

2. Simulação de Monte Carlo

A Simulação de Monte Carlo (SMC) é um método numérico que utiliza um gerador de números aleatórios para simular possíveis valores da variável de interesse e assim obter a estimativa para o valor verdadeiro da mesma.

A introdução do método de Monte Carlo na precificação de derivativos acontece em Boyle(1977). Broadie & Detemple (2004) dividem a aplicação do método em três etapas principais:

- i. Geração de N caminhos aleatórios para cada uma das variáveis de estado;
- ii. Cômputo do *payoff* descontado correspondente a cada um dos N caminhos;
- iii. Cálculo da média dos *payoffs* descontados e do erro padrão da estimativa resultante.

Pela Lei dos Grandes Números, à medida que o número de simulações aumenta, a estimativa converge para o valor verdadeiro da variável. Enquanto que o Teorema Central do Limite fornece a provável magnitude do erro da estimativa após um número finito de simulações.

Apesar de sua simplicidade, o Método de Monte Carlo tem a desvantagem de apresentar uma convergência lenta quando se utiliza de números pseudo – aleatórios gerados computacionalmente através de um algoritmo simples. Neste caso, sua taxa de convergência é $O\left(\frac{1}{\sqrt{N}}\right)$, logo, para se obter uma boa precisão seria necessário aumentar, e muito, o número de caminhos simulados (N) o que, por conseguinte, aumentaria o tempo computacional exigido tornando, em alguns casos, sua aplicação inviável.

A fim de aumentar a eficiência do método recorre-se a utilização das técnicas de redução de variância, ou ainda, pode-se recorrer à utilização da seqüência de números quase-aleatórios em substituição à seqüência de números pseudo-aleatórios, tradicionalmente utilizada.

2.1. Técnicas de Redução de Variância

As técnicas de redução de variância são procedimentos que visam através da diminuição do desvio padrão da amostra aumentar a taxa de convergência do método. Dessa forma, tem-se a introdução de diversas técnicas que diminuem a variância amostral. A seguir faz-se a apresentação de algumas delas.

2.1.1. Variáveis Antitéticas

O método das variáveis antitéticas tenta reduzir a variância através da introdução de pares de números aleatórios negativamente correlacionados pertencentes à mesma distribuição de probabilidade. Assim, se amostrarmos de uma normal padrão uma seqüência de variáveis aleatórias independentes e identicamente distribuídas Z_1, Z_2, \dots basta que introduzamos uma outra seqüência de variáveis independentes e identicamente distribuídas $-Z_1, -Z_2, \dots$. Dessa forma, ao gerarmos uma seqüência estaremos gerando duas, diminuindo pela metade o número de paths simulados.

2.1.2. Variável de Controle

Neste procedimento utiliza-se a informação sobre os erros nas estimativas de quantidades conhecidas para reduzir o erro na estimativa de uma quantidade desconhecida. Dado um mesmo conjunto de simulações determina-se a estimativa de duas variáveis similares, sendo que para uma delas seu valor verdadeiro é conhecido, fato que a elege como a variável de controle. Dessa

maneira, podemos escrever este procedimento como $f_A = \hat{f}_A - (\hat{f}_G - f_G)$, onde f_A é o valor verdadeiro de A, \hat{f}_A é o valor estimado de A, \hat{f}_G é o valor estimado da variável de controle G, e f_G é o valor verdadeiro da variável de controle.

Em Kemna & Vorst(1990), a técnica de variável de controle foi utilizada para determinar o valor de uma opção asiática européia baseada na média aritmética, tendo como variável de controle uma opção congênere baseada na média geométrica cuja solução analítica é conhecida.

2.1.3. *Stratified Sampling* (Amostragem Estratificada)

Esta técnica consiste em dividir a distribuição de probabilidades em n intervalos iguais (no caso, da distribuição uniforme) ou de mesma probabilidade (como no caso da distribuição normal) para, posteriormente, realizar um conjunto de simulações dentro de cada um dos intervalos de modo que $n\%$ da amostra esteja dentro de cada um dos estratos. Este procedimento garante que a distribuição de probabilidade da amostra fique mais próxima da distribuição desejada.

2.1.4. *Latin Hypercube* (Hiper cubo Latino)

Em princípio, podemos utilizar a Amostragem Estratificada em problemas de múltiplas dimensões, mas, na prática, sua utilização não é viável. Já que se estivermos tratando de um problema com K estratos e d dimensões teríamos que ter uma amostra com o tamanho de no mínimo K^d o que torna o procedimento computacionalmente custoso.

Desta forma, o procedimento do Hiper cubo Latino vem resolver este problema específico. Ou seja, em cada uma das dimensões existem K estratos sendo que neste procedimento garante-se que haverá um ponto em cada estrato conjunto. De forma que havendo a projeção desse ponto em determinado eixo de uma dimensão d qualquer, não haja repetição. Isto garante que, ao invés de necessitarmos de uma amostra de tamanho K^d , precisaremos somente uma amostra de tamanho K . Sendo assim, o procedimento do *Latin Hypercube* é visto como uma extensão da Amostragem Estratificada para múltiplas dimensões.

2.1.5. *Importance Sampling* (Amostragem por Importância)

Neste procedimento tenta-se reduzir a variância através da mudança de medida de probabilidade de modo a que dê mais peso para resultados considerados mais importantes, e ou relevantes para o fim a que se propõe, de forma a aumentar a eficiência da amostragem realizada.

2.2. Seqüências de Baixa Discrepância ou Quase-Monte Carlo (QMC)

Joy, Boyle e Tan(1996) mostram que a eficiência do Método de Monte Carlo pode ser aumentada de maneira significativa se os números quase-aleatórios são utilizados neste método em substituição aos números pseudo-aleatórios.

As seqüências de baixa discrepância (ou quase-aleatórias) têm como característica a uniformidade na distribuição dos números, evitando, portanto, problemas comuns à geração de números pseudo-aleatórios, como a ocorrência de lacunas e aglomerações.

Em geral, o método de Quase-Monte Carlo apresenta uma taxa de convergência de $O(1/N)$ estando, teoricamente limitada, à taxa de convergência

de $O\left(\frac{(\ln N)^d}{N}\right)$ onde d representa a dimensão do problema, tendo em vista que estas seqüências perdem sua uniformidade a medida que o número de dimensões aumenta.

Ressalta-se que a utilização de seqüências de baixa discrepância não impede a implementação de técnicas de redução de variância, já que esta combinação poderia promover, em alguns casos, um aumento na precisão do método de simulação (Brotherton-Ratcliffe, 1994). Neste sentido, Frota(2003) observa que as técnicas das variáveis antitéticas e de estratificação não apresentam ganho adicional quando associadas ao QMC.

Os números quase-aleatórios podem ser gerados pelas seqüências de baixa discrepância, tais como as de Halton, Faure e Sobol, descritas brevemente a seguir.

2.2.1. Seqüência de Halton

A seqüência de Halton determina uma base diferente para cada dimensão de um dado problema. A base é determinada da seguinte forma: dada a dimensão d então escolhemos o d -ésimo número primo da série de números primos. Ou seja, dimensão 1 corresponde à base 2; dimensão 2, base 3; dimensão 3, base 5; dimensão 4, base 7; e assim por diante.

Quanto maior a dimensão do problema (número de intervalos de tempo) maior é a base usada, o que aumenta o tamanho do ciclo (quantidade de números que cobrem o intervalo $[0,1)$), acarretando no aumento do tempo computacional expendido.

Com o aumento no número de dimensões do problema, há degradação da uniformidade da seqüência gerada. De acordo com Frota(2003), a perda na uniformidade acontece a partir da oitava dimensão.

2.2.2. Seqüência de Faure

A seqüência Faure usa somente uma base para todas as dimensões. A determinação da base é feita pela escolha do menor número primo, cujo valor é maior ou igual ao número de dimensões do problema. Depois da escolha da base e a determinação dos elementos da seqüência, novas seqüências são geradas para cada uma das outras dimensões a partir da permutação dos elementos da primeira seqüência.

Comparativamente à seqüência de Halton, a base gerada pela seqüência de Faure em altas dimensões é bem menor, o que torna o procedimento de geração de seqüências mais rápido na seqüência de Faure que na de Halton. Além disso, esta seqüência é mais resistente à degradação do que a de Halton, já que a perda de uniformidade acontece a partir da 25ª dimensão.

2.2.3. Seqüência de Sobol

A seqüência de Sobol é mais simples que a Faure, pois, utiliza a base 2 para todas as dimensões, sendo portanto mais rápida, mas o procedimento de reordenação dos elementos dos vetores é mais complexo que a permutação utilizada em Faure, pois utiliza-se de um conjunto de números direcionais. (Glasserman, 2004).

Comparativamente às seqüências de Halton e Faure, a de Sobol é melhor tanto no que diz respeito ao tempo computacional expendido, quanto ao fato de que a perda de uniformidade da seqüência acontece em uma dimensão $d \gg 25$.