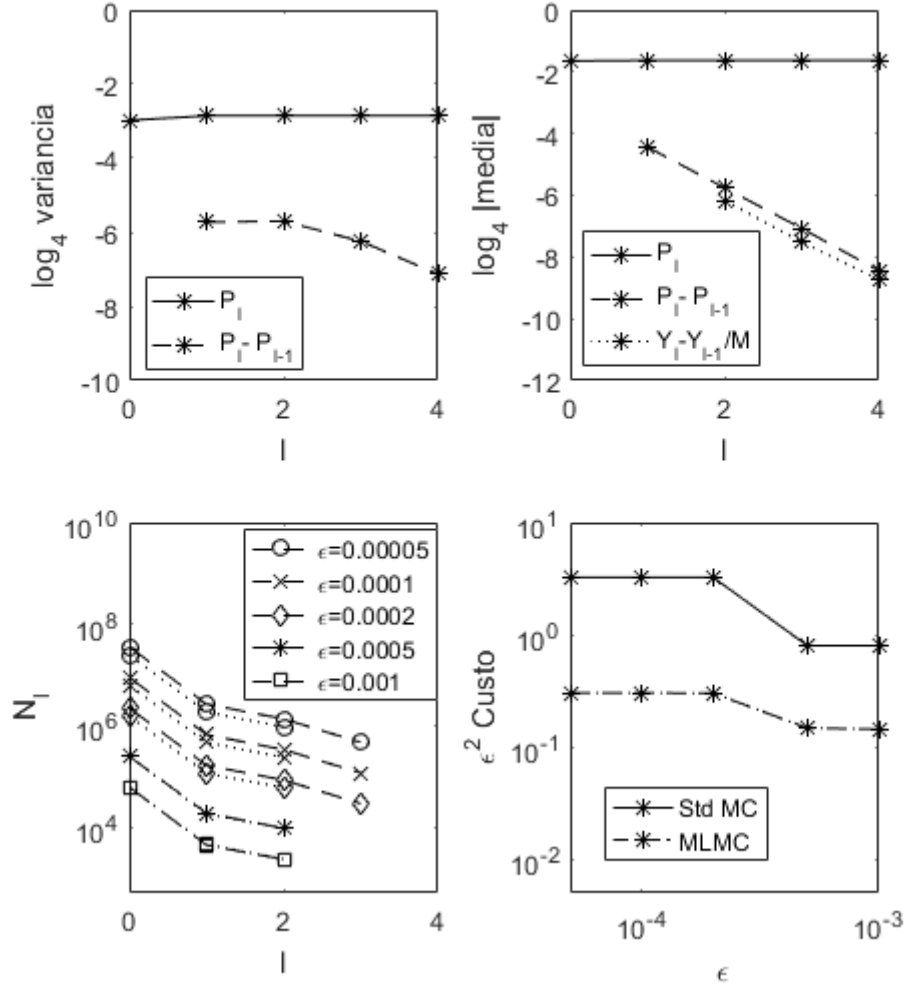
### 3.4.2 Modelo de Heston de volatilidade estocástica

Agora a volatilidade considerada  $\sigma$  anteriormente não será mais constante, também será um processo estocástico, sendo o modelo de preço escrito como

$$\begin{cases} dS = rSdt + \sqrt{V}SdW_2 \\ dV = \lambda(\sigma^2 - V)dt + \xi\sqrt{V}dW_2 \end{cases} \quad (28)$$

Nesse caso, não há uma solução exata para equação diferencial estocástica de preço, desse modo o uso de Monte Carlo auxilia na solução numérica.

Usando esse modelo para a opção de call europeia, e utilizando os mesmo dados da simulação anterior,  $S(0) = 1$ ,  $V(0) = 0.04$ ,  $r = 0.05$ ,  $\sigma = 0.2$ ,  $\lambda = 5$ ,  $\xi = 0.25$  e a correlação entre  $dW_1$  e  $dW_2$ ,  $\rho = -0.5$



Neste caso obtemos gráficos parecidos como os obtidos anteriormente, Porém podemos perceber que a variância cai mais lentamente do que no caso anterior com volatilidade constante.

## 4 Conclusão

Neste trabalho procurou-se o entendimento profundo de todos os assuntos envolvidos na teoria de Equações Diferenciais Estocásticas, assim passou por pontos de teoria de probabilidade avançada com a utilização de teoria de medida, processos aleatórios passando pelo conceito de Movimento Browniano e integração estocástica. Através desse estudo buscamos aplicar o Método de Monte Carlo Multinível, no qual com resultados teóricos e aplicações computacionais conseguimos observar a maior eficiência em relação ao Monte Carlo Clássico.

## 5 Apêndice

```

1 % Funcao [P, Nnivel] = MLMC(M, eps, MLMC_nivel)
2 % P      = valor
3 % Nnivel = numero de amostras em cada nivel
4 % M      = fator em cada nivel
5 % eps    = precisao
6 % MLMC_nivel = estimador da funcao no nivel l
7
8 function [P, Nnivel] = MLMC(M, eps, MLMC_nivel, extrapol)
9
10 L = -1;
11 %numero inicial de amostras para estimar V_nivel, variancia.
12 N = 10000;
13 %inicialmente igual a falso
14 converge = 0;
15
16 while ~converge
17
18 %Estimacao de V_nivel
19 L = L+1;
20 %vetor de valores q calcula soma de  $P_n - P(n-1)$ , soma de  $(P_n - P(n-1))^2$ ,
21 %soma de  $P_n$  e soma de  $P_n^2$ 
22 sums = feval(MLMC_nivel, M, L, N);
23 sum_nivel(1, L+1) = N;
24 sum_nivel(2, L+1) = sums(1);
25 sum_nivel(3, L+1) = sums(2);
26
27 V_nivel = sum_nivel(3, :)./sum_nivel(1, :) - ...
28 (sum_nivel(2, :)./sum_nivel(1, :)).^2;
29
30 % Calculo do otimo Nnivel
31
32 Nnivel = ceil( 2 * sqrt(V_nivel./(M.^(0:L))) * ...
33 sum(sqrt(V_nivel.*(M.^(0:L)))) / eps^2);
34
35 % Atualizacao das somas das amostras
36 for l=0:L
37     dNnivel = Nnivel(l+1)-sum_nivel(1, l+1);
38     if dNnivel>0
39         sums = feval(MLMC_nivel, M, l, dNnivel);
40         sum_nivel(1, l+1) = sum_nivel(1, l+1) + dNnivel;
41         sum_nivel(2, l+1) = sum_nivel(2, l+1) + sums(1);
42         sum_nivel(3, l+1) = sum_nivel(3, l+1) + sums(2);
43     end
44 end
45
46
47 % teste de convergencia
48 if extrapol
49     range = 0;
50     if L>1 & M^L>=16
51         con = M.^range.*( sum_nivel(2, L+1+range)./sum_nivel(1, L+1+range) ...
52             - (1/M)*sum_nivel(2, L+range)./sum_nivel(1, L +range) );
53         converge = (max(abs(con)) < (M^2-1)*eps/sqrt(2)) | (M^L>1024) ;
54     end
55
56

```

```

57     else
58         range = -1:0;
59         if L>1 & M^L>=16
60             con = M.^range.*sum_nivel(2,L+1+range)./sum_nivel(1,L+1+range);
61             converge = (max(abs(con)) < (M-1)*eps/sqrt(2)) | (M^L>1024) ;
62         end
63     end
64 end
65
66
67 % O estimador sera
68 P = sum( sum_nivel(2,1:L+1)./sum_nivel(1,1:L+1) );
69 if extrap
70     P = P + ( sum_nivel(2,L+1)./sum_nivel(1,L+1) ) / (M-1);
71 end
72 end

```

```

1  function preco_call = european_call(r,sigma,T,S0,K)
2
3  % T: maturidade em anos (base 252)
4  % K: preco de exercicio
5  % S0: preco do ativo na data zero
6  % r: taxa de juros continua a.a. (base 252)
7  % sigma: volatilidade a.a. do modelo de Black-Scholes (base 252)
8
9  d1 = (log(S0/K) + (r + 0.5*sigma^2)*T)/(sigma*sqrt(T));
10 d2 = d1 - sigma*sqrt(T);
11 N1 = 0.5*(1 + erf(d1/sqrt(2)));
12 N2 = 0.5*(1 + erf(d2/sqrt(2)));
13
14 preco_call = S0*N1 - K*exp(-r*T)*N2;
15
16 %preco_put = K*exp(-r*T) - S0 + preco_call;

```

```

1  function mlmc_test
2
3  close all; clear all;
4
5  N = 2000000;
6  M = 4;
7
8  % Teste de convergencia
9
10 L = 0:4;
11
12 del1 = [];
13 del2 = [];
14 var1 = [];
15 var2 = [];
16
17 for l = L
18     disp(sprintf('l=%d',l))
19     sums = mlmc_1(M,l,N);
20     del1 = [del1 sums(3)/N ];
21     del2 = [del2 sums(1)/N ];
22     var1 = [var1 sums(4)/N-(sums(3)/N)^2 ];
23     var2 = [var2 sums(2)/N-(sums(1)/N)^2 ];
24 end

```



```

25
26     disp( sprintf( ' value = %f ', del1( length(L) ) ) )
27
28 % Complexidade do MLMC
29
30 Eps = [ 0.001 0.0005 0.0002 0.0001 0.00005 ];
31
32 maxl = 0;
33
34 for extrap = 0:1
35     for i = 1:length(Eps)
36         eps = Eps(i);
37         [P, Nl] = MLMC(M, eps, @mlmc_l, extrap);
38         l = length(Nl)-1;
39         maxl = max(maxl, l);
40         mlmc_cost(i, extrap+1) = (1+1/M)*sum(Nl.*M.^(0:l));
41         std_cost(i, extrap+1) = sum((2*var1((0:l)+1)/eps^2).*M.^(0:l));
42
43         disp( sprintf( ' mlmc_custo = %d, std_custo = %d ', ...
44             mlmc_cost(i, extrap+1), std_cost(i, extrap+1) ) )
45         Nls{i, extrap+1} = Nl;
46     end
47 end
48
49 maxl = max(maxl, 4);
50
51
52 % Calculo do RMQE
53
54 P_exact = european_call(0.05, 0.2, 1, 1, 1);
55
56 for extrap = 0:1
57     for i = 1:length(Eps)
58         eps = Eps(i);
59         for k = 1:10
60             [Pk(k), Nl] = MLMC(M, eps, @mlmc_l, extrap);
61         end
62         RMSE = sqrt( sum((Pk-P_exact).^2) / length(Pk) );
63         disp( sprintf( 'RMQE, eps, razao = %f %f %f ', RMSE, eps, RMSE/eps ) )
64     end
65     disp( ' ' )
66 end
67
68
69
70 %Figuras
71
72 nvert = 2;
73 figure; pos=get(gcf, 'pos'); pos(3:4)=pos(3:4).*[1.0 0.75*nvert];
74
75 set(gcf, 'pos', pos);
76
77 set(0, 'DefaultAxesColorOrder', [0 0 0]);
78 set(0, 'DefaultAxesLineStyleOrder', '-*| - -*| -*| -.*')
79
80 subplot(nvert, 2, 1)
81 plot(L, log(var1)/log(M), L(2:end), log(var2(2:end))/log(M))
82 xlabel('l'); ylabel('log_4 Varianca'); %title( stitle )

```

```

83 legend('P_1', 'P_1-P_{1-1}', 3)
84 axis([0 max1 -10 0])
85
86 subplot(nvert, 2, 2)
87 plot(L, log(abs(del1))/log(M), L(2:end), log(abs(del2(2:end)))/log(M), ...
88      L(3:end), log(abs(del2(3:end)-del2(2:end-1)/M))/log(M))
89 xlabel('l'); ylabel('log_4|Media|'); %title(stitle)
90 legend('P_1', 'P_1-P_{1-1}', 'Y_1-Y_{1-1}/M', 3)
91 axis([0 max1 -12 0])
92
93 set(0, 'DefaultAxesLineStyleOrder', '--o|--x|--d|--*|--s|:o|x|:d|:~|:s');
94
95 if nvert==1
96     print('-deps2c', 'mlmc_gbm1a.eps')
97     figure; pos=get(gcf, 'pos'); pos(3:4)=pos(3:4).*[1.0 0.75];
98     set(gcf, 'pos', pos);
99 end
100
101 subplot(nvert, 2, 2*nvert-1)
102 semilogy(0:length(Nls{5,1})-1, Nls{5,1}, ...
103          0:length(Nls{4,1})-1, Nls{4,1}, ...
104          0:length(Nls{3,1})-1, Nls{3,1}, ...
105          0:length(Nls{2,1})-1, Nls{2,1}, ...
106          0:length(Nls{1,1})-1, Nls{1,1}, ...
107          0:length(Nls{5,2})-1, Nls{5,2}, ...
108          0:length(Nls{4,2})-1, Nls{4,2}, ...
109          0:length(Nls{3,2})-1, Nls{3,2}, ...
110          0:length(Nls{2,2})-1, Nls{2,2}, ...
111          0:length(Nls{1,2})-1, Nls{1,2});
112 xlabel('l'); ylabel('N_1'); %title(stitle)
113 legend('\epsilon=0.00005', '\epsilon=0.0001', '\epsilon=0.0002', ...
114        '\epsilon=0.0005', '\epsilon=0.001', 1)
115
116 axis([0 max1 5e2 1e10])
117
118
119 set(0, 'DefaultAxesLineStyleOrder', '-*|-.*|--*|:*')
120
121 subplot(nvert, 2, 2*nvert)
122 loglog(Eps, Eps.^2.*std_cost(:,1)', Eps, Eps.^2.*mlmc_cost(:,1)')
123 xlabel('\epsilon'); ylabel('\epsilon^2_Custo'); %title(stitle)
124 legend('Std_MC', 'MLMC', 3)
125 axis([5e-5 1e-3 0.005 10])
126
127
128 if nvert==1
129     print('-deps2c', 'mlmc_gbm1b.eps')
130 else
131     print('-deps2c', 'mlmc_gbm1.eps')
132 end
133
134
135 % Estimador do nivel
136
137 function sums = mlmc_l(M, l, N)
138
139 T = 1;
140 r = 0.05;

```

```

141 sig = 0.2;
142
143 nf = M^1;
144 nc = nf/M;
145
146 hf = T/nf;
147 hc = T/nc;
148
149 sums(1:4) = 0;
150
151 for N1 = 1:10000:N
152     N2 = min(10000,N-N1+1);
153
154     X0 = 1;
155
156     Xf = X0*ones(1,N2);
157     Xc = Xf;
158
159     Af = 0.5*hf*Xf;
160     Ac = 0.5*hc*Xc;
161
162     Mf = Xf;
163     Mc = Xc;
164
165     if l==0
166         dWf = sqrt(hf)*randn(1,N2);
167         Xf = Xf + r*Xf*hf + sig*Xf.*dWf;
168         Af = Af + 0.5*hf*Xf;
169         Mf = min(Mf,Xf);
170     else
171         for n = 1:nc
172             dWc = zeros(1,N2);
173             for m = 1:M
174                 dWf = sqrt(hf)*randn(1,N2);
175                 dWc = dWc + dWf;
176                 Xf = Xf + r*Xf*hf + sig*Xf.*dWf;
177                 Af = Af + hf*Xf;
178                 Mf = min(Mf,Xf);
179             end
180             Xc = Xc + r*Xc*hc + sig*Xc.*dWc;
181             Ac = Ac + hc*Xc;
182             Mc = min(Mc,Xc);
183         end
184         Af = Af - 0.5*hf*Xf;
185         Ac = Ac - 0.5*hc*Xc;
186     end
187
188     Pf = max(0,Xf-1);
189     Pc = max(0,Xc-1);
190
191     Pf = exp(-r*T)*Pf;
192     Pc = exp(-r*T)*Pc;
193
194     if l==0
195         Pc=0;
196     end
197
198     sums(1) = sums(1) + sum(Pf-Pc);

```

```

199     sums(2) = sums(2) + sum((Pf-Pc).^2);
200     sums(3) = sums(3) + sum(Pf);
201     sums(4) = sums(4) + sum(Pf.^2);
202 end
203
204 %-----
205
206 function m=mu(x,h)
207
208 m = [ 0.05*x(1,:); ...
209       ((1-exp(-5*h))/h)*(0.04-x(2,:)) ];
210
211 %-----
212
213 function sigdW=sig_dW(x,dW,h)
214
215 dW(2,:) = -0.5*dW(1,:) + sqrt(0.75)*dW(2,:);
216
217 sigdW = [ sqrt(max(0,x(2,:)))*x(1,:).*dW(1,:); ...
218           exp(-5*h)*0.25*sqrt(max(0,x(2,:)))*dW(2,:) ];

```

## 6 Referências Bibliográficas

- [1] Rosenthal, Jeffrey S. - At first look at rigorous probability
- [2] Williams, David - Probability with Martingales
- [3] Kuo, Hui-Hsiung - Introduction to Stochastic Integration
- [4] Evans, Lawrence C. - An Introduction to stochastic Differential Equations
- [5] Sauer, Timothy - Numerical Solution of Stochastic Differential Equations in Finance
- [6] Sauer, Timothy - Computational solution of stochastic differential equations
- [7] Higham, Desmond J. - Black–Scholes Option Valuation for Scientific Computing Students
- [8] Higham, Desmond J. - An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations
- [9] Giles, Mike and Szpruch, Lukasz - Multilevel Monte Carlo methods for applications in finance
- [10] B. Giles, Michael - Multilevel Monte Carlo methods - 2013
- [11] B. Giles, Michael - Multilevel Monte Carlo methods - 2015
- [12] B. Giles, Michael - Multilevel Monte Carlo Path Simulation - 2008
- [13] Higham, Desmond J. - Stochastic ordinary differential equations in applied and computational mathematics
- [14] Glasserman, Paul - Monte Carlo Methods in Financial Engineering (Stochastic Modelling and Applied Probability)