UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

**Aplicação de medidas de causalidade na geração de cenários
de Monte Carlo como alternativa para precificação de
contratos de opções**

Daniel Brignani Rodrigues

Dissertação de Mestrado do Programa de Mestrado Profissional em
Matemática, Estatística e Computação Aplicadas à Indústria (MECAI)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Daniel Brignani Rodrigues

**Aplicação de medidas de causalidade na geração de
cenários de Monte Carlo como alternativa para precificação
de contratos de opções**

Dissertação apresentada ao Instituto de Ciências
Matemáticas e de Computação – ICMC-USP,
como parte dos requisitos para obtenção do título
de Mestre – Mestrado Profissional em Matemática,
Estatística e Computação Aplicadas à Indústria.
VERSÃO REVISADA

Área de Concentração: Matemática, Estatística e
Computação

Orientador: Prof. Dr. Francisco Louzada Neto

USP – São Carlos
Novembro de 2017

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

R696a Rodrigues, Daniel Brignani
Aplicação de medidas de causalidade na geração de cenários de Monte Carlo como alternativa para precificação de contratos de opções / Daniel Brignani Rodrigues; orientador Francisco Louzada Neto. -- São Carlos, 2017.
70 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Mestrado Profissional em Matemática, Estatística e Computação Aplicadas à Indústria) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2017.

1. Causalidade. 2. Séries temporais financeiras. 3. Simulação de Monte Carlo. 4. Precificação de Opções. 5. Projeções de mercado. I. Neto, Francisco Louzada, orient. II. Título.

Daniel Brignani Rodrigues

**On the application of causality measures for Monte Carlo
simulations as alternative to price option contracts**

Master dissertation submitted to the Institute of
Mathematics and Computer Sciences – ICMC-USP, in
partial fulfillment of the requirements for the degree of
the Master – Professional Masters in Mathematics,
Statistics and Computing Applied to Industry. *FINAL
VERSION*

Concentration Area: Mathematics, Statistics and
Computing

Advisor: Prof. Dr. Francisco Louzada Neto

**USP – São Carlos
November 2017**

Este trabalho é dedicado à minha esposa Priscila que além de ser minha alma gêmea, acredita, como eu, que a Ciência (desde seu ensino nas escolas até seu uso para as grandes descobertas da humanidade) é fundamental para melhorar o mundo em que vivemos e deve ser sempre ensinada e defendida por todos.

AGRADECIMENTOS

Gostaria de agradecer primeiramente à minha família, por todo o suporte dado. Agradeço a minha irmã Luiza pelo apoio e pela revisão gramatical da minha dissertação.

Em especial, quero agradecer à minha esposa Priscila, por sempre me incentivar e nunca deixar me desanimar devido ao cansaço, e pela paciência e compreensão necessárias durante todo o tempo que passei na produção deste trabalho.

Também agradeço ao meu orientador Francisco Louzada, aos professores do MECAI e ao ICMC da USP de São Carlos, que se dispuseram a nos orientar e a viajar para São Paulo duas vezes por semana para possibilitar um curso de alto nível de mestrado para alunos como nós, que trabalhamos praticamente o dia todo no mercado financeiro.

Também agradeço ao Grilo, que, desde o curso da graduação, compartilhou muitos conhecimentos e informações comigo, o que permitiu que eu crescesse, tanto na vida acadêmica quanto no mercado de trabalho.

“cum hoc ergo propter hoc”

(Falácia da causa falsa. Significa em tradução livre: “com isto, logo por causa disto”, na qual se supõe que, pelo fato de duas coisas acontecerem juntas, uma é a causa da outra.)

RESUMO

RODRIGUES, D. B. **Aplicação de medidas de causalidade na geração de cenários de Monte Carlo como alternativa para precificação de contratos de opções.** 2017. 70 p. Dissertação (Mestrado – Mestrado Profissional em Matemática, Estatística e Computação Aplicadas à Indústria) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

Este trabalho tem como objetivo utilizar medidas de causalidade entre séries temporais de grandezas financeiras para determinar a dependência entre os ativos do mercado e utilizar as medidas obtidas para fazer inferências sobre a dinâmica desses ativos. Essa metodologia define um previsor para os valores das séries que, juntamente com a determinação das distribuições de probabilidades empíricas dos erros desse previsor por meio do método de Kernel, permite a amostragem aleatória de cenários multivariados, com diversas aplicações. Os ativos considerados para os testes de causalidade são o índice **Ibovespa**, o valor da paridade da moeda dólar-real **USDBRL** (utilizando suas séries de preços e retornos de preços), além da taxa de juros negociada diariamente (**CDI**). O uso do Método de Monte Carlo (MMC) é abordado para a precificação de opções de compra europeias (*calls*) de **USDBRL** e **Ibovespa**, e a comparação dos resultados gerados por essa metodologia com valores calculados pela fórmula de Black-Scholes (método mais utilizado no mercado financeiro, atualmente), evidenciando suas vantagens e desvantagens. Conclui-se, com este estudo, que, por meio da metodologia proposta, é possível replicar alguns comportamentos intrínsecos do mercado (como a observação de tendências nas séries de preços devido a dependências implícitas, e a presença de “caudas pesadas” nas distribuições dos retornos) que são desprezados pela maioria dos modelos paramétricos utilizados hoje, bem como o efeito do uso dessas informações no preço de derivativos.

Palavras-chave: Causalidade, Séries temporais financeiras, Simulação de Monte Carlo, Precificação de Opções, Projeções de mercado.

ABSTRACT

RODRIGUES, D. B. **On the application of causality measures for Monte Carlo simulations as alternative to price option contracts.** 2017. 70 p. Dissertação (Mestrado – Mestrado Profissional em Matemática, Estatística e Computação Aplicadas à Indústria) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

This paper proposes the use of causality measures applied over time-series of financial values to determine the dependency relations between market assets and a way to use the obtained measures to make inferences about the dynamics of these assets. This methodology defines a predictor for values of the time-series that, by determining the empirical probability distributions of the errors generated by this predictor based on the Kernel method, allows a random sampling of multivariate scenarios with many applications. The assets considered for the causality tests are the **Ibovespa** index, the dollar-real parity value **USDBRL** (using their price and price-return series), in addition to the daily traded interest rate (**CDI**). The use of the Monte Carlo Method (MMC) for the pricing of European call options (USDBRL) and **Ibovespa** was discussed, in addition to a comparison of the results generated by this methodology with values calculated by the Black-Scholes formula (currently the most used method by finance institutions), showing its advantages and disadvantages. The conclusion is that, based on the proposed methodology, it is possible to replicate some intrinsic market behaviors (such as the existence of trends in price series, due to implicit dependencies, and the presence of “fat tails” in the distributions of price-returns) that are neglected by most of parametric models, currently, as well as the effect of using this information for pricing derivatives.

Keywords: Causality, Market time-series, Monte Carlo Simulation, Options Pricing, Market Projections.

LISTA DE ILUSTRAÇÕES

Figura 1 – Correlação entre o número de mortes por afogamento e o número de filmes onde Nicholas Cage apareceu: 0.66604.	30
Figura 2 – Correlação entre a taxa de divórcio no Maine e o consumo <i>per capita</i> de margarina: 0.99255.	30
Figura 3 – Correlação entre preços da Coca-Cola (EUA) e da Companhia Paranaense de Energia (Paraná, Brasil): 0.96.	31
Figura 4 – Correlação entre preços do Bank of America (EUA) e da Advantest Corp (Japão): 0.95.	31
Figura 5 – Histograma dos resíduos de USDBRL.	46
Figura 6 – Histograma dos resíduos de retornos de Ibov.	47
Figura 7 – Histograma dos resíduos de retornos de USDBRL.	47
Figura 8 – Histograma dos resíduos de CDI.	48
Figura 9 – Histograma dos resíduos de Ibov.	48
Figura 10 – Média e percentis entre 10% e 90% das trajetórias, comparados ao mercado de USDBRL.	49
Figura 11 – Média e percentis entre 10% e 90% das trajetórias, comparados ao mercado de Ibov.	50
Figura 12 – Média e percentis entre 10% e 90% das trajetórias, comparados ao mercado de USDBRL.	50
Figura 13 – Média e percentis entre 10% e 90% das trajetórias, comparados ao mercado de CDI.	51
Figura 14 – Média e percentis entre 10% e 90% das trajetórias, comparados ao mercado de Ibov.	51
Figura 15 – Preços simulados e reais de opções de compra europeia de USDBRL para prazo $T = 2$ (valores na moeda real).	58
Figura 16 – Preços simulados e reais de opções de compra europeia de USDBRL para prazo $T = 24$ (valores na moeda real).	58
Figura 17 – Preços simulados e reais de opções de compra europeia de USDBRL para prazo $T = 45$ (valores na moeda real).	59
Figura 18 – Preços simulados e reais de opções de compra europeia de USDBRL para prazo $T = 68$ (valores na moeda real).	59
Figura 19 – Preços simulados e reais de opções de compra europeia de IBOV para prazo $T = 12$ (valores na moeda real).	60

Figura 20 – Preços simulados e reais de opções de compra europeia de IBOV para prazo $T = 37$ (valores na moeda real).	60
Figura 21 – Preços simulados e reais de opções de compra europeia de IBOV para prazo $T = 77$ (valores na moeda real).	61
Figura 22 – Preços simulados e reais de opções de compra europeia de IBOV para prazo $T = 122$ (valores na moeda real).	61

LISTA DE TABELAS

Tabela 1 – Valores de $\bar{C}(X_t H^l(X_t)) \pm \sigma(C(X_t H^l(X_t)))$ calculados para $lag\ l$, para cada cenário.	40
Tabela 2 – Valores da correlação de Pearson calculados para $lag\ l$, para cada cenário.	42
Tabela 3 – Valores da correlação de Spearman calculados para $lag\ l$, para cada cenário.	42
Tabela 4 – Variáveis de mercado utilizadas.	44
Tabela 5 – Índices de Geweke calculados para cada teste.	45
Tabela 6 – Variáveis que causam $X_0(0)$	45
Tabela 7 – Valores Iniciais da Simulação, tomados em 28/12/2015.	49
Tabela 8 – Contratos das opções europeias de compra de USDBRL simulados para a data 28/12/2015 (instante $t_0 = 0$).	53
Tabela 9 – Contratos das opções europeias de compra de IBOV simulados para a data 28/12/2015 (instante $t_0 = 0$).	54
Tabela 10 – Preços das opções europeias de compra de USDBRL simulados para a data 28/12/2015.	55
Tabela 11 – Preços das opções europeias de compra de IBOV simulados para a data 28/12/2015.	56
Tabela 12 – Preços das opções europeias de compra de USDBRL Bloomberg para a data 28/12/2015.	57
Tabela 13 – Preços das opções europeias de compra de IBOV Bloomberg para a data 28/12/2015.	57

SUMÁRIO

1	INTRODUÇÃO	21
1.1	O Modelo de Black-Scholes e suas premissas	23
1.2	O método de Monte Carlo	25
1.3	Lema de Itô	26
2	MODELO PROPOSTO	29
2.1	Causalidade	29
2.1.1	<i>Causalidade de Granger</i>	33
2.1.2	<i>Abordagem do problema</i>	34
2.2	Definição de um modelo dinâmico que captura os efeitos medidos da causalidade	36
2.2.1	<i>Mapeamento dos efeitos causais entre multivariáveis e geração de cenários</i>	36
2.2.2	<i>Método Kernel para distribuição de probabilidades</i>	37
2.3	Precificação de derivativos	37
3	RESULTADOS	39
3.1	Testes sobre dados sintetizados	39
3.2	Testes sobre dados de mercado	43
3.2.1	<i>Escolha das variáveis de mercado</i>	43
3.2.2	<i>Causalidade em variáveis de mercado</i>	44
3.3	Cenários simulados e distribuições medidas	46
3.4	Cálculo de preço de derivativo	52
4	DISCUSSÕES E CONCLUSÃO	63
5	CONSIDERAÇÕES FINAIS	67
	REFERÊNCIAS	69

INTRODUÇÃO

Atualmente, a teoria mais utilizada na precificação de opções é a desenvolvida por Fischer Black e Myron Scholes em 1973 (BLACK; SCHOLES, 1973). O modelo de Black-Scholes assume que os retornos gerados pelas variações de preço de um determinado produto respeitam um processo estocástico dado pela seguinte equação diferencial estocástica

$$\frac{dS(t)}{S(t)} = \mu \cdot dt + \sigma \cdot dW(t), \quad (1.1)$$

em que $S(t)$ representa o preço do produto em função do tempo e está definido em um espaço de probabilidades $(\Omega, \mathcal{F}, \mathbb{P})$, μ é a média dos retornos dos preços, também conhecida como *drift*, $dW(t)$ representa um processo de Wiener (também conhecido como Movimento Browniano Geométrico), adaptado à filtração $\{\mathcal{F}(t) : 0 \leq t \leq T\}$, sendo a parte estocástica da equação, e σ é o desvio-padrão do processo de Wiener (conhecido, principalmente no mercado financeiro, como volatilidade do produto).

Na Seção 1.1, encontra-se, em mais detalhes, o desenvolvimento do modelo de Black-Scholes, onde são descritas as premissas utilizadas na dinâmica de preço dos ativos, para deduzir uma relação analítica para precificar um determinado conjunto de opções. Dentre tais premissas, está o fato da série temporal dos retornos dos ativos serem independentes entre si, de forma que a distribuição desses retornos seja gaussiana. Entretanto, observando-se os preços de opções (líquidas) e o comportamento de séries de preços de ativos, é constatado que a distribuição de retornos no tempo apresenta “caudas pesadas”, ou seja, a frequência de eventos extremos não é explicada por uma distribuição gaussiana (BAK; PACZUSKI; SHUBIK, 1997).

Outra hipótese utilizada é a não existência de arbitragem que, mesmo sendo necessária como base de toda teoria de precificação, em casos de mercado com baixa liquidez (comum no mercado brasileiro), existem eventos de arbitragem locais, tais que as ofertas de preço de determinados ativos podem demorar a se formar.

Ao observar, também, os preços de diferentes contratos de opções de compra europeias formados no mercado, nota-se uma dependência com o preço de exercício (*Strike*) K que não reflete a relação descrita pelo modelo. Assim, a metodologia prática (CONT; HAMIDA, 2004) normalmente utilizada para que seja possível a precificação destes derivativos de forma coerente com as observações de mercado, é calibrar o valor do parâmetro volatilidade observada, invertendo-se a equação de Black-Scholes, gerando uma relação $\sigma = \sigma(S(t), t, K)$, de modo a minimizar as diferenças entre os preços observados em mercado e os preços calculados pela fórmula de Black-Scholes. A distorção da volatilidade é o que gera o conhecido *smile* (sorriso) de volatilidade, que, apesar de ser observado, não é intrínseco ao modelo de Black-Scholes.

Essa aproximação motiva tentar encontrar um modelo que explique a dinâmica dos preços de opções praticados no mercado. A proposta deste trabalho é a inserção de um fator de **causalidade** nos processos que regem os preços negociados dos ativos subjacentes aos contratos de opção, de modo que as distribuições geradas por meio de simulações numéricas desses processos sejam condicionadas a fatores econômicos “mensuráveis”. Em outras palavras, é de interesse deste trabalho quantificar as relações de dependência entre grandezas de mercado, aumentando a previsibilidade de agentes que apresentem dinâmica temporal mais complexa, estimando o comportamento de agentes com dinâmica menos complexa. Por exemplo, determinar como as decisões de alteração da taxa de juros básica pela reunião do Copom afetam os valores de produtos de câmbio ou de ações listadas na bolsa pode ser vantajoso para melhor prever estes últimos, caso as estimativas feitas para as primeiras sejam mais confiáveis.

As relações de dependência são determinadas pela causalidade medida entre as séries temporais dos preços dos produtos e fatores econômicos submetidos ao estudo. As estratégias de definir uma métrica de causalidade para a quantificação das dependências entre as negociações dos diversos produtos subjacentes e outros fatores de mercado são discutidas em mais detalhes no Capítulo 2.

Após a identificação das relações de causalidade entre as séries temporais, cenários futuros das variáveis aleatórias, aqui representando os objetos em estudo, são gerados aleatoriamente, de modo que os dados simulados preservam as características de causalidade medidas. Uma abordagem similar para geração de cenários de mercado específico utilizando a relação entre preços de commodities e taxas de câmbio é feita por Rosolen, Araújo e Lyrio (2013).

A estimativa de preços de derivativos pode ser realizada po meio de uma simulação de Monte Carlo, utilizando os cenários simulados como insumos para determinar as trajetórias dos fluxos de caixa (ou *payoffs*) desses derivativos. Na Seção 2.3, uma introdução a esse tipo de produto é feita, onde é discutido como os preços podem ser calculados a partir destas simulações. De modo a comparar os resultados obtidos com os valores observados no mercado, é escolhido um derivativo líquido, cujo preço responda de forma imediata a oscilações de mercado.

Todos os resultados obtidos são apresentados no Capítulo 3, que está dividido entre os testes sobre dados sintetizados realizados para estudo da abrangência e confiabilidade do modelo

proposto, e os testes sobre dados reais observados em mercado, a partir dos quais é realizado o estudo sobre a precificação de derivativos.

Por fim, no [Capítulo 4](#) e no [Capítulo 5](#), encontra-se uma discussão sobre os resultados obtidos neste trabalho, com as críticas em relação às vantagens e limitações das metodologias aplicadas e evoluções no modelo que podem ser estudadas para possíveis trabalhos futuros.

1.1 O Modelo de Black-Scholes e suas premissas

Uma das abordagens para se determinar o preço de um derivativo é utilizar a estratégia do “portfólio replicante”, conforme descrito nas primeiras seções de [Karoui, Jeanblanc-Picquè e Shreve \(1998\)](#). Tal estratégia consiste em definir um portfólio cujo valor, no tempo, replique o derivativo. Dessa forma, pode-se dizer que o custo inicial desse portfólio é equivalente ao preço desse derivativo.

Tomando-se, por exemplo, uma opção de compra europeia (o portador dessa opção pode exercer seu direito à compra da ação somente no instante de vencimento T), com preço de exercício (*strike*) K , que apresenta uma função *payoff* dada por

$$g(S(t)) = \max[S(t) - K, 0]. \quad (1.2)$$

Considera-se, então, um portfólio composto por $\Delta(t)$ ações e um valor $M(t)$, um empréstimo feito à taxa livre de risco r . O valor deste portfólio é dado por

$$\tilde{\Pi}(t) = \Delta(t) \cdot S(t) + M(t). \quad (1.3)$$

Em um mercado livre de arbitragem, existe uma estratégia $\Pi(t)$ que satisfaz a relação de portfólio replicante dada por

$$\Pi(t) - g(S(t)) \equiv 0. \quad (1.4)$$

Também é adotada a premissa da existência de um espaço de probabilidades de risco neutro, ou seja, é possível denotar o empréstimo à taxa livre de risco pela relação

$$M(t) = e^{r \cdot t} \quad (1.5)$$

e reescrever a equação (1.1) como

$$\frac{dS(t)}{S(t)} = r \cdot dt + \sigma \cdot dW(t), \quad (1.6)$$

de modo que a relação

$$\frac{S(t)}{M(t)} = S(0) \cdot e^{\int_0^t \sigma dW(u) - \frac{1}{2} \int_0^t \sigma^2 du} \quad (1.7)$$

é um *martingale*¹.

A solução da equação diferencial (1.6) é conhecida, definindo-se uma condição inicial $S(t = 0) = S(0)$, tal que o preço da ação é conhecido. A relação

$$S(t) = S(0) \cdot e^{(r - \frac{1}{2}\sigma^2) \cdot t + \sigma \cdot W(t)} \quad (1.8)$$

é determinada por meio da aplicação do **Lema de Itô**, apresentado na [Seção 1.3](#).

Define-se que $\Pi(t)$ é um portfólio **autofinanciável**² se satisfaz a seguinte equação diferencial

$$d\Pi(t) = r \cdot (\Pi(t) - \Delta(t) \cdot S(t)) dt + \Delta(t) \cdot dS(t). \quad (1.9)$$

Tal equação tem como solução a relação

$$\frac{\Pi(t)}{M(t)} = \Pi(0) \int_0^t \Delta(u) \cdot d\left(\frac{S(u)}{M(u)}\right), \quad (1.10)$$

que, devido a (1.7) ser um *martingale*, também é um *martingale*.

Assim, o valor inicial do portfólio replicante autofinanciável (e, por construção, o preço da opção) pode ser escrito em função do valor esperado ($\mathbf{E}[\cdot]$) do *payoff*, de acordo com a seguinte relação

$$\Pi(0) = \mathbf{E} \left[\frac{g(S(T))}{M(T)} \middle| \mathcal{F}(t = 0) \right]. \quad (1.11)$$

Substituindo a solução (1.8) em (1.11), tem-se a seguinte relação,

$$\Pi(0) = e^{-r \cdot T} \cdot \mathbf{E} \left[g \left(S(0) \cdot e^{(r - \frac{1}{2}\sigma^2) \cdot T + \sigma \cdot W(T)} \right) \middle| \mathcal{F}(t = 0) \right]. \quad (1.12)$$

Dado que $\frac{W(T)}{\sqrt{T}}$ é uma variável aleatória independente com distribuição gaussiana $G(0, 1)$, com média 0 e variância 1, o valor esperado pode ser escrito de acordo com

$$\bar{x} = \left(r - \frac{1}{2}\sigma^2 \right) \cdot T + \sigma\sqrt{T} \cdot \frac{W(T)}{\sqrt{T}}, \quad (1.13)$$

e

$$\Pi(0) = e^{-r \cdot T} \cdot \int_{-\infty}^{+\infty} g \left(S(0) \cdot e^{\bar{x}} \right) \cdot n \left[\bar{x}; \left(r - \frac{1}{2}\sigma^2 \right) \cdot T; \sigma\sqrt{T} \right] d\bar{x}, \quad (1.14)$$

¹ Processo estocástico $X(t)$ tal que, para um determinado instante τ até o qual o processo estocástico ocorreu, o valor esperado desse processo para qualquer instante $t > \tau$ é seu valor em $t = \tau$, ou seja, $\mathbf{E}[X(t > \tau) | \mathcal{F}(\tau)] = X(\tau)$

² Portfólio em que não existe entrada ou saída de dinheiro, além do valor inicial $\Pi(t = 0)$

em que $n(x; \mu; s)$ é a função de densidade de probabilidade gaussiana, com média μ e variância s^2 , tomada no ponto x .

Aplicando-se a função *payoff* de uma opção de compra europeia, dada pela relação (1.2), o preço C dessa opção é dado por

$$C = S(0) \cdot N(d^+) - K \cdot e^{-r \cdot T} \cdot N(d^-), \quad (1.15)$$

sendo

$$d^\pm = \frac{1}{\sigma \cdot \sqrt{T}} \cdot \left[\ln \left(\frac{S(0)}{K} \right) + \left(r \cdot T \pm \frac{\sigma^2 \cdot T}{2} \right) \right], \quad (1.16)$$

em que T é o tempo até o vencimento da opção e $N(x)$ é a função de distribuição normal acumulada, aplicada no ponto x .

Dentre as premissas adotadas para a obtenção da **fórmula de Black-Scholes** para precificação de opções europeias de compra dada pela relação (1.15), está o fato da dinâmica do preço do ativo objeto seguir a equação diferencial (1.1), que representa que o *log* dos retornos gerados por esse preço são independentes entre si e normalmente distribuídos. Outra premissa adotada é a total ausência de arbitragem do mercado, ou seja, a probabilidade de se encontrar uma estratégia de lucro ou prejuízo certo é zero. Tal premissa é razoável e muito precisa considerando mercados com liquidez muito alta, onde ações de compra e venda de ativos financeiros não influenciem a dinâmica do mercado. Finalmente, a premissa de que a volatilidade σ é constante também é considerada na determinação dos resultados. Essa premissa pode ser relaxada considerando que $\sigma = \sigma(S(t), t)$ pode seguir um processo estocástico, conforme abordado por [Karoui, Jeanblanc-Picqu e Shreve \(1998\)](#).

1.2 O método de Monte Carlo

O nome “Monte Carlo” foi sugerido pelo físico Nicholas Metropolis baseado no cassino com mesmo nome em Mônaco, devido a sua relação com problemas probabilísticos. O método se popularizou nos campos da física na Segunda Guerra Mundial, onde foi determinante no projeto Manhattan e no desenvolvimento da bomba de hidrogênio (ver [Eckhardt \(1987\)](#)).

O Método de Monte Carlo (MMC) é uma técnica computacional baseada em amostragem aleatória de modo a obter soluções numéricas de problemas, a princípio, determinísticos, mas com muitas dimensões, tal que uma solução analítica para esse problema não pode ser (ou ainda não foi) encontrada. Sua importância para diversas áreas da ciência é grande o suficiente para ter incentivado a evolução da computação eletrônica e suas aplicações englobam desde sistemas físicos com muitos graus de liberdade (como sistemas gasosos abordados pela mecânica estatística), passando pelas áreas de ciências sociais (na simulação da interação entre populações para estudo da clusterização observada nos povos) até as áreas de finanças (no estudo do

comportamento das séries temporais). Algumas aplicações específicas são abordadas por [Kroese et al. \(2014\)](#).

O princípio que se baseia o MMC é a lei dos grandes números, tal que valores esperados de variáveis aleatórias, que se resumem a soluções de integrais, podem ser aproximados pelo cálculo de médias sobre amostras aleatórias e independentes. Conhecendo-se as distribuições de probabilidade das variáveis a serem amostradas, o desafio computacional é a geração da maior quantidade de números aleatórios sob tal distribuição, o mais rápido possível (pois as médias resultantes das simulações do MMC convergem infinitamente para o valor esperado).

As simulações numéricas utiliza algoritmos pseudoaleatórios (as sequências começam a se repetir depois de um determinado número de amostragens e são determinísticos dado uma semente inicial). Para a maior parte dos problemas a serem resolvidos utilizando o MMC, essa limitação não causa problemas nos resultados, contanto que alguns cuidados sejam tomados no uso da técnica, como exemplo, limitar as amostras antes que as sequências geradas comecem a se repetir e que os dados gerados não sejam correlacionados.

1.3 Lema de Itô

O **Lema de Itô** é uma extensão da “regra da cadeia” (também conhecida como “técnica da integral por partes”) do cálculo diferencial aplicado a funções determinísticas para funções estocásticas. A identidade foi obtida por Kiyosi Itô em 1944 (ver [Itô \(1944\)](#)), de onde o nome do lema foi definido.

Seja $W(u)$ um processo de Wiener e $V(t) = f(W(t), T - t)$, com $f(x, t)$ duplamente diferenciável em x , diferenciável uma vez em t e definida para todo $x \in (0, \infty)$ e $t \in [0, T]$. Então, o Lema de Itô é descrito pela relação

$$V(t) = V(0) + \int_0^t f_x(x, u) dx(u) - \int_0^t f_t(x, u) du + \frac{1}{2} \int_0^t f_{xx}(x, u) d[f]_u, \quad (1.17)$$

em que as derivadas da função f são representadas por $\frac{df}{dx}(x, t) = f_x(x, t)$, $\frac{d^2f}{dx^2}(x, t) = f_{xx}(x, t)$ e $\frac{df}{dt}(x, t) = f_t(x, t)$. O termo $[f]_t$ é a variação quadrática do processo de Itô $f(x, t)$. Este termo, para o caso de funções determinísticas, é igual a zero, tal que esta relação seja a “regra da cadeia” da integral de Riemann.

Considerando o processo estocástico (1.6) e tomando $f(x) = \log(S)$, as derivadas da função f são dadas por $f_x(S) = S^{-1}$, $f_{xx}(S) = -S^{-2}$ e $f_t(S) = 0$. A variação quadrática de $f(x)$ é $[f]_t = \int_0^t (\sigma \cdot S(u))^2 du$. Assim, pode-se escrever $S(t) = e^{f(x)}$ de acordo com

$$S(t) = e^{\log(S(0)) + \int_0^t \frac{1}{S(u)} \cdot S(u) \cdot \sigma \cdot dW(u) + \frac{1}{2} \int_0^t -\frac{1}{S(u)^2} \cdot (\sigma S(u))^2 du},$$

$$S(t) = S(0) \cdot e^{r \cdot t + \sigma \cdot W(t) - \frac{1}{2} \sigma^2 t}. \quad (1.18)$$

MODELO PROPOSTO

Este capítulo é dividido em três partes: a primeira consiste em propor uma métrica para determinar as relações de causalidade entre duas ou mais séries temporais; em seguida, apresentar uma metodologia de geração de cenários para essas séries que preservam as relações de causalidade medidas será apresentada; e, por fim, será definida uma estratégia de precificação de derivativos baseada nos cenários gerados.

2.1 Causalidade

Os modelos empregados para descrever a dinâmica dos preços e retornos na geração de cenários quando há mais de uma variável não determinística no sistema usualmente utiliza matrizes de covariância entre tais variáveis, de modo que as simulações de eventos futuros tenham essas correlações intrínsecas nos dados (ver [Ortega *et al.* \(2009\)](#) e [Huber e Maric \(2013\)](#)). É de grande interesse a utilização não de correlações entre ativos, mas de uma medida de **causalidade** entre eles. O objetivo é evitar conclusões espúrias em relação ao comportamento condicional das variáveis aleatórias.

Analogamente, conclusões como “assistir filmes do Nicholas Cage aumenta o risco de morte por afogamento”, ou “consumir muita margarina pode ser prejudicial para o casamento” seriam evitadas (ver [Figura 1](#) e [Figura 2](#), onde os dados foram extraídos de <http://www.tylervigen.com/spurious-correlations>).

Figura 1 – Correlação entre o número de mortes por afogamento e o número de filmes onde Nicholas Cage apareceu: 0.66604.

Fonte: Elaborada pelo autor.

Figura 2 – Correlação entre a taxa de divórcio no Maine e o consumo *per capita* de margarina: 0.99255.

Fonte: Elaborada pelo autor.

Em relação a mercado financeiro, o objetivo é responder perguntas como “a ação X subiu 10% devido à queda na paridade da moeda Y, ou a paridade da moeda Y caiu devido à subida da ação X? Ou existe algum outro fator (como a decisão de diminuir a taxa de juros em 0.25%) que

tenha implicado ambos os movimentos?”. Mesmo dados altamente correlacionados podem não representar dependência mútua nenhuma (ver Figura 3 e Figura 4). Como é normalmente

Figura 3 – Correlação entre preços da Coca-Cola (EUA) e da Companhia Paranaense de Energia (Paraná, Brasil): 0.96.

Fonte – <<http://www.businessinsider.com/15-spurious-stock-correlations-2012-4>>

Figura 4 – Correlação entre preços do Bank of America (EUA) e da Advantest Corp (Japão): 0.95.

Fonte – <<http://www.businessinsider.com/15-spurious-stock-correlations-2012-4>>

Nestes exemplos, as correlações são calculadas, como geralmente é feito no mercado financeiro, sobre os retornos diários dos preços das ações. Embora os pares de empresas apre-

sentados não tenham, a princípio, nenhuma relação mútua (sendo o caso da Advantest Corp e do Bank of America empresas cujas ações são negociadas com uma diferença de quase 12h, devido às diferenças de fuso entre Estados Unidos e Japão), suas correlação são altas (maiores que 0.95).

Existem diversas maneiras de determinar, de forma quantitativa, relações de causalidade entre eventos. Neste trabalho, os conceitos de causalidade formalizados por Granger ([GRANGER, 1969](#)), detalhados na [Subseção 2.1.1](#), são discutidos e utilizados. Entretanto, existem muitos trabalhos que usam métodos alternativos ao inicialmente proposto por Granger, mas que mantém, em sua essência, os mesmos conceitos, por exemplo a utilização de co-integração, como proposto por [Bracegirdle e Barber \(2012\)](#), ou, em uma abordagem Bayesiana, a transferência de informação, como discutido em [Zaremba e Aste \(2014\)](#). Tais conceitos não serão tratados nesse trabalho, embora a comparação das eficiências das diferentes medidas de causalidade esteja nas perspectivas futuras deste trabalho.

Na [Subseção 2.1.2](#), é discutida a abordagem dos estudos de causalidade entre séries temporais. Inicialmente, a métrica proposta é testada sobre dados sintetizados para entender quais tipos de relações tal medida pode identificar e quais suas limitações.

Usualmente, as técnicas de medida de causalidade são aplicadas em sistemas onde pode-se controlar as amostras. Um dos casos mais comuns é o estudo da eficácia de um novo composto químico no combate de uma doença: para isso, não basta selecionar um grupo de pessoas (por maior que seja) que possuem determinada doença, e analisar o resultado da utilização desse composto depois de determinado tempo. É necessário criar um grupo de controle de pessoas que possuem a doença, mas que não serão tratadas (placebo), pois uma eventual cura, pode ter sido causada por um fator externo. E mesmo a seleção dos indivíduos de cada grupo deve ser feita de modo a evitar padrões ou tendências (por exemplo, a média de idade de um grupo ser de 50 anos e a média de idade do grupo de controle ser de 20 anos pode criar um viés nos resultados; ou o percentual entre homens e mulheres; ou até mesmo variações da mesma doença em cada indivíduo).

No mercado financeiro, não é possível controlar o ambiente para que experimentos sejam feitos (não é legalmente permitido forçar uma queda no preço de determinada ação para estudo dos efeitos no mercado de câmbio e vice-versa). O desafio é, então, separar as observações de maneira conveniente para que as medidas de causalidade sejam feitas minimizando tanto quanto possível qualquer viés (como evitar usar dados históricos somente em momentos de estresse no mercado), tendo amostras estatisticamente relevantes para os eventos de interesse. Por exemplo, tentar medir a causalidade entre aumentos de taxa de juros e movimentações na taxa de câmbio quando todas as amostras de dados em que as taxas de juros aumentam, o valor do índice de ações também aumenta implica em uma relação possivelmente espúria de causa (pois o índice de ações, e não a taxa de juros, pode estar causando a taxa de câmbio). Dessa forma, a amostra não é estatisticamente relevante para a medida entre esse par.

2.1.1 Causalidade de Granger

Primeiramente, é necessário definir relações de causalidade no que diz respeito a séries temporais. A ideia de causalidade proposta por [Granger \(1969\)](#) e referenciada também por [Eichler \(2007\)](#) se baseia em duas premissas básicas: efeitos são sempre precedidos por causas; e a causa contém informação única sobre o efeito, que não está presente em nenhum outro lugar.

Para formalizar o conceito de causalidade, algumas notações são importantes. Seja X_t^i um vetor (indexado por $i = 1, 2, \dots, M$) de M processos estocásticos com domínio o tempo indexado por t . Denota-se $X^i(k)$ toda a informação gerada pelo processo X_t^i até o instante k , ou seja, $X^i(k) = \{x^i(k), x^i(k-1), x^i(k-2), \dots, x^i(k-\infty)\}$.

Pode-se definir então uma filtragem de X_t^j como $F(X_t^j) = X^i(\tau) \cup X^j(\tau-1)$, com $i \neq j$ (toda a informação gerada pelo processo X_t^j até o instante $\tau-1$ e pelos outros processos até o instante τ).

Seja $F(Y_t^j)$ um subconjunto de $F(X_t^j)$, e $H(X_t^j) = F(X_t^j) - F(Y_t^j)$ toda a informação gerada pelo processo X_t^j até o instante τ excluindo a informação gerada pelo processo Y_t^j até o instante τ .

Seja, agora, um predictor de X_t^i que considera toda a informação do universo até t sendo definido por $P(X_t^i|U_t) = P(X_t^i|F(X_t^i))$. Define-se então uma variável aleatória $\varepsilon(X_t^i|F(X_t^i)) = x^i(t) - P(X_t^i|F(X_t^i))$ como o resíduo desse predictor, com variância dada por $\sigma^2(X_t^i|F(X_t^i))$.

Finalmente, define-se que $Y_t^i \rightarrow X_t^i$ (Y causa X , por Granger) se a relação $\sigma^2(X_t^i|F(X_t^i)) < \sigma^2(X_t^i|H(X_t^i))$. De outra forma, Y causa X se for possível prever X de maneira melhor se toda a informação for usada, do que se ignorar a informação gerada por Y . A mensuração dessa relação de causalidade é feita por meio do Índice de Geweke ([ZAREMBA; ASTE, 2014](#)), definido pela expressão

$$C(X_t^i|H(X_t^i)) = \ln \left(\frac{\sigma^2(X_t^i|H(X_t^i))}{\sigma^2(X_t^i|F(X_t^i))} \right). \quad (2.1)$$

Neste trabalho, para o predictor $P(X_t^i|F(X_t^i))$ é utilizado o vetor auto-regressivo dado por

$$P(X_t^i|F(X_t^i)) = \sum_{j=1}^{j=M} \sum_{p=0}^{p=\infty} \alpha(x_{t-p}^j) \cdot x_{t-p}^j + \varepsilon(X_t^i|F(X_t^i)), \quad (2.2)$$

com $\alpha(x_t^i) = 0$.

Da mesma forma, o predictor $P(X_t^i|H(X_t^i))$ é dado por

$$P(X_t^i|H(X_t^i)) = \sum_{j=1}^{j=M} \sum_{p=0}^{p=\infty} \alpha(x_{t-p}^j) \cdot x_{t-p}^j + \varepsilon(X_t^i|H(X_t^i)), \quad (2.3)$$

com $\alpha(x_t^i) = 0$ e $\alpha(x_{t-p}^j) = 0 \ \forall j, p \text{ tq. } x_{t-p}^j \notin H(X_t^i)$.

2.1.2 Abordagem do problema

A estratégia de medida de causalidade desenvolvida na [Subseção 2.1.1](#) é, primeiramente aplicada sobre séries temporais unidimensionais sintetizadas, permitindo estudos mais aprofundados em relação à abrangência dessa teoria. Importante definir o termo *lag* utilizado a partir desta seção. Tal termo tem a definição de **defasagem** no tempo de uma série discreta. Dessa forma, vetores de dados são gerados sob alguns cenários:

- Cenário 1: relações lineares de dependência

Neste cenário, toma-se a série $X_t = \{x(t), x(t-1), \dots, x(t-p)\}$ tal que:

$x(t) = \sum_{i=1}^p \alpha_x(i) \cdot x(t-i) + \xi(t)$, tal que $\alpha_x(i) \in \mathbb{R}$, i é denominado *lag* e $\xi(t)$ é uma variável aleatória qualquer, que represente um ruído na série¹.

Pode-se então definir alguns casos dentro do Cenário 1:

- Cenário 1.1

$p = 8$, $\alpha_x(1) = -0.7$, $\alpha_x(3) = 0.5$, $\alpha_x(i) = 0.0$ (para $i \notin \{1, 3\}$) e $\xi(t) = U(-0.5, 0.5)$ uma variável aleatória com distribuição uniforme entre -0.5 e 0.5 .

- Cenário 1.2

$p = 8$, $\alpha_x(1) = -0.7$, $\alpha_x(3) = 0.5$, $\alpha_x(i) = 0.0$ (para $i \notin \{1, 3\}$) e $\xi(t) = \frac{\Gamma(2, 0.5) - 1}{4}$ uma variável aleatória com distribuição Gamma, com parâmetros $\alpha = 2$ e $\beta = 0.5$ deslocado de modo que seus valores sejam em torno do valor 0 e de modo que $|X_t| < \infty, \forall t$.

- Cenário 2: relações não-lineares de dependência

Neste cenário, toma-se a série $X_t = \{x(t), x(t-1), \dots, x(t-p)\}$ tal que:

$x(t) = \sum_{i=1}^p \alpha_x(i) \cdot x(t-i)^{\beta_x(i)} + \sum_{j,k=1}^p \alpha_{xx}(j,k) \cdot x(t-j) \cdot x(t-k) + \xi(t)$, tal que $\alpha_x(i) \in \mathbb{R}$, $\alpha_{xx}(j,k) \in \mathbb{R}$, i, j, k são os *lags*, $\xi(t)$ é uma variável aleatória qualquer que represente um ruído na série e $\beta_x(i) \in \mathbb{R}$ é um expoente que, se diferente de um, gera a não-linearidade das dependências da série.

Pode-se então definir alguns casos dentro do Cenário 2:

- Cenário 2.1

$p = 8$, $\alpha_x(1) = -0.7$, $\alpha_x(3) = 0.5$, $\alpha_x(i) = 0.0$ (para $i \notin \{1, 3\}$); $\beta_x(1) = \frac{1}{3}$, $\beta_x(3) = 2.0$, $\beta_x(i) = 0.0$ (para $i \notin \{1, 3\}$); $\alpha_{xx}(j,k) = 0.0$ (para todo (j,k)); e $\xi(t) = U(-0.5, 0.5)$ uma variável aleatória com distribuição uniforme entre -0.5 e 0.5 .

- Cenário 2.2

$p = 8$, $\alpha_x(3) = 0.5$, $\alpha_x(i) = 0.0$ (para $i \notin \{3\}$); $\beta_x(3) = 2.0$, $\beta_x(i) = 0.0$ (para $i \notin \{3\}$); $\alpha_{xx}(1,2) = -0.7$, $\alpha_{xx}(j,k) = 0.0$ (para $(j,k) \notin \{(1,2)\}$); e $\xi(t) = U(-0.5, 0.5)$ uma variável aleatória com distribuição uniforme entre -0.5 e 0.5 .

¹ $\mathbb{R}$ representa o conjunto dos números reais.

– Cenário 2.3

$p = 8$, $\alpha_x(1) = 0.25$, $\alpha_x(2) = -0.25$, $\alpha_x(3) = 0.25$, $\alpha_x(4) = -0.25$, $\alpha_x(i) = 0.0$ (para $i \notin \{1, 2, 3, 4\}$); $\beta_x(1) = \frac{1}{3}$, $\beta_x(2) = \frac{2}{3}$, $\beta_x(3) = 1.0$, $\beta_x(4) = \frac{5}{3}$, $\beta_x(i) = 0.0$ (para $i \notin \{1, 2, 3, 4\}$); $\alpha_{xx}(j, k) = 0.0$ (para todo (j, k)); e $\xi(t) = U(-0.5, 0.5)$ uma variável aleatória com distribuição uniforme entre -0.5 e 0.5 .

– Cenário 2.4

$p = 8$, $\alpha_x(1) = 0.25$, $\alpha_x(2) = -0.25$, $\alpha_x(3) = -0.25$, $\alpha_x(4) = -0.25$, $\alpha_x(i) = 0.0$ (para $i \notin \{1, 2, 3, 4\}$); $\beta_x(1) = \frac{1}{3}$, $\beta_x(2) = \frac{2}{3}$, $\beta_x(3) = 1.0$, $\beta_x(4) = \frac{5}{3}$, $\beta_x(i) = 0.0$ (para $i \notin \{1, 2, 3, 4\}$); $\alpha_{xx}(j, k) = 0.0$ (para todo (j, k)); e $\xi(t) = \frac{\Gamma(2, 0.5)-1}{4}$ uma variável aleatória com distribuição Gamma, com parâmetros $\alpha = 2.0$ e $\beta = 0.5$ deslocado de modo que seus valores sejam em torno do valor 0 e de modo que $|X_t| < \infty$, $\forall t$.

– Cenário 2.5

$p = 8$, $\alpha_x(1) = 0.125$, $\alpha_x(2) = -0.125$, $\alpha_x(3) = 0.125$, $\alpha_x(4) = -0.125$, $\alpha_x(i) = 0.0$ (para $i \notin \{1, 2, 3, 4\}$); $\beta_x(1) = 1$, $\beta_x(2) = 2$, $\beta_x(3) = 3$, $\beta_x(4) = 4$, $\beta_x(i) = 0.0$ (para $i \notin \{1, 2, 3, 4\}$); $\alpha_{xx}(j, k) = 0.0$ (para todo (j, k)); e $\xi(t) = U(-0.5, 0.5)$ uma variável aleatória com distribuição uniforme entre -0.5 e 0.5 .

– Cenário 2.6

$p = 8$, $\alpha_x(1) = 0.125$, $\alpha_x(2) = -0.125$, $\alpha_x(3) = -0.125$, $\alpha_x(4) = -0.125$, $\alpha_x(i) = 0.0$ (para $i \notin \{1, 2, 3, 4\}$); $\beta_x(1) = 1$, $\beta_x(2) = 2$, $\beta_x(3) = 3$, $\beta_x(4) = 4$, $\beta_x(i) = 0.0$ (para $i \notin \{1, 2, 3, 4\}$); $\alpha_{xx}(j, k) = 0.0$ (para todo (j, k)); e $\xi(t) = \frac{\Gamma(2, 0.5)-1}{4}$ uma variável aleatória com distribuição Gamma, com parâmetros $\alpha = 2$ e $\beta = 0.5$ deslocado de modo que seus valores sejam em torno do valor 0 e de modo que $|X_t| < \infty$, $\forall t$.

● Cenário 3: relações sem dependência

Neste cenário, toma-se a série $X_t = \{x(t), x(t-1), \dots, x(t-p)\}$ tal que:

$x(t) = \xi(t)$, tal que $\xi(t)$ é uma variável aleatória qualquer, com média zero.

Pode-se então definir alguns casos dentro do Cenário 3:

- Cenário 3.1 $\xi(t) = U(-0.5, 0.5)$ uma variável aleatória com distribuição uniforme entre -0.5 e 0.5 .
- Cenário 3.2 $\xi(t) = \frac{\Gamma(2, 0.5)-1}{4}$ uma variável aleatória com distribuição Gamma, com parâmetros $\alpha = 2$ e $\beta = 0.5$ deslocado de modo que seus valores sejam em torno do valor 0.
- Cenário 3.3 $\xi(t) = LG(0, 1) - 1$ uma variável aleatória com distribuição Log-Normal com média 0 e variância 1, deslocado de modo que seus valores sejam em torno do valor 0.

Os coeficientes dos cenários são escolhidos de modo que a série X_t não seja divergente e as distribuições de probabilidade das variáveis aleatórias representem um ruído não necessariamente

Browniano e não necessariamente simétrico em torno de uma média igual a 0, como no caso das distribuições Gamma e Log-Normal utilizadas.

Para cada cenário descrito, são geradas 20 (vinte) amostras independentes de 1000 (mil) pontos cada. Assim, para cada amostra, calcula-se os índices de Geweke da seguinte forma

$$C(X_t|H^l(X_t)) = \ln \left(\frac{\sigma^2(X_t|H^l(X_t))}{\sigma^2(X_t|F^l(X_t))} \right), \quad (2.4)$$

em que o subconjunto $F^l = \{X_{t-1}, X_{t-2}, X_{t-3}, X_{t-4}\}$, e o subconjunto $H^l = F^l - \{X_{t-l}\}$ (o subconjunto F^l excluindo X_{t-l}), com $l \in \{1, 2, 3, 4\}$.

Tomando-se, para cada cenário, a média $\bar{C}(X_t|H^l(X_t))$ e o desvio padrão $\sigma(C(X_t|H^l(X_t)))$ do conjunto de índices calculados, define-se que existe causalidade da variável X com lag l , for possível rejeitar a hipótese nula de que o índice de Geweke é igual a 0 com valor- p p (ou grau de confiança $1 - p$), ou seja, se for satisfeita a relação

$$N\left(0; \bar{C}(X_t|H^l(X_t)); \sigma\left(C(X_t|H^l(X_t))\right)\right) < p \quad (2.5)$$

em que $N(x; \mu; s)$ é o valor da distribuição gaussiana com média μ e variância s^2 acumulada no ponto x .

2.2 Definição de um modelo dinâmico que captura os efeitos medidos da causalidade

2.2.1 Mapeamento dos efeitos causais entre multivariáveis e geração de cenários

Tomando os processos da forma X_t^i e todas as relações causais identificadas $(j, k) \rightarrow (i, t)$, tais que $C(X_t^i|H^{t-k}(X_t^j))$ satisfaçam a relação (2.5), sendo $t > k$ se $i = j$ e $t \geq k$ se $i \neq j$.

O previsor dos processos pode, então, ser escrito de acordo com a recursão

$$\hat{x}_t^i = P(X_t^i|H^{t-k}(X_t^j)) = \sum_{j,k} \alpha(x_{t-k}^j) \cdot x_{t-k}^j + \varepsilon(X_t^i|H(X_t^i)), \quad (2.6)$$

tal que os valores de $\alpha(x_{t-k}^j)$ são constantes determinadas pelo método dos mínimos quadrados (MMQ) da seguinte regressão linear

$$\hat{x}_t^i = \sum_{j,k} \left(\alpha(x_{t-k}^j) \cdot x_{t-k}^j + r_k^j \right), \quad (2.7)$$

em que r_k^j são os resíduos resultantes.

Os valores de $\varepsilon(X_t^i | H(X_t^i))$, por sua vez, são simulados por meio da geração de uma variável aleatória com distribuição igual à distribuição do vetor de resíduos r_k^j . A determinação da função densidade de probabilidade é feita de maneira empírica pelo método de Kernel, descrito na [Subseção 2.2.2](#).

Caso sejam identificados pares de relações causais da forma $X_t^j \rightarrow X_t^i$ e $X_t^i \rightarrow X_t^j$, a relação com menor índice de Geweke $C(X_t^i | H^t(X_t^j))$ é desconsiderada.

2.2.2 Método Kernel para distribuição de probabilidades

Para estimar as funções de densidade de probabilidade empíricas das variáveis, o método de Kernel ([AZZALINI, 1981](#)) (da sigla em inglês KDE) é utilizado. Esse método é efetivo em casos de amostras finitas de dados amostrais e resulta em uma função de densidade de probabilidade suave e contínua. Dessa forma, supõe-se que a função de distribuição $f(x)$ de uma variável aleatória $\vec{x}$ seja desconhecida. Tal função pode ser estimada por

$$\hat{f}(x) = \frac{1}{n \cdot h} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right), \quad (2.8)$$

em que n é número de dados na amostra, $K()$ é denominada *kernel*, uma função normalizada, com média zero e h é chamado de *largura da banda*, representando um parâmetro de suavização.

Neste trabalho, a função kernel será definida pela distribuição normal padrão e o parâmetro h é dado por

$$h = \left(\frac{4 \cdot \hat{\sigma}^5}{3 \cdot n} \right)^{\frac{1}{5}}, \quad (2.9)$$

escolhido de acordo com a densidade de dados em cada amostra (rule-of-thumb bandwidth estimator) ([SILVERMAN, 1986](#)) que otimiza a estimativa quando a distribuição real é gaussiana.

2.3 Precificação de derivativos

Considera-se um derivativo com *payoff* descrito por uma função $H(t) = H(X^i(t))$ que depende da trajetória de cada variável X^i até o instante t . O preço, ou valor justo, desse derivativo no instante t_0 é definido como o valor esperado do *payoff* no fim da vida do derivativo T , descontado a valor presente pelo custo de oportunidade (taxa r) ao se comprar esse derivativo pelo mesmo período de acordo com a seguinte relação

$$C(t_0) = \mathbf{E}^{t_0} \left[H(T) \cdot e^{-r(T) \cdot (T - t_0)} \right]. \quad (2.10)$$

O cálculo do valor esperado é feito computacionalmente utilizando o método de Monte Carlo, tal que os cenários das variáveis estocásticas são gerados conforme descrito na [Seção 2.2](#),

para cada cenário, os valores de *payoff* calculados para o instante T e descontados a valor presente, e a média aritmética sobre seus valores é tomada, determinando-se, assim, o preço desse derivativo de acordo com (2.10).

A função de *payoff* $H(t)$ é intrínseca às características do derivativo. Para a comparação dos resultados obtidos com valores praticados no mercado, esse trabalho foca no cálculo de opções de compra europeias (as chamadas *Calls*) de um único ativo subjacente. Esse derivativo possui função *payoff* de acordo com a seguinte expressão

$$H(T) = \max[0, S(T) - K], \quad (2.11)$$

em que K é denominado *Strike* e representa o valor pelo qual o derivativo dá direito ao portador comprar o ativo subjacente à opção no instante T , denominado “Vencimento”. Caso o preço $S(T)$ do ativo seja maior do que K , é vantajoso para o detentor desse direito comprar esse mesmo ativo do vendedor da opção (exercer a opção) e vendê-lo ao mercado, tal que seu lucro seja $S(T) - K$. Caso contrário, esse direito não é exercido e o comprador da opção não tem ganhos ou perdas no instante de vencimento (além do valor pago na ocasião da compra da opção).

RESULTADOS

Neste capítulo, os resultados obtidos pela aplicação do modelo definido no [Capítulo 2](#) são apresentados. Primeiramente, é feito um estudo sobre dados sintetizados para a observação da abrangência do modelo proposto. Em seguida, na [Seção 3.2](#), encontra-se uma apresentação sobre a escolha das variáveis de mercado para aplicação do estudo e os resultados das medidas de causalidade entre as mesmas. Na [Seção 3.3](#), os valores calculados são utilizados para a geração dos cenários de Monte Carlo, e as trajetórias e distribuições construídas são mostradas de forma gráfica. Por fim, na [Seção 3.4](#), é realizado o apreçamento de alguns contratos de opções europeias utilizando os cenários simulados e uma comparação com dados obtidos através de uma fonte de mercado.

3.1 Testes sobre dados sintetizados

De acordo com a metodologia descrita na [Subseção 2.1.2](#), os valores médios do índice de Geweke $\bar{C}(X_t|H^l(X_t))$ e suas respectivas incertezas dentro de um desvio-padrão, representado por $\sigma(C(X_t|H^l(X_t)))$ são determinados e apresentados na [Tabela 1](#).

Aplicando-se o critério do valor-p para determinar existência de causalidade com grau de confiança $1 - p = 99\%$, sendo esse o padrão normalmente adotado pelas instituições financeiras ao mensurar fatores que apresentam incerteza, as relações de causalidade para cada cenário são:

- **Cenário 1.1:**

- $X_{t-1} \Rightarrow X_t$ com índice de Geweke $\bar{C}(X_t|H^1(X_t)) = 0.400$
- $X_{t-3} \Rightarrow X_t$ com índice de Geweke $\bar{C}(X_t|H^3(X_t)) = 0.178$

Tabela 1 – Valores de $\bar{C}(X_t|H^l(X_t)) \pm \sigma(C(X_t|H^l(X_t)))$ calculados para *lag* l , para cada cenário.

Cenário	$l = 1$	$l = 2$	$l = 3$	$l = 4$
1.1	0.400 ± 0.036	0.002 ± 0.001	0.178 ± 0.028	0.001 ± 0.002
1.2	0.403 ± 0.023	0.001 ± 0.001	0.187 ± 0.032	0.001 ± 0.001
2.1	0.371 ± 0.018	0.034 ± 0.015	0.012 ± 0.013	0.000 ± 0.002
2.2	0.001 ± 0.002	0.002 ± 0.002	0.010 ± 0.008	0.001 ± 0.002
2.3	0.151 ± 0.017	0.017 ± 0.009	0.056 ± 0.009	0.029 ± 0.011
2.4	0.255 ± 0.022	0.004 ± 0.006	0.036 ± 0.012	0.002 ± 0.006
2.5	0.016 ± 0.006	0.002 ± 0.002	0.001 ± 0.001	0.001 ± 0.001
2.6	0.015 ± 0.007	0.002 ± 0.002	0.001 ± 0.001	0.001 ± 0.002
3.1	0.001 ± 0.002	0.001 ± 0.002	0.001 ± 0.001	0.001 ± 0.002
3.2	0.001 ± 0.001	0.001 ± 0.001	0.001 ± 0.001	0.001 ± 0.001
3.3	0.000 ± 0.003	0.000 ± 0.002	0.000 ± 0.002	0.000 ± 0.003

Fonte: Elaborada pelo autor.

- **Cenário 1.2:**

- $X_{t-1} \Rightarrow X_t$ com índice de Geweke $\bar{C}(X_t|H^1(X_t)) = 0.403$
- $X_{t-3} \Rightarrow X_t$ com índice de Geweke $\bar{C}(X_t|H^3(X_t)) = 0.187$

- **Cenário 2.1:**

- $X_{t-1} \Rightarrow X_t$ com índice de Geweke $\bar{C}(X_t|H^1(X_t)) = 0.371$

- **Cenário 2.2:**

- Nenhuma relação de causalidade para grau de confiança de 99%

- **Cenário 2.3:**

- $X_{t-1} \Rightarrow X_t$ com índice de Geweke $\bar{C}(X_t|H^1(X_t)) = 0.151$
- $X_{t-3} \Rightarrow X_t$ com índice de Geweke $\bar{C}(X_t|H^3(X_t)) = 0.056$
- $X_{t-4} \Rightarrow X_t$ com índice de Geweke $\bar{C}(X_t|H^4(X_t)) = 0.029$

- **Cenário 2.4:**

- $X_{t-1} \Rightarrow X_t$ com índice de Geweke $\bar{C}(X_t|H^1(X_t)) = 0.255$
- $X_{t-3} \Rightarrow X_t$ com índice de Geweke $\bar{C}(X_t|H^3(X_t)) = 0.036$

- **Cenário 2.5:**

- $X_{t-1} \Rightarrow X_t$ com índice de Geweke $\bar{C}(X_t|H^1(X_t)) = 0.016$

- **Cenário 2.6:**

- Nenhuma relação de causalidade para grau de confiança de 99%

- **Cenário 3.1:**

- Nenhuma relação de causalidade para grau de confiança de 99%

- **Cenário 3.2:**

- Nenhuma relação de causalidade para grau de confiança de 99%

- **Cenário 3.3:**

- Nenhuma relação de causalidade para grau de confiança de 99%

Os dois componentes de dependência linear dos cenários 1.1 e 1.2 podem ser identificados pela métrica de causalidade proposta, independentemente do componente aleatório apresentar distribuição simétrica centrada no 0 (no caso do cenário 1.1) ou não (no caso do cenário 1.2).

Para o cenário 2.1, apenas o componente referente ao elemento de expoente menor que 1, no caso $\frac{1}{3}$, pode ser mensurado, enquanto o componente de dependência quadrática tem o índice de Geweke compatível com o valor 0.

De forma consistente, o componente com dependência quadrática no $lag\ t - 3$, no cenário 2.2, também apresenta medida do índice compatível com 0, assim como no componente cruzado, os índices referentes aos $lags\ t - 1$ e $t - 2$, são compatíveis com o valor nulo.

Nos cenários 2.3 e 2.4, a dependência dos componentes com expoentes $\frac{1}{3}$ e 1 é identificada considerando ambas distribuições do termo aleatório simétrica e não-simétrica. Também, em ambos os casos, a dependência do componente com expoente $\frac{2}{3}$ não pode ser identificada com confiança de 99%. Já a dependência do componente de expoente igual a $\frac{5}{3}$ é identificada apenas no cenário 2.3, onde a distribuição de probabilidades do componente aleatório é simétrica (no caso, uniforme), com índice de Geweke da ordem de metade do índice determinado para o componente linear, com $lag\ t - 3$.

Em relação aos cenários 2.5 e 2.6, tais que os expoentes dos componentes são maiores ou iguais a 1, o índice de Geweke 0.016 medido para o componente com $lag\ t - 1$ mostra dependência apenas no caso do componente aleatório com distribuição simétrica (cenário 2.5). Tal valor está no limiar do grau de confiança de 99% considerado, dado que o mesmo componente linear do cenário 2.6 (com distribuição não-simétrica) possui índice de Geweke de 0.015, ficando fora do intervalo de confiança.

Finalmente, os cenários sintetizados sem nenhuma dependência (cenários 3.1, 3.2 e 3.3) apresentam índices compatíveis com 0 em todos os seus componentes, não gerando um falso positivo, mesmo considerando termos aleatórios com diferentes distribuições (no caso dos cenários 3.2 e 3.3, ambas as distribuições não são simétricas e não são centradas no valor 0).

De forma comparativa, as tabelas 2 e 3 mostram, respectivamente as correlações de Pearson e de Spearman, sobre as mesmas séries onde a causalidade foi medida.

Tabela 2 – Valores da correlação de Pearson calculados para $lag\ l$, para cada cenário.

Cenário	$l = 1$	$l = 2$	$l = 3$	$l = 4$
1.1	-0.638 ± 0.014	0.128 ± 0.019	0.411 ± 0.029	-0.610 ± 0.036
1.2	-0.636 ± 0.016	0.122 ± 0.017	0.418 ± 0.022	-0.606 ± 0.034
2.1	-0.764 ± 0.009	0.613 ± 0.014	-0.417 ± 0.031	0.263 ± 0.045
2.2	-0.039 ± 0.023	-0.035 ± 0.035	0.096 ± 0.040	-0.007 ± 0.040
2.3	0.452 ± 0.023	0.335 ± 0.029	0.332 ± 0.029	-0.087 ± 0.046
2.4	0.517 ± 0.018	0.173 ± 0.030	-0.174 ± 0.037	-0.317 ± 0.042
2.5	0.129 ± 0.025	0.034 ± 0.043	0.022 ± 0.024	0.017 ± 0.029
2.6	0.118 ± 0.028	-0.013 ± 0.028	-0.025 ± 0.027	-0.030 ± 0.033
3.1	0.005 ± 0.037	-0.010 ± 0.032	-0.002 ± 0.031	0.010 ± 0.032
3.2	-0.005 ± 0.027	0.010 ± 0.025	0.018 ± 0.031	0.002 ± 0.034
3.3	-0.001 ± 0.031	0.005 ± 0.032	0.002 ± 0.031	0.009 ± 0.035

Fonte: Elaborada pelo autor.

Tabela 3 – Valores da correlação de Spearman calculados para $lag\ l$, para cada cenário.

Cenário	$l = 1$	$l = 2$	$l = 3$	$l = 4$
1.1	-0.605 ± 0.019	0.135 ± 0.017	0.396 ± 0.027	-0.573 ± 0.038
1.2	-0.599 ± 0.018	0.118 ± 0.019	0.420 ± 0.026	-0.587 ± 0.038
2.1	-0.766 ± 0.009	0.636 ± 0.015	-0.432 ± 0.031	0.295 ± 0.046
2.2	-0.026 ± 0.024	-0.022 ± 0.034	0.097 ± 0.038	-0.003 ± 0.035
2.3	0.442 ± 0.024	0.336 ± 0.032	0.329 ± 0.030	0.099 ± 0.044
2.4	0.541 ± 0.020	0.227 ± 0.023	-0.127 ± 0.031	-0.286 ± 0.037
2.5	0.133 ± 0.024	0.042 ± 0.043	0.030 ± 0.024	0.027 ± 0.028
2.6	0.140 ± 0.027	0.012 ± 0.031	-0.005 ± 0.032	-0.003 ± 0.027
3.1	0.016 ± 0.035	0.000 ± 0.031	0.008 ± 0.029	0.020 ± 0.032
3.2	0.004 ± 0.029	0.010 ± 0.032	0.028 ± 0.030	0.012 ± 0.030
3.3	0.002 ± 0.031	0.009 ± 0.038	0.005 ± 0.023	0.012 ± 0.031

Fonte: Elaborada pelo autor.

Observando-se as correlação de Pearson e Spearman, ambas identificam os mesmos pares de séries que apresentam índice de correlação não nulos, com grau de confiança de 99%, sendo medidas equivalentes entre si, para os cenários considerados. Embora as correlações identifiquem dependências tais que o índice de Geweke não seja suficiente para tal, como no caso do componente com $lag\ t - 2$ do cenário 2.3, dos componentes com $lags\ t - 2$ e $t - 4$ do cenário 2.4 e do componente com $lag\ t - 1$ do cenário 2.6, elas também evidenciam relações **espúrias** de dependência como no caso dos componentes com $lags\ t - 2$ e $t - 4$ dos cenários lineares 1.1 e 1.2 e dos mesmos componentes do cenário não-linear 2.1.

3.2 Testes sobre dados de mercado

3.2.1 Escolha das variáveis de mercado

A teoria descrita nas seções anteriores é aplicada sobre combinações das seguintes séries de mercado:

- **USDBRL**: cotação diária de fechamento da paridade reais (BRL) por unidades de dólares americanos (USD);
- **IBOV**: cotação diária de fechamento do índice de ações brasileiro Ibovespa;
- **DI**: valor de ajuste diário do futuro de Juros em reais com vencimento em Janeiro do ano de 2021 (DIF21);
- **CDI**: taxa do índice *overnight* de juros em reais divulgado diariamente pela CETIP;
- **VOL**: cotação diária de fechamento da volatilidade implícita de opções *at the money* com vencimento em três meses da paridade reais (BRL) por dólares americanos (USD) negociadas internacionalmente;
- **FRW3M**: cotação diária de fechamento do *Forward (NDF)* com vencimento em três meses da paridade reais (BRL) por dólares americanos (USD) negociadas no Brasil;
- **rUSDBRL**: retorno diário da variável USDBRL calculado pela relação

$$rUSDBRL(d-0) = \ln \left(\frac{USDBRL(d-0)}{USDBRL(d-1)} \right);$$
- **rIBOV**: retorno diário da variável IBOV calculado pela relação

$$rIBOV(d-0) = \ln \left(\frac{IBOV(d-0)}{IBOV(d-1)} \right);$$
- **rDI**: retorno diário da variável DI calculado pela relação

$$rDI(d-0) = \ln \left(\frac{1+0.01 \cdot DI(d-0)}{1+0.01 \cdot DI(d-1)} \right);$$
- **rCDI**: retorno diário da variável CDI calculado pela relação

$$rCDI(d-0) = \ln \left(\frac{1+0.01 \cdot CDI(d-0)}{1+0.01 \cdot CDI(d-1)} \right);$$
- **rVOL**: retorno diário da variável VOL calculado pela relação

$$rVOL(d-0) = \ln \left(\frac{VOL(d-0)}{VOL(d-1)} \right);$$
- **rFRW3M**: retorno diário da variável FRW3M calculado pela relação

$$rFRW3M(d-0) = \ln \left(\frac{FRW3M(d-0)}{FRW3M(d-1)} \right).$$

O histórico corresponde ao período entre a data 01/01/2014 até 22/06/2016, correspondendo a 646 observações de cada variável (que considera todos os dias do calendário, exceto fins de semana). Para os dias da semana em que a variável não tem sua cotação divulgada pelo

mercado (dias de feriados, por exemplo), seu valor foi repetido, desde sua última divulgação (para as variáveis de retorno diário, seus valores foram assumidos zero, para esses dias). Neste trabalho, do conjunto de variáveis listado, cinco delas são submetidas ao estudo de causalidade, tal que os testes são realizados conforme segue:

1. **USDBRL**: pode ser causado por USDBRL e IBOV com até 3 (três) dias de *lag*;
2. **IBOV**: pode ser causado por IBOV e USDBRL com até 3 (três) dias de *lag*;
3. **CDI**: pode ser causado por CDI e rUSDBRL com até 3 (três) dias de *lag*;
4. **rUSDBRL**: pode ser causado por rUSDBRL, CDI e rIBOV com até 3 (três) dias de *lag*;
5. **rIBOV**: pode ser causado por rIBOV e rUSDBRL com até 3 (três) dias de *lag*.

Define-se, então, $X_i(l)$ para cada caso acima, representando a i -ésima variável com *lag* l ($X_0(0)$ é a variável para qual o teste está sendo realizado) de acordo com a [Tabela 4](#).

Tabela 4 – Variáveis de mercado utilizadas.

Teste (t)	X_0	X_1	X_2
1	USDBRL	IBOV	-
2	IBOV	USDBRL	-
3	CDI	rUSDBRL	-
4	rUSDBRL	CDI	rIBOV
5	rIBOV	rUSDBRL	-

Fonte: Elaborada pelo autor.

3.2.2 Causalidade em variáveis de mercado

Para cada teste t , os índices de Geweke $C^t(X_0(0)|\bar{X}_i(l))$ são calculados, de acordo com a relação (2.1), para cada par (i, l) , com $l = 0, 1, 2, 3$, $i = 0, 1, 2$, excluindo o caso tal que $(i, l) = (0, 0)$. Os dados são mostrados na [Tabela 5](#).

De modo a determinar as relações de causalidade entre as variáveis acima, será utilizada a expressão (2.5). Entretanto, dado que os índices de Geweke foram determinados utilizando apenas uma amostra para cada teste, os valores de $\sigma(C(X_t|H^l(X_t)))$ e de p são tomados de forma que $N^{-1}(p) \cdot \sigma(C(X_t|H^l(X_t)))$ corresponda a um dado percentil $\mathcal{P} = \mathcal{P}(\bar{p}, \{C(X_t|H^l(X_t))\})$ da lista de todos os índices calculados para aquele teste. Assim, as variáveis que satisfazem as

Tabela 5 – Índices de Geweke calculados para cada teste.

Variável	$C^1(X_0(0) .)$	$C^2(X_0(0) .)$	$C^3(X_0(0) .)$	$C^4(X_0(0) .)$	$C^5(X_0(0) .)$
$X_0(1)$	0.5545	0.6188	0.7005	0.0198	0.0032
$X_0(2)$	0.0002	0.0030	0.0000	0.0111	0.0000
$X_0(3)$	0.0158	0.0000	0.0001	0.0001	0.0010
$X_1(0)$	0.1991	0.1991	0.0009	0.0006	0.2235
$X_1(1)$	0.0575	0.0576	0.0001	0.0001	0.0113
$X_1(2)$	0.0001	0.0056	0.0001	0.0002	0.0000
$X_1(3)$	0.0079	0.0003	0.0004	0.0035	0.0019
$X_2(0)$	-	-	-	0.2217	-
$X_2(1)$	-	-	-	0.0084	-
$X_2(2)$	-	-	-	0.0094	-
$X_2(3)$	-	-	-	0.0045	-

Fonte: Elaborada pelo autor.

condições de causalidade, considerando-se um percentil $\bar{p} = 80\%$ e, ainda, aplicando a condição de que: “*caso sejam identificados pares de relações causais da forma $X_t^j \rightarrow X_t^i$ e $X_t^i \rightarrow X_t^j$, a relação com menor índice de Geweke $C(X_t^i|H^l(X_t^j))$ é desconsiderada*”, são mostradas na Tabela 6.

Tabela 6 – Variáveis que causam $X_0(0)$.

Variável	Causa	Efeito em	$C(X_t H^l(X_t)) - \mathcal{P}(\bar{p}, \{C(X_t H^l(X_t))\})$
$X_0(1)$	IBOV D-1	IBOV	0.4480
$X_0(1)$	USDBRL D-1	USDBRL	0.3838
$X_1(0)$	IBOV	USDBRL	0.0283
$X_0(1)$	CDI D-1	CDI	0.6997
$X_1(0)$	rUSDBRL	CDI	0.0001
$X_0(1)$	rUSDBRL D-1	rUSDBRL	0.0087
$X_1(0)$	rUSDBRL	rIBOV	0.2138
$X_1(1)$	rUSDBRL D-1	rIBOV	0.0016

Fonte: Elaborada pelo autor.

Os previsores para cada variável são calculados, bem como os resíduos das previsões. São construídas as distribuições desses resíduos e das variáveis utilizando o método de Kernel para inferência da distribuição empírica.

São simulados, então, valores das variáveis de acordo com a cronologia das causas e de acordo com o maior índice de Geweke, caso a cronologia seja a mesma entre duas ou mais variáveis por meio do sorteio de números aleatórios respeitando a distribuição de cada variável. Caso não exista causa para a variável sorteada, a distribuição empírica da própria é utilizada.

3.3 Cenários simulados e distribuições medidas

As relações determinadas de causa e efeito entre as variáveis de mercado permitem a definição de previsores para a projeção de seus valores. Dessa forma, o previsor descrito pela relação (2.6) resulta nas seguintes relações que modelam a dinâmica das variáveis estudadas

$$IBOV(t) = 0.999854 \cdot IBOV(t-1) + \varepsilon_{IBOV}, \quad (3.1)$$

$$USDBRL(t) = 0.999285 \cdot USDBRL(t-1) + \varepsilon_{USDBRL}, \quad (3.2)$$

$$CDI(t) = 1.000494 \cdot CDI(t-1) + 0.143212 \cdot rUSDBRL(t) + \varepsilon_{CDI}, \quad (3.3)$$

$$rUSDBRL(t) = -0.058674 \cdot rUSDBRL(t-1) + \varepsilon_{rUSDBRL}, \quad (3.4)$$

$$rIBOV(t) = -0.619217 \cdot rUSDBRL(t) - 0.113892 \cdot rUSDBRL(t-1) + \varepsilon_{rIBOV}. \quad (3.5)$$

Os resíduos das regressões ε são apresentados nos histogramas representados nas Figuras 5 a 9.

Figura 5 – Histograma dos resíduos de USDBRL.

Fonte: Elaborada pelo autor.

Figura 6 – Histograma dos resíduos de retornos de Ibov.

Fonte: Elaborada pelo autor.

Figura 7 – Histograma dos resíduos de retornos de USDBRL.

Fonte: Elaborada pelo autor.

Figura 8 – Histograma dos resíduos de CDI.

Fonte: Elaborada pelo autor.

Figura 9 – Histograma dos resíduos de Ibov.

Fonte: Elaborada pelo autor.

As distribuições dos resíduos são determinadas utilizando o método de Kernel, descrito na [Subseção 2.2.2](#), e o valor da função inversa da distribuição acumulada é calculada para uma discretização uniformemente espaçada de mil (1000) pontos do intervalo fechado entre zero

e um $[0, 1]$. Dessa forma, a variável aleatória ε pode ser simulada por meio da geração de um número aleatório uniformemente distribuído no domínio dessa discretização (ou seja, qualquer valor da discretização tem igual probabilidade de ser sorteado), e do cálculo do valor da função inversa da distribuição acumulada tomada para o ponto sorteado.

Para a construção dos cenários para as variáveis, o estado inicial é tomado a partir da data 28/12/2015, onde os valores iniciais são mostrados na [Tabela 7](#). Cada cenário é simulado dia a dia, até o máximo de $t = 128$ dias aplicando-se os previsores. O número de simulações adotado é mil (1000), tal que a média e o desvio-padrão de cada variável é tomado em alguns instantes t de projeção. Os gráficos contendo tais resultados obtidos, assim como a representação da trajetória real da variável simulada, são apresentados nas Figuras 10 a 14.

Tabela 7 – Valores Iniciais da Simulação, tomados em 28/12/2015.

Variável	Valor
IBOV	43764.34
USDBRL	3.8594
CDI	14.14
rUSDBRL	-0.02118
rIBOV	-0.00571
Fonte – Bloomberg

Figura 10 – Média e percentis entre 10% e 90% das trajetórias, comparados ao mercado de USDBRL.

Fonte: Elaborada pelo autor.

Figura 11 – Média e percentis entre 10% e 90% das trajetórias, comparados ao mercado de Ibov.

Fonte: Elaborada pelo autor.

Figura 12 – Média e percentis entre 10% e 90% das trajetórias, comparados ao mercado de USDBRL.

Fonte: Elaborada pelo autor.

Figura 13 – Média e percentis entre 10% e 90% das trajetórias, comparados ao mercado de CDI.

Fonte: Elaborada pelo autor.

Figura 14 – Média e percentis entre 10% e 90% das trajetórias, comparados ao mercado de Ibov.

Fonte: Elaborada pelo autor.

Diferentemente do que acontece em um movimento Browniano geométrico, em que a trajetória é gerada a partir de variações independentes entre si, pode-se prever tendências nas trajetórias dos ativos de mercado, como observado na [Figura 10](#).

As medidas de causalidade entre as séries temporais induzem tal comportamento de maneira intrínseca fazendo com que os cenários simulados possuam uma propriedade muito observada no mercado financeiro: a expectativa dos agentes de mercado (compradores e vendedores) tem efeito sobre os preços futuros dos ativos.

3.4 Cálculo de preço de derivativo

Conforme descrito na [Seção 2.3](#), o preço vigente de opções de compra europeias $C(t_0 = 0)$ pode ser determinado a partir dos resultados das simulações apresentados na [Seção 3.3](#). Dado que o *payoff* deste derivativo não depende da trajetória do preço do ativo objeto, conforme descrito pela função (2.11), para cada uma das mil trajetórias simuladas para as variáveis de mercado, o *payoff* é calculado no instante de vencimento T da opção.

Para determinar o desconto dado pela expressão $e^{-r(T) \cdot T}$ foi utilizada a trajetória da variável de mercado **CDI** e calculado o fator acumulado diário dado por

$$df(t) = \prod_{k=0}^{t-1} \left(1 + \frac{CDI(k)}{100} \right)^{\frac{1}{252}}. \quad (3.6)$$

Assim, tem-se que o fator de desconto, para o instante T , pode ser escrito conforme a seguinte relação

$$e^{-r(T) \cdot T} = \frac{1}{df(T)}. \quad (3.7)$$

Os contratos de opções simulados possuem as características dadas pela [Tabela 8](#) e [Tabela 9](#).

Tabela 8 – Contratos das opções europeias de compra de USDBRL simulados para a data 28/12/2015 (instante $t_0 = 0$).

Nome do Contrato	Ativo-Objeto	Prazo até Vencimento (dias)	Strike (K)
USDBRL_02_37810	USDBRL	2	3.7810
USDBRL_02_38206	USDBRL	2	3.8206
USDBRL_02_38426	USDBRL	2	3.8426
USDBRL_02_38637	USDBRL	2	3.8637
USDBRL_02_38869	USDBRL	2	3.8869
USDBRL_02_39117	USDBRL	2	3.9117
USDBRL_02_39599	USDBRL	2	3.9599
USDBRL_24_36043	USDBRL	24	3.6043
USDBRL_24_37522	USDBRL	24	3.7522
USDBRL_24_38366	USDBRL	24	3.8366
USDBRL_24_39139	USDBRL	24	3.9139
USDBRL_24_40162	USDBRL	24	4.0162
USDBRL_24_41227	USDBRL	24	4.1227
USDBRL_24_43461	USDBRL	24	4.3461
USDBRL_45_35818	USDBRL	45	3.5818
USDBRL_45_37754	USDBRL	45	3.7754
USDBRL_45_38883	USDBRL	45	3.8883
USDBRL_45_39896	USDBRL	45	3.9896
USDBRL_45_41374	USDBRL	45	4.1374
USDBRL_45_42909	USDBRL	45	4.2909
USDBRL_45_46244	USDBRL	45	4.6244
USDBRL_68_35222	USDBRL	68	3.5222
USDBRL_68_37553	USDBRL	68	3.7553
USDBRL_68_38930	USDBRL	68	3.8930
USDBRL_68_40120	USDBRL	68	4.0120
USDBRL_68_41995	USDBRL	68	4.1995
USDBRL_68_43974	USDBRL	68	4.3974
USDBRL_68_48378	USDBRL	68	4.8378

Fonte: Elaborada pelo autor.

Tabela 9 – Contratos das opções europeias de compra de IBOV simulados para a data 28/12/2015 (instante $t_0 = 0$).

Nome do Contrato	Ativo-Objeto	Prazo até Vencimento (dias)	Strike (K)
IBOV_12_40000	IBOV	12	40000.00
IBOV_12_41000	IBOV	12	41000.00
IBOV_12_42000	IBOV	12	42000.00
IBOV_12_43000	IBOV	12	43000.00
IBOV_12_43764	IBOV	12	43764.34
IBOV_12_44000	IBOV	12	44000.00
IBOV_12_45000	IBOV	12	45000.00
IBOV_12_46000	IBOV	12	46000.00
IBOV_37_40000	IBOV	37	40000.00
IBOV_37_41000	IBOV	37	41000.00
IBOV_37_42000	IBOV	37	42000.00
IBOV_37_43000	IBOV	37	43000.00
IBOV_37_43764	IBOV	37	43764.34
IBOV_37_44000	IBOV	37	44000.00
IBOV_37_45000	IBOV	37	45000.00
IBOV_37_46000	IBOV	37	46000.00
IBOV_77_40000	IBOV	77	40000.00
IBOV_77_41000	IBOV	77	41000.00
IBOV_77_42000	IBOV	77	42000.00
IBOV_77_43000	IBOV	77	43000.00
IBOV_77_43764	IBOV	77	43764.34
IBOV_77_44000	IBOV	77	44000.00
IBOV_77_45000	IBOV	77	45000.00
IBOV_77_46000	IBOV	77	46000.00
IBOV_77_47000	IBOV	77	47000.00
IBOV_77_48000	IBOV	77	48000.00
IBOV_122_40000	IBOV	122	40000.00
IBOV_122_41000	IBOV	122	41000.00
IBOV_122_42000	IBOV	122	42000.00
IBOV_122_43000	IBOV	122	43000.00
IBOV_122_43764	IBOV	122	43764.34
IBOV_122_44000	IBOV	122	44000.00
IBOV_122_45000	IBOV	122	45000.00
IBOV_122_46000	IBOV	122	46000.00
IBOV_122_47000	IBOV	122	47000.00
IBOV_122_48000	IBOV	122	48000.00
IBOV_122_49000	IBOV	122	49000.00
IBOV_122_50000	IBOV	122	50000.00

Fonte: Elaborada pelo autor.

Os respectivos preços $\bar{C}$ determinados por meio da simulação de Monte Carlo são mostrados na [Tabela 10](#) e na [Tabela 11](#).

Tabela 10 – Preços das opções europeias de compra de USDBRL simulados para a data 28/12/2015.

Nome do Contrato	Preço $\bar{C}$	Nome do Contrato	Preço $\bar{C}$
USDBRL_02_37810	0.0743	USDBRL_45_35818	0.1916
USDBRL_02_38206	0.0413	USDBRL_45_37754	0.0757
USDBRL_02_38426	0.0269	USDBRL_45_38883	0.0378
USDBRL_02_38637	0.0163	USDBRL_45_39896	0.0185
USDBRL_02_38869	0.0085	USDBRL_45_41374	0.0056
USDBRL_02_39117	0.0039	USDBRL_45_42909	0.0016
USDBRL_02_39599	0.0006	USDBRL_45_46244	0.0000
USDBRL_24_36043	0.1986	USDBRL_68_35222	0.2139
USDBRL_24_37522	0.0926	USDBRL_68_37553	0.0848
USDBRL_24_38366	0.0505	USDBRL_68_38930	0.0406
USDBRL_24_39139	0.0261	USDBRL_68_40120	0.0201
USDBRL_24_40162	0.0088	USDBRL_68_41995	0.0058
USDBRL_24_41227	0.0021	USDBRL_68_43974	0.0010
USDBRL_24_43461	0.0001	USDBRL_68_48378	0.0000

Fonte: Elaborada pelo autor.

Tabela 11 – Preços das opções europeias de compra de IBOV simulados para a data 28/12/2015.

Nome do Contrato	Preço $\bar{C}$	Nome do Contrato	Preço $\bar{C}$
IBOV_12_40000	3706.04	IBOV_77_42000	3749.66
IBOV_12_41000	2833.74	IBOV_77_43000	3181.41
IBOV_12_42000	2064.04	IBOV_77_43764	2778.73
IBOV_12_43000	1406.71	IBOV_77_44000	2661.21
IBOV_12_43764	993.44	IBOV_77_45000	2197.91
IBOV_12_44000	883.35	IBOV_77_46000	1798.32
IBOV_12_45000	498.94	IBOV_77_47000	1452.05
IBOV_12_46000	251.46	IBOV_77_48000	1155.29
IBOV_37_40000	4201.46	IBOV_122_40000	5622.91
IBOV_37_41000	3457.09	IBOV_122_41000	5000.77
IBOV_37_42000	2784.58	IBOV_122_42000	4415.55
IBOV_37_43000	2194.98	IBOV_122_43000	3866.90
IBOV_37_43764	1797.60	IBOV_122_43764	3476.65
IBOV_37_44000	1685.18	IBOV_122_44000	3361.75
IBOV_37_45000	1264.66	IBOV_122_45000	2900.54
IBOV_37_46000	915.53	IBOV_122_46000	2487.76
IBOV_77_40000	5019.45	IBOV_122_47000	2131.51
IBOV_77_41000	4358.03	IBOV_122_48000	1811.13
IBOV_77_42000	3749.66	IBOV_122_49000	1529.18
IBOV_77_43000	3181.41	IBOV_122_50000	1278.53

Fonte: Elaborada pelo autor.

Como forma de comparação, os valores de mercado das opções foram extraídos da fonte Bloomberg, utilizando a função “OVDV” (no caso das opções de BRLUSD) e “OVME” (no caso das opções de IBOV) do terminal. Os valores encontrados estão listados na [Tabela 12](#) e na [Tabela 13](#). Os gráficos comparativos dos preços em função dos preços de exercício de cada contrato são mostrados para cada vencimento T (representado em dias) de contrato, nas Figuras [15](#) a [22](#).

Tabela 12 – Preços das opções europeias de compra de USDBRL Bloomberg para a data 28/12/2015.

Nome do Contrato	Preço $\bar{C}$	Nome do Contrato	Preço $\bar{C}$
USDBRL_02_37810	0.0900	USDBRL_45_35818	0.4200
USDBRL_02_38206	0.0500	USDBRL_45_37754	0.2600
USDBRL_02_38426	0.0400	USDBRL_45_38883	0.2000
USDBRL_02_38637	0.0300	USDBRL_45_39896	0.1500
USDBRL_02_38869	0.0200	USDBRL_45_41374	0.0900
USDBRL_02_39117	0.0100	USDBRL_45_42909	0.0600
USDBRL_02_39599	0.0000	USDBRL_45_46244	0.0200
USDBRL_24_36043	0.3200	USDBRL_68_35222	0.4200
USDBRL_24_37522	0.2000	USDBRL_68_37553	0.3100
USDBRL_24_38366	0.1500	USDBRL_68_38930	0.2300
USDBRL_24_39139	0.1100	USDBRL_68_40120	0.1700
USDBRL_24_40162	0.0700	USDBRL_68_41995	0.1100
USDBRL_24_41227	0.0400	USDBRL_68_43974	0.0700
USDBRL_24_43461	0.0200	USDBRL_68_48378	0.0200

Fonte – Bloomberg

Tabela 13 – Preços das opções europeias de compra de IBOV Bloomberg para a data 28/12/2015.

Nome do Contrato	Preço $\bar{C}$	Nome do Contrato	Preço $\bar{C}$
IBOV_12_40000	4112.12	IBOV_77_42000	4776.14
IBOV_12_41000	3193.11	IBOV_77_43000	4112.44
IBOV_12_42000	2336.85	IBOV_77_43764	3631.95
IBOV_12_43000	1589.30	IBOV_77_44000	3489.16
IBOV_12_43764	1105.35	IBOV_77_45000	2914.41
IBOV_12_44000	972.76	IBOV_77_46000	2380.80
IBOV_12_45000	512.47	IBOV_77_47000	1914.16
IBOV_12_46000	236.20	IBOV_77_48000	1519.96
IBOV_37_40000	4920.31	IBOV_122_40000	7486.29
IBOV_37_41000	4106.87	IBOV_122_41000	6779.29
IBOV_37_42000	3352.97	IBOV_122_42000	6098.92
IBOV_37_43000	2648.78	IBOV_122_43000	5446.40
IBOV_37_43764	2162.61	IBOV_122_43764	4967.22
IBOV_37_44000	2021.77	IBOV_122_44000	4822.98
IBOV_37_45000	1478.61	IBOV_122_45000	4229.87
IBOV_37_46000	1029.98	IBOV_122_46000	3668.80
IBOV_77_40000	6207.43	IBOV_122_47000	3143.29
IBOV_77_41000	5475.49	IBOV_122_48000	2660.31
IBOV_77_42000	4776.14	IBOV_122_49000	2229.71
IBOV_77_43000	4112.44	IBOV_122_50000	1859.69

Fonte – Bloomberg

Figura 15 – Preços simulados e reais de opções de compra europeia de USDBRL para prazo $T = 2$ (valores na moeda real).

Fonte: Elaborada pelo autor.

Figura 16 – Preços simulados e reais de opções de compra europeia de USDBRL para prazo $T = 24$ (valores na moeda real).

Fonte: Elaborada pelo autor.

Figura 17 – Preços simulados e reais de opções de compra europeia de USDBRL para prazo $T = 45$ (valores na moeda real).

Fonte: Elaborada pelo autor.

Figura 18 – Preços simulados e reais de opções de compra europeia de USDBRL para prazo $T = 68$ (valores na moeda real).

Fonte: Elaborada pelo autor.

Figura 19 – Preços simulados e reais de opções de compra europeia de IBOV para prazo $T = 12$ (valores na moeda real).

Fonte: Elaborada pelo autor.

Figura 20 – Preços simulados e reais de opções de compra europeia de IBOV para prazo $T = 37$ (valores na moeda real).

Fonte: Elaborada pelo autor.

Figura 21 – Preços simulados e reais de opções de compra europeia de IBOV para prazo $T = 77$ (valores na moeda real).

Fonte: Elaborada pelo autor.

Figura 22 – Preços simulados e reais de opções de compra europeia de IBOV para prazo $T = 122$ (valores na moeda real).

Fonte: Elaborada pelo autor.

Observa-se, no resultado das simulações dos contratos de opções, tanto no caso do USDBRL, quanto no caso do IBOV, que os preços determinados pela metodologia proposta

são consistentemente menores que os preços extraídos do mercado, sendo essa diferença mais acentuada no caso do câmbio. A presença de uma tendência nos valores simulados das trajetórias dos ativos desloca a média dos *payoffs* das opções em relação às trajetórias previstas pelo modelo de Black-Scholes. Tal deslocamento da média impacta diretamente o valor das opções e é mais evidente quanto maior for a tendência prevista na trajetória do ativo.

Outro fator relevante para determinar o comportamento dos preços das opções é a distribuição do valor do ativo no instante do vencimento da opção. O modelo de Black-Scholes prevê uma distribuição log-normal dos preços futuros do ativo, consequência da distribuição gaussiana de seus retornos, a qual maximiza a entropia do sistema, enquanto o uso das distribuições empíricas (não-gaussianas) na simulação das trajetórias implicam um grau maior de informação neste sistema.

DISCUSSÕES E CONCLUSÃO

A proposta de usar uma medida de causalidade para aprimorar a previsibilidade da dinâmica de variáveis de mercado tem como vantagem poder encontrar relações de dependência e o sentido que ocorrem. Embora a quantificação dessa medida seja mais complexa, computacionalmente, em comparação ao cálculo da correlação entre duas séries temporais, essa metodologia evita o uso de correlações potencialmente espúrias para se determinar a dinâmica do mercado.

Os cálculos dos índices de Geweke (dos cenários sintetizados e dos dados de mercado), assim como a determinação dos parâmetros dos previsores e as distribuições de suas respectivas variáveis aleatórias são feitos por um programa, desenvolvido para este trabalho na linguagem de programação C, devido à sua flexibilidade e performance de processamento (todos os dados de saída mencionados são gerados em um intervalo de tempo aproximadamente 30 segundos, utilizando uma CPU Intel Centrino Duo de 1.66GHz).

A qualidade das medidas de causalidade está diretamente ligada com a escolha do previsor das variáveis e a relação de dependência entre elas. O uso de uma função linear possibilita a identificação de relações causais, mesmo com dependência não-linear, porém, limitado até certa ordem, inclusive sobre dados que apresentem uma dinâmica estocástica com distribuições não necessariamente simétricas. Dependências de componentes de ordem quadrática ou maior, conforme resultado simulado sobre os dados sintetizados, não podem ser identificados, considerando essa escolha de previsor. Entretanto, tal função, por ser linear, possibilita aumentar o número de variáveis que, hipoteticamente, possuem influência no sistema, com baixo custo computacional, comparando com previsores mais complexos.

A metodologia aplicada para construir cenários com múltiplas variáveis que reflitam as dependências observadas na série temporal dos processos estudados permite explicar a existência de tendências, pois a hipótese de independência dos retornos utilizada no modelo de Black-Scholes não é considerada. Pelo contrário, as medidas de causalidade entre duas séries temporais podem induzir cenários de tendência e que é observado com clareza no cenário de USDBRL

construído. A dependência que uma variável pode apresentar com seu valor num instante imediatamente anterior, por exemplo, consegue replicar comportamentos muito observados em períodos instáveis das séries de preços de mercado, onde grandes retornos são precedidos de grandes retornos. Essas relações implicam uma propriedade na dinâmica das séries temporais, onde a expectativa do mercado para determinado ativo, pode influenciar a dinâmica do próprio ativo.

A utilização da distribuição empírica dos resíduos gerados pelo previsor em uma série histórica para construir a parte estocástica da dinâmica da variável exige alguns cuidados. É suposto, devido a isso, que o comportamento passado de todas as variáveis do sistema é o único determinante de seus comportamentos futuros. A dinâmica de mercado é, normalmente, composta por regimes específicos que acontecem por determinados períodos de tempo. Quanto mais curto for o período de tempo analisado, mais escassos são os dados observados e, portanto, mais difícil estimar uma distribuição empírica. Dessa forma, sendo necessária uma abordagem mais robusta que a discretização do espaço de eventos das variáveis, o método de Kernel possibilita a aproximação da distribuição por uma função contínua, mesmo com dados esparsos no espaço de eventos. Uma distribuição de probabilidades contínua permite uma qualidade maior na determinação da imagem inversa e, devido a isso, a geração de variáveis aleatórias sob essa distribuição.

Pode ser feita uma analogia desse comportamento com a transição de fase de sistemas físicos, onde existe uma dimensão muito grande de agentes para se definir uma dinâmica determinística, sendo a melhor medida desse sistema, uma medida estatística. Dessa forma, assim como todas as metodologias paramétricas de dinâmicas de variáveis de mercado, é necessário que as relações de dependência e as distribuições dos resíduos sejam recalibradas regularmente.

De forma comparativa, no método que utiliza as fórmulas de Black-Scholes, o parâmetro σ de volatilidade é recalibrado pelo menos uma vez por dia.

As variáveis de mercado estudadas nesse trabalho foram escolhidas por três motivos. São, provavelmente, as variáveis do mercado brasileiro mais exaustivamente estudadas, pois representam três grandes indicadores da economia: câmbio (com o valor da paridade dólar-real USDBRL), ações (com o índice de ações mais conhecido da $[B]^3$, o Ibovespa, ou IBOV)¹ e o *juro* (com o valor da taxa anual CDI). Outro motivo é a liquidez das operações que determinam os valores dessas variáveis. O mercado brasileiro carece de liquidez em muitos ativos, o que dificulta o estudo amplo e a produção de trabalhos relacionados aos ativos característicos, como por exemplo, as commodities agrícolas. Dessa forma, escolhendo-se essas variáveis, pode-se assumir uma vantagem estatística nas análises de suas séries temporais. Por fim, já é conhecido que existe uma correlação muito forte (aproximadamente -0.63 no período entre 01/01/2014 e

¹ $[B]^3$ é a antiga BM&F BOVESPA, que teve seu nome alterado após fusão com a empresa Cetip no ano de 2017

22/06/2016) entre os valores de USDBRL e do índice IBOV (mais forte do que dois retornos de dias consecutivos de suas próprias séries históricas).

Nos resultados das trajetórias simuladas para as variáveis de mercado observa-se nos gráficos de suas médias a existência de tendências geradas pelas dependências encontradas.

Observando-se a comparação dos preços calculados via simulação de Monte Carlo, conforme proposto neste trabalho e os preços extraídos de uma fonte de mercado (no caso, são dados disponibilizados pela Bloomberg), nota-se os seguintes pontos: os preços simulados apresentam a tendência de serem menores que os preços de mercado; as diferenças entre os preços simulados e os preços de mercado são maiores no caso da variável USDBRL, em comparação com a variável IBOV; e os dois preços apresentam uma curvatura em função do preço de exercício K .

Em relação aos dois primeiros pontos, essa diferença é caracterizada pela presença de tendência nos preços gerados pela simulação (muito mais acentuada para o USDBRL). Para o cálculo do preço da opção de compra europeia por meio da fórmula de Black-Scholes (que é o valor no qual o preço de mercado está baseado) considera-se que o preço do ativo, no vencimento a opção, descontado a valor presente pela taxa de juros é um martingale, portanto, tem valor esperado igual ao valor atual do ativo. Em outras palavras, as distribuições de probabilidade do preço do ativo, no momento do vencimento do contrato de opção, estão centradas em diferentes pontos, portanto, os valores de *payoff* projetados também são diferentes. Pode-se interpretar as diferenças devido ao maior grau de certeza sobre o preço projetado de um ativo que a metodologia descrita implica.

O comportamento do preço das opções europeias de compra em função do preço de exercício é o responsável por gerar a distorção (conhecido por *smile*) no parâmetro da volatilidade do modelo de Black-Scholes. Essa curvatura observada nos gráficos está presente intrinsecamente nos resultados das simulações, caracterizado pelas distribuições não-gaussianas implícitas nos processos estocásticos gerados. Dessa forma, eventos extremos (que estariam presentes nas caudas das distribuições gaussianas), são considerados com uma probabilidade maior, coerente ao implícito nos preços de mercado das opções.

CONSIDERAÇÕES FINAIS

Neste trabalho, foi definida uma métrica de medir causas entre eventos, utilizando séries temporais. Em seguida, mostra-se como aplicar as medidas de causalidade para simular cenários sobre esses eventos. Essa técnica foi aplicada sobre ativos de mercado, com intuito de estudar as relações de dependência entre alguns casos e por fim, possibilitar a precificação de derivativos. Os contratos de opções de compra do tipo europeias foram utilizadas, devido à facilidade de se comparar com observações do mercado e para mostrar uma possibilidade pouco abordada na literatura científica, principalmente para o mercado brasileiro.

Dessa forma, existem várias possibilidades de evoluir o conteúdo apresentado. Existe grande interesse em aplicar o estudo apresentado sobre dados com maior frequência de observações, ou seja, valores *intraday* de preços de ativos, como uma forma de aprimorar a qualidade e a relevância estatística dos dados, permitindo estudar relações de causas considerando as séries de preços de contratos Futuros BM&F.

Um outro estudo que pode ser realizado tem relação com uma abordagem Bayesiana para as medidas de causalidade. Conforme foi discutido, a eficácia da medida de Granger está diretamente ligada ao preditor escolhido e o tipo de dependência que os processos apresentam. Como forma alternativa, propõe-se utilizar a Transferência de Entropia ([WOLLSTADT *et al.*, 2014](#)), onde se quantifica a informação gerada por um processo, e mede-se o fluxo dessa informação ao considerar um segundo processo que, hipoteticamente, apresenta efeito no primeiro. Essa técnica exige uma amostragem maior de dados para medidas de distribuições empíricas, principalmente para casos multivariáveis. A comparação entre essas duas medidas está, também, nas perspectivas futuras deste trabalho.

Por fim, muitos estudos podem ser feitos, utilizando as técnicas de projeção de cenários descritas neste trabalho, com aplicações alternativas a precificações de derivativos. É de muito interesse das empresas modelos que construam cenários sob diversas circunstâncias para uso em cálculo de Risco de mercado (onde as instituições estão interessadas em saber qual é a exposição

do valor de seus ativos à variação de mercado). Outro exemplo é o interesse cada vez maior das empresas projetarem o montante de impostos a serem pagos (Impostos sobre Renda, PIS, COFINS etc) e, para isso, é necessária a criação de cenários sobre o valor de suas carteiras de ativos.

A importância dada à “informação” vem aumentando nas últimas décadas. Mais valioso ainda é a sua quantificação, em meio a tantos dados que transitam nos canais atuais de comunicação. Existe, portanto, grandes esforços aplicados em melhorar a qualidade e o uso dessas informações. Com esse propósito, este trabalho mostra uma forma de extrair conhecimento sobre dados cujo comportamento, embora não possa ser modelado perfeitamente, garante, da melhor forma possível, a qualidade dessas informações.

REFERÊNCIAS

- AZZALINI, A. A note on the estimation of a distribution function and quantiles by a kernel method. **Biometrika**, Oxford University Press, v. 68, n. 1, p. 326–328, 1981. Citado na página [37](#).
- BAK, P.; PACZUSKI, M.; SHUBIK, M. Price variations in a stock market with many agents. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 246, n. 3-4, p. 430–453, 1997. Citado na página [21](#).
- BLACK, F.; SCHOLES, M. The pricing of options and corporate liabilities. **Journal of Political Economy**, The University of Chicago Press, v. 81, n. 3, p. 637–654, 1973. Citado na página [21](#).
- BRACEGIRDLE, C.; BARBER, D. Bayesian conditional cointegration. **arXiv preprint arXiv:1206.6459**, 2012. Citado na página [32](#).
- CONT, R.; HAMIDA, S. B. Recovering volatility from option prices by evolutionary optimization. 2004. Citado na página [22](#).
- ECKHARDT, R. Stan ulam, john von neumann, and the monte carlo method. **Los Alamos Science**, v. 15, n. 131-136, p. 30, 1987. Citado na página [25](#).
- EICHLER, M. Granger causality and path diagrams for multivariate time series. **Journal of Econometrics**, Elsevier, v. 137, n. 2, p. 334–353, 2007. Citado na página [33](#).
- GRANGER, C. W. Investigating causal relations by econometric models and cross-spectral methods. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 424–438, 1969. Citado nas páginas [32](#) e [33](#).
- Huber, M.; Maric, N. Multivariate distributions with fixed marginals and correlations. **ArXiv e-prints**, nov. 2013. Citado na página [29](#).
- ITÔ, K. 109. stochastic integral. **Proceedings of the Imperial Academy**, The Japan Academy, v. 20, n. 8, p. 519–524, 1944. Citado na página [26](#).
- KAROUÏ, N. E.; JEANBLANC-PICQUÈ, M.; SHREVE, S. E. Robustness of the black and scholes formula. **Mathematical finance**, Wiley Online Library, v. 8, n. 2, p. 93–126, 1998. Citado nas páginas [23](#) e [25](#).
- KROESE, D. P.; BRERETON, T.; TAIMRE, T.; BOTEV, Z. I. Why the monte carlo method is so important today. **Wiley Interdisciplinary Reviews: Computational Statistics**, Wiley Online Library, v. 6, n. 6, p. 386–392, 2014. Citado na página [26](#).
- ORTEGA, J.; PULLIRSCH, R.; TEICHMANN, J.; WERGIELUK, J. A dynamic approach for scenario generation in risk management. **Preprint on arXiv**, v. 904, 2009. Citado na página [29](#).
- ROSOLEN, D.; ARAÚJO, M. V.; LYRIO, M. Previsão dos preços de commodities por meio das taxas de câmbio. **Estudos Econômicos (São Paulo)**, SciELO Brasil, v. 43, n. 4, p. 813–830, 2013. Citado na página [22](#).

SILVERMAN, B. W. **Density estimation for statistics and data analysis**. [S.l.]: CRC press, 1986. v. 26. Citado na página [37](#).

WOLLSTADT, P.; MARTÍNEZ-ZARZUELA, M.; VICENTE, R.; DÍAZ-PERNAS, F. J.; WIBRAL, M. Efficient transfer entropy analysis of non-stationary neural time series. **PLoS One**, Public Library of Science, v. 9, n. 7, p. e102833, 2014. Citado na página [67](#).

ZAREMBA, A.; ASTE, T. Measures of causality in complex datasets with application to financial data. **Entropy**, Multidisciplinary Digital Publishing Institute, v. 16, n. 4, p. 2309–2349, 2014. Citado nas páginas [32](#) e [33](#).