

Noções Breves sobre Simulação

Onofre Alves Simões

Fevereiro de 2021

Índice

1. Introdução	4
2. O Método de Monte Carlo	6
3. Inversão da Função de Distribuição	9
4. Algumas Aplicações	12
5. Princípios Básicos do Método de Monte Carlo	17
5.1 Introdução	17
5.2 Formulação Geral	18
6. Séries Pseudoaleatórias	22
6.1 Números aleatórios	22
6.2 Exemplo de um gerador de NA: o método das congruências	23
6.3 Testes estatísticos	25
6.3.1 Teste de Kolmogorov-Smirnov	26
6.3.2 Testes de sequência	28
7. Geração de Observações de V.A. Correlacionadas	33
7.1 Populações normais multivariadas	33
7.2 Caso geral	34
8. Simulação de Processos Estocásticos	38
7.1 Processos estocásticos em tempo discreto	38
7.2 Processos estocásticos em tempo contínuo	38
9. Métodos Para a Redução do Erro de Estimação	41
Referências	42

1. Introdução

Desde há muito que as técnicas de simulação são uma importante ferramenta para a resolução de problemas. Tratando-se de técnicas extremamente versáteis, podem ser utilizadas em praticamente todos os tipos de sistemas estocásticos, desde as filas de espera aos sistemas financeiros, passando pela gestão de stocks, o desenvolvimento de processos produtivos e de redes de distribuição, ou mesmo pelo cálculo da probabilidade de se conseguir terminar um dado projeto dentro do prazo acordado.

O recurso à simulação torna-se indispensável quando o sistema probabilístico em consideração é demasiado complexo para que o problema em estudo seja solucionado, de forma exata ou, pelo menos, satisfatória, recorrendo a modelos matemáticos e estatísticos. Quando a complexidade inviabiliza esta alternativa, o que é frequente nos modelos financeiros, a simulação é frequentemente a única abordagem praticável.

Um modelo de simulação visa assim “replicar” o comportamento de um determinado sistema, objeto de estudo, com o propósito de procurar detetar as interações existentes entre os vários elementos que o compõem – e que se traduzem sobretudo nas relações inputs/outputs. Por esse motivo, os outputs que tais modelos permitem obter apresentam-se habitualmente em termos de medidas selecionadas, que refletem o desempenho do sistema, face aos diferentes cenários considerados para os inputs. Depois, podem tomar-se decisões sobre o rumo a tomar.

Tradicionalmente, as técnicas de simulação têm sido utilizadas na análise de problemas de dois tipos distintos:

1. Problemas teóricos nas áreas da Matemática, da Física e da Química.
2. Problemas relacionados com aplicações práticas.

Dentro do primeiro grupo, merecem destaque particular:

- A estimação da área limitada por uma curva, incluindo a avaliação de integrais múltiplos. Claro que pode parecer confuso como se há-de estabelecer a ligação entre o cálculo de um integral e uma distribuição de probabilidade, mas a verdade é que se faz.
- A inversão de matrizes.
- A resolução de equações diferenciais parciais.
- O estudo do movimento de partículas num plano.
- O estudo da difusão de partículas.
- A resolução de sistemas de equações lineares.

Quanto ao segundo grupo, é muito comum:

- A simulação de sistemas de inventário, filas de espera, sistemas de distribuição, calendarização de ações de manutenção, ...
- A simulação do funcionamento de uma empresa, do comportamento dos consumidores, da evolução dos capitais necessários para o crescimento da empresa, dos mercados, incluindo os financeiros, da economia, ...
 - A simulação de sistemas sociais e dos comportamentos, ...
 - A simulação de sistemas biomédicos, ...
 - A simulação de estratégias e táticas de combate.

O procedimento é de enorme simplicidade conceptual: consideram-se cenários alternativos para as variáveis de input, simula-se o funcionamento do sistema, dado cada um desses cenários, e analisam-se os outputs obtidos, cenário a cenário. Depois, podem tomar-se decisões sobre o rumo a tomar.

Por exemplo, na simulação do funcionamento de um certo balcão de um banco, dados o horário de atendimento, o número de funcionários e a procura, pode ter-se como objetivo calcular estimativas para o tempo médio de espera dos clientes e o tempo médio em que não há clientes para atender. Na simulação da gestão do stock de um artigo pode procurar avaliar-se o nível médio e o nível máximo de stock e os custos de posse, de encomenda e de ruturas. Na simulação da evolução do preço do ativo subjacente a uma dada opção procura normalmente determinar-se o prémio, ou a probabilidade da opção vir a ser exercida. Na simulação das taxas de juro futuras procura saber-se a probabilidade associada a certos quantitativos de retorno, ...

Uma experiência de simulação diferencia-se de uma experiência laboratorial, na medida em que pode ser inteiramente conduzida no computador. A partir da expressão das interações entre as componentes do sistema por meio de relações matemáticas, é possível recolher a informação necessária de modo muito semelhante ao que sucederia no mundo real (aparte, evidentemente, as simplificações introduzidas no modelo).

A natureza da simulação permite assim ter grande flexibilidade na representação de sistemas complexos, como são via da regra os modelos financeiros. No entanto, o desenvolvimento deste tipo de modelos pode ser muito exigente em termos de recursos.

2. O Método de Monte Carlo

Na chamada simulação de Monte Carlo (MC) os modelos são construídos tendo explicitamente como input variáveis aleatórias, que representam as fontes de incerteza presentes no problema em estudo. Conhecidas as distribuições de probabilidade dessas variáveis aleatórias, é então possível correr o modelo um grande número de vezes, de tal modo que em cada uma das corridas as variáveis aleatórias incluídas assumam particulares valores concretos, ditados pelas respectivas distribuições (ver Glasserman P (2003)).

Ou seja, o que se pretende é que a geração dos particulares valores concretos seja feita de modo a “reconstituir” a distribuição conjunta das variáveis de input.

Por exemplo, admitindo que é necessário gerar observações de um par aleatório (x, y) com f.p conjunta uniforme em $\{1,2,3\} \times \{1,2,3\}$, em 900 000 gerações independentes de observações do par aleatório deve esperar-se que cada uma das nove concretizações possíveis surja 100 000 vezes.

Dependendo do número de variáveis aleatórias input do modelo, e do conjunto dos valores possíveis para cada uma delas, assim poderá ser necessário efetuar milhares ou centenas de milhares de réplicas, antes de a experiência estar completa.

Uma experiência de simulação de Monte Carlo só se considera completa quando a distribuição de probabilidade (empírica) das variáveis output for conhecida com razoável certeza.

Esta necessidade de conhecer a distribuição das variáveis output (impossível de obter exatamente pela técnica tradicional) é muitas vezes incontornável, sobretudo no que diz respeito às caudas, que caracterizam probabilisticamente as situações extremas, normalmente as mais críticas em análises do risco.

Verifica-se assim que o Método de Monte Carlo (MMC) se impõe na simulação de *processos probabilísticos*, uma vez que se baseia muito simplesmente no princípio de recorrer à amostragem para estimar o resultado procurado. Nestas condições, a simulação deve ser tratada como uma *experiência aleatória*.

Ao contrário das soluções que os modelos matemáticos determinísticos fornecem, pontuais e sem que seja possível atribuir-lhes uma probabilidade de concretização, os resultados produzidos quando se corre um modelo de simulação estocástica são *observações de variáveis aleatórias*, com todas as consequências inerentes, estando inclusivamente sujeitas ao erro de amostragem. Isto significa que qualquer inferência relativa ao desempenho do modelo de simulação tem que se sujeitar aos testes estatísticos apropriados.

Como é evidente, o conhecimento das distribuições de probabilidade é a forma mais adequada de descrever o comportamento dos fatores de incerteza existentes e a sua influência sobre os resultados. Alguns exemplos das distribuições mais utilizadas e das aplicações tradicionais:

- *Normal* – Distribuição simétrica completamente especificada pelo conhecimento da média e do desvio padrão e com aplicações tão díspares como o peso e a altura ou as taxas de inflação e os preços da energia.
- *Lognormal* – Com assimetria positiva, é usada para representar grandezas que não assumem valores negativos, mas têm um potencial de crescimento ilimitado, caso do valor de certos ativos imobiliários, de algumas ações ou das reservas de algumas MP.
- *Uniforme* – Outra distribuição simétrica, bastando conhecer o mínimo e o máximo dos valores possíveis; é adequada na modelização de certos custos de fabrico.
- *Triangular* – Caracterizada pelo mínimo, pelo máximo e pela moda, usa-se para a descrição do volume de vendas por unidade de tempo ou do nível do stock de certo tipo de bens num armazém, ao longo do tempo.
- *Exponencial* – Especificada pela média, aplica-se à modelização de tempos de espera por certo tipo de ocorrências, ou das durações de alguns equipamentos.

No decorrer de uma simulação de MC, realiza-se um processo iterado de amostragem casual, que consiste em gerar uma observação de cada uma das variáveis aleatórias consideradas como input e calcular os correspondentes valores das variáveis de output. Estes resultados são então objeto de registo. Depois de um número de iterações considerado suficiente, constrói-se a distribuição das variáveis de output, cuja análise é determinante para alguma eventual tomada de decisão. Consegue-se com este processo um quadro muito completo de toda a situação, que contempla não só aquilo que poderá vir a acontecer, mas também a probabilidade com que poderá vir a acontecer.

As vantagens do MMC, comparativamente a uma análise do tipo determinista, ou de estimação de ponto único, como também se diz, são indiscutíveis:

- Fornece resultados probabilizados.
- Possibilita análises gráficas, pois as múltiplas iterações dão origem a uma enorme quantidade de observações estatísticas, que podem ser objeto dos mais variados tratamentos gráficos, com todas as vantagens interpretativas daí decorrentes.
- Viabiliza análises de sensibilidade, sobretudo quando é importante descobrir quais os fatores que têm maior responsabilidade pelos resultados particularmente gravosos.
- Permite a deteção das combinações de fatores mais arriscadas, pois os analistas conseguem determinar com exatidão os cenários de input associados a certos resultados, informação inestimável para a análise subsequente. Tal não é possível de forma tão integral com os modelos deterministas, a menos que se conheça tão bem o sistema que *a priori* se saiba quais são esses cenários. Mas a probabilidade de alguns serem ignorados é muito maior do que quando se fazem milhares, ou milhões, de corridas do sistema.
- Facilita a correta introdução das relações de interdependência existentes entre as diversas variáveis de input, o que é da maior importância para a precisão dos resultados.

Para além de todas as vantagens enumeradas, acresce que o método de MC é muitas vezes a única ferramenta ao dispor dos analistas financeiros, sobretudo quando se trata de efetuar cálculos complexos na determinação dos preços de certos produtos, ou a avaliação de determinados riscos. Há problemas, com efeito, em que só com uma simulação estocástica bem conduzida é possível chegar a uma solução que mereça alguma confiança por parte dos decisores.

A finalizar este ponto, uma nota histórica. A conceção do MMC é atribuída a um matemático de origem polaca radicado nos Estados Unidos, Stanislaw Ulam. Segundo parece, a ideia ter-lhe-á surgido quando se ocupava com o cálculo da probabilidade de conseguir ganhar um jogo de Solitário. As condições da génese levaram um seu colaborador, Nicholas Metropolis, a dar ao procedimento o nome de Método de Monte Carlo (MMC), em atenção aos múltiplos casinos que há na cidade com o mesmo nome. Os dois publicaram o seu pioneiro trabalho em 1949 (cf. Metropolis e Ulam (1949)).

3. Inversão da Função de Distribuição

Foi atrás amplamente salientado que um dos aspetos dominantes no uso da simulação de MC está associado à necessidade de descrever o problema em estudo também por meio de uma distribuição de probabilidade adequada, da qual as amostras com as observações das variáveis de input são extraídas.

Nos modelos de simulação, o processo de amostragem a partir de qualquer distribuição alicerça-se no uso de números aleatórios no intervalo $[0,1]$, os quais devem satisfazer as seguintes propriedades estatísticas:

1. Ter distribuição uniforme;
 2. Os sucessivos valores têm que ser gerados de uma forma totalmente aleatória, ou seja, as sucessivas observações devem poder associar-se a variáveis aleatórias independentes.
- Dois resultados teóricos muito conhecidos (a chamada transformação uniformizante) fundamentam este uso de números aleatórios no intervalo $[0,1]$ (ver Murteira e Antunes (2012) ou Hogg e Tanis (2014)).

Teorema 1: Se X é v.a. do tipo contínuo, com função de distribuição $F(x)$, então a v.a. $U = F(X)$ tem distribuição $\mathcal{U}(0,1)$.

Teorema 2: Se $F(x)$ é função de distribuição de uma v.a., e se $U \sim \mathcal{U}(0,1)$, então existe uma função $X = \Psi(U)$ com função de distribuição $F(x)$.

Uma outra formulação, que destaca melhor como se há de usar a distribuição $\mathcal{U}(0,1)$ no domínio da simulação de MC, é a seguinte.

Teorema 3: Seja X uma v.a. com função de distribuição $F(x)$, estritamente crescente. Se $U \sim \mathcal{U}(0,1)$, então $F^{-1}(U)$ é uma v.a. identicamente distribuída a X .

A demonstração é imediata, bastando notar que $F^{-1}(u) \leq x \Leftrightarrow U \leq F(x)$, donde resulta que $P[F^{-1}(U) \leq x] = P[U \leq F(x)]$, pois acontecimentos equivalentes têm iguais

probabilidades. Mas, sendo $P[U \leq F(x)] = \int_{-\infty}^{F(x)} f(u)du = \int_0^{F(x)} 1du = F(x)$, então

$P[F^{-1}(U) \leq x] = F(x)$. Fica provado que $P[F^{-1}(U) \leq x] = P[X \leq x]$, ou seja, que $F^{-1}(U)$ é uma v.a. identicamente distribuída a X .

O resultado anterior exige que F seja estritamente crescente, de modo a garantir a existência da inversa. Nos casos em que F não é estritamente crescente, como sucede se X é v.a. discreta, recorre-se à chamada inversa generalizada, em que para cada $u \in [0,1]$ se define $F^{-1}(u) = \inf \{x \in R : F(x) \geq u\}$.

Torna-se assim fácil gerar amostras de qualquer população X : basta gerar amostras aleatórias extraídas da população $U \sim \mathcal{U}(0,1)$ e transformá-las em amostras aleatórias da população X desejada, a partir da relação $X = F^{-1}(U)$.

Claro que, quando não se dispõe da expressão analítica de F , ou esta função não é invertível, é necessário encontrar alternativas. A distribuição Normal de parâmetros μ e σ é uma destas distribuições. Para a resolução do problema há vários métodos, entre os quais o método direto, ou método de Box-Muller (ver, por exemplo, Pidd (2006) ou Taha (2010)).

O método direto inicia-se com a geração de dois números aleatórios com distribuição uniforme em $[0,1]$, representem-se por u_1 e u_2 ;

Prossegue com o cálculo de

$$z_1 = \sin(2\pi u_2) \sqrt{-2 \log_e u_1} \quad \wedge \quad z_2 = \cos(2\pi u_2) \sqrt{-2 \log_e u_1},$$

que se demonstra fornecer duas observações, z_1 e z_2 , da distribuição Normal Standard;

Fica completo com a transformação de z_1 e z_2 nas observações $x_1 = \mu + \sigma z_1$ e $x_2 = \mu + \sigma z_2$, que já são da distribuição $N(\mu, \sigma)$, como se pretendia.

Para a simulação de uma v.a. Y com distribuição log normal (μ, σ) faz-se como que uma “extensão” deste método - pois diz-se que Y tem distribuição log normal (μ, σ) quando $X = \ln Y$ tem distribuição $N(\mu, \sigma)$.

Assim, depois de se obter uma observação x da distribuição normal de parâmetros μ e σ , com o método direto, basta ter em atenção que $y = e^x$ é a observação desejada.

4. Algumas Aplicações

Nas aplicações seguintes, vai recorrer-se ao Excel, mas podia ter-se usado o *Mathematica*, o *Matlab*, o C++, o R ... Relativamente ao Excel, há duas abordagens básicas: a trivial folha de trabalho ou, quando a dimensão/complexidade da situação em estudo não o permite, o uso da linguagem VBA (*Visual Basic for Applications*). Nestes exemplos adopta-se a primeira, que consiste em construir um modelo compacto do problema numa linha da *worksheet* e replicá-lo tantas vezes quantas as necessárias nas linhas seguintes, aproveitando as virtualidades das funções *copy* e *paste* do programa.

Para evitar que o Excel refaça automaticamente a simulação, sempre que se realiza alguma operação na folha de trabalho, pode escolher-se a opção *Manual* no menu das *Calculations*, ou usar a opção *Paste Special... Values and number formats*, no menu *Edit*, para copiar toda a folha com os cálculos para uma nova folha.

Caso 1: No dia 1 de Janeiro de 2011 um investidor aplicou 10000 pelo prazo de 3 anos, nas condições seguintes. A 31 de Dezembro de 2011, 2012 e 2013 será lançado um dado equilibrado: se a pontuação for 5 ou 6, é aplicada a taxa de 5% ao valor acumulado no início do ano; se a pontuação for 1, 2, 3 ou 4, é aplicada a taxa de 0%.

Calcule-se o valor mínimo e também o valor máximo possíveis para o valor acumulado ao fim dos 3 anos e as respetivas probabilidades. Calcule-se ainda a probabilidade de se obter um valor acumulado inferior ao valor acumulado esperado.

(Resolução em aula)

Caso 2: Volte a resolver-se o Caso 1, admitindo agora que o prazo da aplicação é de 5 anos e que, se a pontuação for 5 ou 6, é aplicada a taxa de 5% ao valor acumulado no início do ano; se a pontuação for 2, 3 ou 4, é aplicada a taxa de 0%; e que, se a pontuação for 1, é aplicada a taxa de -2%.

O valor mínimo e o valor máximo são de cálculo imediato:

$$\min A_5 = 10000(0.98)^5 = 9039.208, \text{ com probabilidade } (1/6)^5 = 0.0001286.$$

$$\max A_5 = 10000(1.05)^5 = 12762.81563, \text{ com probabilidade } (1/3)^5 = 0.004115.$$

Quanto ao cálculo de

$$P[A_5 < E[A_5]] = P[A_5 < 10000(1 + E[I])^5] = P[A_5 < 10000(1 + 0,01(3))^5] = P[A_5 < 10684.683],$$

já é tentador aplicar o MMC, embora continue a tratar-se de um problema muito elementar, relativamente ao qual ainda é possível calcular a distribuição exacta de $A_5 = 10000(1 + I_1)(1 + I_2)(1 + I_3)(1 + I_4)(1 + I_5)$.

Definindo as v.a. $I_t, t = 1, \dots, 5$, que representam as taxas de juro a aplicar nos 5 anos (e que, pelas condições do enunciado são claramente i.i.d.), é imediato que as 5 têm função de distribuição idêntica à de uma v.a. I , tal que

$$F(i) = P[I \leq i] = \begin{cases} 0, & i < -2\% \\ 1/6, & -2\% \leq i < 0\% \\ 2/3, & 0\% \leq i < 5\% \\ 1, & i \geq 5\% \end{cases} .$$

Então, de acordo com o que se viu, a cada $u \in [0,1]$ gerado aleatoriamente vai fazer-se corresponder uma das três taxas possíveis, de acordo com a transformação:

$$0 \leq u \leq 1/6, \quad \Rightarrow i = -2\%$$

$$1/6 < u \leq 2/3, \quad \Rightarrow i = 0\% .$$

$$2/3 < u \leq 1 \quad \Rightarrow i = 5\%$$

Para cada trajetória do processo simulam-se as 5 taxas anuais, o que vai permitir obter uma observação de A_5 ; tratando-se de uma simulação de MC repete-se o processo de forma independente tantas vezes quantas as necessárias para, dada a *distribuição* das v.a. input $(I_t, t = 1, \dots, 5)$, se obter a *distribuição* empírica dessa v.a. output (A_5) .

Usando o Excel, tem-se

Comandos nas células A2, C2, E2, G2, I2: =RAND()

Comando em B2: =IF(A2<=1/6;-0,02;IF(A2<=2/3;0; 0,05)) - comandos análogos em D2, F2, H2, J2.

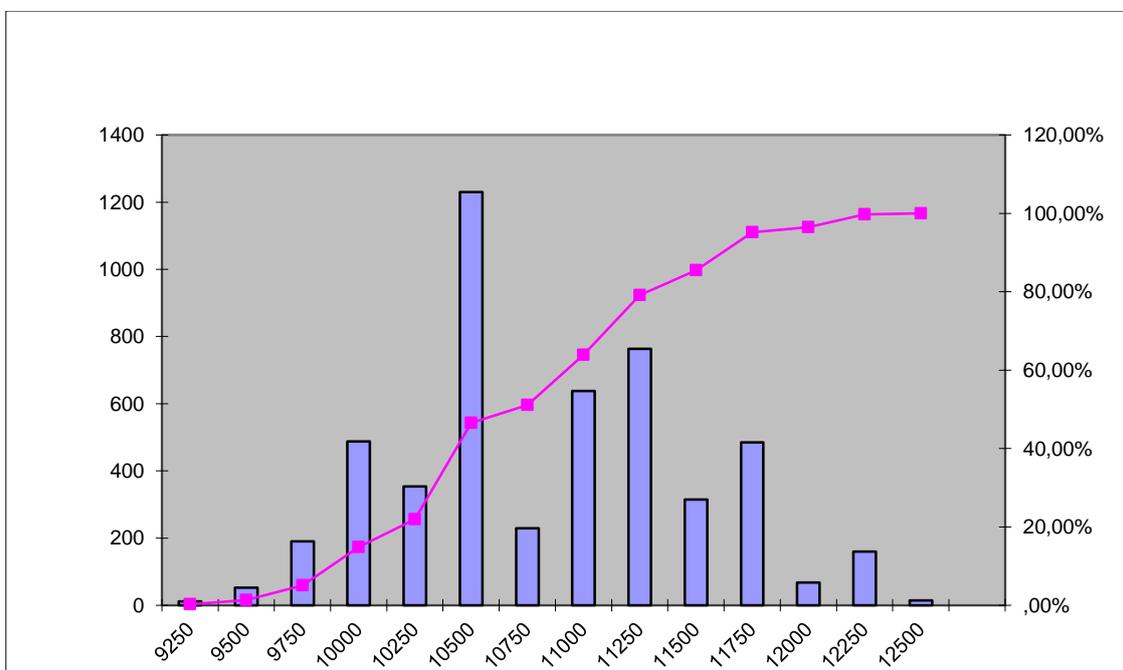
Comando em K2: =10000*(1+B2)*(1+D2)*(1+F2)*(1+H2)*(1+J2)

Aproveitando as funcionalidades do programa, é imediata a replicação de, por exemplo, 5000 trajetórias, gerando assim 5000 observações de A_5 .

	A	B	C	D	E	F	G	H	I	J	K
1		i_1		i_2		i_3		i_4		i_5	A_5
2	0,139307	- 0,02	0,423575	0	0,525444	0	0,205491	0	0,700417	0,05	10290,00000
3	0,609711	0	0,601024	0	0,338185	0	0,258833	0	0,651613	0	10000,00000
...											
5001	0,697201	0,05	0,554567	0	0,339087	0	0,006573	-0,02	0,509787	0	10290,00000

Recorrendo à *Data Analysis*, podem obter-se as estatísticas de sumário, o histograma e um esboço da distribuição empírica de A_5 , com estas particulares 5000 observações.

Mean	10692,50345	Skewness	0,312916964
Standard Error	9,05517534	Range	3539,134025
Median	10588,41	Minimum	9223,6816
Mode	11025	Maximum	12762,81563
Standard Deviation	640,2975888	Sum	53462517,23
Sample Variance	409981,0022	Count	5000
Kurtosis	-0,27668221	Confidence Level (95,0%)	17,7521152



Para calcular uma estimativa da probabilidade pedida, basta converter cada observação de A_5 na observação da correspondente Indicatriz de Bernoulli.

$A_{5_i} < 10684.683 \Rightarrow$ observa-se um 'sucesso'; $A_{5_i} \geq 10684.683 \Rightarrow$ observa-se um 'insucesso';
a estimativa procurada obtém-se dividindo o número de sucessos pelo número de observações (5000). Usando o Excel

Comando em L2: =IF(K2<10884,683;1;0)

	...	K	L
1		A_5	
2		10290,00000	1
3		10000,00000	1
...			
5001		10290,00000	1

Para calcular a estimativa da probabilidade: =sum(L2:L5001)/5000

Com a particular amostra simulada, a estimativa é 0,511 – o histograma já dava alguma indicação sobre o valor aproximado.

Questões que permanecem:

- 5000 será uma dimensão suficiente (note-se que o mínimo nunca chegou a ser observado)?
- Outra amostra com a mesma dimensão forneceria resultados significativamente diferentes destes?

É preciso repetir o processo até que as dúvidas sejam, tanto quanto possível, dissipadas.

Caso 3: Retome-se o Caso 1, admitindo que o dado foi posto de parte e que as melhores previsões indicam que as 3 taxas efectivas anuais vão ter distribuição uniforme no intervalo $[-2\%, 5\%]$. Para além de refazer os cálculos então pedidos, estime-se também a probabilidade de não haver agora qualquer valorização do capital aplicado.

(Resolução em aula)

Caso 4: Considere-se novamente o Caso 3.

4.1 Admita-se que as 3 taxas efectivas anuais terão distribuição exponencial com média igual à da distribuição uniforme Calcule-se agora a probabilidade de a valorização média anual do capital aplicado não exceder 1%.

(Resolução em aula)

Refira-se adicionalmente que o Excel gera observações de uma população com distribuição gama de parâmetros n e $1/\alpha$. (O valor esperado é o produto dos dois). O comando é

=GAMMAINV(rand();n;1/α)

4.2 Admita-se que as 3 taxas efectivas anuais terão distribuição normal com média e variância iguais à da distribuição uniforme. Calcule-se a probabilidade de a valorização média anual do capital aplicado pertencer ao intervalo $[\mu - \sigma, \mu + \sigma]$.

(Resolução em aula; recordar o Método de Box and Muller ou a aproximação dada no texto de apoio ou ver ainda as pp. 36-38 de Korn *et al.*)

Claro que o Excel também gera observações de uma população com distribuição Normal Standard, usando o comando

=NORMSINV(RAND())

E gera observações de uma população com distribuição Normal de média μ e desvio padrão σ , com o comando

=NORMINV(RAND());μ;σ)

4.3 Admita-se que a v.a. $1+i_t, t=1,2,3$, terá distribuição Lognormal com parâmetros $\mu = \sigma = 0.07$. Calcule-se a probabilidade de a valorização média anual do capital aplicado ser superior a $E[i_t]$.

(Resolução em aula)

O Excel também gera observações de uma população com distribuição Lognormal de parâmetros μ e σ (que agora já não correspondem à média nem ao desvio padrão da distribuição, como se sabe). Comando:

=LOGINV(RAND());μ;σ) .

5. Princípios Básicos do Método de Monte Carlo

5.1 Introdução

Para além da transformação uniformizante, de que já se falou, e tendo sempre presente que se está no domínio da amostragem, das distribuições por amostragem e da estimação, isto é, no domínio da Estatística, há alguns princípios básicos fundamentais para um correcto entendimento do MMC.

Como se consegue depreender das aplicações atrás vistas, o objetivo da aplicação do método é muitas vezes calcular as probabilidades de acontecimentos relevantes, o que se consegue calculando os valores esperados de variáveis aleatórias convenientemente definidas.

Este propósito é alcançado a partir do cálculo da média aritmética dos resultados obtidos com um grande número de repetições de experiências aleatórias, conduzidas de modo que em todas elas a distribuição dessas variáveis aleatórias esteja subjacente.

Para ilustrar, recorde-se o Caso 2 em que, para além do valor mínimo e do valor máximo, e respectivas probabilidades (que podem ser calculados sem recorrer ao MMC), se pedia a probabilidade de ser obtido um valor acumulado (v.a. A_5) inferior ao valor acumulado esperado, $E[A_5] = 10\,684,683$ o que já exigia a recolha de uma amostra por simulação de MC. Começou então por se considerar uma amostra com 5000 valores acumulados e a cada uma das 5000 v.a. A_{5i} , associou-se uma v.a. Indicatriz de Bernoulli, X_i ,

$$X_i = \begin{cases} 0, & \text{se } A_{5i} \geq E[A(5)] \\ 1, & \text{se } A_{5i} < E[A(5)] \end{cases}, i = 1, \dots, 5000,$$

em que a probabilidade de sucesso é

$$p = P[X_i = 1] = P[A_{5i} < E[A_5]].$$

Por outro lado, como se sabe,

$$p = E[X],$$

pelo que, para estimar a probabilidade p , é suficiente estimar o valor esperado $E[X]$.

Uma estimativa para p foi assim obtida calculando a média aritmética dos 5000 valores assumidos pelas indicatrizes de Bernoulli, $(x_1, x_2, \dots, x_{5000})$, face à particular amostra $(a_{51}, a_{52}, \dots, a_{55000})$ obtida. Por outras palavras: uma estimativa para p é a média de uma

amostra com 5000 observações extraídas da população $X \sim B(1, p)$, pois todas as variáveis são identicamente distribuídas.

5.2 Formulação Geral

Considere-se uma v.a. X definida num espaço de probabilidade (Ω, \mathcal{F}, P) e seja $(X_n)_{n \in \mathbb{N}}$ uma sucessão de v.a. i.i.d. a X . Seja $E[X] = \mu$ e $Var(X) = \sigma^2$ e admita-se que se quer estimar μ .

Resultados já conhecidos, que é conveniente recordar, para a resolução do problema de estimação em causa:

1. A lei forte dos grandes números (na sua forma mais simples, ou versão de Kolmogorov, que é suficiente no que diz respeito ao MMC), segundo a qual, considerando a medida de probabilidade P , se tem para quase todos os resultados $\omega \in \Omega$

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i(\omega) \xrightarrow{qc} \mu.$$

Por outras palavras, pode dizer-se que a sucessão das médias aritméticas das realizações de X_i converge quase certamente para a média da população X , que se quer estimar. \bar{X}_n é um estimador consistente de μ .

[Note-se que se diz que uma sucessão de v.a. $(X_n)_{n \in \mathbb{N}}$ converge quase certamente para c , $X_n \xrightarrow{qc} c$, quando $X_n - c \xrightarrow{qc} 0$; diz-se que $(X_n - c)_{n \in \mathbb{N}}$ converge quase certamente para zero quando

$$P(A) = 1, A = \{\omega \in \Omega = \Omega \times \Omega \times \Omega \times \dots, : \lim X_n(\omega) = 0\}.$$

2. A média da amostra (ou estimador de Monte Carlo) é um estimador não enviesado da média da população, ou seja, sendo (X_1, X_2, \dots, X_N) uma amostra aleatória recolhida da população X ,

$$E[\bar{X}_N] = E\left[\frac{1}{N} \sum_{i=1}^N X_i\right] = \mu.$$

Mais ainda, a variância do erro de estimação é $Var(\bar{X}_N - \mu) = Var(\bar{X}_N) = \frac{\sigma^2}{N}$.

Daqui se conclui que o estimador de MC goza de boas propriedades. Também se pode concluir que o desvio padrão de \bar{X}_N é de ordem $O(1/\sqrt{N})$, o que indica ser necessário multiplicar a dimensão da amostra em 100 observações para se reduzir o desvio padrão em 0.1. Observa-se assim que a convergência do método é bastante lenta.

3. O uso do desvio padrão do erro cometido como medida da precisão obtida com o MMC pode ser justificado invocando o TLC, pois σ caracteriza a dispersão dos valores da distribuição normal em torno da média.

Assim, nas condições de 2.,

$$\frac{\bar{X}_N - \mu}{\frac{\sigma}{\sqrt{N}}} \xrightarrow{D} \mathcal{N}(0,1) \Leftrightarrow \frac{\sum_{i=1}^N X_i - N\mu}{\sigma\sqrt{N}} \xrightarrow{D} \mathcal{N}(0,1).$$

4. O TLC garante que, para N suficientemente grande, $\bar{X}_N \xrightarrow{D} \mathcal{N}\left(\mu, \frac{\sigma^2}{N}\right)$; tal

conhecimento permite a obtenção de intervalos de confiança para μ . Considerando um nível de confiança aproximado $(1-\alpha)$, vem

$$\left[\frac{1}{N} \sum_{i=1}^N X_i - z_{1-\alpha/2} \frac{\sigma}{\sqrt{N}}, \frac{1}{N} \sum_{i=1}^N X_i + z_{1-\alpha/2} \frac{\sigma}{\sqrt{N}} \right] = \left[\bar{X}_N - z_{1-\alpha/2} \frac{\sigma}{\sqrt{N}}, \bar{X}_N + z_{1-\alpha/2} \frac{\sigma}{\sqrt{N}} \right],$$

onde $z_{1-\alpha/2}$ é o quantil de ordem $1-\alpha/2$ da distribuição Normal-Standard, o que é equivalente a dizer-se que verifica a condição $\Phi\left(z_{1-\alpha/2}\right) = 1-\alpha/2$.

Normalmente, também é necessário estimar o desvio padrão σ da população, para o que se usa como estimador o desvio padrão corrigido da amostra (igualmente um estimador centrado),

$$\bar{\sigma}_N = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X}_N)^2} = \sqrt{\frac{N}{N-1} \left(\frac{\sum_{i=1}^N X_i^2}{N} - \bar{X}_N^2 \right)}.$$

Quando se fixa $\alpha = 0.05$, ou seja, quando se pretende obter um intervalo de confiança a 95%, $z_{1-\alpha/2} = z_{0.975} = 1.96$; usualmente, dado que se trata de um cálculo que envolve aproximações, considera-se então

$$\left[\frac{1}{N} \sum_{i=1}^N X_i - 2 \frac{\bar{\sigma}_N}{\sqrt{N}}, \frac{1}{N} \sum_{i=1}^N X_i + 2 \frac{\bar{\sigma}_N}{\sqrt{N}} \right] = \left[\bar{X}_N - 2 \frac{\bar{\sigma}_N}{\sqrt{N}}, \bar{X}_N + 2 \frac{\bar{\sigma}_N}{\sqrt{N}} \right].$$

[Com $\alpha = 0.1$ - intervalo de confiança a 90% - $z_{1-\alpha/2} = z_{0.95} = 1.645$; com $\alpha = 0.01$ - intervalo de confiança a 99% - $z_{1-\alpha/2} = z_{0.995} = 2.576, \dots$]

OBSERVAÇÃO IMPORTANTE:

Todas as considerações anteriores se podem aplicar, com as necessárias adaptações, ao caso em que é necessário estimar uma quantidade $\theta = E[g(\mathbf{X})]$, sendo $\mathbf{X} = \{X_1, X_2, \dots, X_N\}$ um vetor aleatório em \mathbf{R}^N , $g(\cdot)$ uma função de \mathbf{R}^N em \mathbf{R} e $E[g(\mathbf{X})] < +\infty$. Note-se que \mathbf{X} pode representar uma trajectória de um processo estocástico, por exemplo, o preço de um dado ativo ao longo do tempo; nesse caso, X_i representaria o preço do ativo no momento i .

Atrás, como o objetivo era estimar μ , a média da população, tomou-se

$$g(\mathbf{X}) = \frac{X_1 + X_2 + \dots + X_N}{N},$$

a média dos valores observados da v.a., cujo valor esperado

é exatamente μ . De modo semelhante, quando o objetivo é calcular a probabilidade de se realizar um determinado acontecimento A , expresso em função da v.a. X_i (por vezes, o acontecimento A vem expresso ainda em função dos acontecimentos elementares $\omega \in \Omega$), a questão é resolvida definindo as v.a. Indicatrizes de Bernoulli

$$I_A(X_i) = \begin{cases} 1, & \text{se } X_i \in A \\ 0, & \text{se } X_i \notin A \end{cases},$$

pois da teoria das probabilidades sabe-se que $E[I_A] = P(A)$.

Mais uma vez, e pelas mesmas razões, também agora a estimativa de MC para $P(A)$ se determina calculando a média da amostra $\mathbf{I} = \{I_A(X_1), I_A(X_2), \dots, I_A(X_N)\}$, relativa às observações de $I_A(X_i)$, as quais são obtidas a partir da amostra $\mathbf{X} = \{X_1, X_2, \dots, X_N\}$.

Isto significa que $g(\mathbf{X})$ é igual à proporção de sucessos na amostra das N observações independentes. Como esta proporção não é mais do que a frequência relativa das realizações de A na amostra, vem

$$g(\mathbf{X}) = rf_N(A) = \frac{I_A(X_1) + I_A(X_2) + \dots + I_A(X_N)}{N}.$$

Sabendo-se igualmente ser $Var[I_A] = P(A)(1 - P(A))$, pode tomar-se como estimador da variância $\hat{\sigma}^2 = rf_N(A)(1 - rf_N(A))$. O intervalo de confiança para $P(A)$ com um nível aproximado de 95% é

$$\left[rf_N(A) - 2\sqrt{\frac{rf_N(A)(1 - rf_N(A))}{N}}, rf_N(A) + 2\sqrt{\frac{rf_N(A)(1 - rf_N(A))}{N}} \right],$$

caso particular para as populações de Bernoulli do intervalo anteriormente apresentado

$$\left[\bar{X}_N - 2\frac{\bar{\sigma}_N}{\sqrt{N}}, \bar{X}_N + 2\frac{\bar{\sigma}_N}{\sqrt{N}} \right], \text{ pois } rf_N(A) = \bar{X}_N.$$

Exemplo com os resultados da simulação do preço da *call* europeia:

Intervalo de confiança a 95% para o valor da opção:

$$\left[1,16 - 1,96 \frac{0,9041}{\sqrt{2500}}, 1,16 + 1,96 \frac{0,9041}{\sqrt{2500}} \right] = [1,142, 1,178].$$

Intervalo de confiança a 95% para a probabilidade de se exercer a opção:

$$\left[0,8804 - 1,96 \sqrt{\frac{0,8804(1 - 0,8804)}{2500}}, 0,8804 + 1,96 \sqrt{\frac{0,8804(1 - 0,8804)}{2500}} \right]$$

$$= [0,8676, 0,8931].$$

6. Séries Pseudoaleatórias

6.1 Números Aleatórios

Como se pôde observar nos exemplos estudados, e resulta da própria designação, as simulações estocásticas fazem amplo uso de variáveis aleatórias, pelo que a capacidade de dispor de números aleatórios (NA) com uma particular distribuição se torna indispensável. O grande problema que se coloca é como obter números que sejam realmente aleatórios e imprevisíveis.

Os NA usados nas simulações de Monte Carlo são habitualmente gerados por meio de um algoritmo numérico, o que na realidade os torna valores obtidos de uma forma determinística. No entanto, quando se olha para eles sem se conhecer o algoritmo que os gerou, afiguram-se mesmo como sendo aleatórios. Por essa razão se lhes chama números pseudoaleatórios e conseguem passar a maior parte dos testes estatísticos.

Pelas razões já conhecidas, é particularmente importante a geração de NA com distribuição uniforme no intervalo $[0,1]$, embora a inclusão dos extremos, que formam um conjunto de probabilidade nula, possa ser dispensável.

Existem atualmente muitos métodos para a geração de NA, pelo que só serão apresentados os princípios gerais necessários para a compreensão do modo como os geradores funcionam. Em particular, será introduzido um algoritmo que pode ser implementado com uma simples calculadora e que tem a virtude de tornar familiares os termos técnicos usados.

Antes disso, alguns critérios de qualidade que os geradores de NA devem satisfazer:

1. Requisitos de rapidez e memória

Nas simulações de MC são necessários enormes volumes de dados NA, pelo que é conveniente que estes sejam gerados de forma rápida e eficiente, sem ocuparem demasiada memória.

2. Período suficientemente longo

Devido ao modo como são construídos, e às capacidades dos computadores para representarem números reais, os algoritmos que geram NA funcionam com um conjunto finito de dados. Daí resulta que a sequência de NA gerados acabará por se repetir. O comprimento máximo de números gerados antes da repetição é o chamado período do gerador de NA. O período deve assim ser suficientemente longo de forma a garantir que não se use a mesma sequência mais de uma vez numa experiência de simulação. Uma

regra ditada pela prática propõe que o período não seja inferior ao quadrado da dimensão da amostra, para evitar a presença de correlações.

3. Passar os testes estatísticos à uniformidade e à independência

Os NA devem passar os testes estatísticos da independência e da idêntica distribuição.

4. Possibilidade de replicação

Uma mesma sequência de NA deve poder ser reproduzida tantas vezes quantas as necessárias, de modo a possibilitar a detecção de erros ou a realização de análises de sensibilidade.

Todos os geradores de NA que se baseiam em fórmulas de recorrência podem ser descritos como o quintuplo $(X, \mu, f, \mathcal{U}, g)$, onde:

- X é o conjunto finito dos estados.
- μ é a medida de probabilidade para se selecionar x_0 , o estado inicial do processo, ou semente do gerador de NA.
- f é a função de transição, que descreve o algoritmo, $x_{i+1} = f(x_i)$.
- \mathcal{U} é o espaço dos outputs: normalmente, $[0,1]$, $[0,1[$, $]0,1]$ ou $]0,1[$.
- g é a função de output, que transforma $x_i \in X$ em $u_i \in \mathcal{U}$.

6.2 Exemplo de um gerador de NPA: o Método das Congruências

Dados dois números inteiros positivos x e m , a divisão inteira de x por m tem como resultado dois números: o quociente q e o resto $r, r < x$. Por outras palavras, o número inteiro positivo x admite a decomposição $x = m \times q + r$. Se $r = 0$, diz-se que x é divisível por m .

Definição: Seja $m \neq 0$ um número natural. Dois inteiros positivos x e y dizem-se *congruentes módulo m* se tiverem o mesmo resto na divisão por m . A notação é $x \equiv y \pmod{m}$.

Considerando o algoritmo da divisão, verifica-se que qualquer inteiro positivo é congruente ao seu resto na divisão por um inteiro positivo. Uma vez que para cada m fixado *a priori* o resto obtido é único, pode definir-se uma função que a cada x associa o resto da divisão de x por m . A imagem de x dada por esta função é denotada por $x \pmod{m}$.

Posto isto, considere-se a sucessão de números inteiros definida por recorrência $x_{i+1} = (ax_i + c)(\text{mod } m)$, isto é, x_{i+1} é igual ao resto da divisão inteira de $ax_i + c$ por m . A sucessão começa com um valor inteiro x_0 (a semente, como já se referiu).

A série normalizada vem $u_i = \frac{x_i}{m-1}$. Assim, para este gerador:

- O conjunto finito dos estados é $X = \{0, 1, \dots, m-1\}$.
- A função de transição que descreve o algoritmo é $f(x_i) = (ax_i + c)(\text{mod } m)$.
- O espaço dos outputs é $U = [0, 1]$.
- A função de output, que transforma $x_i \in X$ em $u_i \in \mathcal{U}$, é $g(x_i) = \frac{x_i}{m-1}$.
- Quanto a μ , a medida de probabilidade para se seleccionar x_0 , o estado inicial do processo, ou semente do gerador de NA, varia consoante os casos concretos.

Como é imediato, o modo de construção da sequência faz com que não possam ser gerados mais de m NA. O comprimento do período depende dos valores de a e de c .

É obtido um período máximo quando são satisfeitas as condições seguintes:

- c e m são primos entre si.
- Todo o número primo que é divisor de m também é divisor de $a-1$.
- Se m é divisível por 4, $a-1$ também é divisível por 4.

Estas condições podem ser satisfeitas sob dois tipos de hipóteses:

- Se $c > 0$, m deve ser uma potência de 2, c deve ser ímpar e a deve ser da forma $a = 4k + 1$, k inteiro.
- Se $c = 0$, deve ter-se $x_0 \neq 0$, a^{m-1} deve ser múltiplo de m e $a^j - 1$ não pode ser múltiplo de m , para $j = 1, 2, \dots, m-2$.

Em simulações muito exigentes os geradores congruenciais lineares não devem ser usados, pois apresentam alguns problemas. Por exemplo, se o multiplicador a é pequeno quando comparado com m , verifica-se que um NA muito pequeno é sempre seguido por outro, o que obriga acontecimentos raros a sucederem-se. O período é também insuficiente, em muitos casos. De qualquer modo, são algoritmos rápidos, pouco exigentes em memória e fáceis de implementar e de compreender, vantagens apreciáveis.

Exemplo:

Usando o método das congruências, gere uma série de números pseudo aleatórios usando a seguinte fórmula de recorrência $X_{n+1} = (5X_n + 3)(\text{mod } 8)$ e $X_0 = 2$.

(Note-se que os valores escolhidos para a , c e m garantem que se vai ter período máximo.)

$$X_0 = 2$$

$$X_1 = (5 \times 2 + 3)(\text{mod } 8) = (13)(\text{mod } 8) = 5 \Rightarrow u_1 = \frac{5}{7}$$

$$X_2 = (5 \times 5 + 3)(\text{mod } 8) = (28)(\text{mod } 8) = 4 \Rightarrow u_1 = \frac{4}{7}$$

$$X_3 = (5 \times 4 + 3)(\text{mod } 8) = (23)(\text{mod } 8) = 7 \Rightarrow u_1 = 1$$

$$X_4 = (5 \times 7 + 3)(\text{mod } 8) = (38)(\text{mod } 8) = 6 \Rightarrow u_1 = \frac{6}{7}$$

$$X_5 = (5 \times 6 + 3)(\text{mod } 8) = (33)(\text{mod } 8) = 1 \Rightarrow u_1 = \frac{1}{7}$$

$$X_6 = (5 \times 1 + 3)(\text{mod } 8) = (8)(\text{mod } 8) = 0 \Rightarrow u_1 = 0$$

$$X_7 = (5 \times 0 + 3)(\text{mod } 8) = (3)(\text{mod } 8) = 3 \Rightarrow u_1 = \frac{3}{7}$$

$$X_7 = (5 \times 3 + 3)(\text{mod } 8) = (18)(\text{mod } 8) = 2 \Rightarrow u_1 = \frac{2}{7}$$

6.3 Testes estatísticos

Como já foi dito, os números u_1, u_2, \dots, u_n devem ser observações de v.a. i.i.d. uniformemente distribuídas no intervalo $[0,1]$. Para assegurar que estas propriedades foram alcançadas, são feitos alguns testes. Como é evidente, a maioria dos softwares comerciais de simulação realiza estes testes previamente à comercialização.

Pelo facto de um gerador passar alguns testes, não se pode concluir que é adequado a todo o tipo de simulações. Na verdade, uma vez que os NA na realidade não são verdadeiramente aleatórios, há sempre testes que não conseguem passar. Uma coisa é certa: há alguns testes que nenhum gerador pode falhar, nomeadamente o teste de Kolmogorov-Smirnov ou o teste do Qui.-Quadrado (ver Larson, pp. 527-537), que testam a uniformidade, e o teste das sequências (*run test*, ver Larson pp 518-523), que testa a independência.

Tanto o teste do Qui-Quadrado como o teste de Kolmogorov-Smirnov testam a uniformidade, medindo o grau de aproximação entre a distribuição da amostra fornecida pelo gerador de números aleatórios e a distribuição uniforme teórica. Ambos são concebidos de modo que a hipótese da uniformidade não seja rejeitada quando não existe

uma diferença significativa entre a distribuição da amostra e a distribuição teórica. No ponto seguinte descreve-se de forma muito intuitiva o teste de Kolmogorov-Smirnov, de aplicação mais conveniente neste caso.

Com efeito, podem ser apontadas duas vantagens deste teste em relação ao teste qui-quadrado.

Em primeiro lugar, quando a distribuição populacional é contínua e se conhecem a forma e os parâmetros da sua função densidade de probabilidade, a distribuição da estatística do teste é definida rigorosamente (ao contrário do que sucede com a estatística do teste do qui-quadrado, cuja distribuição é aproximada). Esta vantagem é tanto mais nítida quanto menor for a dimensão da amostra.

Em segundo lugar, o teste K-S é, na maioria das situações, mais potente do que o teste qui-quadrado.

Limitações: exige distribuições populacionais contínuas e completamente especificadas (o que não sucede com o teste do qui-quadrado), bem como um maior esforço computacional.

6.3.1 Teste de Kolmogorov-Smirnov (K-S)

Compara a função de distribuição uniforme contínua (o teste exige distribuições populacionais contínuas e completamente especificadas), $F(u) = u, 0 \leq u \leq 1$, com a distribuição empírica da amostra dada pelo gerador de números aleatórios, u_1, u_2, \dots, u_N , seja $\hat{F}(u) = \frac{\text{número de observações } u_i \leq u}{N}, 0 \leq u \leq 1$. O desvio máximo aceitável é calculado a partir da estatística $D = \max\{|F(u) - \hat{F}(u)|, 0 \leq u \leq 1\}$.

A hipótese nula e a alternativa são formuladas nos seguintes termos:

$$H_0: F(u) = F_0(u) = u, 0 \leq u \leq 1,$$

ou seja, a função de distribuição da população da qual provém a amostra é idêntica a uma função de distribuição que se assume conhecida (e é neste caso a distribuição uniforme em $[0,1]$, $F_0(u) = u, 0 \leq u \leq 1$)

versus

$$H_0: F(u) \neq F_0(u) = u.$$

Note-se que o $\max\{|F(u) - \hat{F}(u)|\}$ procurado não é necessariamente o maior valor que $|F(u) - \hat{F}(u)|$ toma quando se consideram apenas os valores *observados* de X . Dado que

a função $F_0(u)$ é contínua e $\hat{F}(u)$ é uma função em escada, o valor máximo daquela diferença absoluta deve ser procurado na vizinhança de cada valor observado de U .

A distribuição da estatística D é conhecida e está tabelada em função de N e do nível de significância do teste, α . Diz-se que D é uma estatística “*distribution-free*”, no sentido em que só depende da dimensão da amostra, sendo irrelevante a forma da função de distribuição da população.

Para realizar o teste é então necessário determinar o maior desvio absoluto entre $F(u)$ e $\hat{F}(u)$, o que se faz com o seguinte algoritmo:

1. Ordenar de forma crescente os valores da amostra: $u_{(1)} \leq u_{(2)} \leq \dots \leq u_{(N)}$.
2. Calcular $D_{obs}^+ = \max \left\{ \left| \frac{i}{N} - u_{(i)} \right|, 1 \leq i \leq N \right\}$.
3. Calcular $D_{obs}^- = \max \left\{ \left| \frac{i-1}{N} - u_{(i)} \right|, 1 \leq i \leq N \right\}$.
4. Calcular $D_{obs} = \max \{ D_{obs}^+, D_{obs}^- \}$. (De facto, dado que a função $F(u)$ é contínua e $\hat{F}(u)$ é uma função em escada, o valor máximo daquela diferença absoluta deve ser procurado na vizinhança de cada valor observado de U .)
5. Consultando a tabela da distribuição da estatística D , , determinar o valor crítico $D_\alpha(N)$, por nível de significância e dimensão da amostra, e decidir:
 Se $D_{obs} > D_\alpha(N)$, rejeitar a hipótese da uniformidade;
 Se $D_{obs} \leq D_\alpha(N)$, não rejeitar a hipótese da uniformidade.

Exemplo: Verificar se os números 0,154229; 0,045809; 0,394661; 0,714665; 0,959579; 0,372405; 0,928626; 0,855914 (gerados pelo GNA do Excel) não contrariam a hipótese da uniformidade.

$$N = 8$$

$$u_{(1)} \leq u_{(2)} \leq \dots \leq u_{(8)}$$

$$\Rightarrow 0,045809 < 0,154229 < 0,372405 < 0,394661 < 0,714665 < 0,855914 < 0,928626 < 0,959579$$

i	1	2	3	4	5	6	7	8
$u_{(i)}$	0,045809	0,154229	0,372405	0,394661	0,714665	0,855914	0,928626	0,959579
$\frac{i}{8}$	0,125	0,25	0,375	0,5	0,625	0,75	0,875	1
$\left \frac{i}{8} - u_{(i)} \right $	0,079191	0,095771	0,002595	0,105339	0,089665	0,105914	0,053626	0,040421
$\frac{i-1}{8}$	0	0,125	0,25	0,375	0,5	0,625	0,75	0,875
$\left \frac{i-1}{8} - u_{(i)} \right $	0,045809	0,029229	0,122405	0,019661	0,214665	0,230914	0,178626	0,084579

$$D_{obs} = \max \{0,105339; 0,230914\} = 0,230914$$

$D_{0,05}(8) = 0,454$, consultando a tabela da estatística D , com $N = 8$ e $\alpha = 0,05$.

$D_{obs} < D_{0,05}(8) \Rightarrow$ a hipótese da uniformidade não é rejeitada.

6.3.2 Testes de Sequência

Os testes de sequência examinam o arranjo dos números para testar a sua independência. À partida, um conjunto de números aleatórios não deverá ter sequências de mais, nem de menos.

Os testes verificam dois tipos de sequências nas observações realizadas:

1. A variação crescente e decrescente: Uma sequência é definida como sendo uma sucessão crescente ou decrescente de observações.
2. A variação acima e abaixo da média: Uma sequência é definida como sendo uma sucessão de observações (todas) acima, ou abaixo, da média.

6.3.2.1 A variação crescente e decrescente

Considerando a va. X que representa o número de sequências num conjunto de N observações aleatórias, pode deduzir-se que a sua média e variância são respetivamente iguais a

$$\mu_X = (2N - 1)/3$$

$$\sigma_X^2 = (16N - 29)/90$$

Prova-se também que, com N suficientemente grande ($N > 20$),

$$Z_0 = \frac{X - \mu_X}{\sigma_X} \sim N(0,1).$$

Se o gerador for aceitável, então o valor observado da v.a. Z_0 deverá estar próximo de 0, pelo que só se deve rejeitar a hipótese de independência quando isso não acontece de uma forma evidente.

Regra do teste:

Se o valor observado da v.a.

$$Z_0 = \frac{X - \mu_X}{\sigma_X}$$

satisfaz a condição $-z_{\alpha/2} < z_0 < z_{\alpha/2}$, não se rejeita a hipótese de independência.

Exemplo:

Baseado nas sequências de valores crescentes e decrescentes, determinar se a hipótese de independência pode ser rejeitada para $\alpha = 0.05$, tomando o seguinte conjunto de observações de uma v.a. $U(0,1)$.

0.41	0.68	0.89	0.94	0.74	0.91	0.55	0.62	0.36	0.27
0.19	0.72	0.75	0.08	0.54	0.02	0.01	0.36	0.16	0.28
0.18	0.01	0.95	0.69	0.18	0.47	0.23	0.32	0.82	0.53
0.31	0.42	0.73	0.04	0.83	0.45	0.13	0.57	0.63	0.29

As sequências são as seguintes:

+++ - + - + - - - + + - + - - + - + - - - + - - - + - + + - - + + - - - + + -

Há 26 sequências.

Com $N = 40$ e $x = 26$ vem:

$$\mu_X = (2 \times 40 - 1)/3 = 26.33$$

$$\sigma_X^2 = (16 \times 40 - 29)/90 = 6.79$$

$$Z_0 = \frac{(26 - 26.33)}{\sqrt{6.79}} = -0.13$$

Como $|-0.13| < 1.96$, a hipótese de independência não pode ser rejeitada.

6.3.2.2 A variação acima e abaixo da média

Teste muito semelhante ao anterior. Considerando a v.a. Y que representa o número de sequências, agora de valores acima e abaixo da média, e ainda num conjunto de N observações aleatórias, toma-se agora

$$\mu_Y = \frac{2n_1n_2}{N} + \frac{1}{2}$$

$$\sigma_Y^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N - 1)}$$

onde:

n_1 = número de observações acima da média;

n_2 = número de observações abaixo da média;

$N = n_1 + n_2$.

Continua a ser verdade que, com N suficientemente grande ($N > 20$),

$$Z_0 = \frac{Y - \mu_Y}{\sigma_Y} \sim N(0,1).$$

Regra do teste:

Se o valor observado da v.a.

$$Z_0 = \frac{Y - \mu_Y}{\sigma_Y}$$

satisfaz a condição $-z_{\alpha/2} < z_0 < z_{\alpha/2}$, não se rejeita a hipótese de independência.

Exemplo

Tomando as mesmas observações...

0.41	0.68	0.89	0.94	0.74	0.91	0.55	0.62	0.36	0.27
0.19	0.72	0.75	0.08	0.54	0.02	0.01	0.36	0.16	0.28
0.18	0.01	0.95	0.69	0.18	0.47	0.23	0.32	0.82	0.53
0.31	0.42	0.73	0.04	0.83	0.45	0.13	0.57	0.63	0.29

... As sequências são agora as seguintes:

- + + + + + + + - - - + + - + - - - - - + + - - - + + - - - + + - - - + + -

Observa-se:

$n_1 = 18$ (observações acima da média +);

$n_2 = 22$ (observações abaixo da média -);

$N = n_1 + n_2 = 40$;

$y = 17$ sequências

$$\mu_Y = \frac{2n_1n_2}{N} + \frac{1}{2} = 20.3$$

$$\sigma_Y^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N - 1)} = 9.55$$

$|z_0| = 1.07$ e o valor crítico é 1.96; logo, a hipótese de independência não pode ser rejeitada.

7. Geração de Observações de Variáveis Aleatórias Correlacionadas

7.1 Introdução

Em muitas situações, não é possível aceitar que as v.a. de input são todas mutuamente independentes; por exemplo, os retornos de investimentos em diferentes ativos estão muitas vezes correlacionados, ou seja, existe dependência estatística entre as variáveis que os representam.

Só para recordar, sendo X_1 e X_2 v.a., a covariância de X_1 e X_2 define-se

$$Cov(X_1, X_2) = E[(X_1 - \mu_{X_1})(X_2 - \mu_{X_2})] = E[X_1 X_2] - E[X_1]E[X_2].$$

Quanto ao coeficiente de correlação entre X_1 e X_2 , que se demonstra pertencer ao intervalo $[-1, 1]$, define-se

$$\rho_{X_1, X_2} = \frac{Cov(X_1, X_2)}{\sqrt{Var[X_1]Var[X_2]}} \Leftrightarrow Cov(X_1, X_2) = \rho_{X_1, X_2} \sqrt{Var[X_1]Var[X_2]}.$$

Como também se sabe, se X_1 e X_2 são independentes, então $Cov(X_1, X_2)$ e ρ_{X_1, X_2} são nulos; a implicação recíproca não é em geral, verdadeira. [Uma exceção é observada no caso da distribuição normal, em que se pode realmente concluir que variáveis não correlacionadas são independentes.]

Quando se tem um vetor aleatório n -dimensional $\mathbf{X} = (X_1, X_2, \dots, X_n)$, a matriz das variâncias e covariâncias é uma matriz Σ de ordem n , simétrica e semi definida positiva, com elementos não negativos na diagonal principal (as variâncias; as covariâncias estão nas outras posições).

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{bmatrix} = \begin{bmatrix} Var(X_1) & Cov(X_1, X_2) & \dots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Var(X_2) & \dots & Cov(X_2, X_n) \\ \dots & \dots & \dots & \dots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \dots & Var(X_n) \end{bmatrix}.$$

7.2 Populações normais multivariadas

Para começar, considere-se então um vetor $\mathbf{X} = (X_1, X_2, \dots, X_n) \sim N(\mathbf{0}, \Sigma)$.

A questão que se coloca é:

Será possível encontrar uma matriz \mathbf{C} , do tipo $n \times m$, tal que $\mathbf{C}^T \mathbf{Z} \sim \mathbf{N}(\mathbf{0}, \mathbf{\Sigma})$, $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)^T$, $Z_i \sim N(0,1)$ e independentes, $i = 1, \dots, n$?

Quer dizer: pergunta-se se é possível escrever as variáveis dadas, que não são mutuamente independentes, como combinações lineares de variáveis normal-standard que sejam mutuamente independentes?

A resposta é imediata e afirmativa: notando que v.a. resultantes de combinações lineares de v.a. com distribuição normal ainda têm distribuição normal, e que a matriz das variâncias e covariâncias de $\mathbf{C}^T \mathbf{Z}$ é $\mathbf{C}^T \mathbf{C}$, como se prova com toda a facilidade, basta garantir que $\mathbf{C}^T \mathbf{C} = \mathbf{\Sigma}$.

Verifica-se assim que o cálculo da matriz \mathbf{C} implica que se proceda à chamada decomposição de Cholesky da matriz $\mathbf{\Sigma}$ - a decomposição de Cholesky iguala uma matriz \mathbf{A} ao produto $\mathbf{A} = \mathbf{L}^T \mathbf{L}$, sendo \mathbf{L} uma matriz triangular superior (os elementos abaixo da diagonal principal são zeros).

Quando $n = 2$, (distribuição normal bidimensional, ou bivariada, situação muito frequente), tem-se $\mathbf{X} = (X, Y)$, $\boldsymbol{\mu} = [\mu_X, \mu_Y]^T$ e $\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$, que neste caso

particular costuma assumir a forma $\mathbf{\Sigma} = \begin{bmatrix} \sigma_X^2 & \rho \sigma_X \sigma_Y \\ \rho \sigma_X \sigma_Y & \sigma_Y^2 \end{bmatrix}$.

Neste caso elementar, para determinar a matriz \mathbf{C} basta resolver o sistema matricial $\mathbf{C}^T \mathbf{C} = \mathbf{\Sigma}$:

$$\begin{bmatrix} c_{11} & 0 \\ c_{12} & c_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ 0 & c_{22} \end{bmatrix} = \begin{bmatrix} \sigma_X^2 & \rho \sigma_X \sigma_Y \\ \rho \sigma_X \sigma_Y & \sigma_Y^2 \end{bmatrix} \Leftrightarrow \begin{cases} c_{11}^2 & = & \sigma_X^2 \\ c_{11} c_{12} & = & \rho \sigma_X \sigma_Y \\ c_{12} c_{11} & = & \rho \sigma_X \sigma_Y \\ c_{12}^2 + c_{22}^2 & = & \sigma_Y^2 \end{cases} \Leftrightarrow \begin{cases} c_{11} & = & \sigma_X \\ c_{12} & = & \rho \sigma_Y \\ c_{22} & = & \sigma_Y \sqrt{1 - \rho^2} \end{cases}$$

$$\text{Então, } \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ 0 & c_{22} \end{bmatrix} = \begin{bmatrix} \sigma_X & \rho \sigma_Y \\ 0 & \sigma_Y \sqrt{1 - \rho^2} \end{bmatrix} \text{ e } \mathbf{C}^T = \begin{bmatrix} \sigma_X & 0 \\ \rho \sigma_Y & \sigma_Y \sqrt{1 - \rho^2} \end{bmatrix}.$$

Tendo em atenção que as v.a. X e Y têm vetor das médias $\boldsymbol{\mu} = [\mu_X, \mu_Y]^T$, não necessariamente nulas, vem finalmente

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \mathbf{C}^T \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix} + \begin{bmatrix} \sigma_X Z_1 \\ \rho\sigma_Y Z_1 + \sigma_Y \sqrt{1-\rho^2} Z_2 \end{bmatrix} = \begin{bmatrix} \mu_X + \sigma_X Z_1 \\ \mu_Y + \rho\sigma_Y Z_1 + \sigma_Y \sqrt{1-\rho^2} Z_2 \end{bmatrix}.$$

Se, por exemplo, $\mathbf{X} \sim N(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$ é tal que $\mathbf{X} \sim N(5, 17, 49, 144, 0.4)$, tem-se

$$\begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix} = \begin{bmatrix} 5 \\ 17 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 49 & 33,6 \\ 33,6 & 144 \end{bmatrix} \text{ e } \mathbf{C} = \begin{bmatrix} 7 & 4,8 \\ 0 & 1,2\sqrt{84} \end{bmatrix}.$$

Pode escrever-se a combinação

$$\mathbf{X} = \begin{bmatrix} 5 + 7Z_1 \\ 17 + 4,8Z_1 + 1,2\sqrt{84}Z_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 5 \\ 17 \end{bmatrix}, \begin{bmatrix} 49 & 33,6 \\ 33,6 & 144 \end{bmatrix}\right), \quad Z_1, Z_2 \sim N(0,1) \text{ e independentes.}$$

O problema da geração de observações de $\mathbf{X} = (X, Y)$ reduz-se deste modo à geração de observações de $\mathbf{Z} = (Z_1, Z_2)$ e à sua transformação de acordo com os resultados obtidos: para gerar observações de X calcula-se $5 + 7Z_1$ e para gerar observações de Y calcula-se $17 + 4,8Z_1 + 1,2\sqrt{84}Z_2$.

Se $n > 2$ (distribuição normal multidimensional), o procedimento não se altera.

Em resumo: A decomposição de Cholesky é usada no MMC para a simulação de sistemas com múltiplas variáveis correlacionadas. A matriz das variâncias e covariâncias é decomposta de modo a obter-se a matriz triangular inferior \mathbf{C}^T , que aplicada à amostra casual $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)^T$ permite obter uma amostra $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ com a distribuição em causa.

7.3 Caso Geral

Em geral, há duas abordagens distintas para a geração de v.a. correlacionadas, de acordo com as seguintes duas situações:

1. A distribuição conjunta está completamente especificada.
2. Apenas são conhecidas as distribuições marginais e a matriz das variâncias e covariâncias.

1. Distribuição conjunta completamente especificada

Quando se pretende gerar observações de um vetor aleatório $\mathbf{X} = (X_1, X_2, \dots, X_n)$, com dada distribuição conjunta $F(x_1, x_2, \dots, x_n) = P[X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n]$, conhecida,

é possível obter as distribuições marginais e **também as distribuições condicionadas**, e assim gerar as observações de forma sequencial.

Por exemplo, com $n = 2$, $F(x_1, x_2) = F(x_1)F(x_2|x_1)$, pelo que para gerar $\mathbf{X} = (X_1, X_2)$ se gera primeiro X_1 com base em $F(x_1)$ e se gera depois X_2 de forma independente, a partir da distribuição condicionada $F(x_2|x_1)$ - admitindo, claro, que se conseguem efectuar os cálculos e as simulações envolvidos no problema.

Uma situação muito comum em que o método realmente funciona bem é na simulação de processos estocásticos, questão que será abordada mais adiante.

2. Distribuição conjunta desconhecida; apenas se conhecem as distribuições marginais $F_{x_i}(\cdot), i = 1, \dots, n$, e a matriz \mathbf{R} das correlações.

Quando se pretende gerar observações do vetor aleatório $\mathbf{X} = (X_1, X_2, \dots, X_n)$ e apenas se conhecem as distribuições marginais $F_{x_i}(\cdot), i = 1, \dots, n$, e a matriz \mathbf{R} das correlações, pode acontecer que haja mais do que uma distribuição conjunta a partilhar as distribuições marginais e a matriz das correlações que são dadas. Por outro lado, pode ter havido algum erro na especificação e existir inconsistência entre as distribuições marginais $F_{x_i}(\cdot)$ e a matriz \mathbf{R} , caso em que, afinal, não existe nenhuma distribuição conjunta nas condições exigidas.

Admitindo que não há problemas desta natureza, a geração de $\mathbf{X} = (X_1, X_2, \dots, X_n)$ pode fazer-se com base na distribuição normal multivariada (cai-se na secção anterior), dando os seguintes passos:

- Considere-se um vetor $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n) \sim N(\mathbf{0}, \mathbf{\Sigma})$, tal que

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & 1 & \dots & \sigma_{2n} \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & 1 \end{bmatrix}, \quad \text{ou seja,} \quad Z_i \sim N(0,1), \quad i = 1, \dots, n, \Leftrightarrow \text{f.d. de}$$

Z_i é $\Phi(z_i), i = 1, \dots, n$.

Os elementos não principais são desconhecidos e a questão principal é mesmo começar por determiná-los, para depois se poder avançar com o processo de simulação. Note-se que, sendo as variâncias unitárias, $\mathbf{\Sigma}$ acaba por ser também a matriz das correlações.

- Pelo que se viu anteriormente, cada uma das componentes $\Phi(Z_i)$, $i=1,\dots,n$, do vetor aleatório $(\Phi(Z_1), \Phi(Z_2), \dots, \Phi(Z_n))$ tem distribuição $\mathcal{U}(0,1)$.
- Uma vez que as v.a. Z_i são correlacionadas, as v.a. $\Phi(Z_i)$ são igualmente correlacionadas, represente-se a respetiva matriz das correlações por Σ' . Evidentemente, esta matriz também não é conhecida, pois os seus elementos vão depender dos valores σ_{ij} , desconhecidos.
- Fazendo então $X_i = F_{x_i}^{-1}(\Phi(Z_i))$, fica garantido que X_i tem a distribuição marginal pretendida, $F_{x_i}(\cdot)$.
- Sendo as v.a. $\Phi(Z_i)$ correlacionadas, também as v.a. $X_i = F_{x_i}^{-1}(\Phi(Z_i))$ são correlacionadas, represente-se a respetiva matriz das correlações por Σ'' . Os elementos desta matriz serão igualmente funções dos σ_{ij} desconhecidos.
- Recordando que o objetivo é garantir que $\mathbf{X}=(X_1, X_2, \dots, X_n)$ tem distribuições marginais $F_{x_i}(\cdot)$ e matriz das correlações \mathbf{R} , para satisfazer este requisito basta estabelecer a igualdade $\Sigma'' = \mathbf{R}$, um sistema que é, por vezes, de difícil resolução.

- Conhecida $\Sigma = \begin{bmatrix} 1 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & 1 & \dots & \sigma_{2n} \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & 1 \end{bmatrix}$, recorra-se ao processo visto atrás para a geração

de observações do vetor $\mathbf{Z}=(Z_1, Z_2, \dots, Z_n) \sim N(\mathbf{0}, \Sigma)$, onde as correlações já estão integradas, e usem-se essas observações e as relações $X_i = F_{x_i}^{-1}(\Phi(Z_i))$ para simular observações de $\mathbf{X}=(X_1, X_2, \dots, X_n)$, como pretendido.

Exercício: Ilustre o algoritmo acima, admitindo que:

$$f(x_1) = x_1 + \frac{1}{2}, 0 < x_1 < 1.$$

$$f(x_2) = x_2 + \frac{1}{2}, 0 < x_2 < 1.$$

$$\mathbf{R} = \begin{bmatrix} 1 & -\frac{1}{11} \\ -\frac{1}{11} & 1 \end{bmatrix}.$$

8. Simulação de Processos Estocásticos

Seja $X = \{X_t, t \in I\}$ um processo estocástico, isto é uma família de v.a. definidas num espaço de probabilidade (Ω, \mathcal{F}, P) , onde t é um parâmetro que assume valores num conjunto $I \subset \mathbb{R}$, designado conjunto dos índices do processo. Naturalmente, as v.a. X_t que formam o processo estão correlacionadas, não são independentes, pelo que não podem ser geradas de forma independente, a partir das respectivas distribuições. Por exemplo, $X_{t+\varepsilon}$ é quase que completamente determinada por X_t , para valores pequenos de ε . A forma como a questão é resolvida tem que ser adaptada, conforme se trate de um processo estocástico em tempo discreto ou de um processo estocástico em tempo contínuo.

8.1 Processo Estocástico em Tempo Discreto

Seja $X = \{X_t, t \in \{1, 2, \dots, n\}\}$ um processo estocástico em tempo discreto com incrementos independentes, e seja $F_k(\cdot)$ a distribuição do k -ésimo incremento, $X_k - X_{k-1}$. Uma trajectória do processo pode ser obtida, aplicando o seguinte algoritmo:

- Considere-se $X_0(\omega) = 0$.
- Simulem-se observações das v.a. $Y_k(\omega), k = 1, \dots, n$, tais que $Y_k \sim F_k(\cdot)$.
- Faça-se $X_k(\omega) = X_{k-1}(\omega) + Y_k(\omega), k = 1, \dots, n$.

8.2 Processo Estocástico em Tempo Contínuo

Quando $I = [0, T]$ é necessário criar uma grelha suficientemente fina do conjunto dos índices, seja $0 = t_0 < t_1 < \dots < t_n = T$, e seguir o algoritmo anterior, usando, por exemplo, uma interpolação linear nos pontos intermédios.

Para isso, como já não faz sentido continuar a considerar “incrementos independentes”, uma vez que não se está realmente em tempo discreto, impõe-se o conhecimento das sucessivas distribuições condicionadas, representando agora $F_k(\cdot)$ a distribuição da v.a. $Y_k(\omega)$, dada a v.a. $X_{t_{k-1}}$. Nestas condições, uma trajectória do processo pode ser obtida, aplicando o algoritmo:

- Considere-se $X_0(\omega) = 0$.

- Para $k = 1$ até n :
 - Simule-se uma observação da v.a. $Y_k(\omega)$ tal que $Y_k \sim F_k(\cdot)$.
 - Faça-se $X_{t_k}(\omega) = X_{t_{k-1}}(\omega) + Y_k(\omega)$.
 - Para os valores intermédios de t fazer a interpolação

$$X_t(\omega) = X_{t_{k-1}}(\omega) + \frac{t - t_{k-1}}{t_k - t_{k-1}} Y_k(\omega), t \in]t_{k-1}, t_k[.$$

8.3 Exemplos

Caso 5: Considere dois ativos A e B e sejam $S_A(t)$ e $S_B(t)$ os respectivos preços de mercado no momento t . Admita que $S_A \sim GBM(\mu_A, \sigma_A)$, $S_B \sim GBM(\mu_B, \sigma_B)$ e que os dois processos são independentes. Em $t=0$ um certo portfólio é composto por n_A unidades do ativo A e n_B unidades do ativo B . Estime a probabilidade deste portfólio se desvalorizar mais de 10% num horizonte temporal T igual a 0.5 anos, com

$$S_A(0) = 100, S_B(0) = 75, \mu_A = 0.15, \sigma_A = 0.2, \mu_B = 0.12, \sigma_B = 0.18, n_A = n_B = 100.$$

Nota: Neste exemplo, e também no seguinte, só acaba por interessar o valor do processo daqui a meio ano. De qualquer modo, não é difícil ver como seria a geração de toda uma trajetória.

Havendo independência,

$$S_T^A = S_0^A \exp\left\{\left(\mu_A - \sigma_A^2 / 2\right)T + \sigma_A B_A(T)\right\}, B_A(T) \sim N\left(0, (\sqrt{T})^2\right)$$

e

$$S_T^B = S_0^B \exp\left\{\left(\mu_B - \sigma_B^2 / 2\right)T + \sigma_B B_B(T)\right\}, B_B(T) \sim N\left(0, (\sqrt{T})^2\right),$$

bastando gerar observações independentes com distribuição $N\left(0, (\sqrt{T})^2\right)$.

Caso 6: Admita-se agora que os processos $S^A \sim GBM(\mu_A, \sigma_A)$ e $S^B \sim GBM(\mu_B, \sigma_B)$ não são processos independentes, pois sabe-se que os Movimentos Brownianos B_t^A e B_t^B associados a A e B têm coeficiente de correlação $\rho \neq 0$. Voltar a estimar a probabilidade de o portfólio se desvalorizar mais de 10%.

Nestas condições, pode escrever-se

$$S_{t+s}^A = S_t^A \exp\left\{\left(\mu_A - \sigma_A^2 / 2\right)s + \sqrt{s}V_A\right\}$$

$$S_{t+s}^B = S_t^B \exp\left\{\left(\mu_B - \sigma_B^2 / 2\right)s + \sqrt{s}V_B\right\},$$

onde $(V_a, V_b) \sim N(\mathbf{0}, \Sigma)$, $\Sigma = \begin{bmatrix} \sigma_a^2 & \rho\sigma_a\sigma_b \\ \rho\sigma_a\sigma_b & \sigma_b^2 \end{bmatrix}$, caindo-se assim no caso

bidimensional atrás estudado.

9. Métodos Para a Redução do Erro de Estimação

Como já foi referido (ponto 5.2), para se conseguir um aumento de 50% na precisão, isto é, uma redução de 50% no erro cometida com a estimação, é necessário quadruplicar a dimensão da amostra. Por outras palavras, o erro padrão (desvio padrão da amostra) reduz-se a uma taxa igual a apenas a raiz quadrada da dimensão da amostra (não da própria dimensão da amostra), o que motiva os utilizadores a procurar aumentar a precisão por outros meios.

Um dos meios mais usados é o chamado método das variáveis antitéticas, que nalgumas versões permite reduzir para metade a quantidade de NA que é necessário gerar, pois geram-se observações com a sucessão u_1, u_2, \dots, u_N e, numa segunda corrida da simulação, também com a sucessão $1-u_1, 1-u_2, \dots, 1-u_N$. A ideia é de que a amostragem com base em números aleatórios complementares induz a uma correlação negativa entre as variáveis de output das duas corridas, o que permite reduzir a variância.

Por exemplo, considere-se a simulação de observações de uma v.a. de output X com o método das variáveis antitéticas. Seja X_1 o resultado da primeira corrida e X_2 o resultado da segunda corrida.

Como $\bar{X} = \frac{X_1 + X_2}{2}$ é um estimador não enviesado de X e as duas v.a. são negativamente correlacionadas, pode esperar-se que $Var(\bar{X}) = \frac{Var(X_1) + Var(X_2) + 2Cov(X_1, X_2)}{4}$, venha reduzida, o que aumentaria a precisão dos intervalos de confiança para X . No entanto, deve comparar-se com a $Var(\bar{X})$ associada a duas corridas não correlacionadas.

Outros métodos usados: o método das variáveis de controlo, em que se procura integrar no processo de simulação o conhecimento que se tem sobre uma outra variável fortemente correlacionada (*control variate*) com a variável que se quer estimar, de modo a dar atenção apenas à estimação da parcela sobre a qual não há conhecimento; os métodos de quasi-Monte Carlo, que recorrem às chamadas sequências de baixa discrepância, de que são exemplo as sucessões de Halton ou os números de Sobol (ver Korn *et al.*, p 65 e ss. e também pp. 45 e ss.).

Referências

Glasserman P (2003) Monte Carlo Methods in Financial Engineering- Stochastic Modelling and Applied Probability, Springer.

Hogg RV e Tanis A (2014) Probability and Statistical Inference, 9th edition, Prentice-Hall.

Korn R, Korn E, Kroisandt G. (2010) Monte Carlo Methods and Models in Finance and Insurance, Chapman & Hall/CRC Financial Mathematics Series, London.

Larson H (1982) Introduction to Probability Theory and Statistical Inference, 3rd edition, Wiley, John & Sons, Incorporated, NY.

Metropolis N e Ulam S (1949) The Monte Carlo Method, Journal of the American Statistical Association 44 (247): 335–341.

Murteira B e Antunes M (2012) Probabilidades e Estatística, volume I, Escolar Editora.

Pidd M (2006) Computer Simulation in Management Science, 5th edition, John Wiley and Sons Ltd, Chichester.

Taha, H. A. (2010). Operations research: an introduction, 9th edition, Pearson Education.