

COPPEAD/UFRJ

RELATÓRIO COPPEAD Nº 185
APLICAÇÃO DA AMOSTRAGEM DESCRITIVA
NA ANÁLISE DE RISCO

Eduardo Saliby*

Maio de 1987

* Professor adjunto da área de Métodos Quantitativos e pesquisador do Instituto de Pós-Graduação e Pesquisa em Administração da Universidade Federal do Rio de Janeiro - COPPEAD/UFRJ. Engenheiro de Produção pela Escola Politécnica da USP (1971), Mestre em Ciências pela COPPE/UFRJ (1974), PhD pela University of Lancaster, UK (1980).

COPPEAD/UFRJ
MESTRADO EM ADMINISTRAÇÃO

RELATÓRIO TÉCNICO N. 108
APLICAÇÃO DA AMOSTRAGEM DESCRITIVA
NA ANÁLISE DE RISCO

Eduardo Saliby*

Maio de 1987

- * Professor adjunto da área de Métodos Quantitativos e pesquisador do Instituto de Pós-Graduação e Pesquisa em Administração da Universidade Federal do Rio de Janeiro - COPPEAD/UFRJ.
Engenheiro de Produção pela Escola Politécnica da USP (1971), Mestre em Ciências pela COPPE/UFRJ (1974), PhD pela University of Lancaster, UK (1980).

RESUMO

A análise de risco tem por objetivo avaliar o impacto de incertezas na tomada de decisões de investimento com horizontes de médio e longo prazo. Frequentemente, este problema é estudado através da simulação por Monte Carlo, usando-se sua abordagem tradicional: a amostragem aleatória simples.

Mas, ao propormos a amostragem descritiva como o método mais apropriado em simulação, esta sugestão também se estende à análise de risco. A amostragem descritiva representa uma importante mudança conceitual em simulação, pois se baseia numa seleção totalmente determinística e intencional dos valores amostrais.

Este trabalho tem por objetivo mostrar que, também numa análise de risco, a amostragem descritiva é uma melhor abordagem. Neste caso, além de proporcionar estimativas mais precisas, a amostragem descritiva foi de fácil implantação e resultou também numa redução do tempo de processamento.

ABSTRACT

Risk analysis is a method to evaluate the possible outcomes from uncertainties related to a decision of investment. This problem is frequently studied with Monte Carlo simulation, using its traditional approach: simple random sampling.

However, given that descriptive sampling was proposed as the most appropriate sampling method in simulation, this suggestion also extends to risk analysis. Descriptive sampling represents an important conceptual change in simulation, once is based on a fully deterministic and controlled selection of the sample values.

The purpose of the present work is to show that, also in a risk analysis, descriptive sampling is a better approach. In this case, apart from producing more precise estimates, descriptive sampling showed to be easy to implement and, more interesting, the processing time was shortened.

1. INTRODUÇÃO

A análise de risco é uma ferramenta de apoio à tomada de decisão que vem sendo cada vez mais utilizada. Através dela, levamos em conta as incertezas associadas a um projeto de investimentos, avaliando o seu efeito no resultado econômico-financeiro deste projeto. Uma de suas primeiras aplicações é o já clássico artigo de Hertz (1964).

A análise de risco geralmente se aplica a problemas de planejamento envolvendo horizontes de médio a longo prazo. Conforme veremos na seção seguinte, ela se caracteriza pelo uso da simulação por Monte Carlo como abordagem geral de solução. Como consequência, os resultados de uma análise de risco serão imprecisos.

Mas, como uma aplicação da simulação por Monte Carlo, os resultados de uma análise de risco têm parte de sua variabilidade explicada em função dos diferentes conjuntos de entrada gerados a cada corrida (Saliby, 1986). Além disso, esta parcela de variação identificada como o efeito de conjunto, é desnecessária, podendo ser eliminada com o uso da amostragem descritiva (Saliby, 1982).

Ao contrário da amostragem aleatória simples - até o momento a abordagem padrão em simulação por Monte Carlo - a amostragem descritiva baseia-se numa seleção totalmente determinística e controlada dos valores da amostra, sendo aleatória apenas sua seqüência de utilização.

A maior eficiência estatística da amostragem descritiva tem sido verificada em diversos testes, tais como em Ferreira Sobrinho (1982) e Saliby e Funchal (1986). Confirmando sua utilidade, mostraremos que a amostragem descritiva poderá vir a representar um importante aperfeiçoamento da análise de risco. De fato, nossos testes conduziram a resultados mais precisos ao mesmo tempo em que o tempo de processamento foi diminuído.

Como confirmação adicional da validade da amostragem descritiva, foi possível mostrar, através do Modelo Linear de Resposta ou simplesmente MLR (Saliby, 1980), que sua maior precisão decorre realmente do controle do conjunto de entrada. O MLR tem por objetivo compensar os desvios observados entre os momentos das amostras geradas e os respectivos valores teóricos, através de uma regressão linear. Trata-se de um método similar ao das variáveis de controle, uma eficiente técnica de redução de variância descrita, por exemplo, em Bratley, Fox e Schrage (1983).

A importância do MLR prende-se ao fato do seu coeficiente de determinação R^2 , dar-nos uma idéia antecipada da redução de variância esperada quando utilizamos a amostragem descritiva em lugar da amostragem aleatória simples.

A seguir, apresentaremos um breve resumo teórico sobre a análise de risco. Posteriormente, ela será utilizada num problema de natureza simples, porém ilustrativo dos principais conceitos envolvidos.

2. UMA INTRODUÇÃO À ANÁLISE DE RISCO

O objetivo de uma análise de risco é a determinação do perfil de risco associado a um projeto de investimento. O perfil de risco corresponde à distribuição de probabilidades da variável de decisão, geralmente o valor presente do fluxo de caixa descontado (VP) ou a taxa interna de retorno (TIR). Cabe lembrar que, em razão das incertezas quanto aos dados de entrada, a variável de decisão será também de natureza probabilística.

Os métodos tradicionais de avaliação de projetos baseiam-se no uso de um valor único para cada um dos dados de entrada, até mesmo para os fatores sujeitos a incerteza; neste caso, eles são geralmente representados pelo seu valor mais provável. De acordo com Hertz (1964), no entanto, esta prática poderia traduzir situações com uma baixa probabilidade conjunta de ocorrência; além disso, ainda que não seja este o caso, o uso de um valor único não permitiria uma idéia do perfil de risco do projeto. Assim sendo, Hertz sugeriu que estas incertezas fossem incorporadas ao modelo de análise.

Mas, em função da grande dificuldade matemática envolvida na determinação analítica de um perfil de risco, Hertz também sugeriu o emprego da simulação por Monte Carlo como método de solução. Assim, embora uma análise de risco faça geralmente uso da simulação, ela não é uma exigência teórica, mas apenas um método para se estimar o perfil de risco em questão. Cabe ainda mencionar que alguns autores questionam a utilidade da análise de risco, como por exemplo Lewellen e Long (1976); no entanto, apesar de não concordarmos com suas críticas, elas dizem respeito a aspectos conceituais do método e não ao uso de simulação como procedimento de solução.

Uma aplicação da análise de risco envolve as seguintes etapas:

- a) Identificar, dentre os fatores considerados na análise de um projeto de investimento, aqueles cuja incerteza seja relevante para a determinação do valor final da variável de decisão, tais como custos de instalação, custos de matéria prima, níveis de venda, etc.
- b) Obter estimativas da faixa de valores na qual se espera que estes fatores irão variar, e do grau de confiança da ocorrência de valores dentro desta faixa. Por exemplo, podemos prever que a vida útil de um produto irá variar entre 5 e 10 anos, atribuindo-se uma probabilidade de 10% aos valores extremos (5 ou 10 anos), 15% aos valores intermediários (6 ou 9 anos) e 25% aos valores centrais (7 e 8), considerados os mais prováveis.

Em alguns casos, esta atribuição de probabilidades poderá ser feita usando-se uma distribuição teórica conhecida como a normal ou a uniforme. Em outros, como no caso acima, faremos uso de um modelo empírico.

- c) Construção de um modelo de simulação que melhor traduza as conseqüências dos possíveis desdobramentos futuros do projeto. Este modelo deverá levar em conta todos os elementos relevantes ao resultado final do projeto, sejam eles determinísticos ou aleatórios; neste último caso, serão incorporadas as incertezas previamente identificadas e quantificadas através de sua distribuição de probabilidades.
- d) Utilizando a simulação por Monte Carlo, gerar valores para cada uma das variáveis aleatórias do modelo. Estes valores permitirão calcular o valor presente (VP) ou a taxa interna de retorno (TIR) para um possível cenário futuro.
- e) Repetir o procedimento anterior várias vezes de modo a obter uma amostra de valores para VP ou TIR que permitam caracterizar sua distribuição - o perfil de risco - , em particular sua média e desvio padrão.

Como em toda simulação, o volume de cálculos numa análise de risco é elevado. Assim sendo, o uso do computador é imprescindível para se chegar a resultados com uma margem de precisão adequada. Além disso, vem sendo desenvolvidas ferramentas de software voltadas para outras etapas do processo, como por exemplo:

- a) Quantificação de probabilidades através de procedimentos interativos;
- b) Novas linguagens voltadas para um tratamento mais automático deste tipo de problema. Os principais softwares de apoio à decisão, como por exemplo o IFPS e o FCS-EPS, já incorporam rotinas para a análise de risco.

Até o momento, no entanto, desconhecemos qualquer desenvolvimento importante voltado para uma das principais deficiências da análise de risco que é sua baixa precisão. Mesmo nos dias de hoje em que os recursos computacionais são mais baratos, uma análise de risco poderá resultar num tempo de processamento altamente comprometedor.

Diversos trabalhos tratam também da análise de risco, tais como Moore e Thomas (1976), Brauers (1986), Hertz e Thomas (1983 e 1984), Singhvi (1980) e Chapman e Cooper (1985). Apesar das peculiaridades de cada um deles, todos seguem em linhas gerais a metodologia acima descrita, particularmente quanto ao uso da abordagem clássica do método de Monte Carlo que é a amostragem aleatória simples. Portanto, ao propormos o uso da amostragem descritiva em seu lugar, estamos abrindo um novo caminho para o aperfeiçoamento da análise de risco.

Na seção seguinte apresentaremos um exemplo de aplicação da análise de risco, correspondente ao problema utilizado em nossos testes comparativos.

3. O PROBLEMA ESTUDADO

Uma empresa, dentre suas opções de investimento, está considerando a possibilidade de lançamento de um novo produto na área de mercado em que atua.

Para avaliar se este lançamento seria de interesse para a empresa, foram inicialmente identificados e quantificados os principais fatores que poderiam influenciar, de alguma forma, o resultado final do empreendimento. São eles:

INVESTIMENTO INICIAL:

Para a fabricação e colocação do produto no mercado é necessário um desembolso inicial de CZ\$ 1.000.000,00, valor este suficientemente preciso a ponto de ser considerado como certo. Como se trata de um investimento de natureza muito específica, sem utilidade para outros desenvolvimentos, não há valor residual a ser recuperado no futuro.

VIDA MERCADOLÓGICA DO PRODUTO:

O lançamento em questão refere-se a um produto do tipo novidade, para o qual se espera uma vida útil curta em termos de volume de vendas. Com base na experiência e intuição dos seus diretores, a empresa estimou em 60 % a probabilidade do produto permanecer um mínimo de 2 anos no mercado; já uma estimativa mais otimista, de permanecer por 3 anos no mercado, foi avaliada como tendo 40 % de probabilidade. Assim sendo, foi excluída a possibilidade do produto durar mais que 3 anos.

CUSTO UNITÁRIO:

Devido às incertezas que cercam o empreendimento, o custo unitário do produto, excluídas as despesas de investimento, foi também especificado através de uma distribuição de probabilidades. Os possíveis valores e suas respectivas probabilidades são fornecidos abaixo, sendo importante notar que o custo será sempre o mesmo ao longo de toda a vida útil do produto:

CUSTO UNITÁRIO (Cz\$)	PROBABILIDADE (%)
3,00	40
3,50	30
4,00	30

VENDAS ANUAIS:

As vendas do novo produto também variarão de forma aleatória, só que agora, com valores independentes de ano para ano. Com base na experiência passada, determinou-se que as vendas seriam melhor descritas por uma distribuição empírica contínua. Assim sendo, as vendas anuais variariam de forma uniforme num dos 5 intervalos abaixo, que no caso são equiprováveis:

Intervalo	Limites (1000 unidades)	
1	70	a 360
2	360	a 410
3	410	a 490
4	490	a 630
5	630	a 900

Esta distribuição pode ser também descrita, de forma equivalente, através do gráfico de sua função de distribuição, apresentado na figura 1. Tanto no caso da amostragem aleatória simples como na amostragem descritiva, iremos gerar os valores amostrais através da inversa desta função.

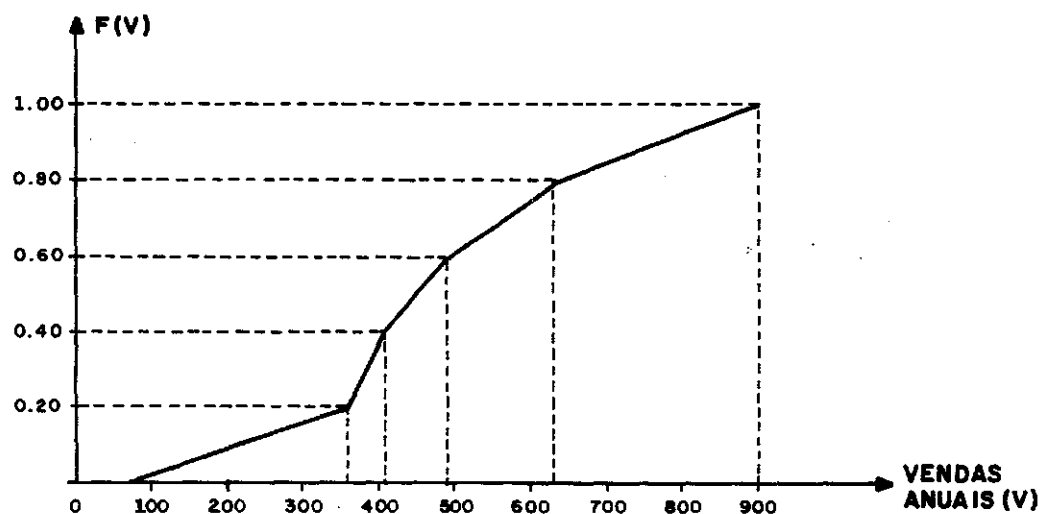


Figura 1. Função de distribuição acumulada das vendas anuais.

PREÇO DE VENDA:

Para efeito do estudo, o preço de venda foi fixado em Cz\$ 5,00 por unidade, independentemente dos demais fatores.

CRITÉRIO DE DECISÃO E TAXA DE DESCONTO:

O critério decisão a ser empregado é o valor presente do fluxo de caixa descontado, VP. Os seus valores serão expressos em Cz\$ 1000,00 .

A taxa de desconto utilizada na avaliação dos projetos da empresa é de 15 % ao ano, antes do imposto de renda.

OUTRAS PREMISSAS:

Considera-se que o produto estará disponível para venda imediatamente após ter sido feito o investimento inicial. Além disso, a título de aproximação, o resultado financeiro será concentrado ao final de cada período anual.

OBJETIVO DA SIMULAÇÃO:

O objetivo do estudo é a determinação do perfil de risco do projeto, com interesse específico na média e no desvio padrão da distribuição do valor presente. Utilizaremos a simulação por Monte Carlo para gerar, a cada corrida, uma amostra de valores

$$VP_i, \quad i = 1, \dots, N$$

e, a partir dela, estimar o valor esperado $E(VP)$ e o desvio padrão $DP(VP)$ através das expressões

$$\overline{VP} = \frac{\sum_{i=1}^N VP_i}{N}, \quad e$$

$$S_{VP} = \left[\frac{\sum_{i=1}^N (VP_i - \overline{VP})^2}{(N-1)} \right]^{1/2}.$$

Em sua abordagem tradicional, a análise de risco faz uso da amostragem aleatória simples para simular os possíveis valores de cada variável aleatória. Assim, sendo NA um "número aleatório" no intervalo $(0,1)$, a tabela 1 ilustra os cálculos envolvidos numa análise de risco definida por apenas $N = 5$ observações de VP .

Tabela 1. Exemplo dos cálculos envolvidos numa análise de risco, para uma corrida com apenas $N = 5$ observações.

Obs	VIDA		CUSTO		VENDAS (1000 unidades)						VALOR PRESENTE
	NA	Va1	NA	Va1	NA	V1	NA	V2	NA	V3	Cz\$ 1000
1	.413	2	.830	4	.654	528	.125	251	--	--	-351
2	.023	2	.261	3	.130	258	.776	613	--	--	376
3	.789	3	.064	3	.411	414	.172	319	.430	422	757
4	.266	2	.628	3.5	.617	502	.193	350	--	--	52
5	.522	2	.942	4	.480	442	.738	587	--	--	-172

Média: $\overline{VP} = 132.4$
D. Padrão: $S = 442.2$
VP

Quanto a sua programação em computador, ela não traz maiores problemas, mesmo para situações mais complexas. A principal dificuldade numa análise de risco, como já mencionamos, reside geralmente no tempo de processamento que pode ser muito elevado em aplicações mais realistas.

Em nossos testes, definimos cada corrida por um conjunto de $N = 100$ observações do valor presente, dando origem a uma única estimativa para cada parâmetro em estudo: $E(VP)$ e $DP(VP)$. Uma característica da simulação reside na variabilidade associada a estas estimativas, uma vez que repetindo-se o procedimento, com diferentes seqüências de números aleatórios, os resultados finais serão também diferentes. Esta variabilidade caracteriza uma imprecisão acerca do parâmetro estimado, sendo por isso indesejável; ela é medida pela variância da estimativa, cujo valor deveria ser mínimo.

Para este nosso problema, dispomos também da solução analítica para o valor esperado de VP, dado por

$$E(VP) = 390.57 ,$$

que corresponde a um resultado exato e, portanto, de variância nula. Interpretando-o, podemos imaginá-lo como sendo a média \overline{VP} que obteríamos, caso tivéssemos uma corrida com N infinito.

Como observação final, note-se que caso pudéssemos calcular analiticamente os demais parâmetros da distribuição de VP, não mais precisaríamos recorrer à simulação como método de solução. Não teríamos assim, qualquer incerteza quanto ao perfil de risco do projeto.

4. RESULTADOS OBTIDOS COM A AMOSTRAGEM ALEATÓRIA SIMPLES

Para uma comparação empírica da precisão das estimativas de ambos os métodos amostrais, definimos um experimento de simulação por um conjunto de $M = 100$ corridas. Cada corrida, por sua vez, foi definida por $N = 100$ observações do valor presente, VP.

Apresentaremos, inicialmente, os resultados da amostragem aleatória simples.

Neste caso, além da média e da variância das estimativas calculadas a partir de 100 corridas, foi também feita uma análise através do Modelo Linear de Resposta (MLR). A tabela 2 apresenta um resumo dos principais resultados obtidos.

Tabela 2. Sumário dos resultados de 100 corridas utilizando a amostragem aleatória simples.

Estimativa	Média	Desvio padrão	R ² do MLR
\overline{VP}	392.82	55.06	94 %
S VP	576.57	44.68	23 %

A análise do MLR, conforme descrita em Saliby (1980), resultou nos coeficientes de determinação R². O MLR é definido por uma equação de regressão que ajusta as estimativas em função dos erros amostrais observados.

Para esta análise, além das estimativas, calculamos também a média e o desvio padrão de cada uma das amostras geradas, isto para cada corrida. A seguir, usando estes valores como entrada, calculamos uma equação de regressão para cada uma das estimativas, no caso consideradas como variável dependente.

Na versão mais simples do MLR, adequada a muitas aplicações, usamos apenas as médias como controle. Neste caso, a regressão tem como variáveis independentes as diferenças observadas entre as médias das amostras geradas e seus respectivos valores teóricos. Já na sua versão completa, usamos também o desvio padrão como controle.

Tendo apenas as médias como controle, e sabendo que os seus valores teóricos são:

$$\begin{aligned} E(\text{VIDA}) &= 2.40 \text{ anos,} \\ E(\text{CUSTO}) &= \text{Cz\$ } 3.45 \text{ e} \\ E(\text{VENDA}) &= 475 \text{ mil unidades anuais,} \end{aligned}$$

apresentamos, como exemplo, a regressão obtida para \overline{VP} :

$$\overline{VP} = 390.19 + 503.17 \cdot (\overline{VIDA} - 2.40) - 865.18 \cdot (\overline{CUSTO} - 3.45) + 3.14 \cdot (\overline{VENDA} - 475)$$

Seu coeficiente de determinação, $R^2 = 0.93$, foi bastante elevado e praticamente igual ao do MLR completo, fornecido na tabela 2.

De acordo com a teoria do MLR (Saliby, 1980), o coeficiente R^2 fornece uma idéia aproximada da contribuição do efeito de conjunto na variação das estimativas. Trata-se, pois, de uma parcela de variação que será eliminada com o uso da amostragem descritiva. Assim sendo, esperamos neste caso uma grande redução de variância para \overline{VP} , enquanto que para $S_{\overline{VP}}$ os ganhos deverão ser relativamente menores mas, ainda assim, relevantes. Conforme veremos na seção seguinte, estas previsões foram plenamente confirmadas.

Ainda com relação ao MLR, ao contrário do que ocorre com o método das variáveis de controle, os coeficientes de regressão são de grande interesse:

O termo constante representa uma estimativa global para o parâmetro em estudo, já corrigida em função dos erros amostrais de cada corrida. No caso, temos

$$\widehat{E(\overline{VP})} = 390.19 ,$$

muito próximo, portanto, do seu valor exato ($E(\overline{VP})=390.57$).

Já com relação aos coeficientes das variáveis independentes, eles indicam a sensibilidade de $E(\overline{VP})$ em função de possíveis variações nos valores teóricos dos momentos considerados como controle. Por exemplo, podemos afirmar que um aumento de 1000 unidades na média anual de vendas resultará num acréscimo de 3.14 em $E(\overline{VP})$. Da mesma forma, sabemos também como $E(\overline{VP})$ irá aumentar caso o produto tenha uma maior média de vida e como irá diminuir caso a média do custo unitário aumente.

Informações sobre a sensibilidade da solução em função dos parâmetros de entrada são, certamente, de grande utilidade prática.

5. UTILIZAÇÃO DA AMOSTRAGEM DESCRITIVA NA ANÁLISE DE RISCO

O experimento anterior foi repetido utilizando-se, agora, a amostragem descritiva. Veremos, inicialmente, os principais detalhes acerca da geração e uso das amostras descritivas para, depois, apresentarmos os principais resultados obtidos.

GERAÇÃO DAS AMOSTRAS DESCRITIVAS

A amostragem descritiva tem por objetivo minimizar o erro amostral associado ao conjunto de valores gerados; na medida em que nossas amostras melhor descrevam as distribuições de probabilidades representadas, mais precisas também serão as estimativas. Em princípio, a amostragem descritiva resulta num controle individual de cada uma das variáveis de entrada, controle este que proporciona substanciais ganhos de precisão.

O presente problema, em função de suas peculiaridades, permite um controle adicional, uma vez que a distribuição conjunta da vida útil e do custo unitário, pode ser representada sem erro. De fato, lembrando que uma corrida é definida por $N=100$ observações de VP e que, para cada observação, usamos um único valor para a vida e o custo, então as amostras destas duas variáveis de entrada serão também definidas por $N=100$ elementos.

A primeira idéia que temos ao empregarmos a amostragem descritiva é a de representar isoladamente cada uma das variáveis. Desta forma, os seus valores amostrais seriam descritos pelas seguintes distribuições de freqüência:

Vida útil	Freqüência	Custo unitário	Freqüência
2	60	3.00	40
3	40	3.50	30
		4.00	30

Apesar desta abordagem ser correta, podemos fazer um melhor uso da amostragem descritiva, estendendo sua idéia à distribuição conjunta das duas variáveis. Assim, em função da sua independência, a amostra descritiva conjunta destas variáveis seria definida pela distribuição de freqüências da tabela 3.

Tabela 3. Freqüência de valores numa amostra descritiva para a distribuição conjunta da vida útil e do custo unitário.

Vida (anos)	Custo unitário (Cz\$)			Total
	3.00	3.50	4.00	
2	24	18	18	60
3	16	12	12	40
Total	40	30	30	100

Definido o conjunto de valores para a vida e o custo unitário, passamos agora ao das vendas anuais. Em primeiro lugar, precisamos determinar o tamanho da amostra a ser gerada.

Voltando por um momento à amostragem aleatória simples, notamos que o total de períodos anuais simulados varia a cada corrida. Assim, se tivermos

$$\begin{aligned} f(2) &= 65 \text{ observações com vida útil de 2 anos, e} \\ f(3) &= 35 \text{ observações com vida útil de 3 anos,} \end{aligned}$$

o tamanho da amostra de vendas anuais será de

$$NV = 2 \times f(2) + 3 \times f(3) = 235 \text{ valores gerados;}$$

mas, como $f(2)$ e $f(3)$ variam a cada corrida, o mesmo se dará com NV.

Um tamanho de amostra variável dificultaria um pouco a utilização da amostragem descritiva; ele também inviabilizaria um controle completo do conjunto de valores gerados, prejudicando assim sua eficiência estatística. Mas, uma feliz consequência do uso da amostragem descritiva neste problema foi que o tamanho da amostra das vendas anuais, variável com a amostragem aleatória simples, transformou-se numa constante. De fato, para cada corrida teremos agora exatamente 60 observações de VP com uma vida de 2 anos e 40 observações com uma vida de 3 anos, resultando num total de

$$NV = 2 \times 60 + 3 \times 40 = 240 \text{ vendas geradas.}$$

Definido o tamanho da amostra, passamos agora à geração dos valores descritivos para as vendas anuais. Temos duas alternativas para fazê-lo:

- a) Gerando três amostras descritivas, uma para cada ano de vida. Neste caso, as amostras para os dois primeiros anos seriam compostas de 100 valores, enquanto que a amostra para o terceiro ano teria apenas 40 valores.
- b) Gerando uma única amostra descritiva que seria utilizada, indistintamente, para os tres anos de vida do produto.

A primeira opção leva-nos a uma maior discretização da distribuição, mas, em compensação, faz com que as vendas de diferentes anos sejam independentes entre si.

Já na segunda opção, os 240 valores que definem uma única amostra descritiva possibilitam uma melhor aproximação da distribuição das vendas. Em contrapartida, ela também resulta numa pequena correlação negativa entre as vendas de diferentes anos, cujo efeito é desprezível na prática.

Assim, embora haja pouca diferença entre as duas alternativas do ponto de vista estatístico, a maior simplicidade computacional levou-nos a optar pelo uso de uma única amostra descritiva. Desta forma, os valores descritivos das vendas foram definidos por

$$V_i = F^{-1} [(i-0.5)/240] , \quad i = 1, \dots, 240 ,$$

onde $F^{-1}(R)$ é a inversa da função de distribuição.

Deve-se notar que a geração dos valores descritivos é análoga à geração dos valores aleatórios. A única diferença reside no fato de que agora, em lugar de uma amostra aleatória de valores uniformes, usamos como argumento da transformada inversa uma amostra descritiva de valores uniformes.

Em termos geométricos, o procedimento acima corresponde a dividir o eixo unitário vertical da figura 1 em $NV = 240$ intervalos iguais, equiprováveis portanto, e tomar o ponto médio de cada um deles para a geração dos valores descritivos.

Quanto à geração das permutações aleatórias, restrita apenas às vendas anuais, ela seguiu o método geral adotado em outros estudos, e descrito com detalhes em Almeida (1983). Em resumo, este método corresponde à simulação de um sorteio de bingo onde n bolas previamente numeradas de 1 a n , são embaralhadas e depois retiradas uma a uma da urna.

UTILIZAÇÃO DA AMOSTRAGEM DESCRITIVA

Replicações independentes, no caso da amostragem descritiva, correspondem ao uso de diferentes permutações aleatórias a cada corrida, mantendo-se sempre o mesmo conjunto de valores de entrada. Desta forma, podemos gerar os valores amostrais ao início de um programa, guardá-los na memória e utilizá-los, repetidamente, a cada corrida.

Tendo os valores descritivos previamente definidos, geramos apenas a sua permutação aleatória a cada corrida, um procedimento geralmente bem mais rápido do que a obtenção dos valores. Por exemplo, a geração de um valor para as vendas envolve o cálculo de uma interpolação linear para a inversa da função de distribuição. No caso da amostragem aleatória simples, esta operação é feita, em média, 240 vezes a cada corrida, para um total de 100 corridas. Com a amostragem descritiva, geramos um único conjunto de 240 valores para todas as corridas do experimento, o que resulta numa sensível redução do tempo total de processamento.

Para cada amostra descritiva gerada usamos, em princípio, dois vetores:

- a) um para guardar os valores amostrais, e
- b) outro para guardar os índices que definem sua permutação aleatória.

Em nossos testes, definimos um vetor de valores para cada variável de entrada. Por outro lado, em função das peculiaridades do problema, somente as vendas tiveram seus valores permutados aleatoriamente, fazendo com que fosse usado um único vetor de índices.

RESULTADOS OBTIDOS COM A AMOSTRAGEM DESCRITIVA

Os resultados de $M = 100$ corridas utilizando a amostragem descritiva são resumidos na tabela 4. Juntamente com a redução de variância obtida em relação ao uso da amostragem aleatória simples, rerepresentamos os coeficientes de determinação do MLR previamente calculados.

Tabela 4. Sumário dos resultados de 100 corridas utilizando amostragem descritiva.

Estimativa	Média	D. padrão	Red. variância	R ² (MLR)
\overline{VP}	390.44	9.30	97 %	94 %
S_{VP}	586.69	33.16	45 %	23 %

Antes de compararmos as variâncias das estimativas, é necessário primeiro observar que suas médias globais foram aproximadamente as mesmas para ambos os métodos. Isto nos permite concluir que nenhum viés relevante foi introduzido ao substituírmos a amostragem aleatória simples pela amostragem descritiva, fato este também constatado nos demais estudos já realizados.

Os resultados da Tabela 4 nos mostram uma grande redução de variância para a estimativa \overline{VP} , conforme prevíamos com a análise do MLR.

Quanto a estimativa S_{VP} , apesar dos ganhos terem sido menores, eles superaram as previsões do MLR, o que não deixa de ser vantajoso. A propósito, tal como ocorreu aqui, cabe notar que o MLR fornece uma estimativa pessimista da redução de variância proporcionada pela amostragem descritiva. Isto se verifica porque, com o MLR, levamos em conta apenas os desvios em relação a média e o desvio padrão da amostra de entrada; enquanto isso, numa amostra descritiva, todos os momentos são controlados.

6. CONSIDERAÇÕES FINAIS

Os resultados obtidos mais uma vez comprovam a validade e a utilidade da amostragem descritiva em simulação. Sua aplicação à análise de risco não só conduziu a estimativas mais precisas como também reduziu o tempo de processamento envolvido. Trata-se de uma importante constatação, uma vez que uma das principais desvantagens da análise de risco reside justamente no volume de processamento que pode ser agora substancialmente reduzido.

Como o exemplo estudado refere-se a um problema didático de natureza simples, é possível que, em aplicações mais realistas da análise de risco, os ganhos com a amostragem descritiva sejam menores. Mas, ainda que estes ganhos não sejam tão elevados, não há motivos para não utilizarmos a amostragem descritiva.

O uso da amostragem descritiva impõe, aparentemente, um elevado grau de artificialidade numa análise de risco, em função do controle da variação de conjunto. Mas, como mostram os resultados deste e dos demais problemas de simulação já testados, isto não leva a qualquer distorção das estimativas.

Como uma consideração final, este estudo ajuda a ilustrar uma das principais diferenças conceituais entre a amostragem descritiva e a amostragem aleatória simples:

A amostragem aleatória simples pode ser vista como um processo sequencial de amostragem, onde a ordem temporal é levada em conta. Neste caso, os elementos são gerados um a um, à medida em que são necessários, podendo o processo ser interrompido a qualquer momento.

Já a amostragem descritiva caracteriza-se por ser um evento único, onde todos os valores amostrais são gerados simultaneamente e em bloco. Esta simultaneidade não mais permite que o processo seja interrompido antes do seu término. Por outro lado, como os valores são utilizados sequencialmente numa simulação, isto nos obriga a definir sua ordem de utilização; para evitarmos o risco de introduzirmos algum tipo de distorção decorrente de uma ordenação arbitrária, fazemos sua permutação aleatória.

REFERÊNCIAS

- ALMEIDA, G.C.P. Geracão de amostras descritivas em simulação. Rio de Janeiro, COPPE/UFRJ, 1983. Tese de mestrado.
- BRATLEY, P.; FOX, B.L.; SCHRAGE, L.E. A guide to simulation. New York, Springer-Verlag, 1983.
- BRAUERS, W.K. Risk, uncertainty and risk analysis. Long Range Planning, 19 (6): 139-143, Dec. 1986. Essay review article.
- CHAPMAN, C.B. & COOPER, D.F. Risk Analysis. In: RAND, G.K. & EGLESE, R.W., ed. Further developments in operational research. Oxford, Pergamon Press, 1985. Cap. 2, p. 12-33.
- FERREIRA SOBRINHO, W.O. Utilização de amostragem descritiva em simulação: em estudo empírico. Rio de Janeiro, COPPEAD/UFRJ, 1982. Tese de Mestrado.
- HERTZ, D. Risk analysis in capital investment, Harvard Business Review, 42 (1): 95-106, Jan./Feb. 1964.
- & THOMAS H. Risk analysis and its applications. Chichester, J. Wiley, 1983.
- & ----- Practical risk analysis, an approach through case histories. Chichester, J. Wiley, 1984.
- LEWELLEN, W.G. & LONG, M.S. Simulation versus single-value estimates in capital expenditure analysis. In: MYERS, S.C., ed. Modern developments in financial management. New York, Holt, Rinehart and Winston, 1976. P. 442-456.
- MOORE, P.G. & THOMAS, H. The anatomy of decisions. Harmondsworth, Penguin, 1976.
- SALIBY, E. A reappraisal of some simulation fundamentals. Lancaster, Universidade de Lancaster, 1980. Tese de Doutorado.
- Uma revisão de fundamentos da simulação: o uso incorreto de amostragem aleatória simples. Pesquisa Operacional, 2 (2): 1-16, dez. 1982b.
- Sources of variation of simulation estimates: an empirical study. Rio de Janeiro, COPPEAD/UFRJ, 1986. (Relatório de Pesquisa, 64)
- & FUNCHAL, G. Estudo da eficiência da amostragem descritiva na simulação de um problema de filas. Pesquisa Operacional, 6 (2), jul./dez. 1986.
- SINGHVI, S.S. A practical approach to risk and sensitivity analyses. Long Range Planning, 13 (1): 12-19, Feb. 1980.