

### 3

## Processos Estocásticos

Um processo estocástico pode ser definido como uma seqüência de variáveis aleatórias indexadas ao tempo e também a eventos. É uma variável que se desenvolve no tempo de maneira parcialmente aleatória e imprevisível. Sob outro ponto de vista, um processo estocástico é uma seqüência de funções mensuráveis, ou seja, uma variável aleatória  $X$  definida num espaço de probabilidade  $(\Omega, P)$  que toma valores num espaço de funções  $F$ .

Tais processos podem ser classificados em estacionários e não estacionários. Em particular, são amplamente utilizados os processos estocásticos estacionários de segunda ordem que são caracterizados por média e a variância constantes no tempo. Em contrapartida, os não estacionários possuem momentos que crescem sem limite, à medida que o tempo passa. Além dessas classificações, os processos estocásticos podem ter o parâmetro tempo contínuo ou discreto.

### Definição Matemática

Segundo Brandão (2000), seja  $W$  um conjunto, onde  $w \in W$  denota um estado da natureza. Seja também uma função  $f$  tal que

$$f : R \times W \rightarrow R \quad f(x, w), x \in R \text{ e } w \in W$$

Dado um  $w \in W$ ,  $f(o, w)$  torna-se função de  $x$ . Valores diferentes de  $w$  e valores diferentes de  $x$ . Assim têm-se duas funções. Quando  $x$  representa o tempo então  $f(x, w1)$  e  $f(x, w2)$  representam duas trajetórias diferentes que dependem de estados do mundo.  $W$  representa a aleatoriedade e a função  $f(x, w)$  é uma função aleatória ou um processo estocástico.

### 3.1

#### **Random Walk**

O processo estocástico conhecido como (*Random Walk*) é considerado um dos processos estocásticos mais simples em tempo discreto e estado discreto. Neste processo tem-se a variável aleatória  $x_t$  e seu valor inicial  $x_0$  é conhecida no instante inicial do processo. Nos instantes sucessivos,  $x_t$  assume saltos de tamanho um para cima ou para baixo, sempre com probabilidade  $\frac{1}{2}$ . Como os saltos são independentes entre si, podemos descrever a dinâmica de  $x_t$  como:

$$x_t = x_{t-1} + \varepsilon_t \quad (3.1)$$

onde  $\varepsilon_t$  é uma variável aleatória com distribuição de probabilidade

$$\text{prob}(\varepsilon_t = 1) = \text{prob}(\varepsilon_t = -1) = \frac{1}{2} \quad (t = 1, 2, 3, \dots) \quad (3.2)$$

A distribuição de probabilidade de  $x_t$  pode ser encontrada a partir da distribuição binomial. Para  $t$  passos, a probabilidade de se terem  $n$  saltos negativos e, conseqüentemente,  $t - n$  saltos positivos será:

$$\binom{t}{n} \left(\frac{1}{2}\right)^n \left(1 - \frac{1}{2}\right)^{t-n} = \binom{t}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^{t-n} = \binom{t}{n} \left(\frac{1}{2}\right)^{n+t-n} = \binom{t}{n} \cdot 2^{-t} \quad (3.3)$$

Assim, se houver  $n$  saltos negativos e  $t - n$  saltos positivos num espaço de tempo  $t$ , o valor de  $x$  será  $-n + (t - n) = t - 2n$ . A probabilidade de  $x$  assumir o valor  $t - 2n$  no tempo  $t$  é dado por:

$$\text{prob}(x = t - 2n) = \binom{t}{n} \cdot 2^{-t} \quad (3.4)$$

Se  $x_0 = 0$  então a esperança de  $x_t$  é igual a zero, pois a probabilidade de descer e de subir é a mesma:

$$E_0(x_t) = 0 \quad (3.5)$$

Verifica-se pela equação 3.1 que o processo é do tipo Markoviano, pois a distribuição de probabilidade de  $X_t$  no instante  $t+1$  depende apenas da distribuição de  $X$  no instante  $t$ .

Logo, pode-se generalizar o processo de duas maneiras: A primeira seria alterando os valores de  $p$  e  $q$ , onde  $q = 1 - p$ . Caso  $p > q$ , o Random Walk possuiria um drift e conseqüente mente,  $E_0(x_t) > 0$ . A segunda maneira para generalizar o processo é tratar o tamanho do salto em cada instante  $t$  como uma variável aleatória contínua. Se a distribuição fosse do tipo normal com média zero e variância  $\sigma$ , o processo é conhecido como processo estocástico em tempo contínuo e estado discreto, exemplo do Processo Auto-regressivo de Primeira Ordem (AR(1)). Ou seja, o Random Walk seria um caso particular do processo auto-regressivo de primeira ordem.

### 3.2

#### Processo Auto-regressivo de Primeira Ordem (AR1)

Segundo Brandão (2000), esse tipo de processo é um processo de reversão à média, isto é, no longo prazo  $x$  tende a um valor constante.

A dinâmica de  $X_t$  pode ser descrita da seguinte forma:

$$x_t = \delta + \rho x_{t-1} + \xi_t \quad (3.6)$$

onde  $\delta$  e  $\rho$  são constantes e  $-1 < \rho < 1$  e  $\xi \sim N(0,1)$

$$E(x_t) = E(\delta + \rho x_{t-1} + \xi_t) \quad (3.7)$$

$$E(x_t) = \delta + \rho x_{t-1} \quad (3.8)$$

Pela equação acima se verifica que o processo é Markoviano. O valor esperado de  $X_t$  no longo prazo será:

$$E_0(x_n) = \frac{\delta}{1-p} \quad (3.9)$$

### 3.3

#### Processo de Markov

Os processos estocásticos anteriores, Random Walk e o AR(1), satisfazem a propriedade de Markov, portanto, são chamados de processo de Markov. Neste processo somente o valor presente de uma variável é relevante para prever o futuro. A vantagem do processo de Markov é que ele simplifica a análise dos processos estocásticos.

Geralmente os preços das ações são modelados usando o Processo de Markov. Se o preço de alguma ação específica segue um processo de *Markov*, previsões sobre a flutuação futura dos preços desta ação não devem levar em conta a flutuação ocorrida na semana, mês ou ano passado.

### 3.4

#### Processo de Wiener

Segundo Dixit and Pindyck (1994), o processo de Wiener, ou Movimento Browniano, é um tipo particular de Processo de Markov em tempo contínuo que apresenta três propriedades interessantes.

- É considerado um processo de Markov em tempo contínuo, ou seja, a distribuição de probabilidades dos valores futuros do processo depende somente do seu valor atual. Para fazer uma previsão do valor futuro da variável basta apenas a sua distribuição de probabilidade e o seu valor atual.
- Como apresenta incrementos independentes, as distribuições de probabilidade para as variações em qualquer instante de tempo são independentes de qualquer outro intervalo de tempo.

- Mudanças no processo sobre qualquer intervalo de tempo são normalmente distribuídas, com variância proporcional, que aumenta linearmente com o intervalo de tempo.

### Definição Matemática

Define-se o processo pela equação abaixo:

$$dx = a(x,t)dt + b(x,t)dz \quad (3.10)$$

Onde  $a(x,t)$  é a função não-aleatória de tendência,  $b(x,t)$  é a função não-aleatória de variância. Ainda segundo Brandão (2000), a relação entre a variação de  $z$  e a variação de  $t$  é dada por:

- 1) A relação entre  $\Delta z$  e  $\Delta t$  é dada por  $\Delta z = \tilde{\varepsilon}_t \sqrt{\Delta t}$ , onde  $\tilde{\varepsilon}_t \sim N(0,1)$
- 2) A variável aleatória  $\varepsilon_t$  não possui correlação serial, isto é,  $E(\varepsilon_t, \varepsilon_s) = 0$ , para  $t \neq s$ . Sendo assim, os valores para dois quaisquer intervalos diferentes são independentes, de forma que  $z(t)$  segue um processo de Markov.

Considerando um intervalo de tempo infinitesimalmente pequeno, ou seja,  $\Delta t \rightarrow 0$  refletindo na derivada  $dt = 0$ , é possível representar o incremento do Processo de Wiener ( $dz$ ) no tempo contínuo como:

$$dz = \varepsilon_t \sqrt{dt} \quad (3.11)$$

Pelo fato de que  $\varepsilon_t \sim N(0,1)$  e utilizando a equação acima, verifica-se que o valor esperado da variação de  $z$  é zero e sua variância é proporcional ao intervalo de tempo da variação ( $dt$ ):

Desse modo:

$$E[dz] = E[\varepsilon_t] \cdot \sqrt{dt} = 0, \text{ pois } E[\varepsilon_t] = 0 \quad (3.12)$$

$$\text{Var}[dz] = \text{Var}[\varepsilon_t \cdot \sqrt{dt}] = (\sqrt{dt})^2 \cdot \text{Var}[\varepsilon_t] = dt \cdot (1)^2 = dt, \text{ pois } DP[\varepsilon] = 1 \quad (3.13)$$

Logo, obtêm-se  $dz \sim N(0, \sqrt{dt})$ .

Utilizando a definição do Processo de Wiener:

$dx = a(x,t)dt + b(x,t)dz$ , onde  $(dz)$  é o incremento e  $a(x,t)$  e  $b(x,t)$  são funções não aleatórias conhecidas. Substituindo-se os parâmetros  $a(x,t)$  e  $b(x,t)$  por, respectivamente,  $\alpha$  (conhecido como parâmetro *drift*) e  $\sigma$  (parâmetro de variância), ambos constantes, chega-se à seguinte equação para o conhecido processo MAB (Movimento Aritmético Browniano):

$$dx = \alpha dt + \sigma dz \quad (3.14)$$

- $\alpha$  representa o parâmetro de tendência no tempo (ou crescimento),
- $\sigma$  o parâmetro de variância, que exprime a incerteza ou ruído do processo, ou seja, determina a amplitude dos choques aleatórios que  $x$  sofre ao longo do tempo e é conhecido como volatilidade e  $x$  é um processo estocástico.

Considerando-se em um intervalo de tempo  $(\Delta t)$ , a mudança em  $x$ , denotada por  $\Delta x$ , tem-se:

$$\Delta x = \alpha \Delta t + \sigma \varepsilon \sqrt{\Delta t} \quad (3.15)$$

Sendo, a média e a variância definidos respectivamente como:

$$E[\Delta x] = \alpha \Delta t \quad (3.16)$$

$$\text{Var}[\Delta x] = \sigma^2 \Delta t \quad (3.17)$$

A soma do componente determinístico (drift) ao componente aleatório normalmente distribuído gera  $(dx)$ . A soma de uma constante com uma variável

aleatória normal resulta numa variável ( $dx$ ) também normal com média  $\alpha$  e variância  $\sigma^2$ .

Muito embora seja utilizado em larga escala para previsão de inúmeros processos, o Movimento Geométrico Browniano pode divergir levando  $x(t)$  para o infinito, de forma que alguns modelos que seguem este processo podem não ser muito realistas. Entretanto essa abordagem é utilizada em larga escala por acadêmicos e profissionais.

No MGB, os parâmetros *drift* e variância são dados por:

$$a(x,t) = \alpha x \quad (3.18)$$

$$b(x,t) = \sigma x \quad (3.19)$$

Substituindo as equações acima na equação 3.10, tem-se:

$$dx = \alpha x dt + \sigma x dz \quad (3.20)$$

A maior parte dos estudos de opções reais assume que os fatores de incerteza (preço do petróleo, taxa de câmbio e as taxas de sondas) seguem um MGB. No MGB, a tendência é um crescimento (ou queda) exponencial e os preços tem uma distribuição log-normal com variância crescendo com o horizonte temporal. Uma representação gráfica deste processo está apresentada na figura abaixo:

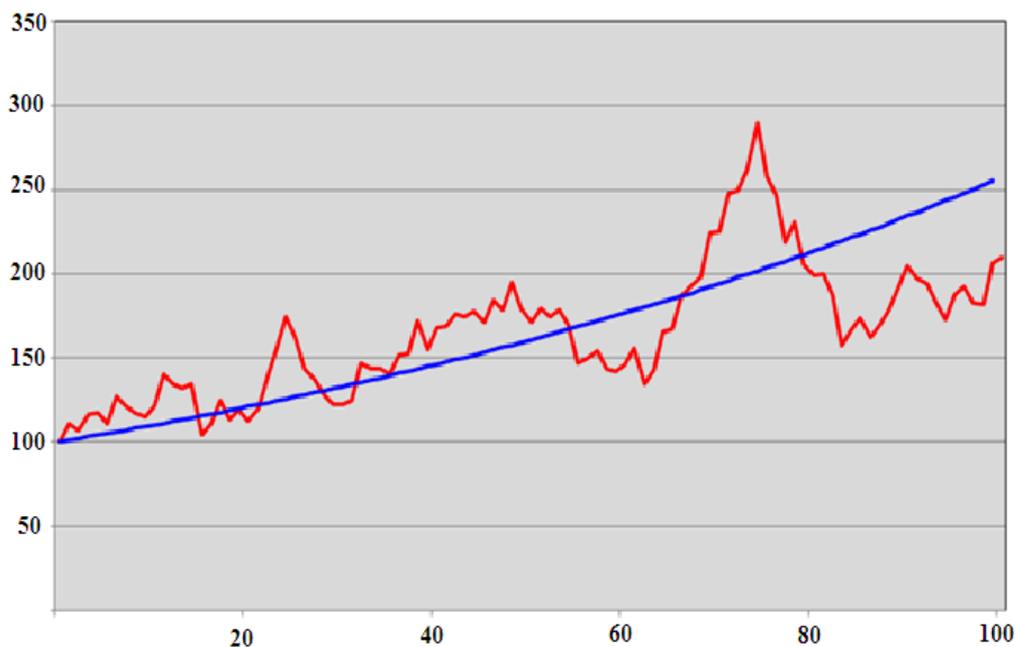


Figura 3.1 – Simulação do Movimento Geométrico Browniano com drift

### 3.5

#### Processo de Reversão a Média

Na abordagem prática, os preços de petróleo são modelados geralmente utilizando o MGB. De acordo com o gráfico anterior, o MGB possui uma tendência de se distanciar do seu ponto inicial com o passar do tempo. Segundo Dias (1996), no caso do processo de reversão à média, a tendência é que os preços devem reverter para um nível de equilíbrio de mercado. Dessa forma, enquanto no curto prazo estes preços podem variar aleatoriamente, no longo prazo eles tendem a retornar para um nível próximo ao seu custo marginal de produção. A grande questão é conseguir determinar com precisão esse movimento.

O processo de reversão à média é um caso particular do processo de Ito. O processo de reversão à média serve para modelar variáveis aleatórias que não seguem um MGB, conforme já foi descrito. Uma dificuldade para este tipo de processo é que ele não possui uma derivada convencional em relação ao tempo, o que pode ser resolvido pelo Lema de Ito, analisado mais adiante.

Segundo Dias (2005), se o preço do petróleo  $P$  estiver distante (acima ou abaixo) de certo preço de equilíbrio ou média de longo prazo  $P'$ , forças de mercado agirão para puxar os preços de volta para o nível de equilíbrio. Este procedimento se dá seguindo a lógica econômica de balanço oferta e demanda, de forma que do lado da oferta, as forças de mercado irão agir para aumentar (se  $P > P'$ ) ou reduzir (se  $P < P'$ ) a produção e o investimento no setor, e do lado da demanda, a mesma tende a cair em casos de altos preços e tende a aumentar em caso de baixos preços. Estes mecanismos de mercado criam uma força de reversão análoga à força de uma mola: ela é mais forte quanto mais longe estiver o preço  $P$  em relação ao seu nível de equilíbrio  $P'$ . Existem diversos modelos de reversão à média. O mais simples deles é o modelo que trabalha com logaritmo dos preços. Neste caso:

$$x = \ln(P) \quad (3.21)$$

$$dx = \eta(\bar{x} - x)dt + \sigma dz \quad (3.22)$$

sendo os elementos que definem a equação diferencial estocástica acima, também conhecida como processo de Ornstein-Uhlenbeck, apresentados a seguir:

- $\eta$  é a velocidade de reversão,
- $\bar{x}$  é o valor de equilíbrio de longo prazo
- $\sigma$  é a volatilidade do processo

A variável  $x$  tem distribuição Normal e, portanto a variável  $P$  tem distribuição lognormal. As equações para a média e variância da variável estocástica  $x$  são dadas pelas expressões a seguir, onde  $x_0 = x(t=0) = \ln(P_0)$ . Diferentemente do MGB, no MRM a variância é limitada. O valor esperado oscila entre o valor inicial  $x_0$  e a média de longo prazo, onde os pesos são taxas de declínios de modo que estes pesos somam um.

Apesar de ser um processo de Markov, a reversão à média não tem incrementos independentes. As propriedades estatísticas do processo de reversão à média são:

- $E[x_t] = \bar{x} + (x_0 - \bar{x})e^{-\eta t}$  (3.23)

- $Var[x_t - \bar{x}] = \frac{\sigma^2}{2\eta}(1 - e^{-2\eta t})$  (3.24)

A figura abaixo mostra graficamente um exemplo típico de um processo de reversão à média:

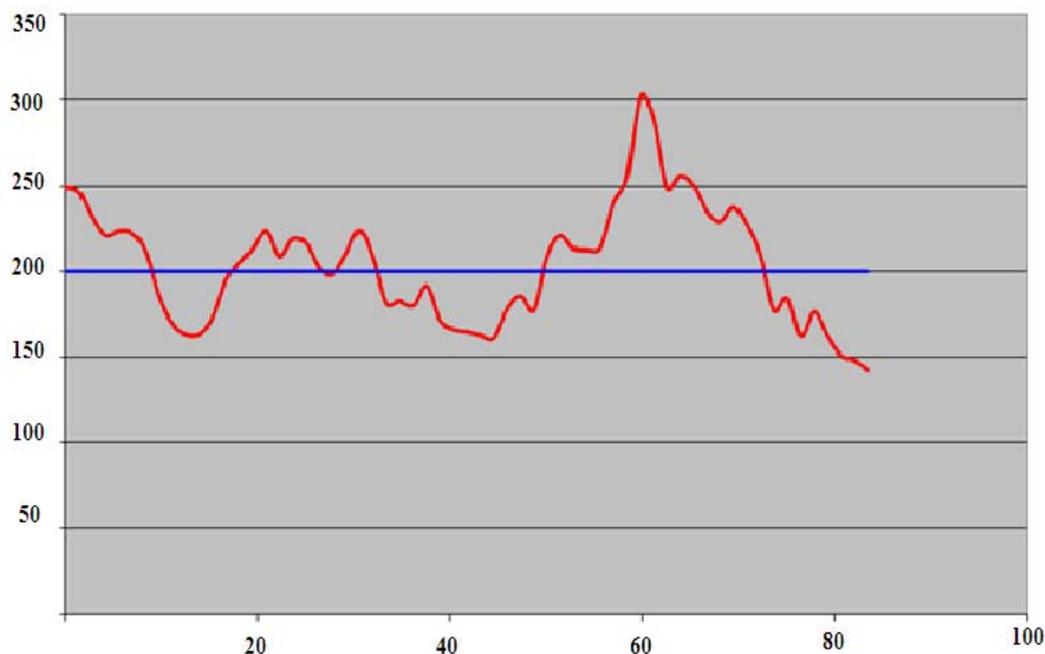


Figura 3.2 – Movimento de Reversão a Média com Nível de Equilíbrio

Embora o MRM possa ser mais adequado para a análise da evolução de preços de (*commodities*) e para taxa de juros, o processo puro de reversão para um nível fixo seria bastante previsível e, portanto, menos confiável do que o MGB. Existe a possibilidade de se combinar o MRM com o MGB para a obtenção de um nível de equilíbrio ou ainda a inclusão do processo de Poisson com saltos, que será discutido a seguir.

Um processo de Itô é um processo contínuo no tempo, mas que não pode ser diferenciado pelas regras ordinárias de cálculo. Entretanto, isto seria essencial para a valoração de uma opção. Sendo assim, faz-se necessário utilizar-se o Lema de Ito, chamado também de Teorema Fundamental do Cálculo Estocástico que será objeto no item a seguir.

### 3.6

#### Lema de Itô

A partir do lema de Ito torna-se viável o cômputo das transformações do processo de Ito. O Lema de Ito tem relativa importância para o cálculo estocástico, assim como a Expansão de Taylor para o cálculo numérico.

Nessa abordagem,  $F(x, t)$  é diferenciável ao menos duas vezes em relação a  $x$ , sendo esta a variável estocástica e uma vez em relação a  $t$ . Utilizando-se o Lema de Ito, esta derivada será:

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} dx^2 \quad (3.25)$$

Nesse caso  $dt^n = 0$  para  $n > 1$ ,  $dz^2 = dt$

sendo

$$dx = a(x, t)dt + b(x, t)dz$$

$$\text{Como } dx^2 = a^2(x, t)dt^2 + b^2(x, t)dt + a(x, t) b(x, t).dt^{3/2} = b^2(x, t)dt \quad (3.26)$$

Substituindo na equação acima obtêm-se pelo Lema de Itô

$$dF = \left[ \frac{\partial F}{\partial t} + a(x, t) \frac{\partial F}{\partial x} + \frac{1}{2} b^2(x, t) \frac{\partial^2 F}{\partial x^2} \right] dt + b(x, t) \frac{\partial F}{\partial x} dz \quad (3.27)$$

Utilizando as variáveis do MGB, simplifica-se a equação abaixo:

$$dF = \left[ \frac{\partial F}{\partial t} + ax \frac{\partial F}{\partial x} + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 F}{\partial x^2} \right] dt + \sigma x \frac{\partial F}{\partial x} dz \quad (3.28)$$

Resolvendo os termos

$$\frac{\partial F}{\partial t} = 0 \therefore \frac{\partial F}{\partial x} = \frac{1}{x} \therefore \frac{\partial^2 F}{\partial x^2} = -\frac{1}{x^2} \quad (3.29)$$

obtêm-se

$$dF = \left(a - \frac{1}{2}\sigma^2\right)dt + \sigma dz \quad (3.30)$$

Percebe-se que a variação de  $F \sim N\left[\left(a - \frac{1}{2}\sigma^2\right), \sigma^2\right]$ , onde  $F = \ln x$ .

Os processos de Wiener e Itô originam os processos descritos até agora. Tais processos são obtidos a partir de funções diferenciáveis dos processos e descrevem os movimentos dos ativos reais. No entanto, alguns processos não são homogêneos no tempo e podem sofrer mudanças abruptas em sua trajetória influenciadas por algum fator exógeno. Tal mudança abrupta pode ser descrita com um salto na trajetória do ativo. O processo de Poisson é utilizado para avaliar este tipo de situação.

### 3.7

#### Processo de Poisson

Em alguns cenários específicos de mercado pode ser mais realista considerar que uma variável econômica segue um processo com saltos discretos. Muitos movimentos exógenos podem ocorrer, de modo a induzir algum movimento abrupto na trajetória dos preços. Esse tipo de situação pode ocorrer em cenários de crise, mudanças inesperadas na política macroeconômica, entrada ou saída de competidores em mercados com poucos competidores, entre outras situações.

Os processos de Poisson são processos estocásticos que possuem a característica de realizar um salto discreto, mas não frequente ao longo do tempo. Tais saltos (*jumps*) podem ser de tamanhos fixos ou aleatórios e o tempo de chegada dos saltos segue uma distribuição de Poisson.

Um processo simples de Poisson é definido pela seguinte equação diferencial:

$$dx = f(x,t)dt + g(x,t)dq \quad (3.31)$$

Onde  $dq$  corresponde ao incremento aleatório, podendo assumir o valor zero ou o valor de um salto de amplitude  $\mu$ , que ocorre com probabilidade  $\lambda dt$ , onde  $\lambda$  é a frequência do processo.

Existem algumas situações que os processos estocásticos se integram. No caso específico dessa dissertação, o preço do petróleo negociado no mercado internacional pode evoluir continuamente segundo MGB ou MRM na maior parte do tempo e, eventualmente, pode sofrer grandes variações instantâneas em decorrência de eventos raros tais como crises financeiras, guerras, dentre outros. O processo heterogêneo abaixo é um misto de Poisson com Itô e pode ser representado matematicamente pela seguinte equação:

$$dx = f(x,t)dt + g(x,t)dz + h(x,t)dq \quad (3.32)$$

O conhecimento de tal processo é importante para valorar opções reais, ainda mais na conjuntura internacional vigente onde as mudanças ocorrem de maneira muito acelerada e de forma quase que imprevisível. O gráfico a seguir ilustra o comportamento de um ativo simulado no processo de Poisson puro.

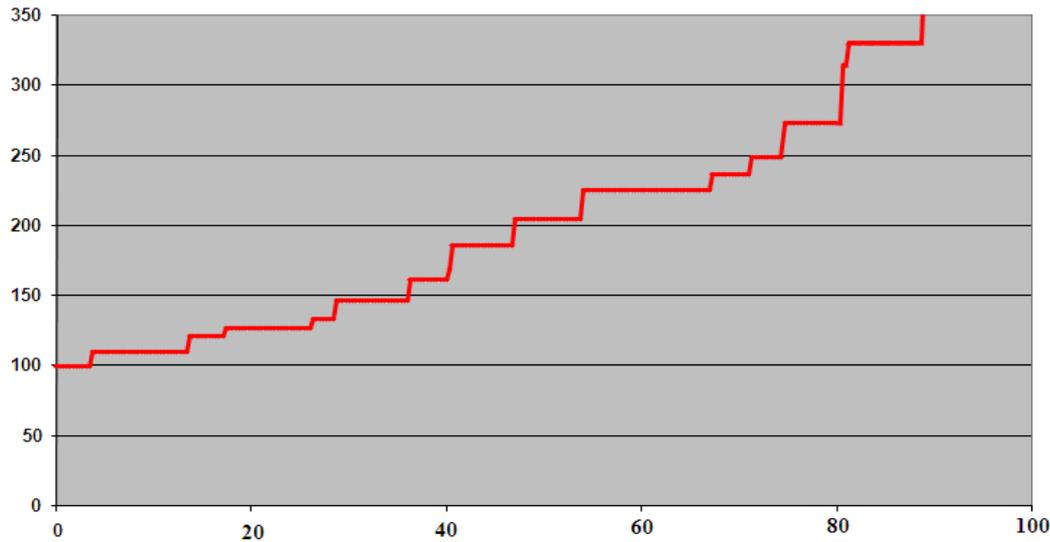


Figura 3.3 – Simulação de um Processo de Poisson Puro

### 3.8

#### Métodos Numéricos

Serão apresentados os principais métodos numéricos: Modelo Binomial, Diferenças Finitas e Simulação de Monte Carlo e Quase Monte Carlo

#### 3.8.1

##### Modelo Binomial

O método binomial exposto abaixo foi baseado na abordagem de Cox, Ross e Rubinstein (1979) e é considerado um dos métodos mais utilizados para a avaliação de ativos. A idéia básica do método é a neutralidade ao risco. A metodologia discretiza o processo de neutralidade ao risco representado pela EDP de Black e Scholes, e aplica a programação dinâmica para determinar o valor do ativo. Esta técnica de valoração de opções deu origem ao termo que se conhece como árvore binomial.

A figura abaixo ilustra uma árvore binomial para uma ação S em três períodos:

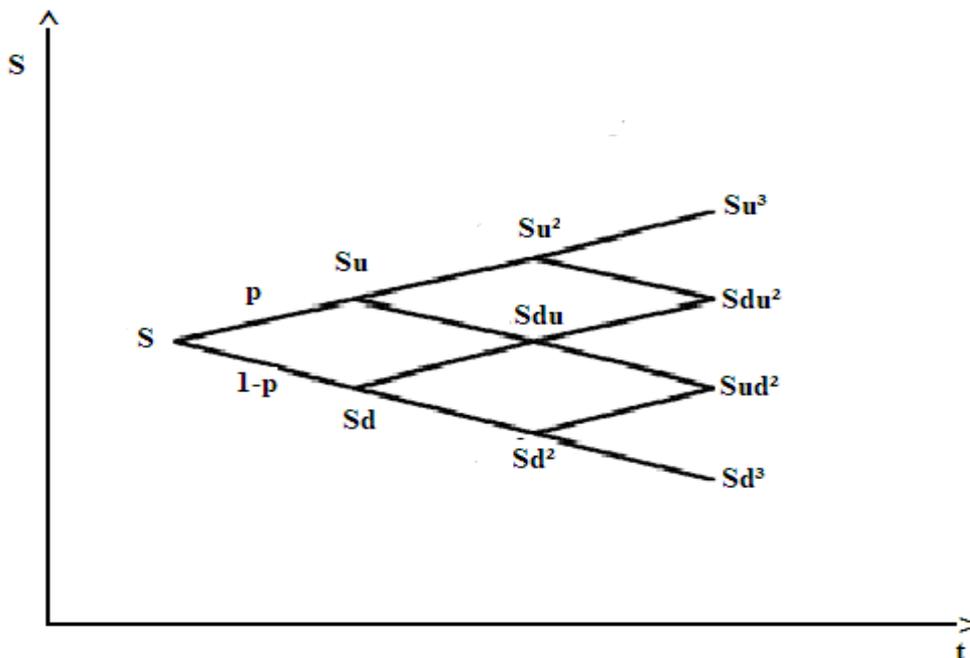


Figura 3.4 – Exemplo de uma árvore binomial com 3 períodos de tempo

Nesta técnica utilizam-se as árvores binomiais que foram utilizadas pelo fato de o preço do ativo subjacente seguir um processo binomial multiplicativo em períodos discretos. A distribuição de probabilidades do ativo em cada período, suposta log-normal, foi aproximada por uma distribuição do tipo binomial e esta representa as diferentes trajetórias que poderão ser seguidas pelo preço da ação ao longo do tempo.

As premissas assumidas nessa abordagem são de que a vida útil da opção é dividida em  $M$  períodos de tempo e considera-se que o preço do ativo é negociado somente nesses períodos. Como ilustra a figura acima, a árvore é construída partindo-se de um valor inicial  $S$ , gerando apenas dois preços possíveis ( $S_u$  e  $S_d$ ) no segundo período, três preços possíveis ( $S_u^2$ ,  $S_{ud}$  e  $S_d^2$ ) no terceiro período e assim por diante, até que a vida útil seja atingida.

Com a abordagem neutra ao risco, as preferências do investidor ao risco não necessitam ser levadas em conta, já que se pode criar uma carteira dinâmica livre de risco em cada instante  $t$ . Suponha o processo a seguir:

$$dS = \alpha S dt + \sigma S dz \quad (3.30)$$

Onde  $\alpha$  é a taxa de crescimento de S,  $\sigma$  é a volatilidade de S e  $dz$  é o incremento de Wiener. No processo neutro ao risco  $\alpha$  pode ser substituído por  $r - \delta$ , onde  $r$  é a taxa livre de risco. Os parâmetros  $u$ ,  $d$  e  $p$  são dados pelas equações a seguir:

$$u = \sigma\sqrt{dt} \quad (3.31)$$

$$d = \frac{1}{u} \quad (3.32)$$

$$p = \frac{1}{2} + (\alpha - \frac{1}{2}\sigma^2)\sqrt{dt} \quad (3.33)$$

Sendo assim, o algoritmo de programação dinâmica é aplicado ao longo da árvore binomial. Em cada um dos nós terminais a remuneração da opção é calculada de acordo com a equação a seguir:

$$F_t(S_t) = \max(S_t - I, 0) \quad (3.34)$$

As seqüências de decisões que podem ser adotadas em cada período de tempo são decompostas em duas partes: decisão imediata e decisão futura. A decisão ótima é aquela que maximiza o valor presente líquido. A decomposição do problema em duas decisões é fundamentada no Princípio de Bellman. Para o problema de avaliação de uma opção de espera, a Equação de Bellman é dada a seguir:

$$F_t(S_t) = \max \left\{ S_t - I, \frac{1}{(1+r)} E_t[F_{t+1}(S_{t+1})] \right\} \quad (3.35)$$

onde,  $F_t$  representa o valor da oportunidade de investimento no tempo  $t$ ,  $S_t$  o valor presente dos fluxos de caixa no tempo  $t$ ,  $I$  o custo de investimento,  $r$  a taxa de desconto livre de risco e  $E_t[F_{t+1}]$  o valor esperado da oportunidade de investimento em  $t+1$ , condicionado às informações em  $t$ . O valor esperado da oportunidade de investimento pode ser calculado, pois o modelo binomial disponibiliza as probabilidades de transição entre dois período de tempo consecutivos.

Como exemplo, considere que em  $t+1$ , a remuneração da opção, quando o valor do ativo passa de  $S_t$  para  $uS_t$ , é dado por  $F_{t+1,u}(S_{t+1}) = \max(uS_t - I, 0)$ , e  $F_{t+1,d}(S_{t+1}) = \max(dS_t - I, 0)$  quando o valor do ativo passa de  $S_t$  para  $dS_t$ . Já o valor da opção em  $t$  é obtido usando a equação anterior e calculando o valor esperado da oportunidade de investimento em  $t+1$ . Estes cálculos são feitos em todos os nós da árvore, do período  $T - \Delta t$  até 0 onde o preço da oportunidade de investimento ( $F_0$ ) é calculado.

$$F_t(S_t) = \max \left\{ S_t - I, \frac{1}{(1+r)} [pF_{t+1,u} + (1-p)F_{t+1,d}] \right\} \quad (3.36)$$

### 3.8.2

#### Modelo de Diferenças Finitas

O método de diferenças finitas avalia um derivativo através da solução da equação diferencial a que o derivativo satisfaz, convertendo-a numa lista de equações diferenciais que são resolvidas iterativamente. Para ilustrar o método, analisa-se uma opção de venda do tipo americana sobre um ativo que não paga dividendos. A equação diferencial que a opção deve satisfazer é:

$$\frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} - rf = 0 \quad (3.37)$$

Deve-se escolher um número de intervalos de tempo entre o instante inicial  $t$  e a data de vencimento da opção ( $T$ ). Assim, o horizonte de tempo será dividido em intervalos de magnitude  $\frac{(T-t)}{N}$ , com  $(N+1)$  instantes de tempo  $(t, t + \Delta t, t + 2\Delta t, \dots, T)$ .

E ainda, é preciso discretizar o preço do ativo subjacente. Para isso, define-se inicialmente um preço limite ( $S_{\max}$ ) ( $S_{\max}$ ) suficientemente grande de modo a garantir que a probabilidade do preço do ativo ultrapassar este valor seja nula. Assim, o grid dos preços da ação será dividido em intervalos de

magnitude  $\Delta S = \frac{(S_{m\acute{a}x})}{M}$ , e teremos  $(M+1)$  preos possveis para o ativo  $(0, \Delta S, 2\Delta S, \dots, S_{m\acute{a}x})$ .

A figura abaixo mostra o exemplo de um grid para a abordagem por diferenas finitas:

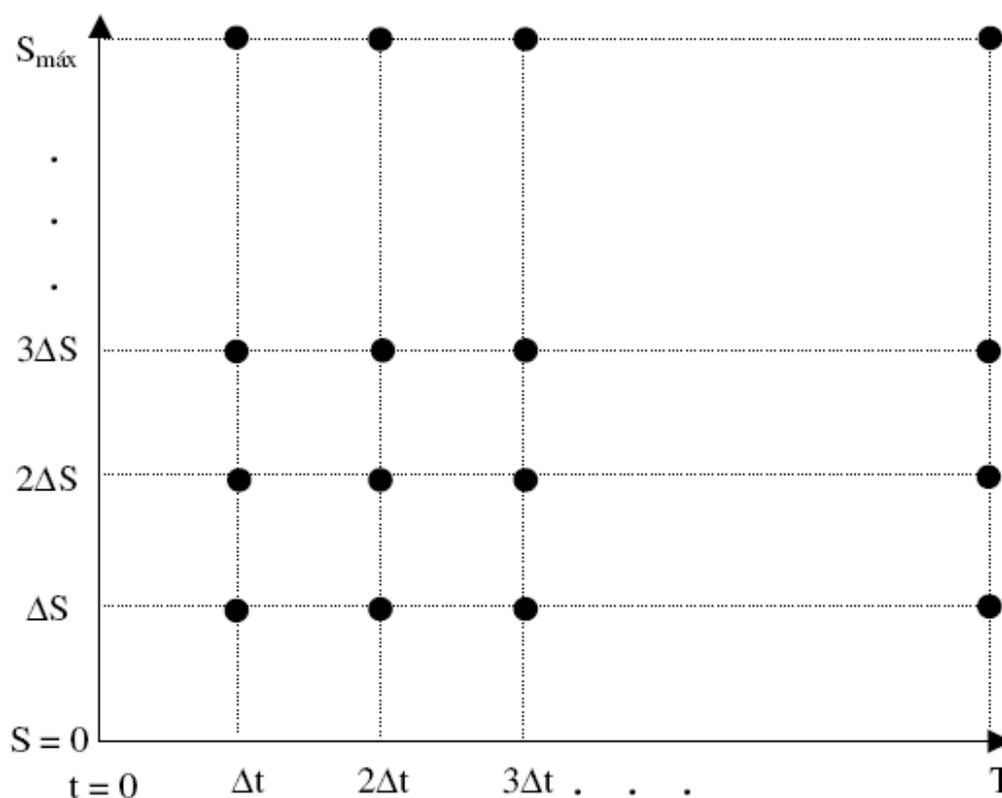


Figura 3.5 – Grid Diferenas Finitas

O grid  formado por um total de  $(M + 1)(N + 1)$  pontos. O ponto  $(i, j)$  corresponde ao ponto no instante  $i\Delta t$  com a ao a um preo  $j\Delta S$ . O mtodo das diferenas finitas pode ser aplicado de forma explcita ou implcita. A grande desvantagem da abordagem implcita em relao a explcita deve-se a necessidade da resoluo de  $(M-1)$  equaes simultaneamente para obteno dos valores da opo em cada instante de  $t$ . No entanto, sempre apresenta resultados convincentes, pois, converge para a soluo da equao diferencial quando  $\Delta S$  e  $\Delta t$  tendem a zero.

A abordagem pelo mtodo explcito  mais simples que pelo mtodo implcito. Porm, o modo explcito nem sempre converge para a soluo da

equação diferencial. Este problema de instabilidade depende dos intervalos  $\Delta S$  e  $\Delta t$  usados para a formação do grid.

### 3.8.3

#### Métodos de Simulação

A seguir serão apresentadas as principais técnicas de simulação. O objetivo é abordar as técnicas de simulação utilizadas nesse trabalho bem como mostrar suas principais aplicações na geração de cenários futuros para os principais fatores de risco.

#### 3.8.3.1

##### Método Tradicional de Monte Carlo

A simulação de Monte Carlo foi introduzida pela primeira vez por Boyle (1977) na avaliação de opções. Tradicionalmente a simulação era vista como um método inadequado para avaliar opções americanas, por aquela ser um método progressivo e a avaliação de opções retroativa. Nos últimos anos vários autores propuseram diferentes métodos que tornaram possível a avaliação de opções americanas, utilizando a simulação: desde Tilley (1993), Barraquand e Martineau (1995); e Longstaff e Schwartz (2001), este último que dos vários métodos, foi o que ganhou particular relevância e será abordado nessa dissertação.

A Simulação de Monte Carlo é uma ferramenta de simulação estatística que utiliza métodos de amostragem para resolver problemas de natureza estocástica ou determinística. É considerado um método apropriado para resolver os problemas de dimensão alta e/ou parâmetros estocásticos e, muitas vezes, é utilizada para calcular o valor esperado de uma variável que é função de várias variáveis estocásticas e que não pode ser tratada analiticamente.

A metodologia da simulação de Monte Carlo utiliza pontos aleatórios para fazer o cálculo numérico de uma integral, ou seja, utiliza uma seqüência de números pseudo-aleatórios para obtenção da área sob a função de estudo como mostra o gráfico abaixo:

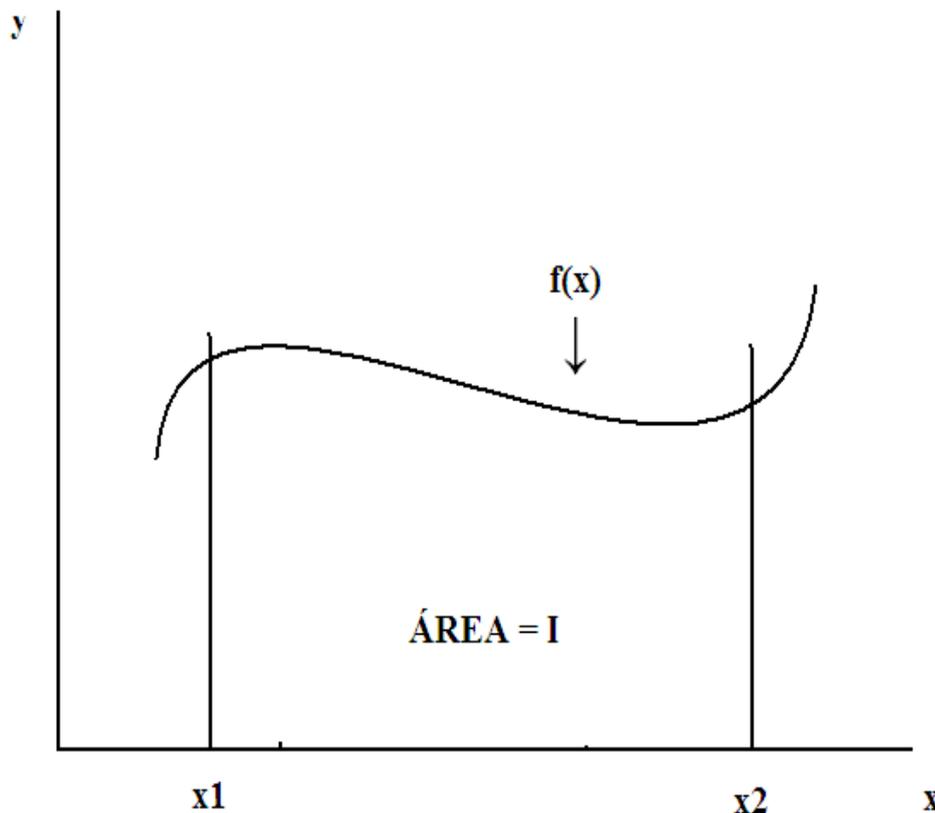


Figura 3.6 – Área da integral na Simulação de Monte Carlo

Para obter o valor da integral correspondente à área descrita acima, utiliza-se de um método de soma que gera uma boa aproximação desse valor. A estimação de Monte Carlo vai aproximar a integral  $I$  do seu valor real, através da média de uma amostragem suficientemente grande e aleatória de  $x_i$ 's.

A fórmula de cálculo é descrita da seguinte maneira:

$$I = \int_{x_1}^{x_2} f(x)dx \approx \frac{(x_2 - x_1)}{N} \sum_{i=1}^N f(x_i) \quad (3.51)$$

O método de Monte Carlo pode ser visto como um problema de cálculo integral. O objeto principal de qualquer simulação está no cômputo de uma ou mais expectativas da forma  $E[\varphi(X)]$ . Em geral, este método envolve problemas

de cálculo de integrais multidimensionais, que pode ser expresso pela seguinte função de acordo com a aproximação de Monte Carlo:

$$\theta = \int \varphi(x)f(x)dx \approx \frac{1}{N} \sum_{\substack{i=1 \\ x_i \in B}}^N \varphi(x_i) \quad (3.52)$$

onde,  $x_i$  é uma amostra aleatória e independente de  $N$  números da função  $f(x)$  oriundos de uma amostra aleatória da distribuição uniforme no intervalo de  $[0,1]$ s . A simulação de Monte Carlo (SMC) geralmente segue os seguintes passos:

- Modelagem das variáveis de entrada importantes, incluindo uma descrição das interdependências entre diferentes variáveis ao longo do tempo;
- Especificação das distribuições de probabilidade para cada uma das variáveis de entrada, com base num histórico de dados ou no conhecimento e sensibilidade dos profissionais envolvidos;
- Obtenção da amostra aleatória (usando um gerador de números aleatórios) a partir da distribuição de probabilidades das variáveis de entrada, possibilitando assim o cálculo dos fluxos de caixa líquidos de cada período e o respectivo VPL do projeto para a amostra considerada;
- O processo é repetido diversas vezes, obtendo-se para cada vez que se repete o processo um VPL para o projeto. Ao final, uma distribuição de probabilidades do VPL do projeto pode ser gerada.

A utilização da SMC é especialmente adequada para opções dependentes de múltiplas variáveis de estado ou opções que dependem do caminho.

As vantagens principais da simulação de Monte Carlo são:

- Fácil implementação. O cálculo requer apenas uma amostra aleatória de valores possíveis para  $x$  para se chegar o seu equivalente em  $f(x)$ .
- Meio rápido e prático de cálculo da integral quando se trabalha com problemas de alta dimensão.

- O desvio-padrão não depende da dimensão da integral como a maioria das outras técnicas dependem.
- O método de Monte Carlo também permite uma aplicação prática muito grande a problemas da vida real, justamente na simulação de valores possíveis que alguma variável poderá vir assumir.

Já as principais desvantagens e demais críticas a essa metodologia são feitas principalmente por Niederreiter (1992) apud Krykova (2003). São elas:

- a simulação de Monte Carlo converge devagar para o verdadeiro valor – da ordem de  $\frac{1}{\sqrt{N}}$  – onde N é o tamanho da amostra simulada. Isso significa que para reduzir o erro pela metade é necessário que aumente o tamanho da amostra em quatro vezes.
- Os erros nas “caudas” das distribuições geradas por Monte Carlo podem ter o seu valor sobre estimado devido ao desenvolvimento do integrando da função.
- A simulação necessita que o analista arbitre um valor inicial para o processo ou utilize um gerador automático para tal função

Na tentativa de melhorar os resultados da SMC, pode-se aumentar o número de simulações, o que por sua vez pode elevar demasiadamente o custo computacional deste processo. Existem algumas técnicas de redução de variância que aumentam a eficiência da SMC sem um acréscimo significativo do custo computacional.

### 3.8.3.1.1

#### **Geração de Números Aleatórios**

Este é a etapa que requer a análise mais cuidadosa. Gerar números aleatórios no Excel utilizando, por exemplo, a função padrão randn(), pode resultar em algumas distorções em termos de resultados como afirma Albuquerque (2008). O autor afirma que o gerador Excel não consegue aproximar os valores nos pontos extremos da distribuição. É importante ressaltar que essa

diferença entre os dois inversores pode levar a problemas na correta determinação dos pontos ao longo da curva de distribuição de probabilidades acarretando e imprecisão nos valores gerados e conseqüentemente da análise via tal método.

### 3.8.3.1.2

#### Tipos de Geradores de Números Aleatórios

Existem basicamente três tipos de geradores de números aleatórios: pseudo-aleatórios, quasi-aleatórios e aleatórios. Os números aleatórios são de difícil obtenção o que restringe seu uso prático. Os números quasi-aleatórios, apesar de convergirem mais rapidamente para problemas de baixa dimensão, apresentam deficiências a medida que se aumentam as dimensões. Porém uma boa solução nesse caso seria a utilização do gerador aleatório híbrido. Para o programa dessa dissertação, foram utilizados números pseudo-aleatórios.

### 3.8.3.1.3

#### Convergência

Nas seqüências Pseudo-Aleatórias o estimador converge à taxa  $\frac{1}{N}$ . Onde  $N$  é o número de simulações. Na precificação de derivativos se forem necessários um grande número de intervalos até o vencimento a dimensão cresce, o que pode se tornar um problema para as seqüências quase-aleatórias. A grande vantagem das pseudo-aleatórias é que o erro independe da dimensão. A desvantagem é a baixa velocidade de convergência (por exemplo, se quisermos diminuir o erro pela metade deve-se quadruplicar o número de simulações).

Nas seqüências quase aleatórias o limite máximo da taxa de convergência do estimador é  $\frac{\log(N)^d}{N}$ .

Quando a dimensão  $d$  é baixa esta seqüência apresenta resultados muito superiores aos números pseudo-aleatórios, com o estimador convergindo aproximadamente a taxa  $1/N$ .

Contudo, quando a dimensão cresce, podem ocorrer problemas como ilustrado no gráfico abaixo, que mostra a irregularidade da distribuição de duas seqüências adjacentes de alta dimensão.

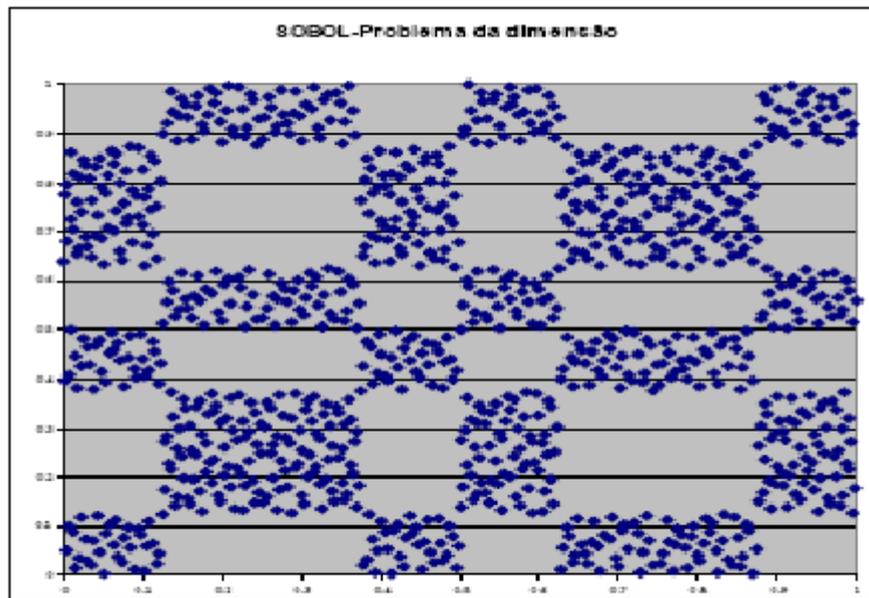


Figura 3.7 – Problemas com alta dimensão

O gráfico acima ilustra a seqüência de Sobol e sua utilização em problemas de dimensão elevada.

A construção de seqüências aleatórias é feita através de um algoritmo que gera uma distribuição uniforme no intervalo (0,1). Existem basicamente 3 métodos que transformam uma distribuição uniforme em normal padrão : Estatística, Box Miller e Parametrização. Escolheu-se o Box Miller neste trabalho porque se comporta bem em seqüências pseudo-aleatórias. Há também a possibilidade de utilização da inversão de Moro que se mostra bastante eficaz na transformação. Tanto Box Miller quanto moro foram implementados no programa em VBA da dissertação.

### 3.8.4

#### Aproximação de Box & Muller

Um importante método usado para gerar dados da distribuição normal é o de Box-Müller, que é baseado na generalização do método da transformação de probabilidade para mais de uma dimensão. A distribuição normal padrão dentro da família normal, definida pelos seus parâmetros  $\mu$  e  $\sigma^2$ , é aquela com média nula e variância unitária. Esta densidade será referenciada por  $N(0: 1)$ . Essa deve ser uma função externa escrita pelo usuário. Os argumentos  $\mu$  e  $\sigma^2$  da função são a média e o desvio padrão da distribuição normal que pretende-se gerar. A função de densidade da normal é:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3.53)$$

O método é baseado na generalização do método da transformação de probabilidades para mais de uma dimensão. Considere  $p$  variáveis aleatórias  $X_1, X_2, \dots, X_p$  com função de densidade conjunta  $f(x_1, x_2, \dots, x_p)$  e  $p$  variáveis aleatórias  $Y_1, Y_2, \dots, Y_p$ , funções de todos os  $X$ 's, então a função de densidade conjunta dos  $Y$ 's é:

$$f(y_1, \dots, y_p) dy_1 \dots dy_p = f(x_1, \dots, x_p) \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \dots & \frac{\partial x_1}{\partial y_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_p}{\partial y_1} & \dots & \frac{\partial x_p}{\partial y_p} \end{vmatrix} dx_1 \dots dx_p \quad (3.54)$$

em que  $J = \left| \frac{\partial(x)}{\partial(y)} \right|$  é o Jacobiano da transformação dos  $X$ 's em

relação aos  $y$ 's. considerando a transformação (3.54) para gerar dados  $y$  normais com densidade dadas por (3.53) devidamente adaptada para acomodar  $y$  e não  $x$ . Para que seja aplicada a transformação de Box-Müller, é preciso considerar duas variáveis aleatórias uniformes entre 0 e 1, representadas por  $x_1$  e  $x_2$  e duas funções delas, representadas por  $y_1$  e  $y_2$  e dadas por:

$$y_1 = \sqrt{-2 \ln x_1} \cos(2\pi x_2) \quad (3.55)$$

$$y_2 = \sqrt{-2 \ln x_1} \text{sen}(2\pi x_2) \quad (3.56)$$

Assim, a função de densidade conjunta de  $y_1$  e  $y_2$  é dada por  $f(x_1; x_2) | J |$ , sendo, portanto:

$$f(y_1; y_2) = \left[ \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{y_1^2}{2}\right)} \right] \left[ \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{y_2^2}{2}\right)} \right] \quad (3.57)$$

Desde que a densidade conjunta acima seja o produto de duas normais independentes, pode-se afirmar que as duas variáveis geradas são normais padrão independentes.

Assim, utiliza-se esse resultado para gerar variáveis aleatórias normais. A dificuldade é apenas computacional, pois a utilização de funções trigonométricas como seno e co-seno pode tornar lento o algoritmo gerado. Uma saída utilizada na prática é ao invés de considerar os valores das variáveis aleatórias uniformes  $x_1$  e  $x_2$  de um quadrado de lado igual a 1 (quadrado unitário), tomarmos  $u_1$  e  $u_2$  como coordenadas de um ponto aleatório em um círculo unitário (de raio igual a 1). A soma de seus quadrados  $R^2 = u_1^2 + u_2^2$  é um valor de uma variável aleatória uniforme que pode ser usada como  $x_1$ . Já o ângulo que o ponto  $(u_1; u_2)$  que determina em relação ao eixo  $u_1$ , pode ser usado como um ângulo aleatório dado por  $\theta = 2\pi x_2$ . A vantagem da não utilização direta da expressão (3.55) refere-se ao fato do co-seno e do seno poderem ser obtidos alternativamente por:  $\cos(2\pi x_2) = \frac{u_1}{\sqrt{R^2}}$  e  $\text{sen}(2\pi x_2) = \frac{u_2}{\sqrt{R^2}}$ , evitando assim as funções trigonométricas. Na figura abaixo ilustra-se os conceitos apresentados e denominamos o ângulo que o ponto  $(u_1; u_2)$  determina em relação ao eixo  $u_1$  por  $\theta$ .

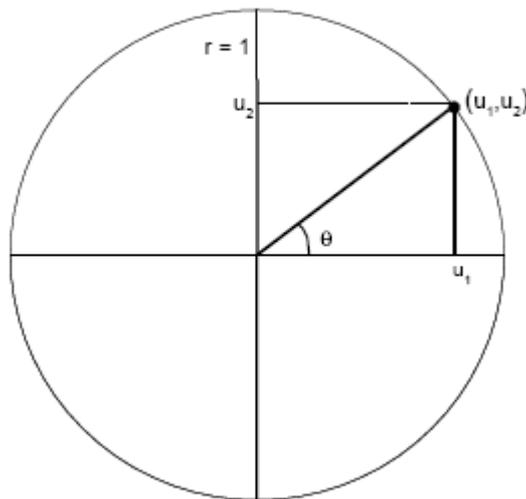


Figura 3.8 – Círculo unitário onde um ponto aleatório  $(u_1; u_2)$  determina em relação ao eixo  $u_1$

### 3.8.5

#### Técnicas de Redução de Variância

Existem diversas técnicas utilizadas para diminuir os erros nas estimativas ao utilizar-se o método de simulação para precificar opções. O erro é mensurado do pela seguinte equação:  $\frac{\sigma_{sim}}{\sqrt{N_{sim}}}$ . Quanto menor o desvio da simulação, maior a precisão dos resultados. Esse desvio deve estar associado à variância dos valores da opção obtidos através de repetição das simulações um grande número de vezes, sendo assim um desvio padrão populacional e não amostral. Vários métodos foram desenvolvidos no intuito de reduzir esse desvio padrão e assim diminuir também o número de simulações necessárias a determinado nível precisão. Pode-se citar: Variáveis Antitéticas, Estratificação, Variáveis de controle, Importance Sampling, dentre outros. A seguir será abordada a principal técnica utilizada nessa aplicação que são as variáveis antitéticas.

### 3.8.5.1

#### Variáveis Antitéticas

Esta abordagem é a mais simples e também a mais utilizada para reduzir a variância de uma estimativa. Utiliza-se o conceito de que uma trajetória de preços e sua imagem possuem a mesma probabilidade de ocorrência. Ou seja, deve-se gerar uma variável estocástica negativamente correlacionada à variável de estado do ativo objeto. Assim, cada trajetória deve ser associada a um par de seqüências, isto é, duas trajetórias negativamente correlacionadas. Supondo que uma variável aleatória  $Z_i$  tenha uma distribuição normal padronizada, então  $-Z_i$  também terá a mesma distribuição. Desse modo o mesmo preço obtido utilizando  $Z_i$  pode ser obtido com  $-Z_i$ , de modo que para gerar N caminhos aleatórios basta apenas gerar N/2 números aleatórios melhorando a rapidez na convergência como no exemplo abaixo:

Exemplo Numérico:

Dados do Modelo

Investimento : \$ 1000

Taxa de Dividendos do Projeto: 5% ao ano

Valor do Projeto: \$ 1000

Taxa Livre de Risco: 10% ao ano

Volatilidade do MGB: 25%

Intervalo de Tempo: 3 anos

Quadro 3.1 – Exemplo Numérico da utilização da Variável Antitética

Número de trajetórias	Opção MC	Desvio MC	Opção Anti	Desvio Anti
100	173,34	19,90	175,27	13,39
500	170,40	11,31	171,39	6,15
1.000	170,90	7,59	171,75	5,53
2.000	170,73	6,21	170,76	4,70
5.000	170,20	3,59	170,22	3,19
10.000	170,05	2,61	170,10	2,09

Acima, verificou-se a partir da utilização do programa desenvolvido em VBA pelo autor que à medida que aumenta o número de trajetórias o valor da opção converge para um nível de preço. Em cada item da simulação a introdução de variáveis antitéticas contribuiu para a diminuição do desvio quadrático médio.