



Carolina Ramos Simões

Estimação de funções densidade de probabilidade neutras face ao risco: uma aplicação a opções S&P 500

Dissertação de Mestrado em Métodos Quantitativos em Finanças, orientada pelos Professores Doutores Ana Margarida Monteiro e Helder Sebastião e apresentada ao Departamento de Matemática da Faculdade de Ciências e Tecnologia e à Faculdade de Economia da Universidade de Coimbra.

Outubro de 2017



Estimação de funções densidade de probabilidade neutras face ao risco: uma aplicação a opções S&P 500

Carolina Ramos Simões



UNIVERSIDADE DE COIMBRA

Mestrado em Métodos Quantitativos em Finanças

Master in Quantitative Methods in Finance

Dissertação de Mestrado | MSc Dissertation

October 2017

Agradecimentos

Um sincero obrigado aos Professores Doutores Ana Margarida Monteiro e Helder Sebastião, pelo desafio que me propuseram e por todo o tempo despendido durante a realização desta dissertação.

Aos meus amigos, por serem essenciais na minha vida, agradeço por todo o apoio e compreensão e por acompanharem ativamente todos os passos do meu percurso.

Por último, um profundo agradecimento à minha família, em especial aos meus pais e à minha irmã, por nunca terem deixado de acreditar em mim.

Resumo

Com esta dissertação pretende-se estudar e implementar vários métodos de estimação de funções densidade de probabilidade neutras face ao risco, identificando as perspectivas dos investidores quanto à evolução do mercado. Assim, este trabalho incide sobre preços teóricos de opções europeias, gerados pelo modelo de Black-Scholes, e sobre preços de mercado de opções do mesmo estilo sobre o índice S&P 500, transacionadas no *Chicago Board Options Exchange* (CBOE) e relativas ao dia 4 de janeiro de 2016 para diferentes maturidades.

Inicialmente consideraram-se preços teóricos, gerados pelo modelo de Black-Scholes, e compararam-se, através da análise gráfica e do cálculo do *root mean square error* (RMSE), com os preços estimados por diferentes métodos. Os métodos utilizados neste estudo foram a mistura de duas distribuições lognormais, a expansão de Edgeworth e o método da volatilidade implícita proposto por Shimko. Seguidamente, estimaram-se os parâmetros relativos aos diferentes métodos e através deles produziram-se as funções densidade de probabilidade neutras face ao risco. Após esta estimação, efetuou-se a análise das perspectivas de mercado, utilizando opções sobre o índice S&P 500 para sete maturidades diferentes. Deste modo, compararam-se os preços das opções de mercado com os preços estimados pelo modelo de Black-Scholes, pela mistura de duas distribuições lognormais, pela expansão de Edgeworth e pelo método da volatilidade implícita proposto por Shimko. De seguida, estimaram-se os parâmetros dos diferentes métodos e, através destes, produziram-se as funções densidade de probabilidade neutras face ao risco. Por fim, analisaram-se algumas variáveis estatísticas das funções densidade de probabilidade obtidas, tais como a média, a variância, o coeficiente de assimetria e o excesso de curtose, de modo a retirar conclusões acerca das expectativas dos investidores em relação à evolução do mercado.

Conteúdo

Lista de Figuras	ix
Lista de Tabelas	xi
1 Introdução	1
2 Função densidade de probabilidade neutra face ao risco: uma revisão de literatura	3
2.1 Modelo de Black-Scholes	5
2.2 Modelos paramétricos utilizados na estimação da função densidade de probabilidade neutra face ao risco	8
2.2.1 Modelo de Heston	8
2.2.2 Mistura de Distribuições Lognormais	9
2.3 Modelos semi-paramétricos utilizados na estimação da função densidade de probabilidade neutra face ao risco	11
2.3.1 Mistura de funções hipergeométricas confluentes	11
2.3.2 Expansão de Edgeworth	12
2.4 Modelos não paramétricos utilizados na estimação da função densidade de probabilidade neutra face ao risco	14
2.4.1 Método de Kernel	14
2.4.2 Método da Volatilidade Implícita proposto por Shimko	15
3 Métodos de estimação das funções densidade de probabilidade neutras face ao risco: análise comparativa	17
3.1 Metodologia para a estimação da função densidade de probabilidade neutra face ao risco	18
3.2 Aplicação dos diferentes métodos e resultados	19
3.2.1 Variação das funções densidade de probabilidade de distribuições lognormais com a maturidade	19
3.2.2 Estimação da função densidade de probabilidade neutra face ao risco utilizando dados teóricos	21
3.2.3 Estimação da função densidade de probabilidade neutra face ao risco utilizando dados de mercado: opções sobre o índice S&P 500	25
4 Conclusão	43

Bibliografia

45

Lista de Figuras

3.1	Estimação das funções densidade de probabilidade lognormais teóricas, referentes ao conjunto A e quatro maturidades diferentes (1 mês, 3 meses, 6 meses e 12 meses).	20
3.2	Comparação entre preços de opções <i>call</i> gerados de acordo com a fórmula de Black-Scholes e preços de opções <i>call</i> gerados pelos métodos da MLN, da EW e de Shimko, para o conjunto B de dados.	22
3.3	Estimação das funções densidade de probabilidade neutras face ao risco, para dados teóricos de Black-Scholes, referentes ao conjunto B.	23
3.4	Comparação entre preços de opções <i>call</i> gerados de acordo com a fórmula de Black-Scholes e preços de opções <i>call</i> gerados pelos métodos da MLN, da EW e de Shimko, para o conjunto C de dados.	24
3.5	Estimação das funções densidade de probabilidade neutras face ao risco, para dados teóricos de Black-Scholes, referentes ao conjunto C.	25
3.6	Comparação entre preços de opções <i>call</i> de mercado e preços de opções <i>call</i> estimados pelos modelos de BS, MLN, EW e Shimko para a primeira maturidade (11 dias).	27
3.7	Comparação entre preços de opções <i>call</i> de mercado e preços de opções <i>call</i> estimados pelos modelos de BS, MLN, EW e Shimko para a segunda maturidade (46 dias).	28
3.8	Comparação entre preços de opções <i>call</i> de mercado e preços de opções <i>call</i> estimados pelos modelos de BS, MLN, EW e Shimko para a terceira maturidade (74 dias).	29
3.9	Comparação entre preços de opções <i>call</i> de mercado e preços de opções <i>call</i> estimados pelos modelos de BS, MLN, EW e Shimko para a quarta maturidade (102 dias).	30
3.10	Comparação entre preços de opções <i>call</i> de mercado e preços de opções <i>call</i> estimados pelos modelos de BS, MLN, EW e Shimko para a quinta maturidade (165 dias).	31
3.11	Comparação entre preços de opções <i>call</i> de mercado e preços de opções <i>call</i> estimados pelos modelos de BS, MLN, EW e Shimko para a sexta maturidade (256 dias).	32
3.12	Comparação entre preços de opções <i>call</i> de mercado e preços de opções <i>call</i> estimados pelos modelos de BS, MLN, EW e Shimko para a sétima maturidade (347 dias).	33
3.13	Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à primeira maturidade (11 dias).	36
3.14	Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à segunda maturidade (46 dias).	37
3.15	Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à terceira maturidade (74 dias).	37

- 3.16 Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à quarta maturidade (102 dias). 38
- 3.17 Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à quinta maturidade (165 dias). 38
- 3.18 Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à sexta maturidade (256 dias). 39
- 3.19 Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à sétima maturidade (347 dias). 39

Lista de Tabelas

2.1	Restrições dos parâmetros utilizados na mistura de funções hipergeométricas confluentes para a obtenção de funções densidade de probabilidade conhecidas.	12
3.1	Origem dos preços e métodos utilizados na estimação das FDPN.	17
3.2	Média, variância, coeficiente de assimetria e excesso de curtose das funções densidade de probabilidade lognormais teóricas obtidas para quatro maturidades diferentes (1 mês, 3 meses, 6 meses e 12 meses).	20
3.3	Parâmetros dos dois conjuntos de dados utilizados na geração pelo modelo de Black-Scholes de preços teóricos	21
3.4	RMSE - Preços de opções estimados para o conjunto B.	22
3.5	RMSE - Preços de opções estimados para o conjunto C.	23
3.6	Sumário dos dados utilizados para as diferentes maturidades.	26
3.7	RMSE de opções <i>call</i> sobre o índice S&P 500.	34
3.8	RMSE de opções <i>put</i> sobre o índice S&P 500.	34
3.9	Síntese de resultados do RMSE para as opções <i>call</i> e <i>put</i> sobre o índice S&P 500, referentes as diferentes maturidades.	34
3.10	Parâmetros extraídos na estimação das FDPN segundo o modelo de Black-Scholes, utilizando opções sobre o índice S&P 500.	35
3.11	Parâmetros extraídos na estimação das FDPN segundo o método da mistura de duas distribuições lognormais, utilizando opções sobre o índice S&P 500.	35
3.12	Parâmetros extraídos na estimação das FDPN segundo o método da expansão de Edgeworth, utilizando opções sobre o índice S&P 500.	35
3.13	Parâmetros extraídos na estimação das FDPN segundo o método de Shimko, utilizando opções sobre o índice S&P 500.	36
3.14	Média, variância, coeficiente de assimetria e excesso curtose da FDPN gerada pelo modelo de Black-Scholes.	40
3.15	Média, variância, coeficiente de assimetria e excesso de curtose da FDPN gerada pela mistura de duas distribuições lognormais.	40
3.16	Média, variância, coeficiente de assimetria e excesso de curtose da FDPN gerada pela expansão de Edgeworth.	40
3.17	Média, variância, coeficiente de assimetria e excesso de curtose da FDPN gerada pelo método de Shimko.	41

Capítulo 1

Introdução

Os derivados financeiros são contratos financeiros subscritos sobre outro ativo, designado por ativo subjacente. A sua existência é impulsionada pela necessidade de proteção contra o risco de alterações do preço do ativo subjacente. Deste modo, a sua importância tem crescido consideravelmente nos últimos trinta anos e a modelação dos seus preços tornou-se uma questão central da Matemática Financeira. É, neste contexto, que o preço dos derivados financeiros, como são exemplos os contratos de opções, os contratos *forward* e os contratos de futuro, fornecem informação importante acerca das expectativas dos agentes económicos quanto à evolução do preço do ativo subjacente. Assim, através dos preços dos contratos de opções e, considerando a hipótese de neutralidade face ao risco, é possível estimar a função densidade de probabilidade neutra face ao risco para o preço do ativo subjacente, na data de vencimento do contrato. Esta função permite ilustrar a distribuição das probabilidades de ocorrência atribuídas a eventos futuros e através desta caracterizar o perfil das expectativas dos agentes económicos relativamente à evolução do preço do ativo subjacente. Na literatura são abordados, extensivamente, diversos métodos utilizados na estimação da função densidade de probabilidade neutra face ao risco, o que demonstra a importância do tema agora em análise.

A dissertação que agora se apresenta incide na estimação, através de alguns métodos, das funções densidade de probabilidade neutras face ao risco, utilizando preços teóricos de opções, gerados pelo modelo de Black-Scholes, e preços de opções sobre o índice S&P 500, transacionadas no *Chicago Board Options Exchange* (CBOE). Deste modo, iniciamos o estudo expondo as abordagens consideradas nesta dissertação, analisando as suas potencialidades e limitações. Seguidamente analisamos alguns desses métodos e retiramos conclusões acerca das expectativas dos agentes económicos quanto à evolução do mercado.

O tema agora em estudo é abordado, nesta dissertação, ao longo de quatro capítulos. Após a presente introdução, onde se apresenta genericamente o tema, elaboramos, no capítulo dois, uma síntese bibliográfica relativa aos principais estudos efetuados nesta área. Inicialmente apresentamos alguns conceitos e resultados básicos. De seguida focamos a nossa análise no modelo mais conhecido na determinação de preços de opções, o modelo de Black-Scholes [4]. Este modelo apresenta limitações, principalmente no que respeita aos seus pressupostos. Tal incentivou os investigadores a desenvolverem métodos mais complexos, pelo que são apresentadas algumas das abordagens desenvolvidas na tentativa de ultrapassar tais limitações.

No terceiro capítulo utilizamos alguns dos métodos de estimação da função densidade de probabilidade neutra face ao risco, apresentados no segundo capítulo. Iniciamos este capítulo explicando o processo de estimação e as funções do programa R utilizadas para tal. Seguidamente observamos o comportamento das funções densidade de probabilidade de distribuições lognormais, como proposto por Black-Scholes [4], fazendo variar apenas o tempo até à maturidade. Estimamos ainda as funções densidade de probabilidade neutras face ao risco através dos diferentes métodos, utilizando dados teóricos e dados de mercado relativos ao índice S&P 500. Pelo modelo de Black-Scholes geramos preços de opções teóricas e estimamos, através desses preços, os parâmetros que produzem as funções densidade de probabilidade neutras face ao risco relativas aos diferentes métodos em estudo - mistura de duas distribuições lognormais [3], expansão de Edgeworth [16] e método da volatilidade implícita proposto por David Shimko [20]. Para a estimação das funções densidade de probabilidade neutras face ao risco usando dados de mercado, são aplicados o modelo de Black-Scholes e os métodos acima referidos. Esta estimação permite-nos extrair as expectativas de mercado inerentes aos preços das opções sobre o índice S&P 500.

No último capítulo comentamos os resultados obtidos ao longo desta dissertação e sugerimos áreas de investigação futura relativas a este tema.

Capítulo 2

Função densidade de probabilidade neutra face ao risco: uma revisão de literatura

Os derivados financeiros [15] podem ser definidos como instrumentos que resultam de contratos celebrados e valorizados em relação a um determinado ativo subjacente e fixam no presente o preço de uma transação que poderá ocorrer no futuro. Os contratos *forward*, os contratos de futuros e os contratos de opções são alguns exemplos de derivados financeiros.

Os contratos *forward* [11] [15] são transacionados em mercados descentralizados e informais, onde todos os termos do contrato são alvo de negociação bilateral. Assim, a flexibilidade permitida, proveniente da negociação do preço feita por ambas as partes, permite eliminar o risco inerente a alterações futuras e imprevisíveis do preço no mercado "à vista" do ativo subjacente. No entanto, a reduzida formalidade dos mercados torna o processo negocial mais moroso, reduz a transparência e a liquidez da transação, eleva os custos de informação e o risco de incumprimento e potencia ineficiências no sistema de preços.

Tal como acontece nos contratos *forward*, os contratos de futuros [11] [15] são acordos para entrega diferida. No entanto, estes são padronizados e reversíveis, caracterizando-se pela fungibilidade e elevada negociabilidade. A fungibilidade manifesta-se em contratos que, por terem as mesmas características contratuais, ou seja, o mesmo ativo subjacente e a mesma data de entrega, são substitutos perfeitos. A negociabilidade é uma propriedade de contratos negociados em mercados formais, organizados e centralizados, que tende a reduzir os custos de informação e do processo negocial e a aumentar a sua liquidez, eficiência e transparência, garantindo as melhores e as mais justas condições para todos os agentes económicos. As características contratuais subjacentes aos contratos de futuros encontram-se plenamente padronizadas, ficando apenas a cargo das partes contratantes a negociação do seu preço.

Os contratos de opções [11] [15] são acordos onde o comprador adquire o direito, mas não o dever, de comprar (no caso das opções do tipo *call*), ou vender (no caso das opções do tipo *put*) uma determinada quantidade do ativo subjacente, a um preço previamente estabelecido numa data futura (opções de estilo europeu) ou durante um certo período de tempo (opções de estilo americano). O comprador paga um dado valor ao vendedor, que é obrigado a exercer o contrato se o comprador assim

o desejar. O preço a pagar pela celebração deste contrato é conhecido por prémio da opção; o valor ou preço acordado no contrato para a venda ou compra futura designa-se por preço de exercício ou *strike price*; a data limite para a celebração da compra ou venda acordada denomina-se por maturidade, data de exercício ou data de vencimento. Quem compra ou detém uma opção assume uma posição longa; por sua vez, quem vende ou subscreve assume uma posição curta. As opções podem ser negociadas em mercados formais ou informais. Nesta dissertação focamos o nosso estudo em contratos de opções de estilo europeu, ou seja, em opções que apenas podem ser exercidas pontualmente no final da sua maturidade.

Os contratos de opções são semelhantes aos contratos *forward* e aos contratos de futuros, no entanto, como o vendedor fica sujeito à vontade do comprador exercer ou não o seu direito, verifica-se assimetria nos termos de contrato. Isto acontece porque, racionalmente, o comprador apenas exerce a opção se o preço do ativo subjacente for superior ao preço de exercício, no caso das opções *call*, e o inverso, no caso das opções *put*. Caso o comprador decida não exercer a opção, o prejuízo que assume iguala o valor do prémio pago pela opção. Os ganhos e as perdas do vendedor são simétricos ao do comprador. Sejam $V_C(S_T, T, X)$ e $V_P(S_T, T, X)$, respetivamente, os preços das opções *call* e *put* europeias na data de vencimento T ; seja X o preço de exercício ou *strike price* e S_T o preço do ativo subjacente em T . Os perfis de ganhos das opções europeias, *call* e *put*, na data de expiração são dados, respetivamente, por [11] [15]:

$$V_C(S_T, T, X) = \max\{S_T - X, 0\}, \quad (2.1)$$

$$V_P(S_T, T, X) = \max\{X - S_T, 0\}. \quad (2.2)$$

O preço dos derivados financeiros, mais especificamente o preço das opções, permite obter informação acerca das expectativas dos investidores quanto ao preço do ativo subjacente na maturidade. Assim, é possível estimar a função densidade de probabilidade implícita num contexto de neutralidade face ao risco, considerando preços de opções relativas a um determinado ativo subjacente [12]. Neste contexto, é permitida a atribuição de probabilidades a uma vasta gama de acontecimentos futuros caracterizados por alterações do preço do ativo subjacente. No entanto, o desvio destas probabilidades implícitas, relativamente às probabilidades reais de mercado, depende do nível de aversão ao risco dos agentes económicos [3].

Existem vários métodos considerados na literatura, para extrair a função densidade de probabilidade neutra face ao risco (doravante designada por FDPN), estimada a partir de um conjunto de preços de opções observados no mercado; para tal podem ser consideradas três abordagens distintas: modelos paramétricos, modelos não paramétricos e modelos semi-paramétricos. Os modelos paramétricos consistem na estimação de parâmetros consistentes com os preços observados no mercado, considerando uma forma específica para a função densidade, por exemplo uma mistura de duas distribuições lognormais. A estimação por modelos não paramétricos apresenta mais flexibilidade pelo facto de não exigir restrições de parâmetros nem de uma especificação *a priori* para a estrutura da FDPN. A estimação por modelos semi-paramétricos combina especificidades dos dois tipos de estimação anteriores [19].

O princípio de neutralidade face ao risco [10] permite modelar o preço corrente de opções *call* e *put* europeias, no instante t , como o valor esperado do ganho futuro descontado à taxa de juro sem risco. Assim, verifica-se que:

$$V_C(X, t, T) = e^{-r(T-t)} \int_X^{+\infty} q(S_T)(S_T - X) dS_T, \quad (2.3)$$

$$V_P(X, t, T) = e^{-r(T-t)} \int_0^X q(S_T)(X - S_T) dS_T, \quad (2.4)$$

onde r é a taxa de juro sem risco com capitalização contínua e q representa a FDPN para o preço do ativo subjacente na maturidade, S_T . A fim de garantir que q representa uma função densidade de probabilidade é necessário que esta seja sempre não negativa e que o integral em todo o seu domínio seja igual a um.

Breedon e Litzenberger [5] demonstraram que seria possível estimar, a partir do preço de opções *call* europeias, a FDPN do preço do ativo subjacente, na data de vencimento do contrato. Assim, obtiveram:

$$\left. \frac{\partial^2 V_C(X, t, T)}{\partial X^2} \right|_{X=S_T} = e^{-r(T-t)} q(S_T). \quad (2.5)$$

A Equação (2.5) indica que a segunda derivada da função que modela os preços de uma opção *call* europeia (Equação (2.3)), em ordem ao seu preço de exercício, é proporcional à FDPN do ativo subjacente na maturidade.

A existência de um ambiente de neutralidade face ao risco está diretamente relacionada com a inexistência de oportunidades de arbitragem [19]. Diz-se que existem oportunidades de arbitragem quando existe uma probabilidade positiva de obter um retorno positivo, sem existir risco de perda de capital [11]. Assim, para obtermos FDPN viáveis é necessário assegurar que os preços das opções *call* não contenham oportunidades de arbitragem. De facto, verifica-se que a existência destas pode gerar valores negativos na FDPN [3]. Para evitar a existência de oportunidades de arbitragem deve ser garantido que a função definida pelos preços de opções *call* europeias, em ordem ao seu preço de exercício, é positiva, estritamente convexa e estritamente decrescente [7].

2.1 Modelo de Black-Scholes

O modelo de Black-Scholes [4] é considerado o modelo de determinação de preços teóricos de opções mais conhecido nos mercados financeiros. Este modelo é aplicável a opções com probabilidade nula de exercício antecipado, ou seja, em particular a opções *call* e *put* europeias.

Tal como acontece na grande parte dos modelos de avaliação financeira, assumimos que os mercados são perfeitos, contínuos e competitivos. Estes pressupostos asseguram que os ativos subjacentes são perfeitamente divisíveis e estão sempre disponíveis para troca, não existem taxas ou custos de transação, as taxas de juro ativas e passivas têm o mesmo valor, não há restrições na venda a descoberto (*short selling*) e que os investidores são *price-takers*.

O modelo de Black-Scholes assume que o preço do ativo subjacente segue um movimento Browniano geométrico (MBG). Deste modo, é considerado que o preço do ativo subjacente no instante t , S_t , segue uma lei de distribuição lognormal e satisfaz a equação estocástica:

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad (2.6)$$

onde dS_t são as alterações instantâneas do preço do ativo subjacente, dt representa um intervalo de tempo infinitesimal e W_t é um processo de Wiener ou do movimento Browniano. O parâmetro μ , conhecido por tendência ou *drift*, está associado ao retorno esperado do ativo. Tanto este parâmetro como a volatilidade do preço do ativo, σ , são considerados constantes ao longo de toda a vida da opção e para todos os preços de exercício.

Aplicando o Lema de Itô à Equação (2.6) obtemos que $\ln(S_T)$ segue uma distribuição normal de média $\alpha = \ln(S_t) + (\mu - \frac{1}{2}\sigma^2)(T - t)$ e desvio-padrão $\beta = \sigma\sqrt{T - t}$.

A equação de derivadas parciais de Black-Scholes modela o preço de uma opção do tipo *call* ou *put*, V , em função do preço do ativo subjacente, S_t , e do tempo t . Esta equação é representada pela seguinte expressão:

$$\frac{\partial V}{\partial t}(S_t, t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2}(S_t, t) + rS_t \frac{\partial V}{\partial S_t}(S_t, t) - rV(S_t, t) = 0. \quad (2.7)$$

Através da Equação (2.7) é possível verificar que o preço da opção depende da taxa de juro sem risco, r , que é conhecida e constante no tempo, e da volatilidade do preço do ativo subjacente, σ , mas não depende do *drift*, μ . Numa realidade onde os investidores são neutros ao risco, o retorno esperado de todos os títulos coincide com a taxa de juro sem de risco, ou seja, $\mu = r$. Tal acontece porque os investidores neutros ao risco não necessitam de um prémio que os induza a correr riscos e também porque o valor atual de cada *cash-flow* num mundo neutral face ao risco pode ser obtido descontando o seu valor esperado à taxa de juro sem risco [15].

De modo a encontrar a solução da equação de derivadas parciais de Black-Scholes (Equação (2.7)), apresentamos as suas condições finais e de fronteira. Como foi visto anteriormente, o valor de uma opção, *call* e *put*, na maturidade T , é dado, respetivamente, pela Equação (2.1) e pela Equação (2.2). Estas equações são consideradas as condições finais, também conhecidas por *payoffs*. As condições de fronteira são impostas quando S_t tende para 0 e quando tende para $+\infty$.

No caso das opções *call*, sabe-se que, quando $S_t \rightarrow +\infty$, a opção vai ser exercida, visto que se garante que $S_t > X$. O valor da opção *call* é, então, dado por:

$$V_C(S_t, t) \approx S_t - Xe^{-r(T-t)}, \quad t \geq 0. \quad (2.8)$$

Obtemos, assim, a primeira condição de fronteira para as opções *call*:

$$\lim_{S_t \rightarrow +\infty} [S_t - V_C(S_t, t)] = Xe^{-r(T-t)}, \quad t \geq 0. \quad (2.9)$$

A segunda condição de fronteira para opções *call* é definida para $S_t \rightarrow 0$ e é dada por:

$$\lim_{S_t \rightarrow 0^+} V_C(S_t, t) = 0, \quad t \geq 0. \quad (2.10)$$

Esta condição resulta do facto do preço da opção ser nulo se o valor do ativo subjacente também o for, ou seja, a opção não tem valor.

Analogamente, para as opções *put*, quando $S_t \rightarrow +\infty$, temos a seguinte condição de fronteira:

$$\lim_{S_t \rightarrow +\infty} V_P(S_t, t) = 0, \quad t \geq 0, \quad (2.11)$$

que traduz o facto da opção não ser exercida quando o valor do ativo subjacente é muito superior ao preço de exercício. Quando $S_t \rightarrow 0$, temos:

$$\lim_{S_t \rightarrow 0^+} [V_P(S_t, t) + S_t] = X e^{-r(T-t)}, \quad t \geq 0, \quad (2.12)$$

visto que quando $S_t = 0$, o preço da opção *put* vai igualar o preço de exercício descontado à taxa de juro sem risco, r . Isto permite-nos obter as fórmulas de preços de Black-Scholes, respetivamente, para opções *call* e *put*:

$$V_C(S_t, t) = S_t N(d_1) - e^{-r(T-t)} X N(d_2), \quad S_t > 0, \quad t \in [0, T], \quad (2.13)$$

$$V_P(S_t, t) = e^{-r(T-t)} X N(-d_2) - S_t N(-d_1), \quad S_t > 0, \quad t \in [0, T], \quad (2.14)$$

com

$$d_1 = \frac{\ln(S_t/X) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \quad (2.15)$$

$$d_2 = d_1 - \sigma\sqrt{T-t},$$

e onde $N(d)$ representa a função de distribuição acumulada de uma variável normal padronizada. Ou seja, é a probabilidade de uma variável que segue a lei de distribuição normal padronizada, $N(0, 1)$, ser menor que d , tal que $d \in \mathbb{R}$,

$$N(d) = P(Z \leq d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^d e^{-\frac{1}{2}x^2} dx. \quad (2.16)$$

A expressão $N(d_2)$ presente na Equação (2.13) representa a probabilidade da opção ser exercida num mundo neutral face ao risco [15]. Mais ainda, e como visto anteriormente, também pela análise das Equações (2.13), (2.14) e (2.15) podemos observar a ausência do parâmetro do *drift*, μ . Este facto revela que o valor de uma opção não depende das preferências de risco dos investidores.

As fórmulas de Black-Scholes apresentadas não consideram o pagamento de dividendos. Assumindo que o ativo subjacente paga uma taxa de dividendo contínua, δ , as fórmulas de Black-Scholes para opções do tipo *call* e *put* são dadas por:

$$V_C(S_t, t) = S_t e^{-\delta(T-t)} N(d_1) - e^{-r(T-t)} X N(d_2), \quad S_t > 0, \quad t \in [0, T], \quad (2.17)$$

e

$$V_P(S_t, t) = e^{-r(T-t)} X N(-d_2) - S_t e^{-\delta(T-t)} N(-d_1), \quad S_t > 0, \quad t \in [0, T], \quad (2.18)$$

onde

$$d_1 = \frac{\ln(S_t/X) + (r - \delta + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}, \quad (2.19)$$

$$d_2 = d_1 - \sigma\sqrt{T - t}.$$

Nas fórmulas de Black-Scholes existe uma variável que não é diretamente observável no mercado. Essa variável é a volatilidade do preço do ativo, σ , e portanto é necessário estimá-la. O método de estimação consiste em, utilizando valores observados no mercado, resolver a Equação (2.13), para o caso das opções *call*, ou a Equação (2.14), para o caso das opções *put*, de modo a encontrar σ pelo método de tentativa e erro [15], designando este valor por volatilidade implícita.

Evidências empíricas indicam que os pressupostos do modelo não se verificam na prática e a volatilidade não é nem constante no tempo, nem constante para diferentes preços de exercício [11]. Assim, maioria das vezes, verifica-se que o gráfico das volatilidades implícitas em função do preço de exercício tem o aspecto de um sorriso (conhecido como sorriso da volatilidade ou *volatility smile*). Isto significa que as opções *at-the-money* têm valores mais baixos para as volatilidades implícitas do que as opções *in-the-money* e *out-of-the-money* [7]. Se o modelo de Black-Scholes estivesse correto iríamos obter uma linha paralela ao eixo dos x (eixo dos preços de exercício). Na prática, é possível observar que a FDPN obtida, utilizando as volatilidades implícitas, não coincide com a função densidade de uma distribuição lognormal, como defende este modelo [22].

2.2 Modelos paramétricos utilizados na estimação da função densidade de probabilidade neutra face ao risco

Os modelos paramétricos ou assumem um processo para o preço do ativo subjacente ou consideram uma forma funcional para a função densidade, dependendo estes modelos de um número reduzido de parâmetros. Estes modelos permitem a extração da FDPN a partir de um conjunto de distribuições estatísticas e de um conjunto de dados observados no mercado [19].

Nesta secção apresentamos algumas características referentes a dois modelos paramétricos de estimação da FDPN: o modelo de Heston e a mistura de distribuições lognormais.

2.2.1 Modelo de Heston

O modelo de Heston [14] veio introduzir um elevado nível de realismo e complexidade porque, ao contrário do que se verifica no modelo de Black-Scholes, assume que a volatilidade não é constante. Assim, este modelo considera que a volatilidade é variável ao longo do tempo e comporta-se de acordo com um processo estocástico. Os autores definem que o preço e a variância do ativo seguem, respetivamente, os seguintes processos estocásticos:

$$dS_t = \mu S_t dt + S_t \sqrt{v_t} dW_1, \quad (2.20)$$

e

$$dv_t = k(\theta - v_t)dt + \sigma\sqrt{v_t}dW_2, \quad (2.21)$$

onde k e θ são parâmetros constantes positivos. O preço do ativo, S_t , segue um movimento Browniano geométrico, com a volatilidade σ a variar com o tempo; a variância no instante t , v_t , segue um processo de Cox-Ingersoll-Ross [9]. O parâmetro μ , designado *drift*, é a taxa de retorno instantânea esperada do ativo subjacente, θ corresponde à média de longo prazo da variância v_t , k é a taxa de reversão à média, ou seja, é a taxa a que v_t retorna à sua média de longo prazo e σ é a volatilidade instantânea do processo da variância. Quando $\sigma = 0$, a volatilidade é determinística, e os retornos continuamente compostos dos ativos seguem uma distribuição normal e quando $\sigma > 0$ ocorre um aumento da curtose nos retornos do ativo. W_1 e W_2 são dois processos de Wiener com coeficiente de correlação dado por ρ .

O coeficiente de correlação ρ afeta positivamente a assimetria do retorno dos ativos, influenciando a forma da FDPN. Isto significa que se $\rho > 0$ e o preço do ativo subir então a volatilidade do preço do ativo vai aumentar e vai produzir uma cauda direita mais pesada na FDPN. Se $\rho < 0$, uma diminuição do preço do ativo vai causar uma diminuição na volatilidade e o peso da cauda esquerda da FDPN vai aumentar.

2.2.2 Mistura de Distribuições Lognormais

A mistura de distribuições lognormais como processo de estimação para a FDPN foi proposta primeiramente por Bahra [3] e Melick e Thomas [18]. Bahra [3] utilizou a mistura de duas distribuições lognormais para obter a FDPN, usando opções sobre matérias-primas, transacionadas na LIFFE (*London International Financial Futures and Options Exchange*), e opções sobre divisas transacionadas na Bolsa de Valores de Filadélfia (PHLX). Melick e Thomas [18] estudaram o impacto da Guerra do Golfo Pérsico nos preços de opções americanas sobre contratos de futuros sobre petróleo, utilizando uma mistura de três distribuições lognormais.

Como tem sido documentado na literatura, as distribuições do preço do ativo financeiro podem encontrar-se na vizinhança das distribuições lognormais. Deste modo, os autores consideram plausível utilizar a soma ponderada das funções densidade de probabilidade de m distribuições lognormais diferentes, ou seja:

$$q(S_T) = \sum_{i=1}^m [\theta_i L(\alpha_i, \beta_i; S_T)], \quad (2.22)$$

onde $L(\alpha_i, \beta_i; S_T)$ representa a i -ésima função densidade de probabilidade de uma distribuição lognormal, com parâmetros $\alpha_i \in \mathbb{R}$ e $\beta_i > 0$. Os pesos das distribuições, $\theta_1, \dots, \theta_m$, devem satisfazer a condição $\sum_{i=1}^m \theta_i = 1$, $\theta_i > 0$ para $i = 1, \dots, m$.

Como referido anteriormente, o modelo de Black-Scholes impõe que a FDPN de S_T , $q(S_T)$, seja uma distribuição lognormal com parâmetros α e β , ou, alternativamente, que a FDPN do retorno do ativo subjacente siga uma distribuição normal de parâmetros μ e σ . Assim, e segundo Bahra [3]:

$$L(\alpha_i, \beta_i; S_T) = \frac{1}{S_T \beta_i \sqrt{2\pi}} e^{-\frac{(\ln(S_T) - \alpha_i)^2}{2\beta_i^2}}, \quad (2.23)$$

onde

$$\alpha_i = \ln(S_i) + \left(\mu_i - \frac{1}{2} \sigma_i^2 \right) (T - t) \quad \text{e} \quad \beta_i = \sigma_i \sqrt{T - t}, \quad i = 1, \dots, m. \quad (2.24)$$

A mistura de duas distribuições lognormais introduz flexibilidade na captura de algumas características dos dados, ajustando-se a diferentes cenários possíveis, inclusive a distribuições bimodais, caudas pesadas e assimetria na distribuição. A mistura de duas distribuições lognormais é descrita por cinco parâmetros: dois parâmetros para cada distribuição lognormal (α_1, β_1 correspondente à primeira distribuição e α_2, β_2 para a segunda), e um parâmetro de ponderação, θ , que descreve o peso relativo da primeira distribuição, com $\theta \in [0, 1]$.

Dado o pressuposto de $q(S_T)$ apresentado na Equação (2.22), temos, neste caso particular, a seguinte relação:

$$q(S_T) = \theta L(\alpha_1, \beta_1; S_T) + (1 - \theta) L(\alpha_2, \beta_2; S_T). \quad (2.25)$$

Através das Equações (2.3) e (2.4) obtemos as seguintes equações para os preços de opções *call* e *put* europeias dadas, respetivamente, por:

$$V_C(X, t, T) = e^{-r(T-t)} \int_X^{+\infty} [\theta L(\alpha_1, \beta_1; S_T) + (1 - \theta) L(\alpha_2, \beta_2; S_T)] (S_T - X) dS_T \quad (2.26)$$

e

$$V_P(X, t, T) = e^{-r(T-t)} \int_0^X [\theta L(\alpha_1, \beta_1; S_T) + (1 - \theta) L(\alpha_2, \beta_2; S_T)] (X - S_T) dS_T. \quad (2.27)$$

Bahra [3] deduziu as seguintes soluções analíticas para as Equações (2.26) e (2.27):

$$V_C(X, t, T) = e^{-r(T-t)} \left[\theta \left(e^{\alpha_1 + \frac{1}{2} \beta_1^2} N(d_1) - XN(d_2) \right) + (1 - \theta) \left(e^{\alpha_2 + \frac{1}{2} \beta_2^2} N(d_3) - XN(d_4) \right) \right], \quad (2.28)$$

$$V_P(X, t, T) = e^{-r(T-t)} \left[\theta \left(-e^{\alpha_1 + \frac{1}{2} \beta_1^2} N(-d_1) + XN(-d_2) \right) + (1 - \theta) \left(-e^{\alpha_2 + \frac{1}{2} \beta_2^2} N(-d_3) + XN(-d_4) \right) \right], \quad (2.29)$$

com

$$\begin{aligned} d_1 &= \frac{-\ln(X) + \alpha_1 + \beta_1^2}{\beta_1}, \\ d_2 &= d_1 - \beta_1, \\ d_3 &= \frac{-\ln(X) + \alpha_2 + \beta_2^2}{\beta_2}, \\ d_4 &= d_3 - \beta_2. \end{aligned} \quad (2.30)$$

As estimações dos parâmetros α_1 , β_1 , α_2 , β_2 e θ , são obtidas pela minimização da distância entre os preços teóricos de opções *call*, apresentados na Equação (2.28), e opções *put*, apresentados na Equação (2.29), gerados pela mistura de duas distribuições lognormais e os preços de mercado, no intervalo dos preços de exercício considerados. O problema de minimização é dado por:

$$\min_{\alpha_1, \alpha_2, \beta_1, \beta_2, \theta} \left[\sum_{i=1}^n (V_C(X_i, t, T) - \hat{V}_{C_i})^2 + \sum_{i=1}^n (V_P(X_i, t, T) - \hat{V}_{P_i})^2 + \left(\theta e^{\alpha_1 + \frac{1}{2}\beta_1^2} + (1 - \theta) e^{\alpha_2 + \frac{1}{2}\beta_2^2} - e^{r(T-t)} S_t \right)^2 \right], \quad (2.31)$$

sujeito a $\beta_1, \beta_2 > 0$ e $0 \leq \theta \leq 1$, ao longo do intervalo de preços de exercício observados.

O problema de minimização inclui, simultaneamente, o preço das opções *call* e *put*, pois estes são avaliados sob o mesmo tipo de distribuição. Na Equação (2.31), $e^{\alpha_i + \frac{1}{2}\beta_i^2}$, para $i = 1, 2$, representa a média da FDPN da distribuição lognormal i . Deste modo, verifica-se que a média ponderada é representada pela média da mistura das duas FDPN. O terceiro termo da equação, $e^{r(T-t)} S_t$ estabelece que o valor esperado da FDPN deve ser igual ao preço futuro do ativo subjacente e é introduzido para garantir que não existem oportunidades de arbitragem.

A principal crítica feita a este método é que os preços apresentam normalmente uma assimetria negativa enquanto que as distribuições lognormais apresentam assimetrias positivas [12].

2.3 Modelos semi-paramétricos utilizados na estimação da função densidade de probabilidade neutra face ao risco

Os modelos semi-paramétricos são caracterizados por incluírem especificidades dos modelos paramétricos e dos modelos não paramétricos, sendo que o número de parâmetros a ser estimado é superior ao número de parâmetros dos modelos paramétricos e inferior ao dos modelos não paramétricos.

Neste ponto são apresentados dois modelos semi-paramétricos para a estimação da FDPN: a mistura de funções hipergeométricas confluentes e a expansão de Edgeworth.

2.3.1 Mistura de funções hipergeométricas confluentes

A mistura de funções hipergeométricas confluentes permite a estimação da função densidade de probabilidade sem que seja necessário assumir uma forma específica para a distribuição. Este método consiste no uso de uma fórmula que engloba funções densidades de várias distribuições conhecidas como a normal, gama, gama inversa, weibull, pareto e mistura destas funções densidade de probabilidade.

Os pioneiros na utilização deste método na estimação da FDPN foram Abadir e Rockinger [1]. Estes autores representaram as densidades utilizando as seguintes funções hipergeométricas confluentes, designadas também por funções *Kummer*:

$${}_1F_1 \equiv \sum_{j=0}^{+\infty} \frac{(\alpha)_j s^j}{(\beta)_j j!} \equiv 1 + \frac{\alpha}{\beta} s + \frac{\alpha(\alpha+1)}{\beta(\beta+1)} \frac{s^2}{2} + \dots \quad (2.32)$$

onde

$$(a)_j \equiv (a)(a+1)\cdots(a+j-1) = \frac{\Gamma(a+j)}{\Gamma(a)}, \quad (2.33)$$

sendo Γ a função gama e $-\beta \notin \mathbb{N} \cup 0$. As funções *Kummer* podem assumir, consoante os valores dados aos seus parâmetros, diferentes classes de funções densidade de probabilidade [6].

No contexto da estimação da FDPN, considera-se o preço de uma opção *call* europeia como uma mistura de duas funções hipergeométricas confluentes:

$$V_C(X) \equiv c_1 + c_2X + \mathbb{I}_{X>m_1} a_1(X-m_1)^{b_1} {}_1F_1(a_2; a_3; b_2(X-m_1)^{b_3}) + (a_4) {}_1F_1(a_5; a_6; b_4(X-m_2)^2), \quad (2.34)$$

onde \mathbb{I} representa a função indicatriz e $-a_3, -a_6 \notin \mathbb{N} \cup \{0\}$, $b_2, b_4 \in \mathbb{R}^-$.

Os parâmetros $a_1, a_2, a_3, a_4, a_5, a_6, b_1, b_2, b_3, b_4, c_1, c_2, m_1$ e m_2 podem ser estimados minimizando o quadrado dos desvios entre os preços das opções *call* teóricas ou de mercado e os preços das opções *call* dados pela Equação (2.34). Assumindo que estamos perante um mundo neutral face ao risco, é possível calcular a FDPN, aplicando a fórmula de Breeden e Litzenberg apresentada na Equação (2.5), à fórmula do preço das opções *call* (Equação (2.34)).

A imposição de restrições a alguns dos parâmetros permite-nos obter funções densidade de probabilidade conhecidas. Alguns exemplos são apresentados na Tabela 2.1.

Tabela 2.1 Restrições dos parâmetros utilizados na mistura de funções hipergeométricas confluentes para a obtenção de funções densidade de probabilidade conhecidas.

Parâmetros	Leis de distribuição		
	Normal	Gama	Weibull
a_1	0	$\frac{(-b_2)^{b_1-1}}{\Gamma(b_1+1)}$	-1
a_2	-	$b_1 - 1$	$\frac{1}{b_3}$
a_3	-	$b_1 + 1$	$\frac{1}{b_3} + 1$
a_4	$\frac{1}{2\sqrt{-b_4\pi}}$	0	0
a_5	$-\frac{1}{2}$	-	-
a_6	$\frac{1}{2}$	-	-
b_1	-	-	1
b_3	-	1	-

2.3.2 Expansão de Edgeworth

Neste ponto apresentamos alguns aspetos relativos ao método desenvolvido por Jarrow e Rudd [16] que utiliza a expansão de Edgeworth no cálculo da FDPN. Este método capta desvios da curtose e da assimetria em relação à função distribuição lognormal. De facto, quando a curtose e a assimetria tomam valores padronizados para a função distribuição lognormal, o resultado corresponde ao obtido

pelo modelo de Black-Scholes. Corrado e Su [8] aplicaram o método baseado na expansão de Edgeworth a opções sobre o índice S&P 500.

A expansão de Edgeworth é de algum modo semelhante ao desenvolvimento em série de Taylor de uma função, no entanto, a expansão de Edgeworth é aplicada a funções densidade. Para obter a expansão de Edgeworth consideramos: uma variável aleatória S_T , uma função distribuição (cdf) Q , uma função distribuição aproximada L , uma função densidade de probabilidade (pdf) q , uma função densidade de probabilidade aproximada l e que a função característica de Q é dada por $\phi_Q(u) = \int_{-\infty}^{+\infty} e^{ius} q(s) ds$.

Considerando que os primeiros n momentos $\mu_j(Q)$ existem então os primeiros $n - 1$ cumulantes $\kappa_j(Q)$ também existem e são definidos pelo somatório:

$$\log(\phi_Q(u)) = \sum_{j=1}^{n-1} \kappa_j(Q) \frac{(iu)^j}{j!} + o(u^{n-1}). \quad (2.35)$$

Sendo assim, se a função característica $\phi_Q(u)$ for conhecida é possível obter os cumulantes considerando o desenvolvimento do seu logaritmo na vizinhança de $\mu = 0$.

As relações dadas entre os momentos centrados e os cumulantes, até à quarta ordem, são dados por:

$$\begin{aligned} \kappa_1(Q) &= \mu_1(Q), \\ \kappa_2(Q) &= \mu_2(Q), \\ \kappa_3(Q) &= \mu_3(Q), \\ \kappa_4(Q) &= \mu_4(Q) - 3\mu_2^2(Q). \end{aligned} \quad (2.36)$$

Sendo que o primeiro cumulante corresponde à média, o segundo à variância, o terceiro a uma medida de enviesamento e o quarto a uma medida de curtose. A mesma notação é utilizada para a função distribuição aproximada L .

Jarrow e Rudd [16] mostraram que, depois da imposição de que o primeiro momento da função densidade aproximada iguala o primeiro momento da verdadeira função densidade, a função densidade de probabilidade implícita pode ser definida como:

$$\begin{aligned} q(S_T) &= l(S_T) + \frac{\kappa_2(Q) - \kappa_2(L)}{2!} \frac{d^2 l(S_T)}{dS_T^2} - \frac{\kappa_3(Q) - \kappa_3(L)}{3!} \frac{d^3 l(S_T)}{dS_T^3} + \\ &+ \frac{\kappa_4(Q) - \kappa_4(L) + 3(\kappa_2(Q) - \kappa_2(L))^2}{4!} \frac{d^4 l(S_T)}{dS_T^4} + \varepsilon(S_T), \end{aligned} \quad (2.37)$$

em que o termo $\varepsilon(S_T)$ representa os restantes termos do desenvolvimento.

Jarrow e Rudd [16] sugeriram que a função aproximada utilizada fosse uma distribuição lognormal $l(S_T)$ e impuseram a condição $\kappa_2(Q) = \kappa_2(L)$. Assim, a partir da Equação (2.37), os autores deduziram a fórmula que permite obter o valor de uma opção *call*, $V_C(Q)$. A fórmula é dada por:

$$V_C(Q) = V_C(L) - e^{-r(T-t)} \frac{\kappa_3(Q) - \kappa_3(L)}{3!} \frac{dl(X)}{dS_T} + e^{-r(T-t)} \frac{\kappa_4(Q) - \kappa_4(L) + 3(\kappa_2(Q) - \kappa_2(L))^2}{4!} \frac{d^2l(X)}{dS_T^2} + \varepsilon(X), \quad (2.38)$$

onde $V_C(L)$ é dado pela fórmula de Black-Scholes, apresentada anteriormente na Equação (2.13).

2.4 Modelos não paramétricos utilizados na estimação da função densidade de probabilidade neutra face ao risco

Os métodos não paramétricos são ferramentas analíticas, que providenciam uma forma eficaz de encontrar estrutura de conjuntos de dados observados, sem a imposição de um modelo paramétrico particular. Estes métodos são particularmente eficazes quando os modelos paramétricos estandardizados são inadequados.

Nesta secção apresentamos o método de Kernel e o método da volatilidade implícita proposto por Shimko.

2.4.1 Método de Kernel

Ait-Sahalia e Lo [2] utilizaram um método não paramétrico, baseado em regressões estatísticas associadas a funções Kernel. Este método consiste em utilizar preços de mercado, relativos a um determinado período temporal, para estimar uma fórmula não paramétrica de preços das opções \hat{V}_C e, em seguida, diferenciar este estimador duas vezes em ordem a X de modo a obter a densidade de acordo com Breeden e Litzenberger [5].

Para estimar $V_C(\cdot)$ de forma não paramétrica, os autores consideraram um conjunto de dados históricos de preços de opções $\{V_{C_i}\}$ com características $Z_i \equiv [S_{t_i}, X_i, (T-t)_i, r_{t_i, (T-t)_i}, \delta_{t_i, (T-t)_i}]$. De seguida, os autores tentaram encontrar uma função V_C (e não um conjunto de parâmetros) que fosse tão próxima de $\{V_{C_i}\}$ quanto possível. Para encontrar esta solução, os autores utilizaram um método estatístico conhecido como regressão de Kernel. Este método é não paramétrico porque produz um estimador sem requerer que a função $V_C(\cdot)$ seja parametrizada por um número finito de parâmetros.

O estimador Kernel utilizado é o estimador de *Nadaraya-Watson*:

$$\hat{V}_C(Z) = \frac{\sum_{i=1}^n K\left(\frac{Z-Z_i}{h}\right)V_{C_i}}{\sum_{i=1}^n K\left(\frac{Z-Z_i}{h}\right)}, \quad (2.39)$$

onde K é uma função Kernel e h é um parâmetro de suavização que define a largura do intervalo referente à função Kernel utilizada. O parâmetro h determina o grau de suavização da estimativa, aproximando-se de zero à medida que aumenta a dimensão da amostra, n .

Este procedimento implica que o preço de uma opção com características Z seja calculado através de uma média ponderada de preços observados, V_{C_i} , atribuindo maior peso a opções com características Z_i próximas das características Z da opção cujo preço está a ser estimado. Quando h tende para zero a FDPN torna-se menos suave nos valores próximos de Z_i e é dado mais peso a valores da variável aleatória Z_i próximos de Z .

Depois de obtida a estimativa para a função dos preços das opções, obtém-se a FDPN através da aplicação da Equação (2.5), apresentada por Breeden e Litzenberger.

Métodos baseados em regressões Kernel são intensivos em termos de dados e, devido ao intervalo limitado de preços de exercício disponíveis num determinado instante do tempo, são, geralmente, de difícil implementação [21].

2.4.2 Método da Volatilidade Implícita proposto por Shimko

Os métodos da volatilidade implícita consistem em estabelecer uma aproximação para o sorriso da volatilidade, por exemplo, através de um polinómio. Esta abordagem substitui a extrapolação da FDPN a partir dos preços das opções, como visto nos métodos anteriores, pela extrapolação da FDPN através das volatilidades implícitas.

Shimko [20] foi um dos impulsionadores desta metodologia, propondo um método de ajuste de curvas ao sorriso da volatilidade, ajustando um polinómio quadrático à curva de volatilidade implícita, através do método dos mínimos quadrados. O autor apresenta a função da estrutura da volatilidade, σ , modelada como uma função parabólica dos preços de exercício, no intervalo dos preços de exercício transacionados:

$$\sigma = A_0 + A_1X + A_2X^2. \quad (2.40)$$

Fora deste intervalo assume-se que a volatilidade implícita é constante. Esta equação traduz que a volatilidade de uma opção i depende do preço de exercício X_i , com $i = 1, \dots, m$, onde m representa o número de preços observados.

Shimko [20] utilizou a paridade *put-call* para fazer uma regressão linear simples sobre o preço de exercício e encontrar os valores para a taxa de dividendo, δ , e para a taxa de juro, r :

$$V_C - V_P = S_t e^{-\delta(T-t)} - X e^{-r(T-t)}. \quad (2.41)$$

Após obtidos estes valores, iremos utilizar preços de opções *call* diretamente observados no mercado para calcular as volatilidades implícitas obtidas através das fórmulas de Black-Scholes. Em seguida, o autor propõe suavizar essas volatilidades implícitas de acordo com a relação quadrática apresentada na Equação (2.40).

A volatilidade suavizada pode ser usada para encontrar os preços de opções *call* suavizadas (usando Black-Scholes) e estes podem ser diferenciados para encontrar os valores da função de distribuição acumulada e da FDPN para cada valor possível do preço de exercício.

A FDPN, $q(X)$, calculada através da Equação (2.5), e a função de distribuição acumulada, $Q(X)$, são dadas, respetivamente, por:

$$q(X) = -n(d_2)[d_{2x} - (A_1 + 2A_2X)(1 - d_2d_{2x}) - 2A_2X], \quad (2.42)$$

e

$$Q(X) = 1 + Xn(d_2)(A_1 + 2A_2X) - N(d_2), \quad (2.43)$$

onde d_1 e d_2 são apresentados na Equação (2.15), N é a função de distribuição normal acumulada e

$$\begin{aligned} n(d_2) &= \frac{1}{\sqrt{2\pi}} e^{-d_2^2/2}, \\ d_{1x} &= -\frac{1}{Xv} + \left(1 - \frac{d_1}{v}\right) (A_1 + 2A_2X), \\ d_{2x} &= d_{1x} - (A_1 + 2A_2X), \\ v &= \sigma\sqrt{T-t}. \end{aligned} \quad (2.44)$$

Dado que o intervalo de preços de exercício válido é limitado, a distribuição implícita apenas será calculada para os valores entre o preço de exercício mais baixo e o preço de exercício mais alto. Shimko resolveu este problema ajustando uma distribuição lognormal a cada cauda de tal modo que o integral da função densidade de probabilidade, $q(S_T)$, sobre todo o domínio fosse igual a um. Uma das desvantagens do método é que a função obtida pode tomar valores negativos, não sendo por isso uma função densidade.

Capítulo 3

Métodos de estimação das funções densidade de probabilidade neutras face ao risco: análise comparativa

Neste capítulo estudamos o comportamento de diversos métodos de estimação da FDPN, que foram apresentados na secção anterior, testando a sua aplicabilidade a dados teóricos e a dados reais. Para tal recorreremos a diferentes abordagens: modelo de Black-Scholes (BS), mistura de duas distribuições lognormais (MLN), expansão de Edgeworth (EW) e método de Shimko. O desempenho destes métodos será testado usando preços de opções teóricos, gerados pelo modelo de Black-Scholes, e através de preços de opções sobre o índice de mercado *Standard and Poor's 500* (S&P 500), cuja transação ocorre no mercado da bolsa de opções de Chicago, *Chicago Board Options Exchange* (CBOE). A Tabela 3.1 indica quais os métodos utilizados na estimação das FDPN, de acordo com a origem dos preços.

Tabela 3.1 Origem dos preços e métodos utilizados na estimação das FDPN.

Origem dos preços das opções	Métodos de estimação
Preços teóricos gerados pelo modelo de Black-Scholes	- Mistura de duas distribuições lognormais
	- Expansão de Edgeworth
	- Método da volatilidade implícita proposto por Shimko
Preços de opções sobre o índice S&P 500	- Modelo de Black-Scholes
	- Mistura de duas distribuições lognormais
	- Expansão de Edgeworth
	- Método da volatilidade implícita proposto por Shimko

Para reduzir a ordem de grandeza dos dados e evitar problemas de estimação da FDPN, utilizamos a transformação dos dados tal como foi proposta por Fengler e Hin [13]. Assim, para os dados teóricos, consideramos o escalonamento dos dados, designado por *moneyness*, fazendo a seguinte redução da dimensionalidade: $X' = X/S_t$ e $S'_t = 1$. No estudo efetuado com dados de mercado, consideramos que o ativo subjacente é um contrato *forward* com a mesma maturidade que o contrato de opção. A

utilização do contrato *forward* como ativo subjacente permite negligenciar a taxa de dividendo. Assim, considerando que o preço do contrato *forward* é dado por F_t^T , aplicamos o seguinte escalonamento aos dados: $X' = X/F_t^T$ e $F_t^{T'} = 1$. Este escalonamento é designado por *forward-moneyness*.

Com o objetivo de obter uma análise mais detalhada do desempenho dos diferentes métodos, e de modo a possibilitar uma comparação mais precisa do comportamento dos modelos, examinamos os enviesamentos dos preços das opções. Existem diversos instrumentos estatísticos que permitem fazer esta comparação, no entanto nesta dissertação, utilizamos o *root mean square error* (RMSE) por ser bastante intuitivo e de fácil aplicação e cuja expressão matemática é:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (V(X_i, t, T) - V_i)^2}. \quad (3.1)$$

O RMSE é uma medida de erro da qualidade global do ajuste do método estimado, comparando os preços de opções, *call* ou *put*, $V(X_i, t, T)$, estimados pelos diferentes métodos, com os valores teóricos ou de mercado, V_i . Quanto menor for o valor obtido no RMSE, mais próximos são os valores estimados dos valores reais.

3.1 Metodologia para a estimação da função densidade de probabilidade neutra face ao risco

Na análise teórica, o modelo de Black-Scholes (BS) é utilizado para gerar preços teóricos de opções, sendo que a FDPN associada é a função densidade lognormal teórica. No caso das opções de mercado, utilizamos a função *extract.bsm.density* do *package RND* do programa R, de forma a estimar a FDPN associada a este modelo. Os *inputs* desta função são os valores referentes às opções de mercado, considerados para uma determinada maturidade. Do processo de estimação extraem-se os parâmetros que permitem definir a densidade lognormal, conforme o modelo de Black-Scholes. Seguidamente calculam-se os preços das opções *call* e *put*, de acordo com os parâmetros obtidos anteriormente. Para este fim, utilizamos a função *price.bsm.option* do *package RND* do R, que calcula os preços das opções de acordo com as Equações (2.17), (2.18) e (2.19). Uma vez obtidos os valores das opções calculados pelo modelo Black-Scholes, comparamos os mesmos com os preços das opções de mercado, através da análise gráfica e do RMSE. Deste modo, é possível produzir a FDPN, utilizando os valores estimados anteriormente e usando a função *dlnorm*, pertencente ao *package RND* do R.

Para a estimação do método que implementa a mistura de duas distribuições lognormais (MLN) é utilizada a função *extract.mln.density* do *package RND* do programa R. Esta função utiliza os valores teóricos ou os valores de mercado considerados para uma determinada maturidade de modo a obter os parâmetros que optimizam a Equação (2.31). Os *outputs* da função são: o peso relativo de uma das funções densidade, θ , sendo que o peso da outra função é $(1 - \theta)$, a média e o desvio-padrão das duas funções lognormais, α_1 e β_1 para a primeira e α_2 e β_2 para a segunda. De seguida, utilizamos a função *price.mln.option* do *package RND* do R para calcular os preços das opções estimados pela mistura de duas distribuições lognormais, utilizando os parâmetros obtidos anteriormente. Esta função calcula os preços segundo as Equações (2.28), (2.29) e (2.30). Estes preços podem ser comparados

com os preços teóricos ou de mercado, através da análise gráfica e do RMSE. Assim, o gráfico da FDPN é obtido recorrendo à função *dmln* do *package RND* do software R.

Outro método utilizado no nosso estudo é a Expansão de Edgeworth (EW), implementado através da função *extract.ew.density* do *package RND* do R. Esta função utiliza dados de opções *call* e devolve a volatilidade, σ , a curtose e a assimetria dos dados. Depois de estimados estes parâmetros iremos utilizamo-los no cálculo do preço de opções pelo método da Expansão de Edgeworth, como indicado na Equação (2.38). Este cálculo requer o uso da função *price.ew.option* incluído no *package RND* do R. Por análise gráfica e do cálculo do RMSE, é possível comparar os preços das opções estimadas com os preços das opções teóricas ou de mercado. A FDPN é traçada recorrendo à ajuda da função *dew* do *package RND* do R e dos valores estimados anteriormente.

O último método testado é o método da volatilidade implícita proposto por Shimko. A implementação deste é feita através da função *extract.shimko.density* do *package RND* do R. Deste modo, é possível estimar os parâmetros da função quadrática, A_0 , A_1 e A_2 , apresentada na Equação (2.40). Seguidamente é nosso objectivo estimar o preço das opções através deste método e compará-los com os preços teóricos ou com os preços de mercado utilizados neste estudo. Para obter o preço das opções *call* e *put* recorreremos à função *price.shimko.option* do *package RND* do R. Seguidamente, através da análise gráfica e do cálculo do RMSE, comparamos esses preços com os utilizados no nosso estudo. A FDPN é calculada a partir dos parâmetros previamente estimados e da função *dshimko* do *package RND* do R.

3.2 Aplicação dos diferentes métodos e resultados

3.2.1 Variação das funções densidade de probabilidade de distribuições lognormais com a maturidade

Com o objetivo de estudar a variação das funções densidade de probabilidade de distribuições lognormais (LN) teóricas, como proposto pelo modelo de Black-Scholes, consideramos quatro diferentes maturidades, *ceteris paribus*. Para tal definimos um conjunto A de dados onde arbitramos que a taxa de juro, r , toma o valor de 5%, a volatilidade, σ , de 20% e a taxa de dividendo, δ , de 10%. O preço do ativo subjacente, S , tem um valor inicial de 1000 e são considerados onze preços de exercício equidistantes, a variar entre 500 e 1500. O estudo é efetuado para quatro maturidades diferentes: um mês, três meses, seis meses e doze meses.

As diferentes estimações das funções densidade de probabilidade de distribuições lognormais teóricas estão representadas na Figura 3.1. De modo a ser possível fazer um estudo mais pormenorizado apresentamos na Tabela 3.2 a média, a variância, o coeficiente de assimetria e o excesso de curtose das quatro funções densidade de probabilidade de distribuições lognormais teóricas.

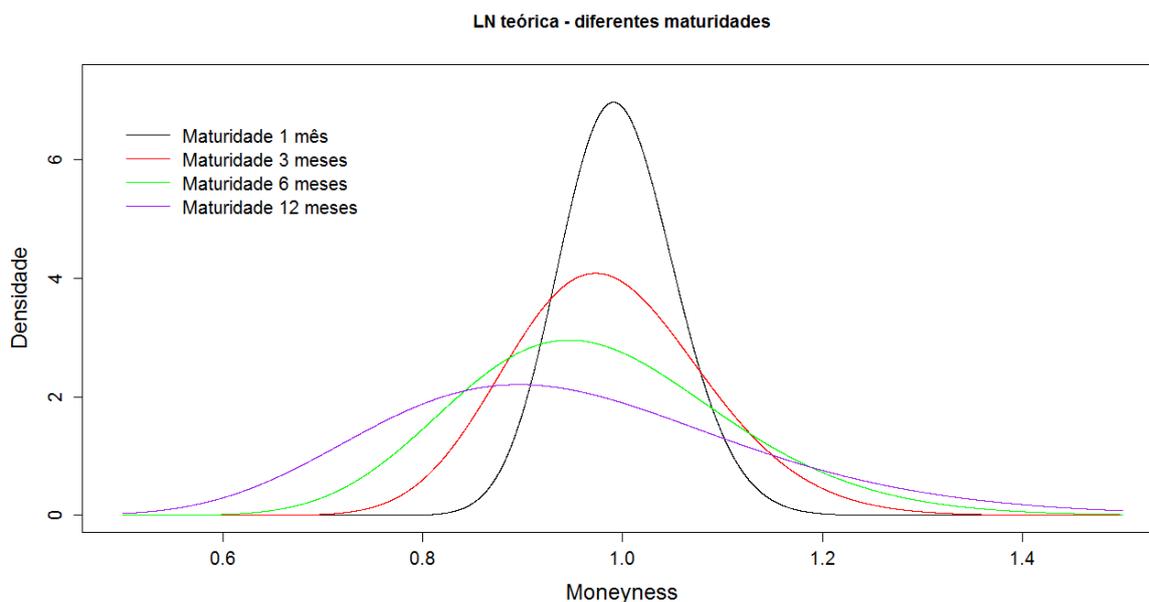


Fig. 3.1 Estimação das funções densidade de probabilidade lognormais teóricas, referentes ao conjunto A e quatro maturidades diferentes (1 mês, 3 meses, 6 meses e 12 meses).

Tabela 3.2 Média, variância, coeficiente de assimetria e excesso de curtose das funções densidade de probabilidade lognormais teóricas obtidas para quatro maturidades diferentes (1 mês, 3 meses, 6 meses e 12 meses).

Maturidade	Média	Variância	Assimetria	Excesso de Curtose
1 mês	0.9958	0.0033	0.1735	0.0536
3 meses	0.9876	0.0098	0.3001	0.1536
6 meses	0.9748	0.0189	0.3785	0.1219
12 meses	0.9459	0.0332	0.3886	-0.1533

A diminuição da média e o aumento da variância com o aumento do tempo até à maturidade são resultados possíveis de observar pela análise da Figura 3.1. Na Tabela 3.2 observa-se que o coeficiente de assimetria é positivo para todas as maturidades, ou seja, as caudas direitas de todas as funções densidade de probabilidade apresentadas são mais pesadas do que as caudas esquerdas. Este resultado traduz a existência de uma maior probabilidade de retorno significativo positivo, do que de um retorno significativo negativo. Este coeficiente aumenta com o aumento do tempo até à maturidade da opção. Este aumento do coeficiente de assimetria significa que os investidores especulam que a probabilidade de obter um retorno positivo é cada vez maior.

O excesso de curtose é uma medida padronizada que indica o grau de concentração dos valores da distribuição em torno do centro dessa mesma distribuição. Quando o excesso de curtose é positivo a distribuição diz-se leptocúrtica e indica que a função tem caudas mais pesadas que as caudas de uma função distribuição normal com a mesma média e a mesma variância. Isto traduz-se numa menor probabilidade de ocorrência de acontecimentos extremos. Quando o excesso de curtose é negativo a distribuição diz-se platicúrtica e função tem caudas mais "finas" que as caudas da função distribuição normal. A distribuição designa-se mesocúrtica quando o excesso de curtose é nulo.

Através da análise da Figura 3.1 e da Tabela 3.2 observa-se também que, por o ativo estar mais exposto à incerteza com o aumento do tempo até à maturidade, a probabilidade de se assumirem valores mais afastados do inicial é maior.

Foi feito um estudo análogo utilizando os restantes métodos e foram obtidos resultados semelhantes das FDPN com a variação do tempo até à maturidade.

3.2.2 Estimação da função densidade de probabilidade neutra face ao risco utilizando dados teóricos

Utilizando dois conjuntos diferentes de dados (B e C), para o mesmo ativo subjacente, geramos, pelo modelo de Black-Scholes, preços teóricos de opções. Deste modo, arbitramos que o preço do ativo subjacente, S_t , no instante inicial é 1000 e que o conjunto B de dados é composto por quinze preços de exercício, compreendidos entre 50 e 2850, igualmente espaçados com intervalo de 200 e é caracterizado por valores elevados da taxa de juro sem risco, r , do tempo para a maturidade, $T - t$, da volatilidade, σ , e da taxa de dividendo, δ . O conjunto C de dados é constituído por quinze preços de exercício, entre 993 e 1007, igualmente espaçados e é caracterizado por valores reduzidos destes parâmetros. Na Tabela 3.3 apresentam-se os valores dos parâmetros arbitrados para os conjuntos B e C.

Tabela 3.3 Parâmetros dos dois conjuntos de dados utilizados na geração pelo modelo de Black-Scholes de preços teóricos

Parâmetros	Conjunto B	Conjunto C
S_t	1000	1000
r	90%	0.96%
$T - t$	1 ano	15 dias
σ	70%	1%
δ	90%	0.2%

Utilizando os conjuntos de dados B e C caracterizados na Tabela 3.3 estimamos preços de opções *call* e *put* de acordo com as fórmulas de Black-Scholes, pela função *price.bsm.option* do *package RND*.

Como foi referido anteriormente, uma das hipóteses do modelo de Black-Scholes é que o preço do ativo subjacente segue um movimento Browniano geométrico e, portanto, a FDPN associada a preços gerados por este modelo segue uma distribuição lognormal. Deste modo, é nosso objetivo comparar a FDPN gerada por uma distribuição lognormal teórica com as FDPN estimadas pelos outros três métodos de estudo: mistura de duas distribuições lognormais (MLN), expansão de Edgeworth (EW) e método de Shimko.

Iniciamos o nosso estudo estimando preços de opções de acordo com os métodos em estudo para o conjunto de dados B. Deste modo, pela análise da Figura 3.2, é possível comparar os preços de opções *call* estimados pelo modelo de Black-Scholes com os preços de opções *call* estimados pelos métodos MLN. Observa-se que os preços de opções *call* estimados pelos três modelos, originam boas aproximações dos preços de opções *call* estimados pelo modelo de Black-Scholes.

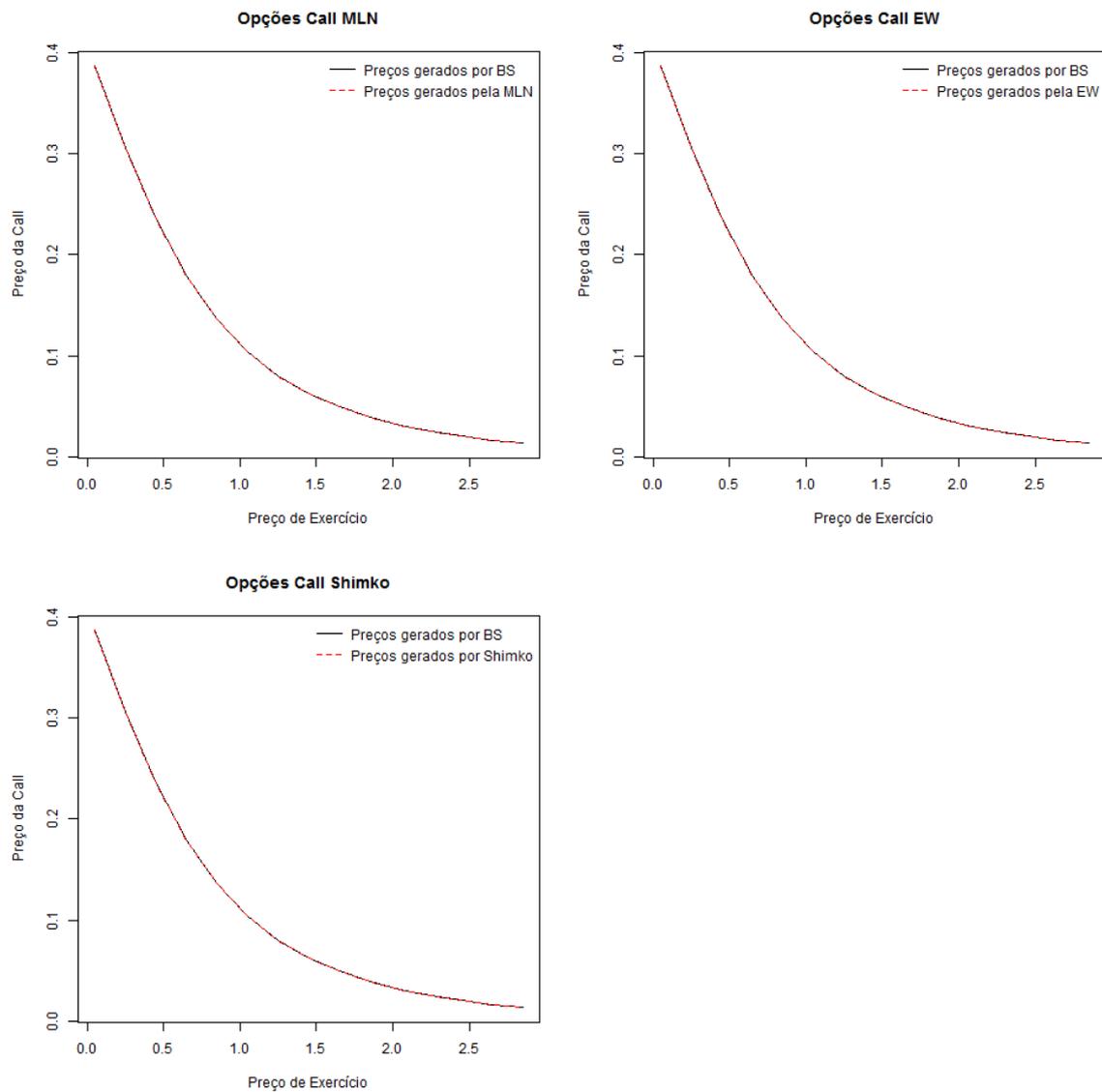


Fig. 3.2 Comparação entre preços de opções *call* gerados de acordo com a fórmula de Black-Scholes e preços de opções *call* gerados pelos métodos da MLN, da EW e de Shimko, para o conjunto B de dados.

Para uma análise mais aprofundada calculamos o RMSE, cujos resultados são apresentados na Tabela 3.4.

Tabela 3.4 RMSE - Preços de opções estimados para o conjunto B.

Preços de Black-Scholes	MLN	EW	Shimko
Opções Call	2.225×10^{-6}	2.729×10^{-17}	5.346×10^{-7}
Opções Put	2.143×10^{-6}	7.344×10^{-17}	5.346×10^{-7}

Deste modo, torna-se claro que, para o conjunto B, o método que apresenta o melhor ajuste, tanto para opções *call* como para opções *put*, é a expansão de Edgeworth, por apresentar o menor RMSE.

Estamos, agora, em condições de apresentar as FDPN geradas pelos diferentes métodos. As diferentes FDPN podem ser observadas na Figura 3.3.

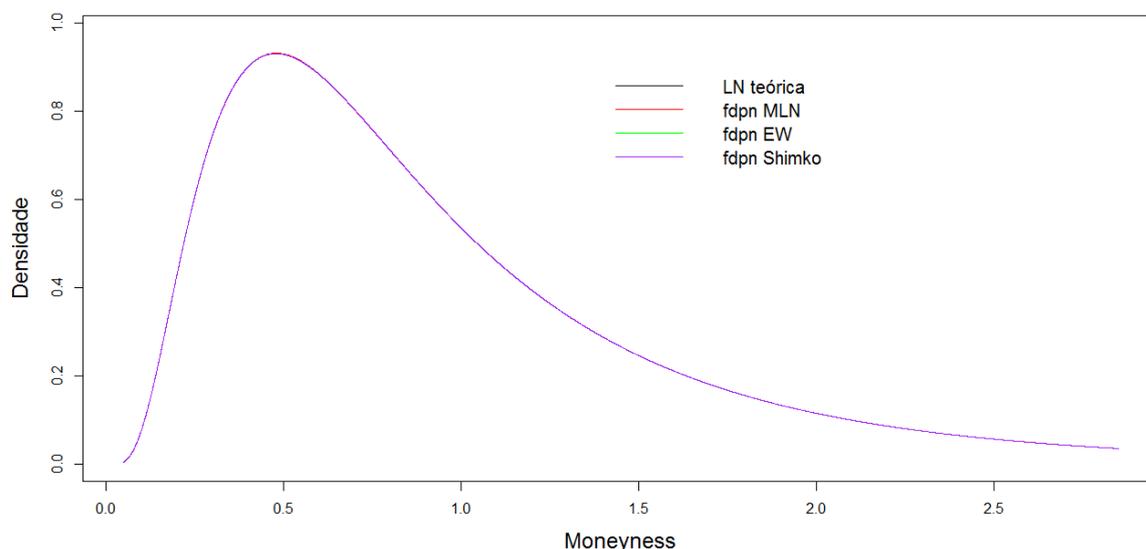


Fig. 3.3 Estimação das funções densidade de probabilidade neutras face ao risco, para dados teóricos de Black-Scholes, referentes ao conjunto B.

Constata-se na Figura 3.3 que as FDPN são coincidentes o que indica que todos os métodos geram aproximações quase perfeitas relativamente à FDPN teórica, ou seja, à função densidade de probabilidade da distribuição lognormal gerada pelo modelo de Black-Scholes. Todas as estimações respeitam, por isso, a assimetria positiva, tendo uma cauda direita da função densidade de probabilidade mais pesada que a cauda esquerda. Isto significa que os investidores esperam que a probabilidade de haver um retorno positivo do ativo subjacente seja superior à probabilidade de haver um retorno negativo. Esta conclusão vai ao encontro do esperado, visto que as funções densidade lognormais são caracterizadas por terem uma assimetria positiva.

Iremos agora estudar o conjunto C de dados. Analogamente ao que fizemos para o conjunto B, escalonamos os dados e comparamos os preços de opções gerados pelo modelo de Black-Scholes com os preços estimados pelos restantes métodos em estudo. Os resultados encontram-se na Figura 3.4 e na Tabela 3.5.

Tabela 3.5 RMSE - Preços de opções estimados para o conjunto C.

Preços de Black-Scholes	MLN	EW	Shimko
Opções Call	6.022×10^{-7}	3.593×10^{-6}	6.561×10^{-5}
Opções Put	4.729×10^{-7}	3.593×10^{-6}	2.876×10^{-3}

No conjunto de dados C, apesar de não existir um método que, de forma evidente melhor se ajuste aos preços de opções estimados pelo modelo de Black-Scholes, como aconteceu para o conjunto B, conclui-se que a MLN é o método que apresenta menor desvio.

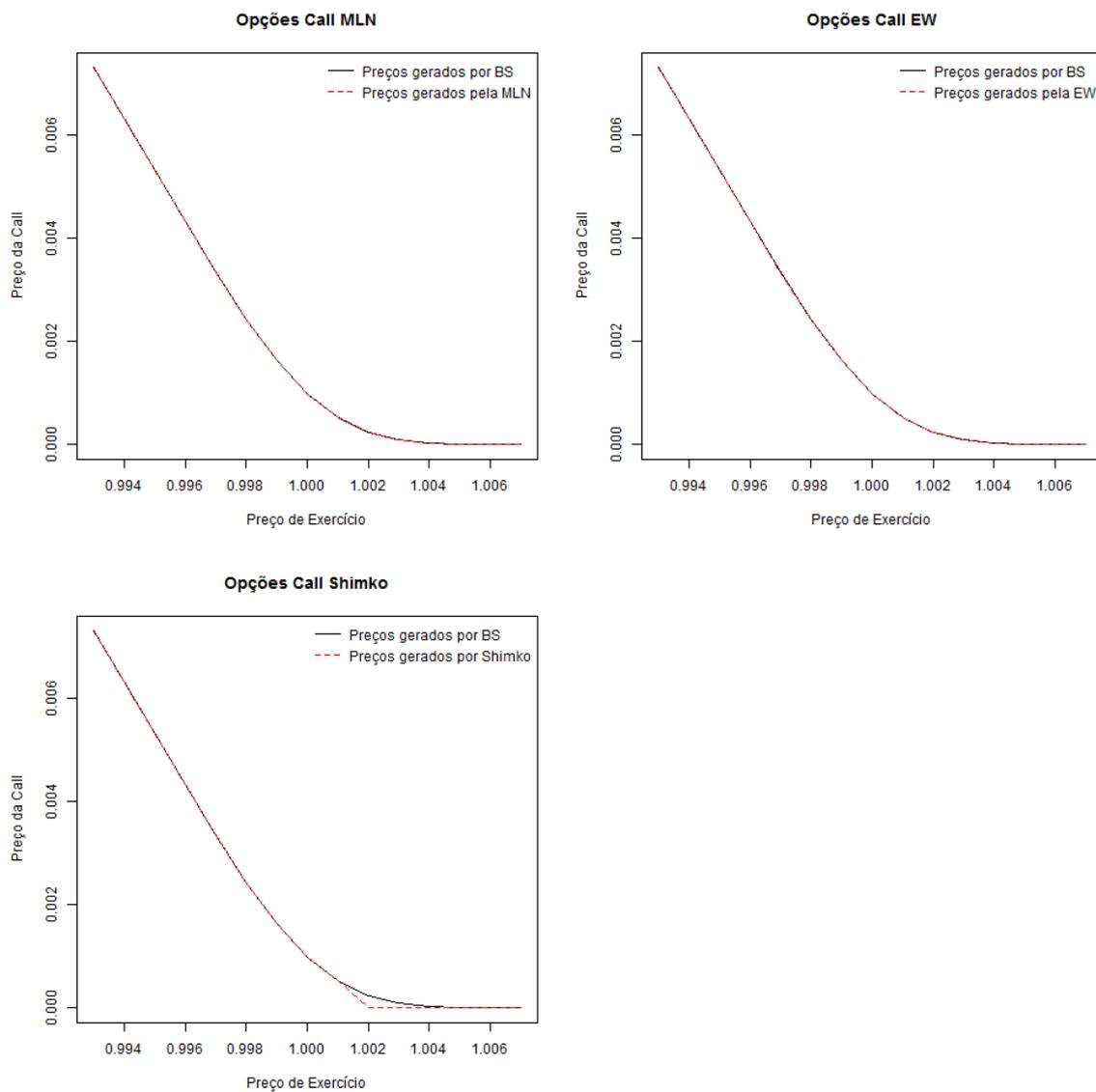


Fig. 3.4 Comparação entre preços de opções *call* gerados de acordo com a fórmula de Black-Scholes e preços de opções *call* gerados pelos métodos da MLN, da EW e de Shimko, para o conjunto C de dados.

As estimações das FDPN para os diferentes métodos encontram-se representadas na Figura 3.5, observando-se aí, que as curvas são praticamente coincidentes. No entanto, o sinal do coeficiente de assimetria não se revela tão intuitivo como no caso anterior.

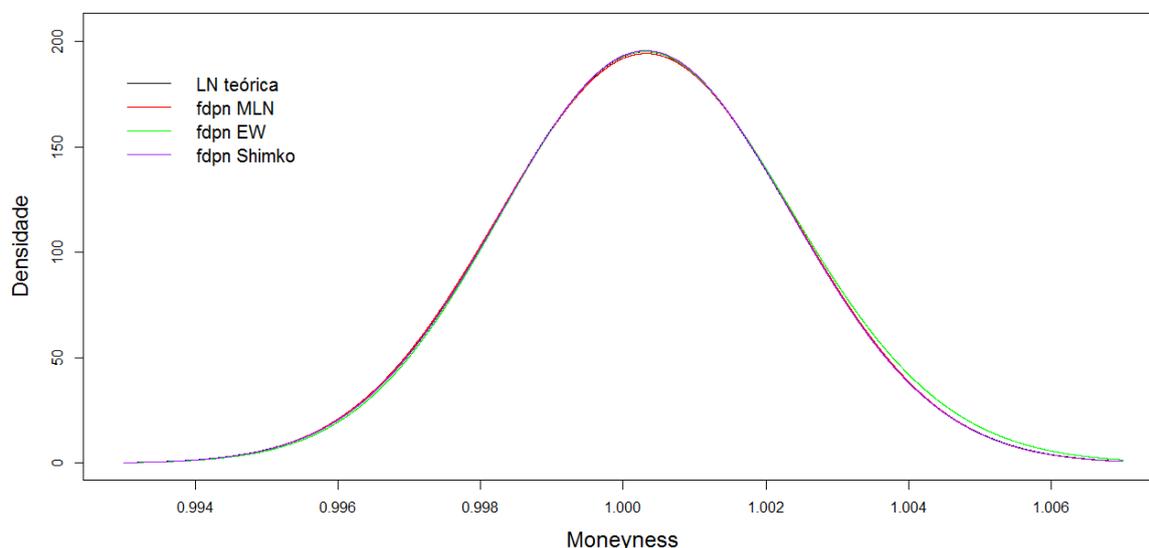


Fig. 3.5 Estimação das funções densidade de probabilidade neutras face ao risco, para dados teóricos de Black-Scholes, referentes ao conjunto C.

3.2.3 Estimação da função densidade de probabilidade neutra face ao risco utilizando dados de mercado: opções sobre o índice S&P 500

O índice *Standard and Poor's* 500 (S&P 500) é composto por 500 ações de empresas líderes e é considerado uma referência no mercado de ações americano. As opções sobre este índice são ativamente negociadas no *Chicago Board Options Exchange* (CBOE).

As opções S&P 500 são particularmente utilizadas para a estimação da FDPN não só porque o índice subjacente é aceite como uma aproximação para o portefólio de mercado americano mas também devido à sua elevada liquidez. Os dados utilizados referem-se a opções de estilo europeu e a sua data de vencimento corresponde à terceira sexta-feira do mês de vencimento da opção.

Vamos considerar uma única data inicial, 4 de janeiro de 2016, escolhida aleatoriamente, para ilustrar a extração da FDPN prática. Vamos considerar sete datas de vencimento distintas - 15 de janeiro (11 dias), 19 de fevereiro (46 dias), 18 de março (74 dias), 15 de abril (102 dias), 17 de junho (165 dias), 16 de setembro (256 dias) e 16 de dezembro de 2016 (347 dias).

De acordo, por exemplo com Jondeau *et al.* [17], utiliza-se a média aritmética dos preços de fecho diário de compra (*bid*) e de venda (*ask*), designada de cotação, para calcular o valor do índice pois, por estes serem revistos com mais frequência, torna-se mais fiável que utilizar o preço de transação. O preço *bid* é o preço mais alto que um investidor está disposto a pagar pela opção e o preço *ask* é o preço mais baixo a que um investidor está disposto a vender essa opção. Estes valores são retirados, para cada preço de exercício, no final do dia de transação.

Os nossos dados foram selecionados de modo a respeitarem o princípio de não arbitragem [7], ou seja, são escolhidas as opções que respeitem a não negatividade, a convexidade e a monotonia

decrecente dos preços das opções *call*, em ordem ao preço de exercício. De modo a evitar opções mal cotadas, foram excluídas também as opções que têm um preço *bid* igual a zero. A contagem dos dias utilizada é a convenção *atual/360* pois esta convenção está associada à anualização das taxas de juro. As taxas de juro relativas ao dia 4 de janeiro foram retiradas do site da Reserva Federal do Banco Central dos Estados Unidos da América e interpoladas para cada período de tempo.

Como referido anteriormente, iremos considerar que o ativo subjacente é um contrato *forward*. Assim, iremos calcular o preço implícito *forward*, para cada data de expiração. Este calculo é feito, para cada maturidade, através da paridade *put-call* do par *put-call* mais *at-the-money*, ou seja, aquele cuja diferença, em termos absolutos, entre o preço da opção *call* e da opção *put* é menor [13]. Assim, iremos calcular, para cada maturidade:

$$F_t^T = X^{PC} + (V_C^{PC} - V_P^{PC})e^{r(T-t)}, \tag{3.2}$$

onde F_t^T é o preço implícito *forward* em t para a maturidade T , X^{PC} é o preço de exercício referente ao par *put-call* mais *at-the-money* e V_C^{PC} e V_P^{PC} são, respetivamente, os preços das opções *call* e *put* do par considerado.

Os valores para as diferentes maturidades encontram-se expressos na Tabela 3.6. Nela apresentamos a data até à maturidade, ou seja, $T - t$ em dias, de n observações, X^{min} e X^{max} referem-se, respetivamente, aos preços de exercício mínimo e máximo, V_C^{min} e V_C^{max} são, respetivamente, os preços mínimo e máximo das opções *call*, V_P^{min} e V_P^{max} são, respetivamente, os preços mínimo e máximo das opções *put*, F_t^T é o preço implícito *forward* em t para a maturidade T e r é a taxa de juro sem risco, em percentagem, interpolada para cada maturidade.

Tabela 3.6 Sumário dos dados utilizados para as diferentes maturidades.

$T - t$ (dias)	n	X^{min}	X^{max}	V_C^{min}	V_C^{max}	V_P^{min}	V_P^{max}	F_t^T	r (%)
11	45	1720	2175	0.075	296.700	0.475	158.600	2016.50	0.0623
46	58	1605	2225	0.200	408.700	2.175	212.850	2012.25	0.1833
74	76	1540	2250	0.425	472.450	3.850	240.650	2009.30	0.2067
102	36	1425	2300	0.375	585.500	3.800	292.000	2007.71	0.2560
165	44	1325	2425	0.500	680.300	5.650	421.700	2002.15	0.4450
256	46	1300	2500	0.725	702.250	11.050	501.300	1995.98	0.5407
347	52	1275	2575	1.025	724.800	16.200	579.500	1990.80	0.6013

Pela análise desta tabela, é possível verificar que o aumento da maturidade provoca uma diminuição do preço implícito *forward*; tal significa que o mercado está em *backwardation*. Este termo é utilizado quando a curva a prazo dos preços dos contratos *forward* tem uma inclinação descendente. Este fenómeno está intimamente relacionado com a diminuição da procura do ativo subjacente ou com a perceção de escassez do ativo no presente, pelo que os agentes económicos que o detêm só estão dispostos a vendê-lo caso recebam um prémio adicional para a sua entrega imediata.

A utilização de um intervalo de preços de exercício crescente com a maturidade deve-se ao facto de as opções estarem mais expostas a variações de preços mais acentuadas, tanto positivas

como negativas, quanto maior for o tempo até à maturidade. Este fenómeno está relacionado com a possibilidade de existirem alterações cada vez mais substanciais do preço do ativo com o aumento do tempo até à maturidade e é intuitivamente comparável com o aumento da variância, vista no caso teórico.

Seguidamente, escalonamos os dados segundo Fengler e Hin [13] e estimamos os preços das opções *call* e *put* de acordo com os métodos em estudo. Nas Figuras 3.6, 3.7, 3.8, 3.9, 3.10, 3.11 e 3.12 são comparados os preços das opções *call* estimados pelos métodos com os preços das opções *call* sobre o índice S&P 500 para as diferentes maturidades e nas Tabelas 3.7 e 3.8 comparamos o RMSE obtidos para os dois tipos de opções.

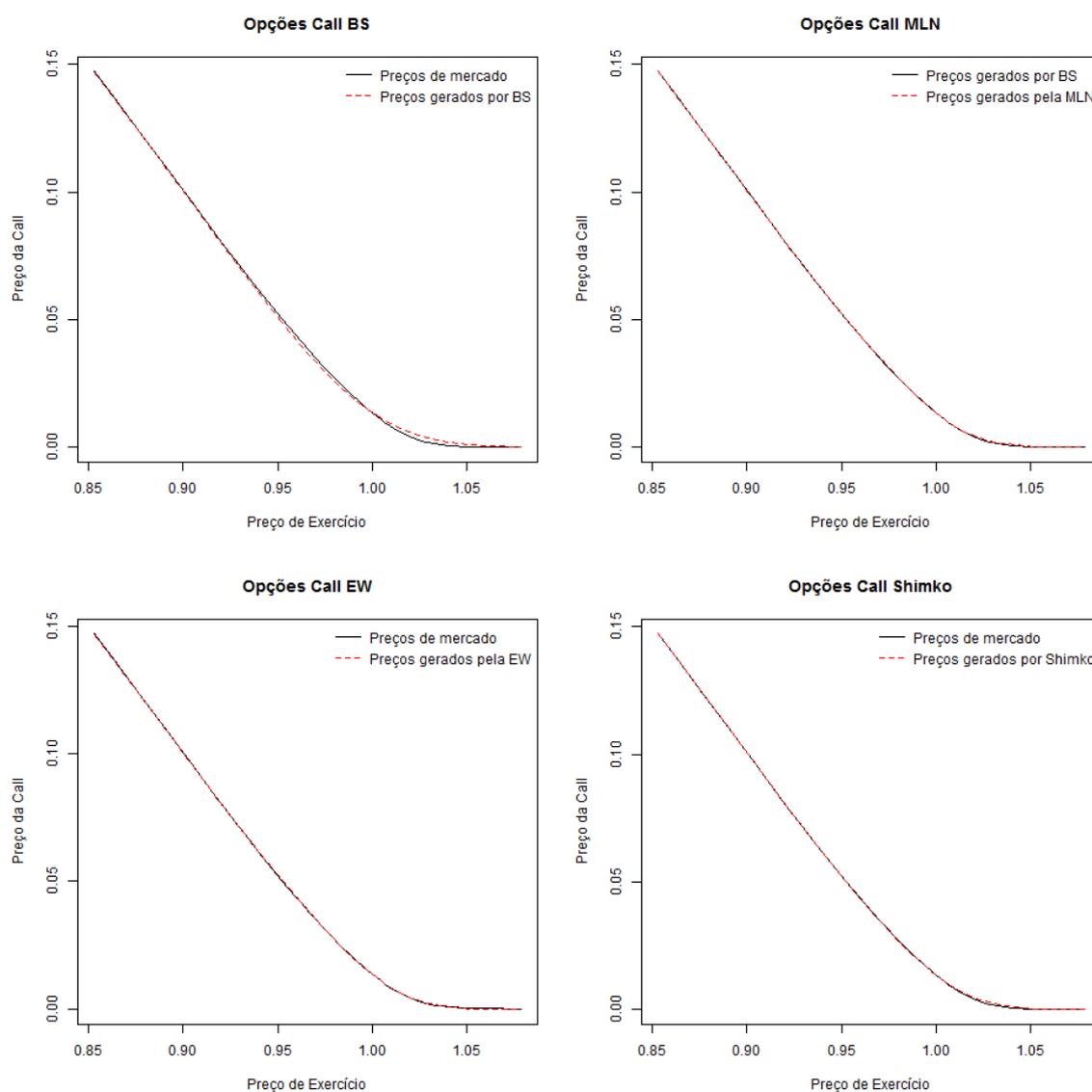


Fig. 3.6 Comparação entre preços de opções *call* de mercado e preços de opções *call* estimados pelos modelos de BS, MLN, EW e Shimko para a primeira maturidade (11 dias).

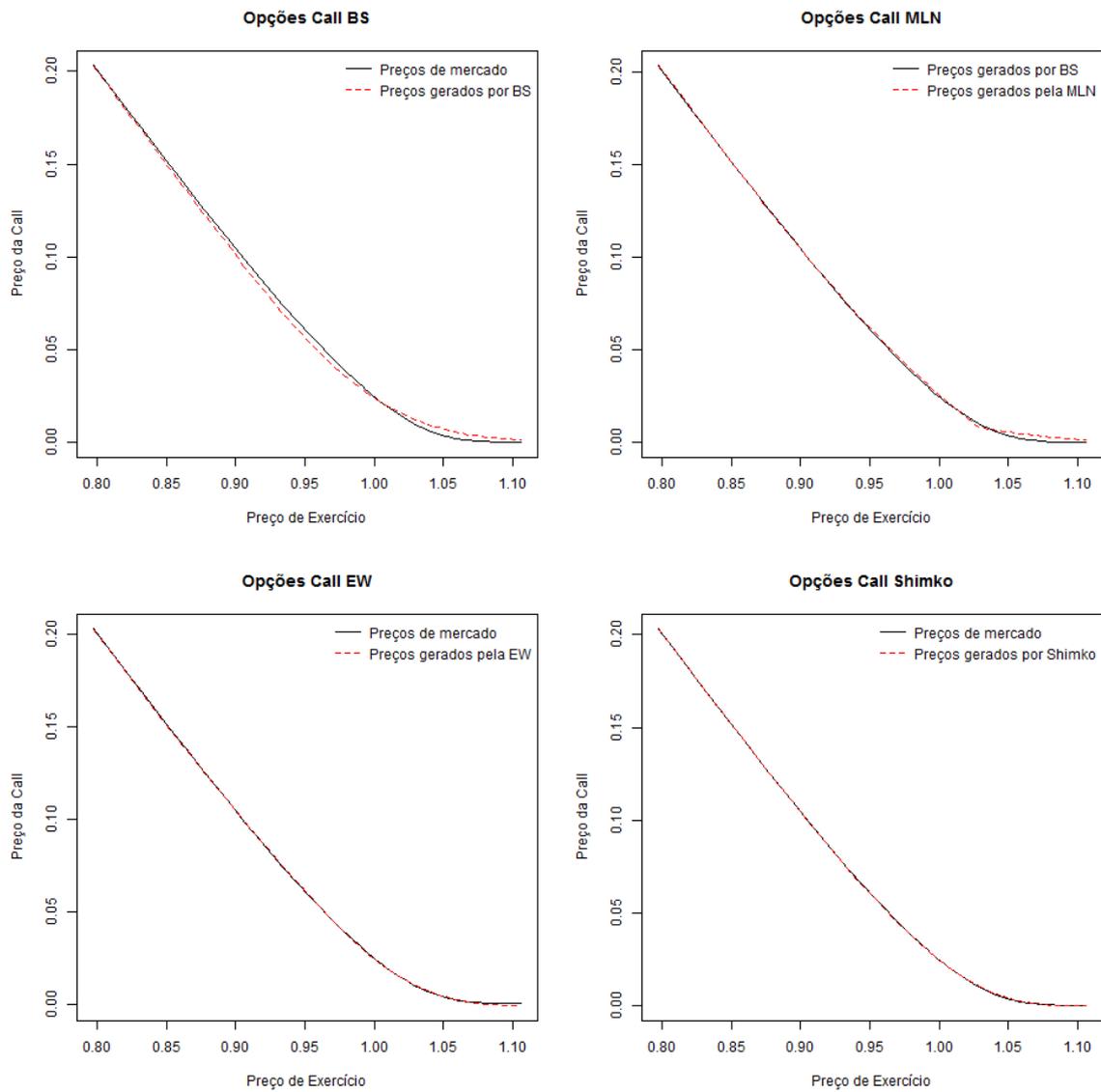


Fig. 3.7 Comparação entre preços de opções call de mercado e preços de opções call estimados pelos modelos de BS, MLN, EW e Shimko para a segunda maturidade (46 dias).

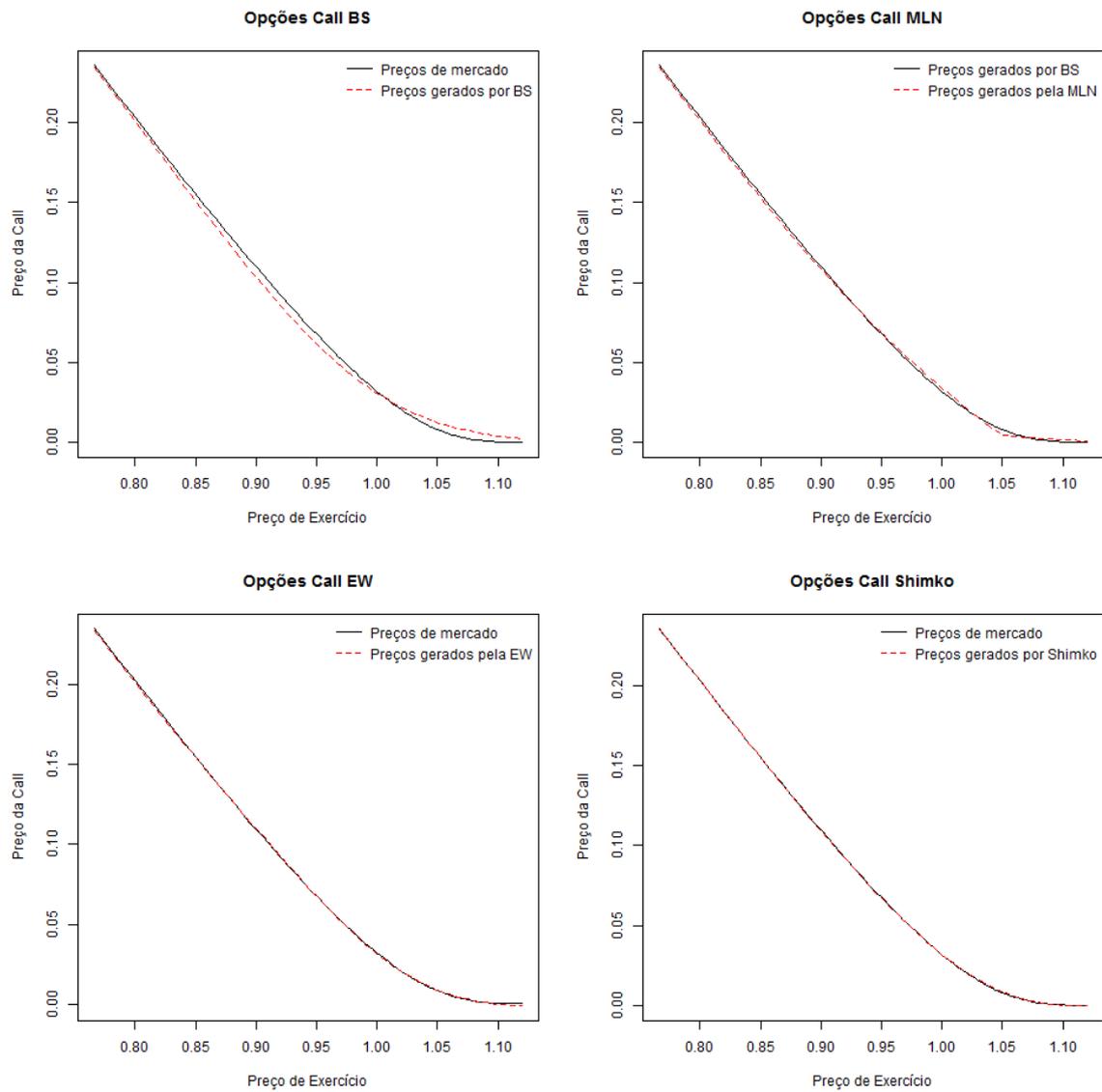


Fig. 3.8 Comparação entre preços de opções *call* de mercado e preços de opções *call* estimados pelos modelos de BS, MLN, EW e Shimko para a terceira maturidade (74 dias).

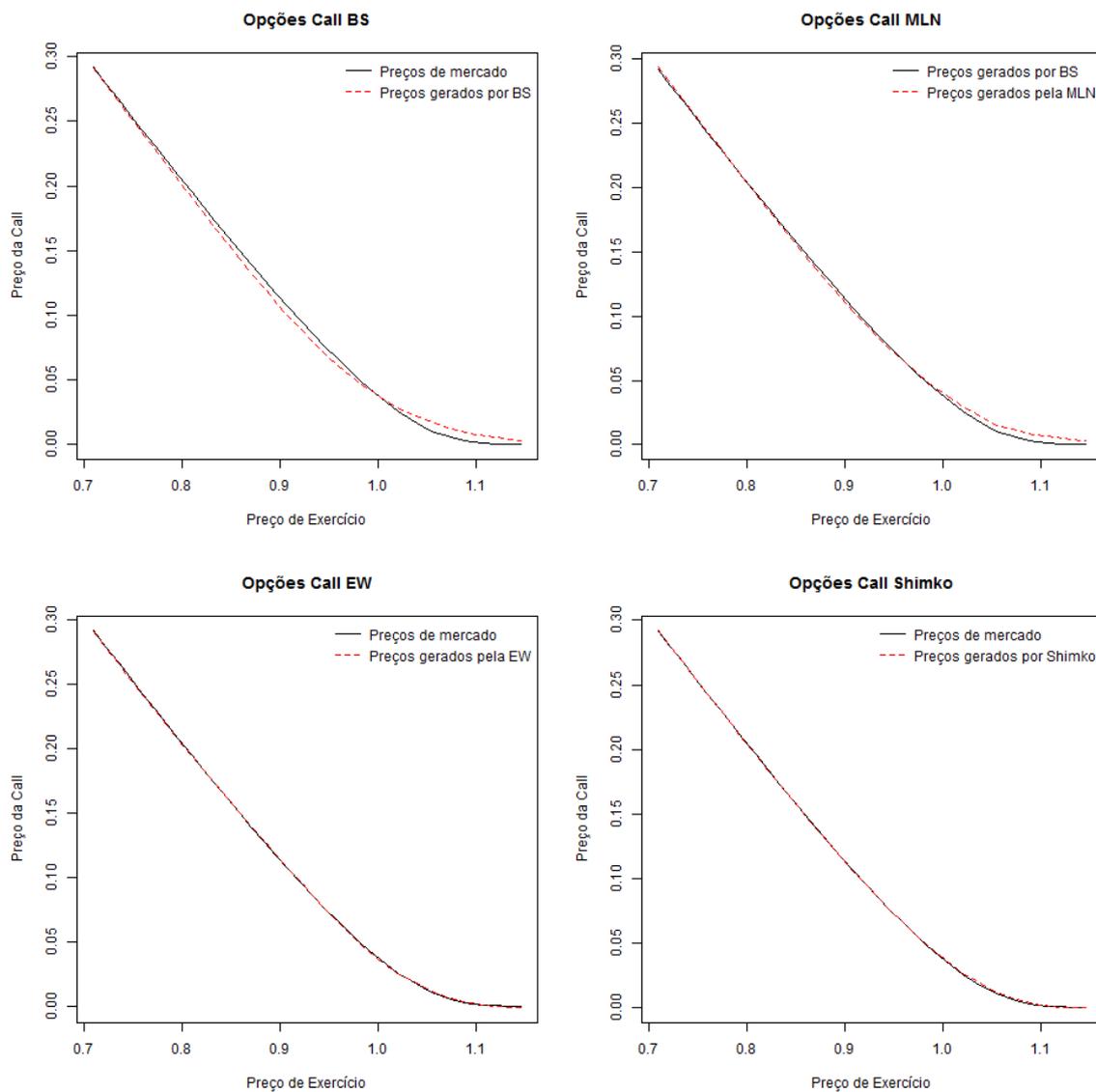


Fig. 3.9 Comparação entre preços de opções *call* de mercado e preços de opções *call* estimados pelos modelos de BS, MLN, EW e Shimko para a quarta maturidade (102 dias).

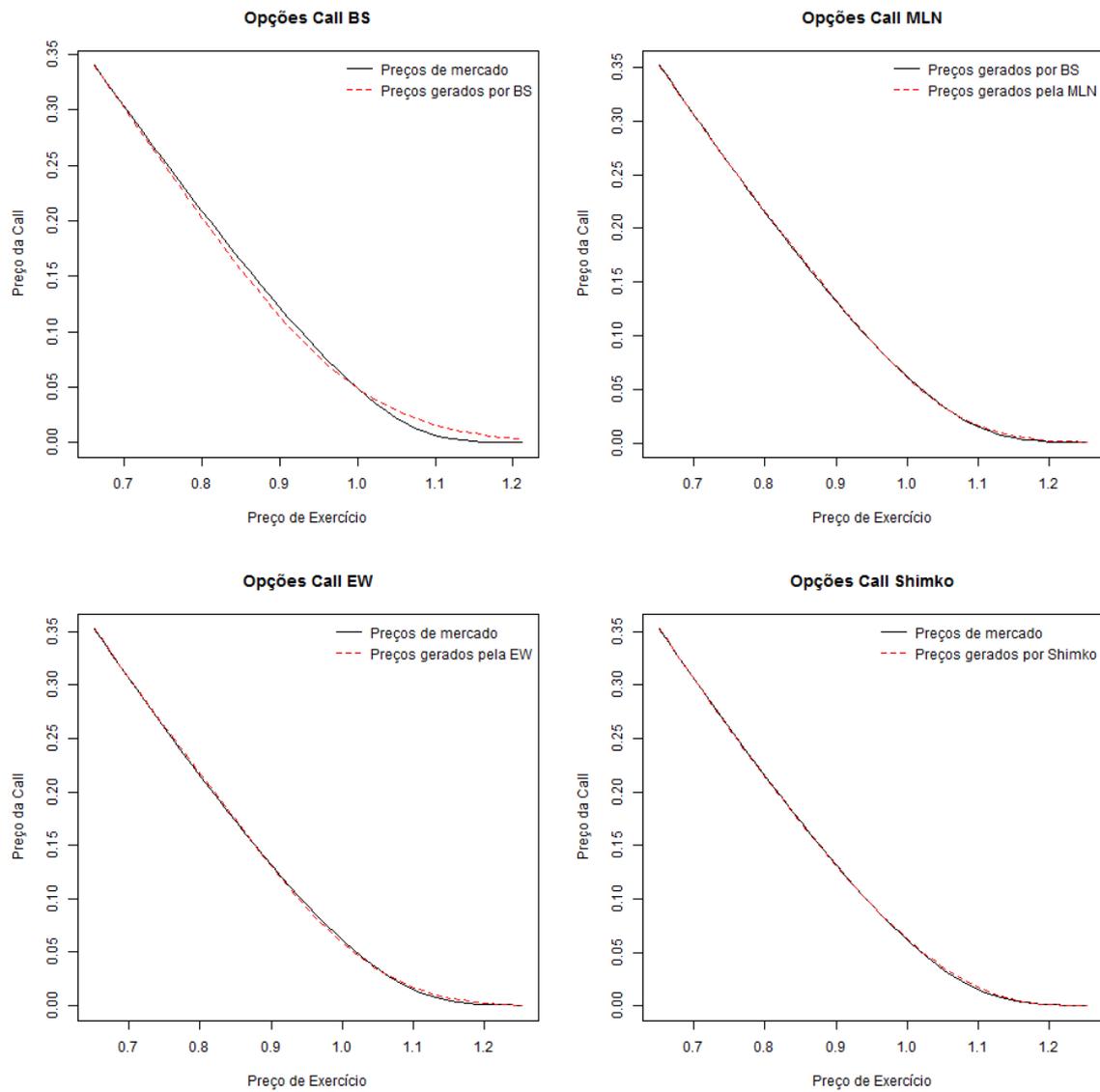


Fig. 3.10 Comparação entre preços de opções *call* de mercado e preços de opções *call* estimados pelos modelos de BS, MLN, EW e Shimko para a quinta maturidade (165 dias).

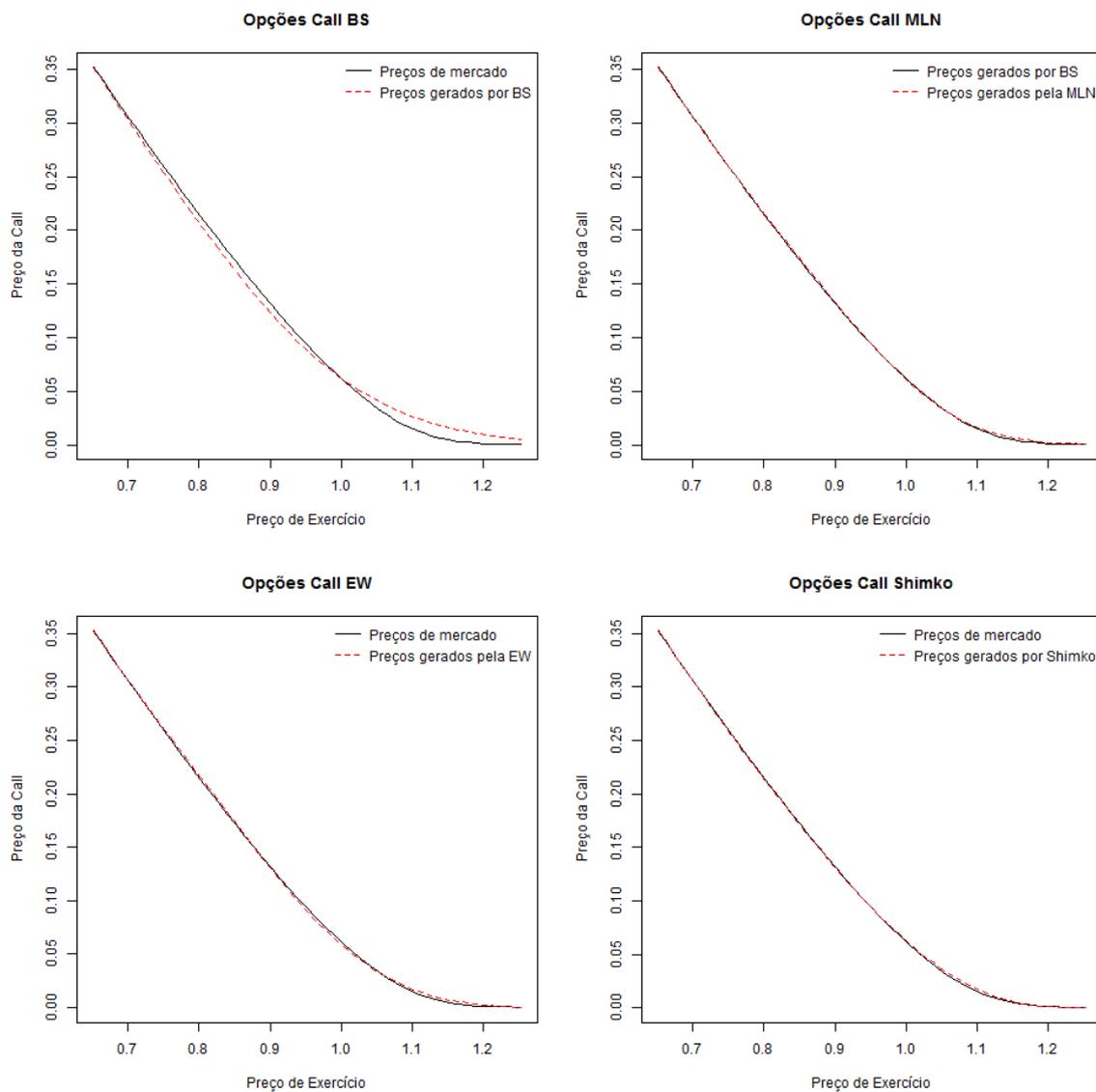


Fig. 3.11 Comparação entre preços de opções *call* de mercado e preços de opções *call* estimados pelos modelos de BS, MLN, EW e Shimko para a sexta maturidade (256 dias).

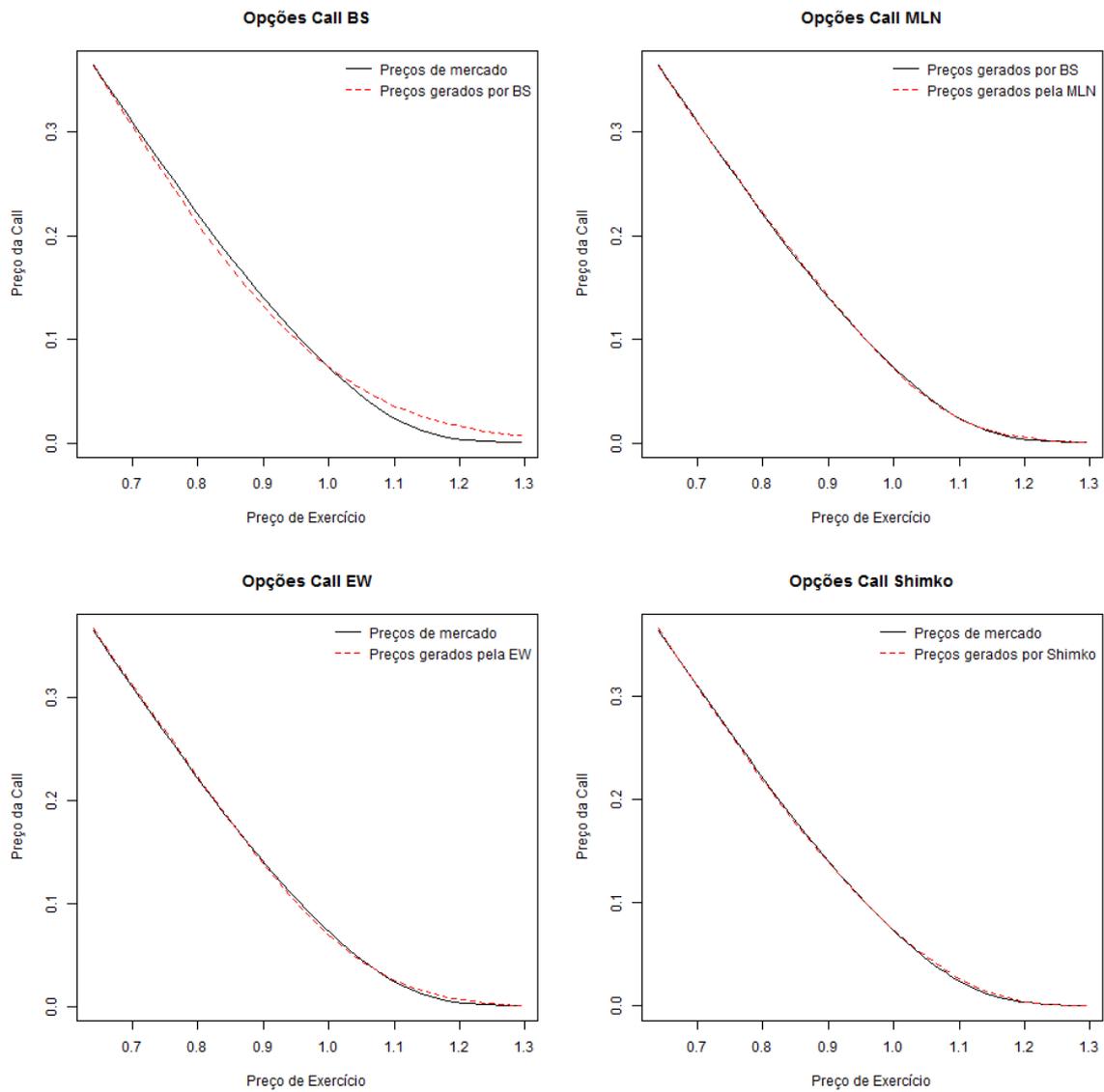


Fig. 3.12 Comparação entre preços de opções *call* de mercado e preços de opções *call* estimados pelos modelos de BS, MLN, EW e Shimko para a sétima maturidade (347 dias).

Tabela 3.7 RMSE de opções *call* sobre o índice S&P 500.

Maturidade (dias)	BS	MLN	EW	Shimko
11	1.102×10^{-3}	1.372×10^{-4}	1.615×10^{-4}	2.362×10^{-4}
46	2.856×10^{-3}	1.353×10^{-3}	4.369×10^{-4}	2.792×10^{-4}
74	4.153×10^{-3}	1.637×10^{-3}	6.318×10^{-4}	3.053×10^{-4}
102	4.932×10^{-3}	3.026×10^{-3}	8.114×10^{-4}	4.876×10^{-4}
165	6.024×10^{-3}	8.721×10^{-4}	1.037×10^{-3}	1.161×10^{-3}
256	7.586×10^{-3}	1.089×10^{-3}	1.778×10^{-3}	1.238×10^{-3}
347	8.608×10^{-3}	1.194×10^{-3}	2.482×10^{-3}	1.539×10^{-3}

Tabela 3.8 RMSE de opções *put* sobre o índice S&P 500.

Maturidade (dias)	BS	MLN	EW	Shimko
11	1.076×10^{-3}	1.582×10^{-4}	1.888×10^{-4}	2.051×10^{-4}
46	2.939×10^{-3}	1.117×10^{-3}	6.049×10^{-4}	4.137×10^{-4}
74	4.282×10^{-3}	1.683×10^{-3}	8.505×10^{-4}	4.957×10^{-4}
102	5.246×10^{-3}	3.783×10^{-3}	1.274×10^{-3}	8.189×10^{-4}
165	6.752×10^{-3}	9.998×10^{-4}	2.373×10^{-3}	2.249×10^{-3}
256	8.861×10^{-3}	1.269×10^{-3}	4.130×10^{-3}	4.073×10^{-3}
347	1.033×10^{-2}	1.460×10^{-3}	6.033×10^{-3}	6.101×10^{-3}

Observa-se que as opções estimadas pelo modelo de Black-Scholes são as que apresentam maior enviesamento e que este se agrava com o aumento da maturidade. Este fenómeno verifica-se em todos os métodos e é compatível com a perspetiva de que a incerteza aumenta com o tempo até à maturidade. Verifica-se também nas Tabelas 3.7 e 3.8 que as melhores estimações são as geradas pelos métodos da mistura de duas distribuições lognormais e pelo método de Shimko. Estes resultados são sintetizados na Tabela 3.9.

Tabela 3.9 Síntese de resultados do RMSE para as opções *call* e *put* sobre o índice S&P 500, referentes as diferentes maturidades.

Maturidade (dias)	Método onde se observou	
	Menor Enviesamento	Maior Enviesamento
11	MLN	BS
46	Shimko	BS
74	Shimko	BS
102	Shimko	BS
165	MLN	BS
256	MLN	BS
347	MLN	BS

Estamos então, em condições de gerar as FDPN relativas aos diferentes métodos. No entanto, antes disso vamos apresentar os parâmetros obtidos através estimação dos modelos, que serão utilizados para a estimação das FDPN. A nomenclatura utilizada é a mesma que foi considerada anteriormente.

Tabela 3.10 Parâmetros extraídos na estimação das FDPN segundo o modelo de Black-Scholes, utilizando opções sobre o índice S&P 500.

Maturidade (dias)	BS	
	μ	σ
11	1.801×10^{-4}	0.1951
46	4.125×10^{-4}	0.1686
74	4.586×10^{-4}	0.1704
102	-1.757×10^{-4}	0.1771
165	-1.749×10^{-3}	0.1792
256	-3.945×10^{-3}	0.1792
347	-6.898×10^{-3}	0.1808

Tabela 3.11 Parâmetros extraídos na estimação das FDPN segundo o método da mistura de duas distribuições lognormais, utilizando opções sobre o índice S&P 500.

Maturidade (dias)	MLN				
	θ	α_1	α_2	β_1	β_2
11	0.3275	-3.148×10^{-2}	1.451×10^{-2}	3.980×10^{-2}	2.007×10^{-2}
46	0.4302	2.737×10^{-2}	-2.270×10^{-2}	1.265×10^{-9}	8.385×10^{-2}
74	0.4747	4.836×10^{-2}	-5.055×10^{-2}	2.246×10^{-9}	9.945×10^{-2}
102	0.1808	5.383×10^{-2}	-1.344×10^{-2}	1.399×10^{-9}	0.1039
165	0.2425	-0.1901	4.900×10^{-2}	0.1694	6.243×10^{-2}
256	0.2626	-0.2361	6.462×10^{-2}	0.2027	7.591×10^{-2}
347	0.2715	-0.2803	7.705×10^{-2}	0.2320	8.804×10^{-2}

Tabela 3.12 Parâmetros extraídos na estimação das FDPN segundo o método da expansão de Edgeworth, utilizando opções sobre o índice S&P 500.

Maturidade (dias)	EW		
	$\kappa_2(Q)$	$\kappa_3(Q)$	$\kappa_4(Q)$
11	0.1942	-1.1148	0.2779
46	0.1755	-1.4402	3.606×10^{-2}
74	0.1794	-1.5510	1.727×10^{-4}
102	0.1803	-1.4874	7.626×10^{-2}
165	0.1794	-1.2752	2.858×10^{-7}
256	0.1768	-1.0300	8.728×10^{-7}
347	0.1746	-0.8085	9.543×10^{-8}

Tabela 3.13 Parâmetros extraídos na estimação das FDPN segundo o método de Shimko, utilizando opções sobre o índice S&P 500.

Maturidade (dias)	Shimko		
	A_0	A_1	A_2
11	2.4123	-3.5195	1.3010
46	0.4836	9.328×10^{-2}	-0.4035
74	0.5105	-2.325×10^{-2}	-0.3105
102	0.4423	3.173×10^{-2}	-0.2944
165	0.2317	0.3005	-0.3486
256	0.2317	0.2087	-0.2594
347	0.2250	0.1541	-0.1986

Relativamente aos parâmetros estimados podem ser retiradas algumas conclusões. O aumento da maturidade provoca uma tendência de redução do retorno esperado do ativo, estimado pelo modelo de Black-Scholes, μ . Verifica-se também que a volatilidade do preço do ativo estimada pelo modelo de Black-Scholes, σ , é aproximadamente constante e bastante próxima da variância estimada pelo método da expansão de Edgeworth, representada por $\kappa_2(Q)$. Como visto anteriormente, $\kappa_3(Q)$ representa uma medida de enviesamento e através da análise dos valores deste parâmetro verifica-se que este é sempre negativo, o que pode significar que os investidores estão receosos, esperando uma probabilidade de um retorno negativo superior a uma probabilidade de retorno positivo.

Seguidamente apresentamos os gráficos das FDPN geradas pelos quatro métodos, para as diferentes maturidades, bem como a média, a variância, o coeficiente de assimetria e o excesso de curtose de cada FDPN apresentadas nas Tabelas 3.14, 3.15, 3.16 e 3.17.

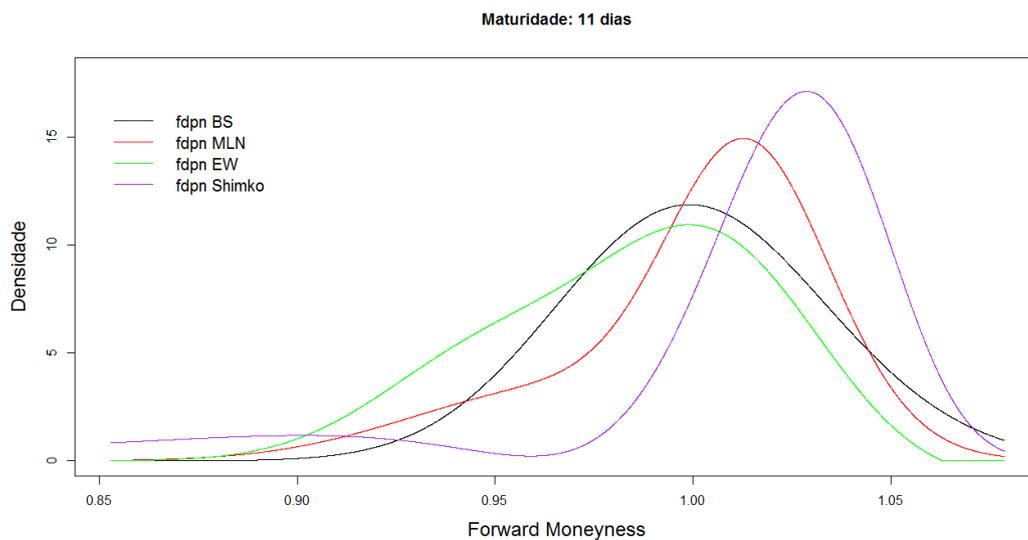


Fig. 3.13 Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à primeira maturidade (11 dias).

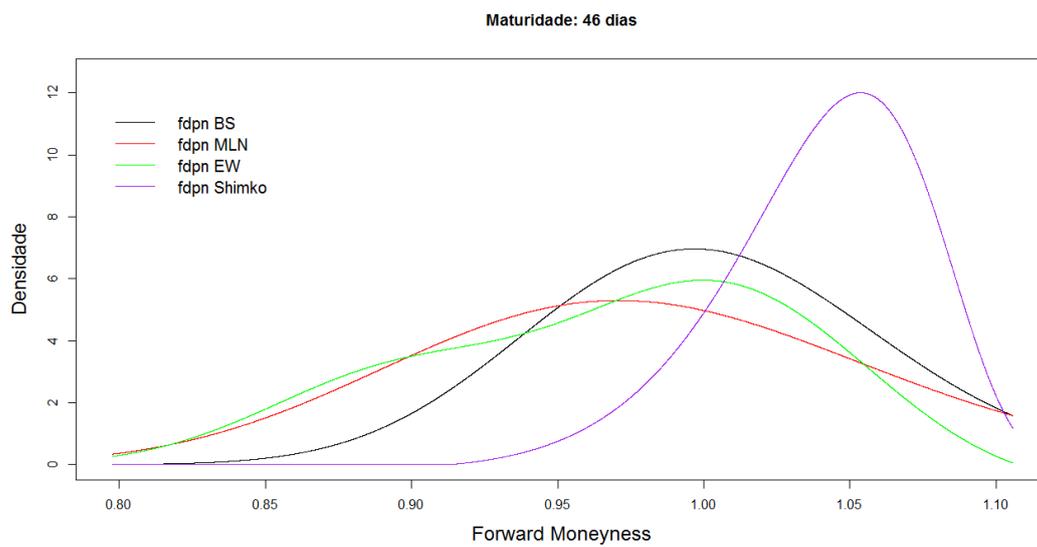


Fig. 3.14 Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à segunda maturidade (46 dias).

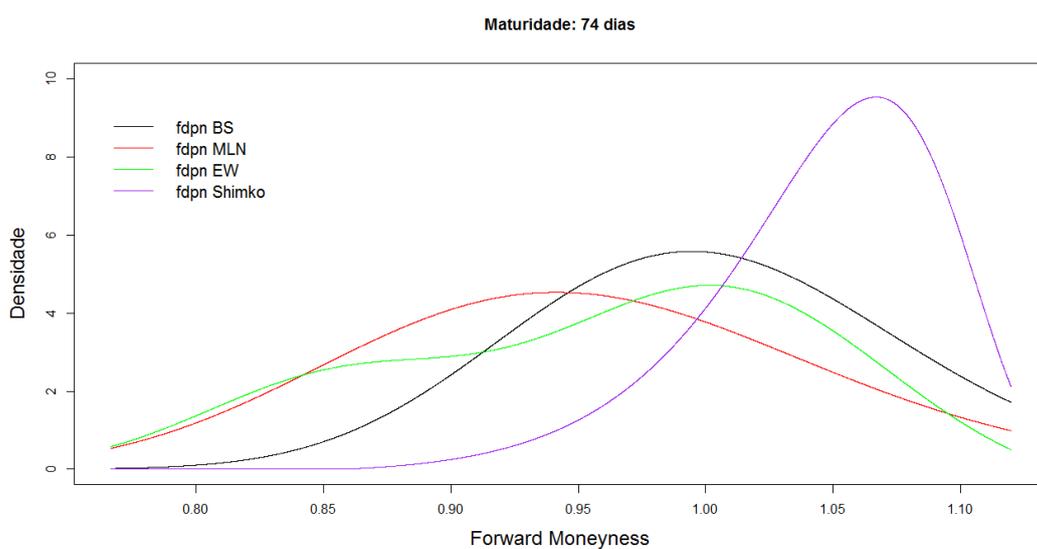


Fig. 3.15 Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à terceira maturidade (74 dias).

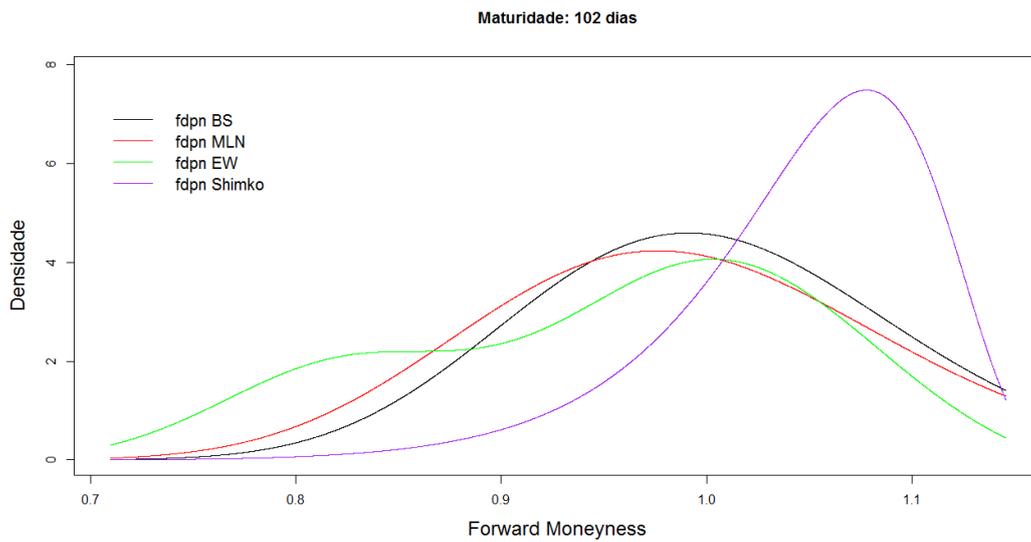


Fig. 3.16 Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à quarta maturidade (102 dias).

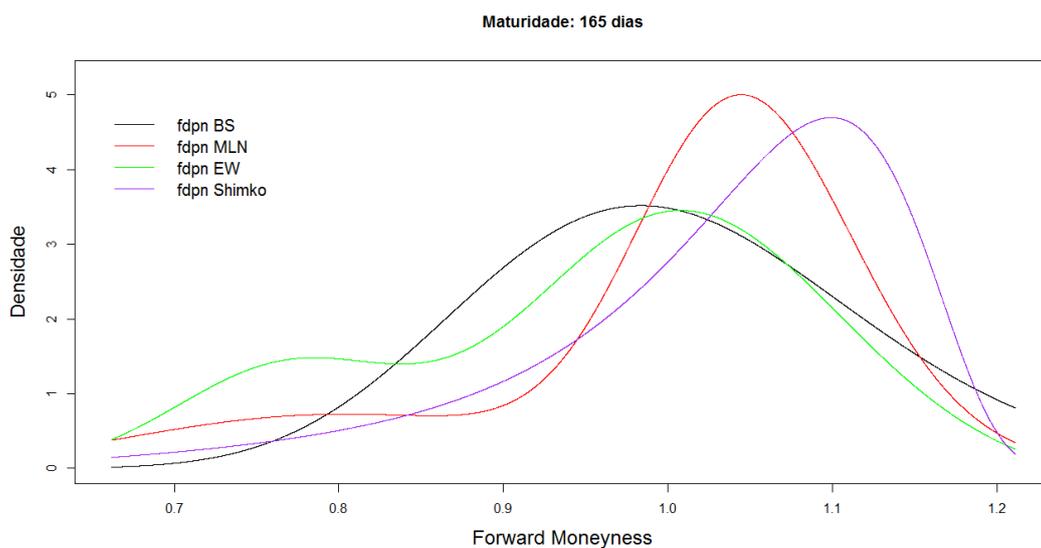


Fig. 3.17 Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à quinta maturidade (165 dias).

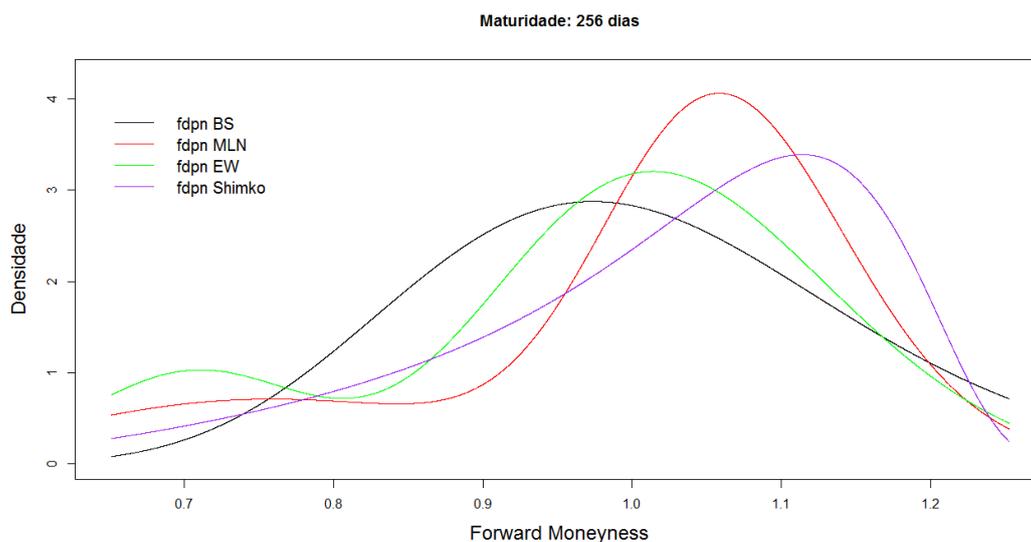


Fig. 3.18 Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à sexta maturidade (256 dias).

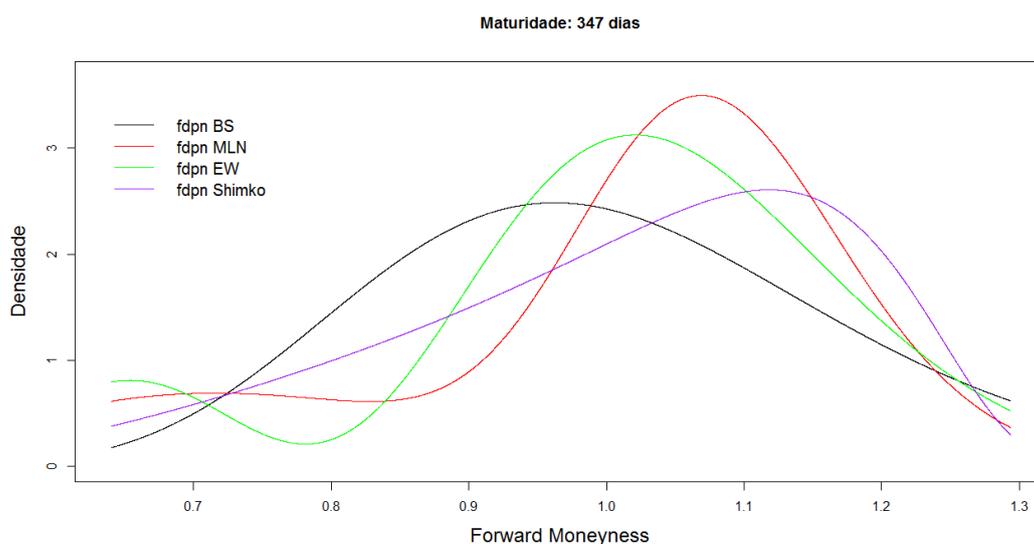


Fig. 3.19 Estimação das funções densidade de probabilidade neutras face ao risco, para dados de mercado sobre o índice S&P 500, relativos à sétima maturidade (347 dias).

Tabela 3.14 Média, variância, coeficiente de assimetria e excesso curtose da FDPN gerada pelo modelo de Black-Scholes.

Maturidade (dias)	BS			
	Média	Variância	Assimetria	Excesso de Curtose
11	0.9995	1.066×10^{-3}	-7.764×10^{-2}	-0.2770
46	0.9955	2.875×10^{-3}	-0.1959	-0.4049
74	0.9914	4.333×10^{-3}	-0.2362	-0.4528
102	0.9889	6.362×10^{-3}	-0.2135	-0.4743
165	0.9898	1.105×10^{-2}	-9.132×10^{-2}	-0.5284
256	0.9854	1.611×10^{-2}	-3.945×10^{-2}	-0.6285
347	0.9828	2.117×10^{-2}	2.894×10^{-2}	-0.6989

Tabela 3.15 Média, variância, coeficiente de assimetria e excesso de curtose da FDPN gerada pela mistura de duas distribuições lognormais.

Maturidade (dias)	MLN			
	Média	Variância	Assimetria	Excesso de Curtose
11	0.9999	1.196×10^{-3}	-0.8834	0.7626
46	0.9698	4.628×10^{-3}	-0.1288	-0.6565
74	0.9476	6.365×10^{-3}	1.473×10^{-2}	-0.6979
102	0.9752	7.428×10^{-3}	-0.1835	-0.5309
165	1.0067	1.280×10^{-2}	-1.0483	0.7430
256	1.0153	1.709×10^{-2}	-0.8920	0.3226
347	1.0249	2.083×10^{-2}	-0.7849	0.1317

Tabela 3.16 Média, variância, coeficiente de assimetria e excesso de curtose da FDPN gerada pela expansão de Edgeworth.

Maturidade (dias)	EW			
	Média	Variância	Assimetria	Excesso de Curtose
11	0.9833	1.277×10^{-3}	-0.3696	-0.4112
46	0.9660	4.229×10^{-3}	-0.3004	-0.6852
74	0.9568	6.851×10^{-3}	-0.2760	-0.8258
102	0.9509	1.004×10^{-2}	-0.3271	-0.7799
165	0.9570	1.572×10^{-2}	-0.3846	-0.6622
256	0.9821	1.940×10^{-2}	-0.5052	-0.3518
347	1.0163	1.973×10^{-2}	-0.5868	0.2356

Tabela 3.17 Média, variância, coeficiente de assimetria e excesso de curtose da FDPN gerada pelo método de Shimko.

Maturidade (dias)	Shimko			
	Média	Variância	Assimetria	Excesso de Curtose
11	1.0140	1.827×10^{-3}	-1.9369	3.6583
46	1.0401	1.151×10^{-3}	-0.5767	1.752×10^{-2}
74	1.0444	2.005×10^{-2}	-0.7774	0.3856
102	1.0451	3.784×10^{-2}	-1.0295	1.3172
165	1.0333	1.126×10^{-2}	-1.0071	0.7276
256	1.0250	1.747×10^{-2}	-0.6945	-0.1981
347	1.0186	2.321×10^{-2}	-0.4595	-0.6135

Nas Figuras 3.13, 3.14, 3.15, 3.16, 3.17, 3.18 e 3.19 observa-se que, quando consideramos dados de mercado, as FDPN apresentam uma grande variedade de formas. Este fenómeno deve-se ao facto dos diferentes métodos atribuírem diferentes probabilidades às variações do preço do ativo subjacente.

Da análise das Tabelas 3.14, 3.15, 3.16 e 3.17 ressaltam tendências que podemos generalizar. Verifica-se um aumento da variância das FDPN com o aumento da maturidade que se justifica pelo facto de à medida que aumenta o período temporal, aumentar também a dispersão dos dados. Este resultado não é verificado no método de Shimko. Observa-se também que quase todas as FDPN têm assimetria negativa. Isto revela um certo receio por parte dos investidores que esperam uma probabilidade de retorno negativo superior à probabilidade de retorno positivo.

Capítulo 4

Conclusão

Os preços das opções são ferramentas essenciais para traçar o perfil das expectativas dos agentes económicos relativamente à evolução do preço do ativo subjacente. É na análise dessas expectativas que as funções densidade de probabilidade neutras face ao risco para o preço de um ativo financeiro têm um papel importante. Isto deve-se ao facto de estas traduzirem as probabilidades que o mercado atribui às variações futuras do preço do ativo subjacente.

Black e Scholes [4] apresentaram um modelo em que é imposto que o preço do ativo subjacente siga uma distribuição lognormal. No entanto, as restrições necessárias para que este modelo seja viável não são verificadas na prática. Isto obrigou a que os investigadores procurassem formular novos métodos que permitissem encontrar estimações mais consistentes da função densidade de probabilidade neutra face ao risco.

Nesta dissertação foram apresentados alguns dos métodos elaborados por diferentes autores e foi feita uma análise empírica com alguns desses métodos. Deste modo, foram escolhidos para estudo o modelo de Black-Scholes, pela importância que apresenta, um modelo paramétrico (mistura de duas distribuições lognormais), um modelo semi-paramétrico (expansão de Edgeworth) e um modelo não paramétrico (método da volatilidade implícita proposto por Shimko). Para este feito utilizámos, como dados de entrada da estimação da função densidade de probabilidade neutra face ao risco, preços de opções gerados pelo modelo de Black-Scholes e preços de opções de mercado sobre o índice S&P 500, para diferentes maturidades. De modo a evitar problemas de otimização relativos à estimação dos parâmetros necessários para produzir as funções densidade de probabilidade neutras face ao risco, tanto na análise teórica como na de mercado, foi aplicado o escalonamento dos dados conforme o proposto por Fengler e Hin [13].

Iniciámos a nossa análise com o estudo de funções densidade de probabilidade de distribuições lognormais, como proposto pelo modelo de Black-Scholes, observando o seu comportamento a alterações do tempo até à maturidade. Através desta análise concluímos que, à medida que aumenta o tempo até à maturidade, aumenta também a dispersão dos dados e a sua concentração à esquerda. O aumento da dispersão dos dados, que pode ser observado através do aumento da variância e resulta do aumento da incerteza provocado pelo aumento do horizonte temporal. O aumento da concentração dos dados à esquerda, traduzido pela diminuição do valor da média e pelo aumento da assimetria positiva, pode revelar que, teoricamente, com o aumento da maturidade, os investidores irão atribuir

uma probabilidade superior à diminuição do preço do ativo subjacente na maturidade do que a um aumento desse preço.

Depois de observados estes factos teóricos utilizámos dois conjuntos de dados teóricos discrepantes. Para estes dois conjuntos foram gerados preços de opções, conforme o modelo de Black-Scholes. Para cada conjunto, foi estabelecida as comparações entre a função densidade de probabilidade de uma distribuição lognormal como proposto por Black-Scholes, e as funções densidades de probabilidade neutras face ao risco geradas pela mistura de duas distribuições lognormais, pela expansão de Edgeworth e pelo método de Shimko. Através deste estudo foi possível concluir que, quando considerados dados teóricos, os métodos produzem funções densidade de probabilidade neutras face ao risco consistentes.

A análise de dados de mercado S&P 500 requer cuidado adicional no tratamento dos dados, nomeadamente de modo a que seja garantida a ausência de oportunidades de arbitragem [7]. Não obstante isto, esta redução da base de dados pode ser desvantajosa por ignorar informações, possivelmente importantes, contidas nos preços das opções. Os resultados obtidos pelo cálculo do *root mean square error* (RMSE) indicam-nos que os métodos em estudo geram preços de opções próximos dos preços observados no mercado. Verificamos também que as funções densidade de probabilidade neutras face ao risco obtidas pelos métodos apresentam diversidade de formas, atribuindo diferentes probabilidades aos mesmos eventos. No entanto, é possível observar que estas, maioritariamente, apresentam assimetria negativa, ao contrário do verificado no caso teórico, o que pode significar que os investidores esperam retornos negativos com probabilidade superior a retornos positivos. Verifica-se também que, com o aumento do prazo de exposição à incerteza, a variância aumenta, o que indica que a probabilidade atribuída a valores extremos torna-se cada vez maior.

A fiabilidade dos modelos está condicionada por um número elevado de factores endógenos e exógenos que poderão alterar possíveis interpretações. De entre estes fatores podemos evidenciar a quantidade e a qualidade da informação disponível, o grau de liquidez, o nível de perfeição dos mercados, aspectos institucionais e técnicos do mercado e desequilíbrios entre os agentes económicos que nele atuam.

Esta dissertação teve como objetivo estudar a aplicabilidade de alguns métodos a opções sobre um dos índices mais transacionados no mercado, o índice S&P 500. Deste modo, consideramos que seria interessante, num estudo futuro, analisar a aplicabilidade destes métodos a outros mercados de índices, como por exemplo o NASDAQ, bem como utilizar opções sobre outros ativos subjacentes, como mercadorias, taxas de juro ou divisas.

Bibliografia

- [1] Abadir, K. M. and Rockinger, M. (2003). Density functionals, with an option-pricing application. *Econometric Theory*, 19:778–811.
- [2] Aït-Sahalia, Y. and Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *The Journal of Finance*, 53(2):499–547.
- [3] Bahra, B. (1997). Implied risk-neutral probability density functions from option prices: theory and application. *Bank of England Working Paper*.
- [4] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654.
- [5] Breeden, D. T. and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in options prices. *Journal of Business*, 51(4):621–651.
- [6] Bu, R. and Hadri, K. (2007). Estimating option implied risk-neutral densities using spline and hypergeometric functions. *Econometrics Journal*, 10:216–244.
- [7] Cornuejols, G. and Tütüncü, R. (2007). *Optimization Methods in Finance*. Cambridge University Press.
- [8] Corrado, C. J. and Su, T. (1996). S&p500 index option tests of jarro and rudd's approximate option valuation formula. *The Journal of Futures Markets*, 16(6):611–629.
- [9] Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407.
- [10] Cox, J. C. and Ross, S. A. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3:145–166.
- [11] Cvitanović, J. and Zapatero, F. (2004). *Introduction to the Economics and Mathematics of Financial Markets*. Massachusetts Institute of Technology Press.
- [12] Duan, J.-C., Härdle, W. K., and Gentle, J. E. (2012). *Handbook of Computational Finance*. Springer-Verlag Berlin Heidelberg.
- [13] Fengler, M. R. and Hin, L.-Y. (2015). Semi-nonparametric estimation of the call-option price surface under strike and time-to-expiry no-arbitrage constraints. *Journal of Econometrics*, 184(2):242–261.
- [14] Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343.
- [15] Hull, J. C. (2002). *Options, Futures and Other Derivatives*. Prentice Hall.
- [16] Jarrow, R. and Rudd, A. (1982). Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics*, 10(3):347–369.

-
- [17] Jondeau, E., Poon, S.-H., and Rockinger, M. (2007). *Financial Modeling Under Non-Gaussian Distributions*. Springer-Verlag London.
- [18] Melick, W. R. and Thomas, C. P. (1997). Recovering an asset's implied pdf from option prices: An application to crude oil during the gulf crisis. *The Journal of Financial and Quantitative Analysis*, 32(1):91–115.
- [19] Monteiro, A. M., Tütüncü, R. H., and Vicente, L. N. (2008). Recovering risk-neutral probability density functions from options prices using cubic splines and ensuring nonnegativity. *European Journal of Operational Research*, 187(2):525–542.
- [20] Shimko, D. (1993). Bounds of probability. *Risk*, 6(4):33–37.
- [21] Syrdal, S. A. (2002). *A Study of Implied Risk-neutral Density Functions in the Norwegian Option Market*. Norges Bank.
- [22] Taylor, S. J. (2007). *Asset Price Dynamics, Volatility and Prediction*. Princeton University Press.