INPE-500-LAFE
NAS

MANUAL DE ANÁLISE DE DECISÕES

Heiko Humann
Jonas de Oliveira Junior
Lauro Tadeu Guimarães Fortes
Nívea Teixeira Batista

Julho 1974

cc.: 100

SERVIÇO PÚBLICO FEDERAL
CONSELHO NACIONAL DE PESQUISAS
INSTITUTO DE PESQUISAS ESPACIAIS
São José dos Campos - Estodo de S. Paulo - Brasil

MANUAL DE ANÁLISE DE DECISÕES

Este relatório foi apresentado, como parte dos requisitos necessários à obtenção do Grau de Mestre em Ciências na área de Análise de Sistemas e Aplicações pelo Instituto de Pesquisas Espaciais, por Heiko Humann, Jonas de Oliveira Junior, Lauro Tadeu Guimarães Fortes, Nívea Teixeira Batista, tendo como orientador Dr. José Eugênio Guisard Ferraz.

Sua publicação foi autorizada pelo abaixo assinado,

Fde Mendonça
Fernando de Mendonça
Diretor Geral

PRESIDÊNCIA DA REPÚBLICA
CONSELHO NACIONAL DE PESQUISAS
INSTITUTO DE PESQUISAS ESPACIAIS
São José dos Campos - Estado de S. Paulo - Brasil

EM 04 DE junho DE 19 74

Os abaixo assinados, Membros da Banca Examinadora, APROVAM este trabalho no seu conteúdo, estilo e qualidade.

	NOME	ASSINATURA
1.	Dr.José E.Guisard Ferraz	
2.	Dr. Miguel Taube Netto	
3.	Dr. Derli C. Machado da Silva	
4.		
5.		
6.		
7.		
8.		
9.		
10.		

PARTICIPANTE(S): Autores

1.	NIVEA TEIXEIRA BATISTA	
2.	Heiko Humann	
3.	Lauro Tadeu G. Fontes	
4.	Jonas de Oliveira Junior	
5.		
6.		

São José dos Campos, 04 de junho de 19 74

Ricardo G. R. Palmeira

Chefe da Divisão de Ensino

AGRADECIMENTOS

Ao Instituto de Pesquisas Espaciais, na pessoa do seu Diretor Geral, Dr. Fernando de Mendonça, pela oportunidade da realização deste trabalho.

Ao Dr. José Eugênio Guisard Ferraz pela orientação e incentivo.

Aos Drs. Ward Edwards e Miguel Taube Netto pelos ensinamentos e sugestões.

À Standard Ogilvy & Mather, na pessoa do Sr. Jerry Wayne Selph, pela oportunidade de realização do estudo apresentado no Capítulo VI.

A Eric Jan Roorda pela colaboração e implementação do Programa de Processamento de Árvores de Decisão.

Às Srtas. Teresa Tokiko Takahashi, Maria Lourdes Martins de Carvalho e Nilze Maria Brandão, pelos serviços de datilografia.

SUMÁRIO

O presente manual tem por finalidade a apresentação dos conceitos básicos de Análise de Decisões.

Na sua elaboração, tentou-se obter um texto que fosse rigoroso o suficiente para ser adotado como texto de um Curso do Programa de Mestrado e, ao mesmo tempo, fosse acessível a leitores não especialistas no assunto como, por exemplo, Administradores e Gerentes.

No primeiro capítulo procura-se, com um exemplo simples, motivar o leitor para o problema de Análise de Decisão, sendo então apresentados os conceitos fundamentais, como: Árvore de Decisão, Atos, Consequências, Valor Esperado, Valor da Informação, etc.

Nos capítulos seguintes, são abordados, em detalhe, os principais aspectos da Teoria de Decisão, ou seja: Probabilidade e Utilidade.

Inicialmente, é feita uma revisão dos conceitos de Probabilidade e introduzida a noção de Probabilidade Subjetiva. Em seguida, são apresentadas as técnicas para levantamento das curvas de distribuição de probabilidades subjetivas, bem como de sua eventual revisão (do ponto de vista Bayesiano).

Nos capítulos dedicados à Teoria de Utilidade, é feita inicialmente a apresentação dos axiomas básicos da mesma. A seguir são propostos vários métodos práticos de levantamento de curvas de utilidade. Em particular, é estudado o problema de Análise Multidimensional.

Uma segunda parte do manual é dedicada à apresentação de exemplos de aplicação das técnicas anteriormente estudadas.

A área de levantamento de distribuição de Probabilidade Subjetiva é exemplificada com um estudo do impacto de uma campanha de publicidade sobre a procura de um certo produto.

Como exemplo de aplicação da Teoria de Utilidade, é apresentado um modelo de avaliação de universidades.

Em apêndice, é apresentado um programa para processamento de Árvores de Decisão.

ABSTRACT

The fundamental concepts involved in Decision Analysis are presented rigorously enough to be used as a textbook in a graduate course, but also accessible to the non-specialized reader.

The first chapter motivates the reader (through the use of a simple example) to the kind of problems Decision Analysis deals with; it also presents basic concepts such as decision trees, acts, consequences, expected value and value of information.

In the remaining chapters it is attempted to explain with more detail the most important subjects concerned with Decision Analysis, that is, probability and utility. At first the fundamental notions of probability theory are presented introducing the reader to the idea of subjective probability; the subject if further expanded upon by investigating the general techniques used in obtaining the distribution curves of subjective probabilities and revising subjective probabilities in light of new information using the bayesian approach. Then, in chapters dedicated to utility theory the reader is introduced to the basic utility theory axioms and some practical methods of encoding utility curves. The problem of Multidimensional Utility Analysis is also studied.

Finally an application of these techniques to real situations is shown through: (a) probability encoding procedures applied to study the effects of an advertising campaign on the volume of sales of a certain product; and (b) the use of utility theory to construct an evaluation procedure for universities.

A computer program to solve decision trees is presented as an appendix.

ÍNDICE

LISTA DE FIGURAS

LISTA DE TABELAS

PARTE I

INTRODUÇÃO

CAPÍTULO I

INTRODUÇÃO À ANÁLISE DE DECISÕES

1.1 - SITUAÇÃO DO PROBLEMA

Diariamente nos defrontamos com a necessidade de tomar decisões em que, nem sempre, as consequências da ação que adotamos são conhecidas com certeza. Frequentemente, nossa decisão deve ser feita antes que algum evento aleatório ocorra, o que gera incerteza sobre o resultado que obteremos, pois estes dependerão desses eventos. Nossa decisão consiste em selecionar um curso de ação e somente o acaso decidirá sobre os resultados.

Assim, um problema de decisão existe quando uma escolha deve ser feita entre ações alternativas com consequências incertas que dependem de eventos aleatórios, ou seja, sobre os quais o decisor não tem controle; a esses eventos chamaremos estados da natureza.

Em administração, decisões sob incerteza são bastante comun. Podemos dizer que quase todas as decisões importantes em administração são tomadas sob condições de incerteza. Por exemplo, temos o caso do fabricante que deve decidir sobre o lançamento de um novo produto, não conhecendo ao certo qual será a demanda por esse produto. Um outro caso, no qual também existe incerteza sobre a demanda, é o problema de

quanto estocar de um certo produto, em um supermercado. Se o estoque é menor que a demanda, o vendedor perde a oportunidade de vender, deixa de obter lucros, além dos fregueses ficarem insatisfeitos. Por outro lado, se ele estoca em excesso, podem advir prejuízos. Inúmeros outros exemplos podem ser citados de situações em que o decisor precisa selecionar um curso definido de ação entre vários disponíveis, sem saber se terá um bom resultado ou não.

Neste ponto é importante salientar que existe uma distinção entre uma má decisão e um mau resultado. Pode acontecer que, mesmo que uma boa decisão seja feita, um resultado satisfatório não seja obtido, já que as consequências da decisão são controladas pelo acaso. Assim, vemos que a incerteza é quase sempre um elemento chave na estrutura do problema, ou seja, todo problema de decisão envolve pelo menos um elemento de incerteza. Naturalmente, há casos em que ela é tão pequena que pode ser ignorada. Neste caso, o problema é tratado como uma decisão determinística e para resolvê-lo métodos matemáticos têm sido desenvolvidos, dos quais o mais difundido é a Programação Matemática.

Para análise de problemas em que a incerteza envolvida é muito grande, usamos a Análise de Decisões. Portanto, podemos dizer que ela tem por objetivo melhorar o processo de tomada de decisões, através de uma análise sistemática e racional dos problemas.

Vimos que os resultados de uma decisão dependem de qual evento, dentre um conjunto deles, ocorre, ou seja, de qual estado da natureza é verdadeiro. Geralmente, o decisor possui apenas alguma informação sobre a possibilidade de cada um desses eventos ocorrer. Essa informação possibilita a ele atribuir uma probabilidade a cada um dos possíveis estados da natureza. Assumindo que o ato ótimo, para o decisor, depende das probabilidades de que cada consequência associada a ele venha a ocorrer e das preferências do decisor com relação a essas consequências, determinamos esse ato ótimo.

Um aspecto importante é que, quase sempre, é possível ao responsável pela decisão melhorar seu conhecimento sobre os estados da natureza através da aquisição de novas informações. Para obter essas informações, o decisor terá um certo custo. Um dos objetivos da análise é determinar até que ponto é válido pagar por essa informação, ou seja, qual o valor econômico da informação adicional obtida através de experimentação ou amostragem.

Podemos, portanto, dividir o processo de Análise de Decisões em três fases, o que pode ser melhor visualizado na figura 1.1.

Com base nas informações disponíveis, são estabelecidas, na primeira fase, as relações entre as variáveis: tanto as controladas pelo decisor quanto aquelas sobre as quais ele não exerce controle. As primeiras são as chamadas variáveis de decisão e as seguintes variáveis

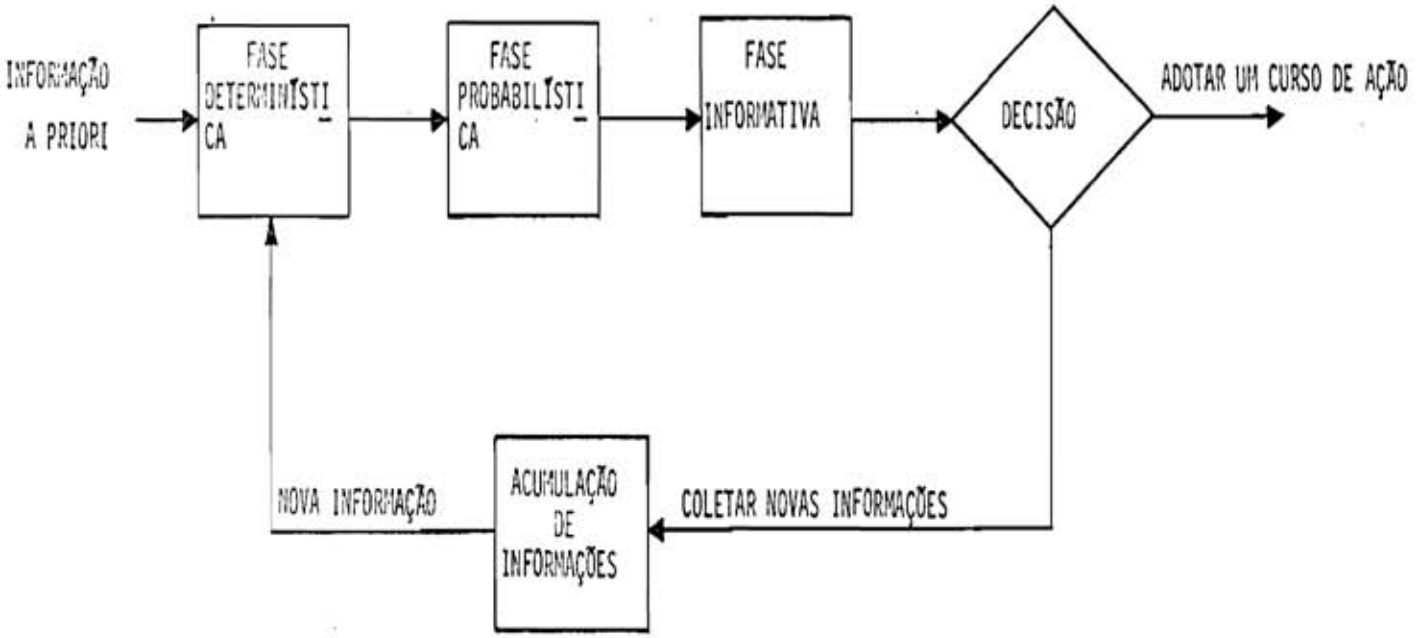


Figura 1.1 - Ciclo da Análise de Decisões.

de estado.

Nesta fase, é feita a sensitividade determinística. Incorporando-se medidas de preferência num modelo determinístico construído para o problema em análise, obtemos um certo valor. Testes de sensitividade consistem, então, em fixar as variáveis de estado, variando-se apenas uma e verificando-se a variação no valor que é fornecido pelo modelo. Isto nos permite saber até que ponto este valor é sensível a variações nas variáveis de estado, ou seja, quais dessas variáveis são críticas para o valor em questão. Esses testes são feitos para todas as variáveis de estado, sendo interessante também variar duas ou mais conjuntamente. Com isso, podemos eliminar todas aquelas variáveis às quais os resultados não são muito sensíveis.

Na fase probabilística, o decisor quantifica a incerteza existente atribuindo distribuições de probabilidade às variáveis de estado. Análise de sensitividade estocástica é feita nesta fase, admitindo-se que uma das variáveis de estado é conhecida e determinando-se distribuições de probabilidade condicionais para as demais variáveis. O efeito dessa variável no resultado é então medido. Segundo Ronald Howard [19] esta análise é de grande importância pois é ela que nos dirá se a obtenção de novas informações sobre esta variável é necessária. Além disso, ela pode também mostrar que variáveis consideradas críticas na fase anterior podem perder parte de sua importância num ambiente não determinístico.

Finalmente, na terceira fase, é determinado o valor econômico de se obter informações adicionais sobre as variáveis de estado. A comparação do valor de um experimento com seu custo permite ao decisor determinar se ele deve obter novas informações ou se deve adotar um determinado curso de ação. Se ele decide pela primeira alternativa, a nova informação deve ser incorporada à estrutura do problema, repetindo se o ciclo.

Nas próximas seções, serão introduzidos vários conceitos e, para facilitar seu entendimento, usaremos o seguinte exemplo:

Um fabricante deve decidir se fabrica ou não determinado produto. Caso ele decida pela fabricação, terá um lucro de Cr$300.000,00 se o produto for aceito pelo mercado consumidor e um prejuízo de Cr$ 100.000,00, caso não seja aceito. Baseando-se nas informações disponíveis sobre o mercado, ele acredita que há 70% de chance do produto ser aceito.

Com um determinado custo, o fabricante pode realizar uma pesquisa de mercado que lhe dará maiores informações sobre a probabilidade do produto ser aceito, embora essas informações não sejam totalmente confiáveis. No entanto, resultados de pesquisas feitas anteriormente pela mesma organização que conduzirá a pesquisa, mostram que ela acerta em 90% das vêzes.

O problema, então, consiste em determinar a sequência ótima de decisões e, através da determinação de ganhos adicionais, quanto o fabricante pode pagar pela pesquisa.

Um ponto que cabe ressaltar é que trataremos apenas de decisões individuais. Como decisão individual é considerada aquela em que há apenas um interesse em jogo, isto é, há apenas um decisor. Este pode ser um indivíduo, um grupo deles ou uma organização. Decisões em que há vários decisores com interesses conflitantes, denominadas "decisões em grupo", não serão tratadas neste manual.

1.2 - DADOS BÁSICOS

Os componentes básicos de um problema de decisão são:

a) Um conjunto $A = \{a_1, a_2, \ldots, a_n\}$ dos possíveis atos, onde ato é toda escolha disponível ao decisor à qual está associada diretamente uma recompensa.

b) O conjunto $S = \{s_1, s_2, \ldots, s_k\}$ dos estados da natureza.

c) O conjunto $C = \{c_{ij}\}$, $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, n$ das consequências ou recompensas, que dependerão de que estado ocorre e do ato adotado. Assim, c_{ij} é a recompensa obtida para o ato j se s_i ocorre.

O conjunto C é uma matriz onde as linhas estão relacionadas aos estados e as colunas aos atos alternativos que o decisor pode adotar, como se pode ver na tabela abaixo:

TABELA 1.1

Matriz de Recompensas

Estado	ATOS				
	a_1	a_2	a_3	...	a_n
s_1	c_{11}	c_{12}	c_{13}	...	c_{1n}
s_2	c_{21}	c_{22}	c_{23}	...	c_{2n}
⋮	⋮	⋮	⋮	⋮	⋮
s_k	c_{k1}	c_{k2}	c_{k3}	...	c_{kn}

d) Um conjunto $E = \{e_0, e_1, \ldots, e_m\}$, cujos elementos são os experimentos que podem ser realizados pelo decisor com o fim de melhorar seu conhecimento sobre os estados da natureza. O elemento e_0 representa a opção de não realizar experimentos.

e) O conjunto $R = \{r_1, r_2, \ldots, r_p\}$ dos resultados experimentais.

Para nosso exemplo, os dados básicos são:

Atos { a_1: não fabricar o produto
a_2: fabricar

Estados { s_1: o produto é aceito
s_2: o produto não é aceito

As consequências podem ser vistas na tabela 1.2.

TABELA 1.2

Matriz de Recompensas - Exemplo

Estados	ATOS	
	a_1	a_2
s_1	0	300.000
s_2	0	-100.000

Experimentos { e_0: não realizar pesquisa
e_1: realizar pesquisa

Resultados Experimentais { r_1: o produto será aceito
r_2: não será aceito

1.3 - ÁRVORES DE DECISÃO

O fluxo de decisão no nosso exemplo seria: o fabricante deve adotar um ato, sem realizar a pesquisa, ou fazer a pesquisa de mercado. Se ele decide pela primeira alternativa e escolhe a_1 ou a_2, o acaso decidirá qual evento ocorre. Caso ele realize a pesquisa, a decisão seguinte é do acaso: a pesquisa conclui que o produto será aceito ou não. Então, o fabricante opta por a_1 ou a_2 e temos, daí para a frente, um fluxo semelhante ao de não realizar a pesquisa.

Um fluxo de decisão pode ser visto melhor por meio de um diagrama denominado árvore de decisão, denominação proveniente da estrutura ramificada do gráfico.

Para a construção de uma árvore de decisão, as seguintes convenções são usadas:

a) O símbolo □ representa uma decisão a ser tomada pelo decisor.

b) O símbolo ○ representa uma decisão do acaso.

Para o nosso exemplo, temos o diagrama apresentado na figura 1.2.

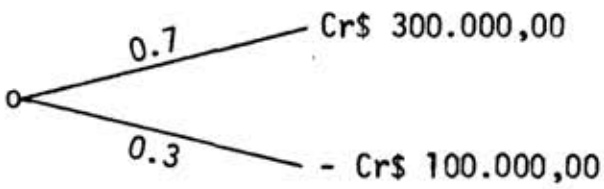
A todos os ramos oriundos de nós onde há decisão do aca

so, devemos associar probabilidades.

Esse diagrama, embora só seja útil em problemas onde há poucas alternativas de ação e poucos estados, facilita enormemente a análise, porque possibilita ao decisor ter sempre em mente a estrutura lógica do problema.

1.4 - DETERMINAÇÃO DA DECISÃO ÓTIMA

Na figura 1.2, notamos que, se o decisor não realiza a pesquisa e fabrica o produto, decisão que ele toma baseando-se na sua informação atual sobre os gostos do consumidor, ele chega ao ponto A. Neste ponto, o que ele tem a sua frente é uma espécie de jogo em que, se o produto for aceito, ele ganha um certo valor e, caso contrário, tem uma perda. Tal "jogo" é chamado uma loteria. Assim, nosso fabricante tem a seguinte loteria:



Um outro exemplo de loteria seria o caso em que temos, numa caixa, 10 bolas, das quais 8 são azuis e 2 brancas. Retiramos uma bola ao acaso; se ela é azul ganhamos Cr$ 10,00, caso contrário, perdemos Cr$ 5,00. Essa loteria pode, então, ser representada por:

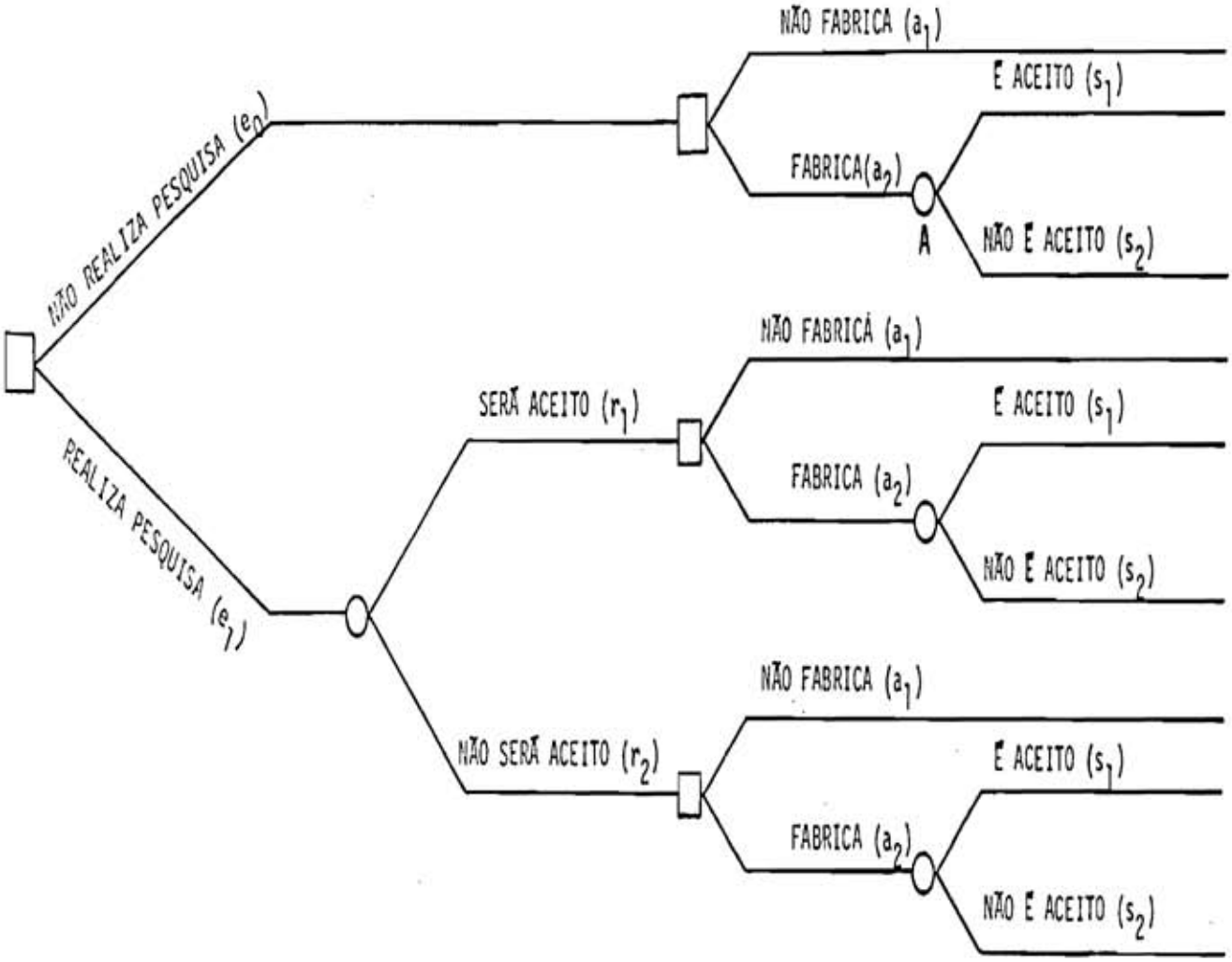
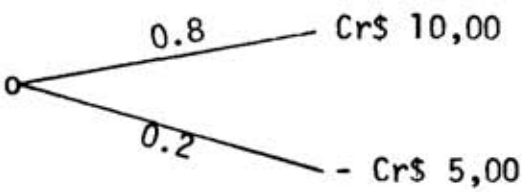


Figura 1.2 - Árvore de Decisões.



Para chegar-se à alternativa ótima, todos os cursos de ação disponíveis ao decisor são analisados separadamente determinando-se um valor para cada um deles. Estes valores são, então, comparados, sendo considerada ótima a alternativa com o maior valor. Nos itens seguintes, veremos os vários aspectos sob os quais atos alternativos podem ser analisados.

1.4.1 - Atribuição de Probabilidades

Para a análise que faremos, necessitamos das distribuições de probabilidade tanto dos resultados experimentais, quanto dos estados da natureza.

Assim, podemos sumarizar as informações disponíveis na tabela 1.3, abaixo:

TABELA 1.3

Probabilidades Condicionais

Resultados da Pesquisa	ESTADO REAL	
	s_1	s_2
r_1	0.9	0.1
r_2	0.1	0.9

O elemento na linha k, coluna i, representa a probabilidade do resultado da pesquisa ser r_k dado que o estado é s_i. Assim, o número 0.9 na primeira linha e primeira coluna representa a probabilidade da pesquisa concluir que o produto será aceito, dado que ele realmente o será.

Temos, ainda, que:

$$P(s_1) = 0.7$$
$$P(s_2) = 0.3$$

As probabilidades acima são denominadas "probabilidades a priori" dos estados s_1 e s_2, porque são aquelas de que se dispõe antes que qualquer informação adicional seja obtida. Usando as novas informações, podemos revisar as probabilidades a priori dos estados, valendo-nos da regra de Bayes.

Assim, se o resultado r_k é obtido e queremos determinar a probabilidade do estado ser s_i, temos:

$$P(s_i|r_k) = \frac{P(r_k|s_i) \cdot P(s_i)}{P(r_k)} \quad , \quad \begin{matrix} i = 1, 2 \\ k = 1, 2 \end{matrix}$$

e $$P(r_k) = P(r_k|s_1) \cdot P(s_1) + P(r_k|s_2)\, P(s_2)$$

$P(r_k|s_i)$ é encontrado na tabela 1.3.

$P(s_i|r_k)$ é chamada probabilidade "a posteriori" do estado s_i, ou seja, é a probabilidade revisada com base na informação fornecida pelo experimento.

Então, temos:

$$P(r_1) = 0,9 \times 0,7 + 0,1 \times 0,3 = 0,66$$

$$P(r_2) = 1 - P(r_1) = 0,34$$

$$P(s_1|r_1) = \frac{0,9 \times 0,7}{0,66} = 0,95$$

$$P(s_2|r_1) = 1 - P(s_1|r_1) = 0,05$$

$$P(s_1|r_2) = \frac{0,1 \times 0,7}{0,34} = 0,2$$

$$P(s_2|r_2) = 1 - P(s_1|r_2) = 0,8$$

1.4.2 - Valor Monetário Esperado

Para qualquer loteria $L \equiv (c_1p_1, c_2p_2, \ldots, c_np_n)$, onde c_i é o i-ésimo resultado possível e p_i a probabilidade desse resultado, $i = 1, 2, \ldots, n$, definimos seu valor esperado como:

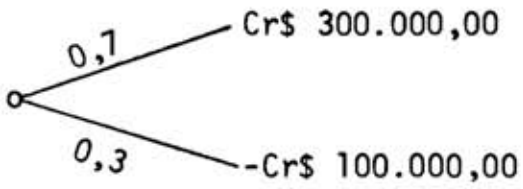
$$\overline{V}(L) = \sum_{i=1}^{n} c_i p_i$$

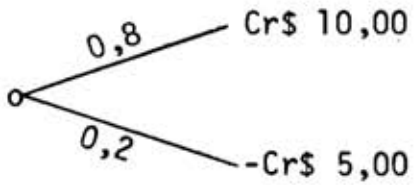
Por exemplo, o valor esperado da loteria L_1 abaixo é:

$$\overline{V}(L_1) = 0{,}7\ (\text{Cr\$ } 300.000{,}00) + 0{,}3\ (-\text{Cr\$ } 100.000{,}00)$$

$$\therefore \quad \overline{V}(L_1) = \text{Cr\$ } 180.000{,}00$$

onde

$L_1 \equiv$ 

A loteria $L_2 \equiv$ 

tem valor esperado $\overline{V}(L_2) = 0{,}8\ (\text{Cr\$ } 10{,}00) + 0{,}2\ (-\text{Cr\$ } 5{,}00)$

$\therefore \quad \overline{V}(L_2) = \text{Cr\$ } 7{,}00.$

Para qualquer ato a_j, definimos, então, seu Valor Monetário Esperado, como:

$$\overline{VM}(a_j) = \sum_{i=1}^{k} p_i\ c_{ij}$$

onde $p_i \equiv$ probabilidade de que o estado s_i ocorra.

$c_{ij} \equiv$ i-ésima consequência associada ao ato j.

Se, na figura 1.2, incorporarmos as recompensas e as probabilidades e calcularmos o valor esperado para cada ato, podemos, através da comparação desses valores, determinar o ato ótimo. Este é definido como aquele que tem o maior valor.

Assim, se $\overline{VM}^*$ é o valor esperado do ato ótimo, temos que:

$$\overline{VM}^* = \max_j \overline{VM}(a_j) = \max_j \sum_i p_i c_{ij}$$

Na figura 1.3, temos a mesma árvore anterior, agora com as probabilidades nos vários ramos onde há incerteza e as consequências associadas a cada par (ato, estado).

Vemos que, não realizando a pesquisa, o lucro é máximo se o fabricante decide por a_2, pois:

$$\overline{VM}(a_1) = 0$$

$$\overline{VM}(a_2) = \text{Cr\$ } 180.000,00$$

e $\overline{VM}^* = \max(0, \text{Cr\$ } 180.000,00) = \text{Cr\$ } 180.000,00$

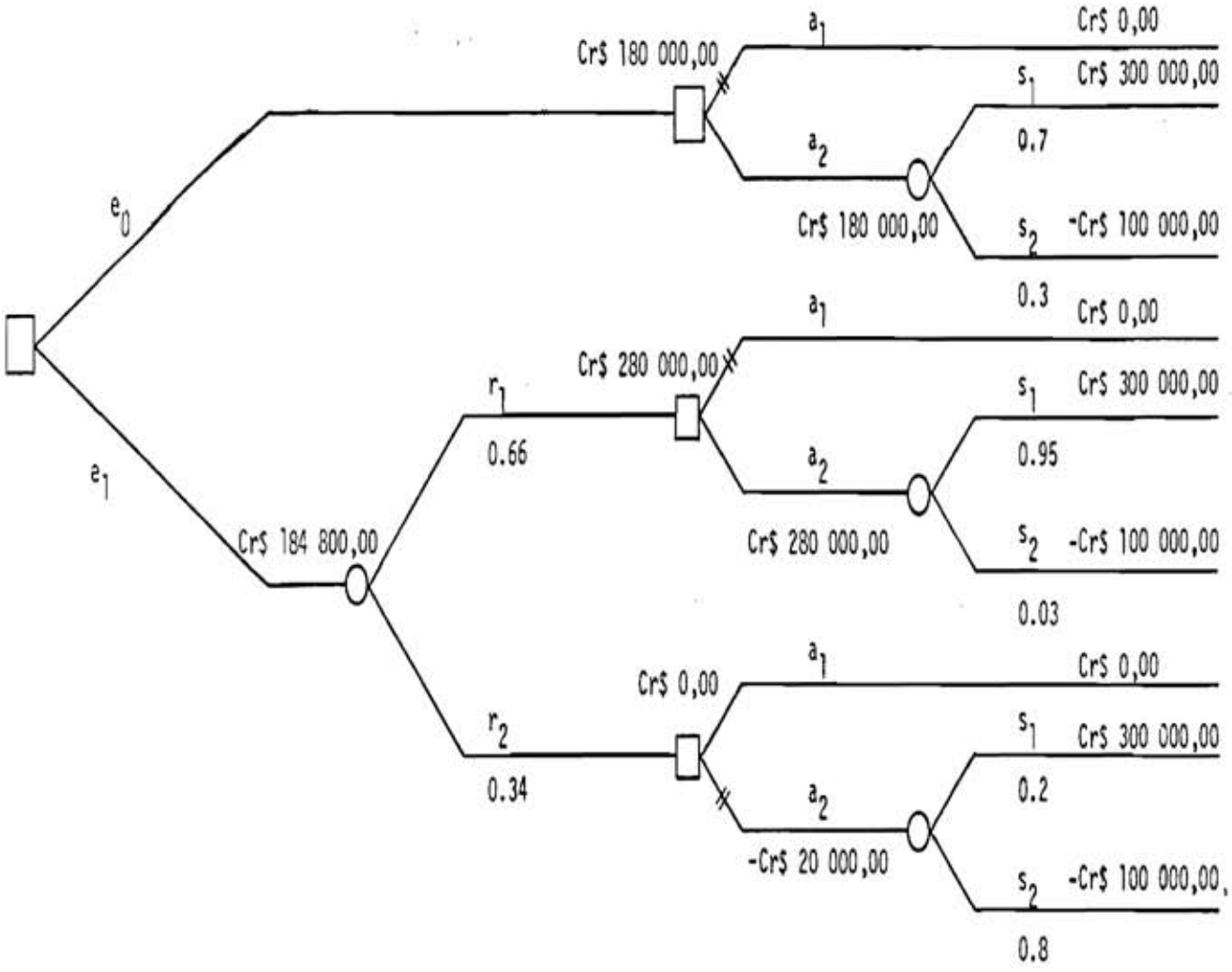


Figura 1.3 - Árvore de Decisões Processada.

Portanto, o ato ótimo é a_2, a_1 sendo eliminado como se pode ver na figura 1.3.

Se a pesquisa é realizada, o resultado desta não está sob o controle do decisor sendo, portanto, incerto. Se o resultado é r_1, isto é, a pesquisa diz que o produto será aceito, o ato ótimo será a_2, já que:

$$\overline{VM}(a_1) = \text{Cr\$ } 0{,}00$$

$$\overline{VM}(a_2) = \text{Cr\$ } 280.000{,}00$$

e $$\overline{VM}^* = \text{Cr\$ } 280.000{,}00$$

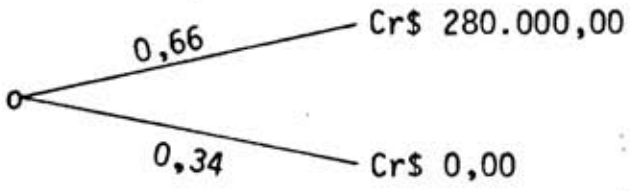
Se a pesquisa conclui que o produto não será aceito, o ato ótimo é não fabricar o produto (a_1), pois:

$$\overline{VM}(a_1) = \text{Cr\$ } 0{,}00$$

$$\overline{VM}(a_2) = -\text{Cr\$ } 20.000{,}00$$

e $$\overline{VM}^* = \text{Cr\$ } 0{,}00$$

Temos, neste ponto, uma loteria, qual seja:



O valor esperado dessa loteria é:

$$0{,}66 \times Cr\$\ 280.000{,}00 + 0{,}34 \times Cr\$\ 0{,}00 = Cr\$\ 184.800{,}00$$

que representa o ganho médio obtido antes de ser desembolsado o custo da pesquisa, quando esta é realizada.

1.4.3 - Valor da Informação Imperfeita

Antes de decidir se conduz ou não um experimento destinado a melhorar sua informação sobre os estados da natureza, o decisor necessita algum conhecimento dos ganhos adicionais decorrentes das informações que poderá obter com o experimento. Somente comparando estes acréscimos com o custo do experimento é que ele tem condições de decidir. Consideremos o caso, por exemplo, do lançamento de uma nave espacial. Naturalmente, não é sensato nem econômico realizar vários lançamentos, ou mesmo um, para se ter uma idéia da probabilidade de sucesso, dado o alto custo do experimento. Em tais casos, o que comumente se faz é desenvolver a pesquisa voltada para o assunto. Desejamos, então, determinar quanto investir em pesquisa e, no caso de realizar experimentos, quanto pagar por essa realização.

Isto é feito, como dissemos, comparando o acréscimo no ganho esperado, decorrente da melhoria de nossa informação sobre determinado evento, com o custo do experimento ou da pesquisa. Assim, definimos o Valor da Informação Imperfeita como a diferença entre o ganho esperado quando se realiza o experimento e aquele que se espera obter sem experimentação.

No nosso exemplo, o que queremos é determinar até quanto vale pagar pela pesquisa de mercado. Não realizando a pesquisa, o fabricante tem um lucro esperado de Cr$ 180.000,00 ao passo que, com a pesquisa, seu ganho médio é Cr$ 184.800,00. O valor da pesquisa é, então, Cr$ 4.800,00 e o fabricante não deverá pagar mais do que esse valor por ela.

Assim, se o custo da pesquisa é menor que Cr$ 4.800,00, o fluxo ótimo de decisão é realizar a pesquisa; se o resultado indicar que o produto será aceito, então o produto deve ser fabricado; caso contrário, o ato ótimo é a_1.

Se o custo é maior que Cr$ 4.800,00, a pesquisa não deve ser realizada e o ato ótimo é a_2.

1.4.4 - Valor Esperado da Informação Perfeita

Suponhamos que o fabricante tenha a possibilidade de saber qual estado é real. Conhecendo isto, ele adota o ato ótimo. Por exemplo, se s_1 fosse o estado real ele escolheria a_2, tendo um lucro de Cr$ 300.000,00; dado que s_2 ocorrerá, ele escolhe a_1, seu lucro sendo nulo. Como s_1 ocorre com probabilidade 0,7 e s_2 com probabilidade 0,3, seu ganho médio, será:

$$0,7 \times \text{Cr\$ } 300.000,00 + 0,3 \times \text{Cr\$ } 0,00 = \text{Cr\$ } 210.000,00$$

Se o fabricante decide apenas com base na informação atual, sem realizar a pesquisa, seu ganho médio será Cr$ 180.000,00. A diferença entre o valor médio que ele espera obter quando toda a incerteza é eliminada e o valor esperado do ato ótimo, sem experimentação, é chamada de Valor Esperado da Informação Perfeita ($\overline{VIP}$).

No caso, temos:

$$\overline{VIP} = \text{Cr\$ } 210.000,00 - \text{Cr\$ } 180.000,00 = \text{Cr\$ } 30.000,00$$

Se essa informação, que lhe possibilita o conhecimento certo do estado real, tiver um preço inferior a Cr$ 30.000,00 ele a comprará. Portanto, este valor é o máximo que ele paga pela informação. Deve-se notar que, como essa informação possibilita ao decisor conhecer

com certeza os resultados da decisão que tomar, ela vale mais que qualquer outro tipo de informação. Assim, temos que

$$\overline{VIP} \geq \text{Valor da Informação Imperfeita}$$

Por qualquer informação a que possa ter acesso, o fabricante pagará no máximo Cr$ 30.000,00, que é o quanto vale para ele saber ao certo se seu produto será aceito ou não.

Para determinar $\overline{VIP}$ existe um outro método que veremos na seção seguinte.

1.4.5 - Perdas de Oportunidade

Com muita frequência, é mais simples determinar atos ótimos através das "perdas de oportunidade". Estas são definidas como a diferença entre o lucro condicional do ato ótimo, dado um evento particular e o lucro condicional de qualquer outro ato, dado o mesmo evento. Lucro condicional é aquele que é obtido se ocorre um evento particular e é definido para cada par (ato, estado). Dado que um certo evento ocorre, o decisor adota o ato cujo valor é máximo. Adotando outro ato (não ótimo), ele perde a oportunidade de obter um lucro maior e essa quantia que ele deixa de ganhar é chamada perda de oportunidade.

Na tabela 1.4, temos os lucros condicionais, para o ramo

e_o do nosso exemplo. Para cada estado s_i, o lucro do ato ótimo está assinalado com um asterisco.

TABELA 1.4

Lucros Condicionais

Estados	ATOS	
	a_1	a_2
s_1	0	300.000*
s_2	0*	-100.000

Perdas de oportunidade podem ser calculadas pela fórmula:

$$PO_{ij} = (\max_j c_{ij}) - c_{ij}$$

onde $PO_{ij} \equiv$ perda de oportunidade em que se incorre se o ato j é adotado e ocorre s_i.

Ainda para o ramo e_0, as perdas de oportunidade estão mostradas na tabela 1.5.

TABELA 1.5

Perdas de Oportunidade

Estados	ATOS	
	a_1	a_2
s_1	300.000	0
s_2	0	100.000

Dado que o produto não é aceito (s_2), o ato ótimo é não fabricá-lo (a_1); adotando este ato, o decisor tem uma perda nula, ao passo que, adotando a_2, sua perda será:

$$0 - (-100.000) = 100.000$$

Da mesma forma foram calculados os demais elementos, devendo-se notar que, para um ato ótimo, a perda de oportunidade é sempre nula e que, para os demais atos, ela é positiva.

Assim, vemos que a perda de oportunidade é a perda decorrente de se tomar uma decisão errada, isto é, aquela em que se incorre em virtude da má ou da pouca informação.

Como as perdas de oportunidade para um ato são condicionais a um certo estado, ou seja, são definidas para cada um dos estados, definimos a Perda de Oportunidade Esperada para um ato a_j, como:

$$\overline{P\,O}\,(a_j) = \sum_{i=1}^{k} PO_{ij}\, p_i = \sum_{i=1}^{k} (\max_j c_{ij} - c_{ij})\, p_i$$

onde $\overline{P\,O}\,(a_j)$ ≡ perda de oportunidade esperada do ato a_j.

PO_{ij} ≡ perda de oportunidade para o ato a_j, dado o estado s_i.

p_i ≡ probabilidade de s_i ser o estado real.

Para nosso exemplo, temos:

$$\overline{PO}\,(a_1) = 0{,}7\ (Cr\$\ 300.000{,}00) + 0{,}3\ (Cr\$\ 0{,}00) = Cr\$\ 210.000{,}00$$

$$\overline{PO}\,(a_2) = 0{,}7\ (Cr\$\ 0{,}00) + 0{,}3\ (Cr\$\ 100.000{,}00) = Cr\$\ 30.000{,}00$$

Na seção 1.4.4, em que definimos valor esperado da informação perfeita ($\overline{VIP}$), dissemos que havia uma outra maneira de definir $\overline{VIP}$, que é:

$$\overline{VIP} = \min_j \overline{PO}\,(a_j), \qquad j = 1, 2, \ldots, n$$

Assim, o Valor Esperado da Informação Perfeita é igual à menor perda de oportunidade esperada. No caso, esta é Cr$ 30.000,00, como já tínhamos visto.

Podemos concluir que o ato ótimo é aquele cuja perda de oportunidade esperada é mínima. Isto equivale a dizer que o ato ótimo é

o que tem o máximo valor monetário esperado, já que a soma do valor monetário esperado com a perda de oportunidade esperada é igual para todos os atos, ou seja:

$$\overline{VM}\,(a_j) + \overline{PO}\,(a_j) = k', \quad \forall_j$$

onde k' é o valor que se espera obter, em média, quando toda incerteza é eliminada,

$$k' = \sum_{i=1}^{k} (\max_j c_{ij})\, p_i \,.$$

Isto pode ser visto se lembrarmos que

$$\overline{PO}\,(a_j) = \sum_{i=1}^{k} (\max_j c_{ij} - c_{ij})\, p_i =$$

$$= \sum_{i=1}^{k} (\max_j c_{ij})\, p_i - \sum_{i=1}^{k} c_{ij}\, p_i$$

$$\therefore \quad \overline{PO}\,(a_j) = k' - \overline{VM}\,(a_j) \Longrightarrow \overline{VM}\,(a_j) + \overline{PO}\,(a_j) = k$$

No nosso exemplo, temos:

$\overline{PO}\,(a_1)$ = Cr\$ 210.000,00 $\qquad$ $\overline{VM}\,(a_1)$ = Cr\$ 0,00

$\overline{PO}\,(a_2)$ = Cr\$ 30.000,00 $\qquad$ $\overline{VM}\,(a_2)$ = Cr\$ 180.000,00

$$\overline{PO}\ (a_1) + \overline{VM}\ (a_1) = Cr\$\ 210.000,00$$

$$\overline{PO}\ (a_2) + \overline{VM}\ (a_2) = Cr\$\ 210.000,00$$

1.4.6 - Equivalentes Certos

Vemos, pela figura 1.2, que uma decisão complexa pode ser decomposta em decisões mais simples, até chegarmos a decisões elementares.

Tomemos, por exemplo, um ramo da árvore de decisões da figura 1.2. Seja o ramo e_0, representado na figura 1.4, abaixo:

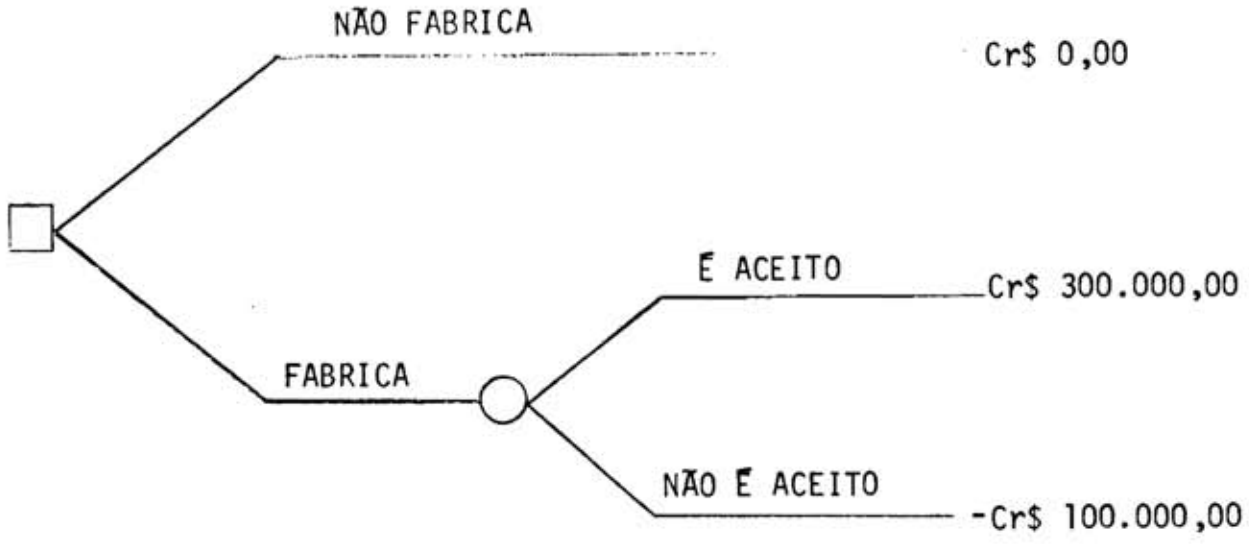


Figura 1.4 - Ramo e_0 da árvore de decisões do exemplo apresentado.

Para chegar a uma decisão final, o decisor precisa resolver todas as decisões deste tipo que porventura venham a aparecer.

O critério anteriormente apresentado para resolver tal tipo de decisão é que o decisor procura maximizar o valor monetário esperado, ou seja, de posse das probabilidades do seu produto ser aceito ou não, calcula o valor esperado da decisão "fabricar"; se este valor esperado for positivo, toma esta decisão; caso contrário, a decisão ótima é não fabricar o produto.

Na prática, poucas pessoas concordariam com tal critério. Os fatores que influenciam uma decisão são inúmeros e vão muito além dos simples valores monetários envolvidos. Podemos notar que, mesmo se a decisão de fabricar for tomada corretamente (no caso de seu valor esperado ser positivo), ela pode levar a um grande prejuízo, no caso do produto não ser aceito. Isto pode influir nos demais negócios da firma e na situação do fabricante que certamente levará este fato em conta ao tomar sua decisão.

O problema com que se defronta nosso fabricante neste ponto é o seguinte: será que a decisão de fabricar tem para ele, um valor igual ao valor esperado ? Em geral, a resposta é não. Os fatores acima mencionados fazem com que, para o fabricante, a decisão tenha um valor diferente do seu valor esperado; em geral, menor.

A este valor que, para o decisor, é equivalente à loteria que é constituida pela decisão de fabricar o produto, chamamos "Equivalente Certo" da loteria.

Quando o equivalente certo é menor que o valor monetário esperado, dizemos que o decisor tem aversão ao risco. Se é igual, trata se de um decisor "médio" ou indiferente ao risco. E, nos raros casos em que o equivalente certo é maior que o valor esperado, dizemos que o decisor tem atração pelo risco.

Nossa regra de decisão deve, então, ser alterada do seguinte modo:

Colocado frente a uma decisão elementar do tipo apresentado na figura 1.4, o decisor deve avaliar o seu equivalente certo para a loteria (levando em conta as recompensas, as probabilidades e suas atitudes frente ao risco). Caso este valor seja positivo, ou seja, maior do que a recompensa da outra decisão possível (não fabricar), ele deve decidir pela fabricação do produto.

Quando trabalhamos com equivalentes certos, o processamento é idêntico ao que se faria se estivéssemos trabalhando com valores esperados. A única diferença é que, nos diversos nós de decisão do acaso, ao invés de calcularmos o valor esperado, usamos equivalentes certos.

O que vemos é que, de uma maneira geral, as pessoas diferem em sua atitude frente ao risco e que seus padrões de valores também são diferentes. Assim, nem sempre valores monetários constituem um guia para ação. Então, é preciso uma medida de valor que leve em conta as preferências do indivíduo, bem como seu comportamento diante do risco. A tal medida denominaremos utilidade e esta será definida para cada recompensa associada a um ato. O critério de decisão será maximizar não o valor monetário esperado, mas a utilidade esperada.

Os Capítulos IV e V serão dedicados à teoria da utilidade.

1.5 - CONCLUSÕES

O princípio básico da análise que apresentamos é o de que, ao invés de tomar uma decisão no início do processo, o decisor define que ato adotaria se determinado evento tivesse ocorrido. Consideremos o ramo e_1 do exemplo que apresentamos anteriormente. Se o fabricante realiza uma pesquisa de mercado e esta conclui que o produto será aceito, que curso de ação ele adotaria ? Naturalmente, aquele que tem para ele o maior valor, ou seja, admitindo que este valor é $\overline{VM}$, ele adotaria a_2 (fabricar o produto). Se a pesquisa revelasse que o produto não seria aceito, o ato ótimo seria a_1 (não fabricar), já que, se ele adotasse a_2, teria um prejuízo esperado de Cr$ 20.000,00.

Através da eliminação de atos não ótimos e da atribuição

de valores aos vários nós do acaso, chegamos ao início da árvore, onde temos uma decisão entre alternativas simples cujos valores já estão determinados. A alternativa escolhida será aquela que tiver o maior valor, seja este o valor monetário esperado ou a utilidade esperada. Vemos, assim, que a análise é feita a partir dos extremos à direita da árvore até o início da mesma e a solução que obtemos é uma sequência ótima de decisões. Tal análise é denominada "análise em forma extensiva" e, segundo alguns, possui um sério inconveniente que é o de requerer a atribuição de probabilidades a todos os ramos em que ocorrem eventos aleatórios. Nem sempre é possível associar probabilidades "objetivas" a certos tipos de eventos e variáveis aleatórias. Usamos o termo "objetivas" para probabilidades atribuídas com base na repetição de um experimento, ou seja, probabilidades como limite da frequência relativa de ocorrência de um certo evento. Frequentemente, dispomos apenas de uma pessoa com informações sobre a ocorrência desse evento e usamos tais informações para determinar quão provável é essa ocorrência. Falamos, então, em "probabilidades subjetivas" e, como essas são estimativas grosseiras, há pessoas que não se sentem bem usando a análise extensiva, preferindo um outro tipo de análise denominada "análise em forma normal". Essa análise não requer a atribuição de probabilidades a todos os nós do acaso; ela trabalha com intervalos dentro dos quais podem estar as probabilidades necessárias. Para cada intervalo possível, determina-se uma "estratégia" ótima, onde estratégia é uma regra de decisão que diz ao decisor o que fazer em qualquer ponto, dado qualquer fato passado. Um exemplo de estratégia é, no caso do nosso fabricante: não realizar a pesquisa e

fabricar o produto a qual denotaremos por (e_0, a_2); outro exemplo seria $[e_1, (r_1, a_2), (r_2, a_1)]$, que significa realizar a pesquisa e, se o resultado for r_1, fabricar o produto; se for r_2, não fabricar.

Usando a análise normal, testamos a sensitividade do ato ótimo às probabilidades. No entanto, podemos também fazer isto com a análise extensiva. Variando, por exemplo, as probabilidades e realizando toda a análise novamente, podemos verificar se uma dada probabilidade é crítica, no sentido de que altera o ato ótimo. Isto talvez mostre que é interessante e válido realizar novos experimentos que aumentem a informação sobre o evento, permitindo, assim, revisar as probabilidades. Sobre análise em forma normal ver Raiffa [40] e Schlaifer [47].

Nos Capítulos II e III faremos, respectivamente, uma breve revisão dos conceitos de probabilidade e a apresentação de técnicas para codificação de incerteza, isto é, para determinação de distribuições de probabilidades subjetivas.

No início deste seção, dissemos que, partindo do final da árvore para o início, a cada nó do acaso, considerado como uma loteria, associamos um valor que é ou o valor monetário esperado, ou o valor dessa loteria para o decisor (equivalente certo) ou a utilidade esperada da loteria. Quando substituímos cada consequência por sua utilidade, processamos a árvore da mesma maneira que quando trabalhamos com os próprios valores monetários. O Capítulo IV será dedicado ao estudo da

teoria da utilidade e, no Capítulo V, apresentaremos técnicas para levantamento de curvas de utilidade.

PARTE II

TEORIA DA DECISÃO

CAPÍTULO II

REVISÃO DOS CONCEITOS DE PROBABILIDADES

2.1 - PROBABILIDADES "A PRIORI"

O termo "probabilidade a priori" foi usado por Jacob Bernoulli para definir probabilidades.

No lançamento de uma moeda honesta, se nos perguntarem qual a probabilidade de sair cara numa jogada, nossa intuição nos diz que esta probabilidade é $\frac{1}{2}$. Da mesma forma, se alguém nos perguntar sobre a probabilidade de obtermos um "dois" quando um dado honesto é lançado, responderemos que esta é igual a $\frac{1}{6}$. No entanto, se nos pedirem para justificar porque respondemos $\frac{1}{2}$ ou $\frac{1}{6}$, certamente nossa justificativa não será precisa.

É nesse sentido que é usado a denominação "probabilidades a priori" pois, ao respondermos que a probabilidade de sair "cara" no lançamento de uma moeda é $\frac{1}{2}$, estamos supondo que todos os resultados possíveis (cara e coroa) são igualmente prováveis de ocorrer.

O matemático francês P. S. Laplace [28] formulou este conceito, afirmando que "a teoria das chances consiste em reduzir todos os eventos do mesmo tipo a um certo número de casos igualmente possí

veis". Isso, em outras palavras, significa "reduzí-los de forma a ficarmos igualmente indecisos com relação à sua ocorrência e também de maneira que se possa determinar o número de casos favoráveis do evento cuja probabilidade se quer determinar. A razão entre esse número e o de casos possíveis é a medida da probabilidade, que consiste numa relação cujo numerador é o número de casos favoráveis e cujo denominador é o número de todos os casos possíveis".

Esse enfoque de probabilidade apresenta duas características:

1ª) É possível prever a ocorrência de qualquer um dos resultados do experimento. Por exemplo, no lançamento de um dado, admitindo que este é honesto, ao sermos interrogados sobre a probabilidade de obtermos a face "dois", dizemos que essa é $\frac{1}{6}$.

2ª) Ele se baseia em raciocínio abstrato, não dependendo da realização do experimento.

2.2 - PROBABILIDADE COMO LIMITE DA FREQUÊNCIA RELATIVA

Um outro enfoque de probabilidade é aquele baseado na repetitividade de um experimento.

Seja novamente o experimento "lançar uma moeda". Defina

mos os eventos: C "aparece cara" e K "aparece coroa". O experimento foi repetido 200 vezes, e os resultados obtidos acham-se apresentados na tabela 2.1, onde n é o número de lançamentos, n_C o número de vezes que o evento C ocorreu e f_C a frequência relativa da ocorrência de C.

TABELA 2.1

Resultados de sucessivos lançamentos de uma moeda

n	n_C	$f_C = \frac{n_C}{n}$
10	6	0,60
30	17	0,56
50	23	0,46
80	43	0,54
100	49	0,49
140	72	0,51
180	88	0,49
200	102	0,51

Na figura 2.1, temos a frequência relativa da ocorrência de C em relação ao número de jogadas. Notamos que f_C flutua consideravelmente quando n é pequeno e, à medida que n aumenta, ela tende a se estabilizar em torno do valor 0,5. Isto significa que a frequência relativa possui a característica denominada regularidade estatística, quando n é grande.

O valor em torno do qual f_C tende a se estabilizar é a

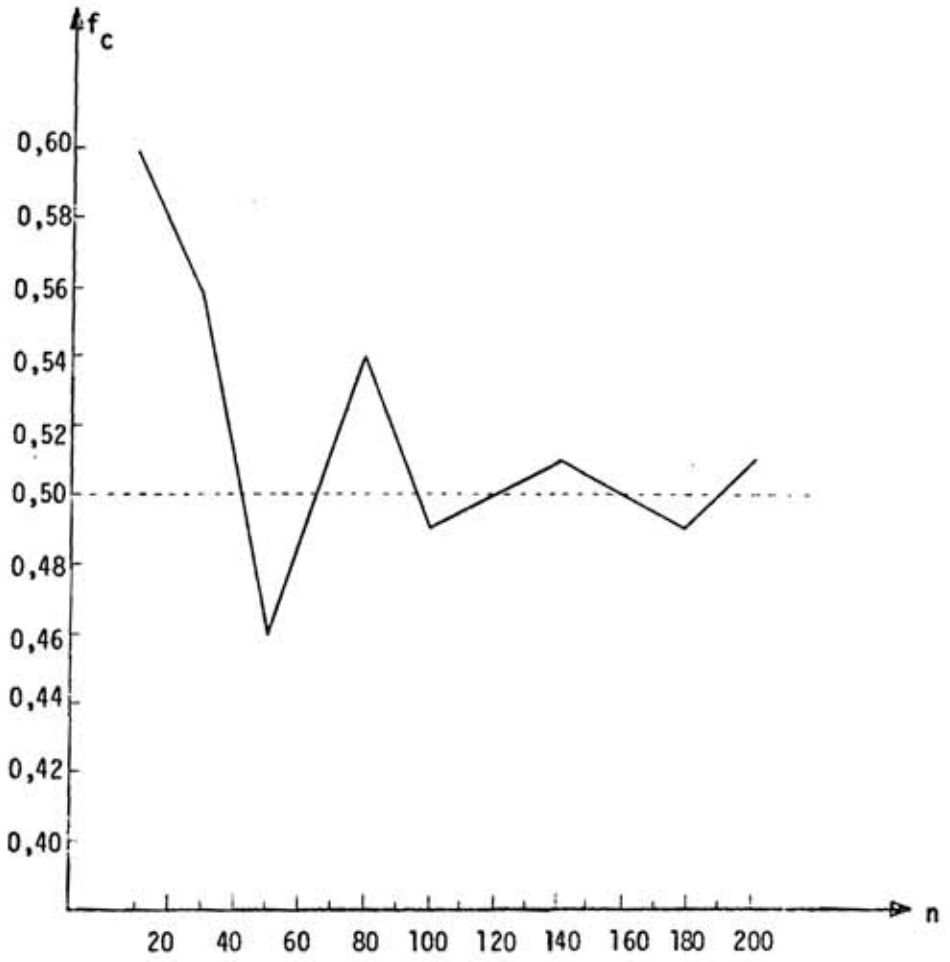


Figura 2.1 - Frequência Relativa da Ocorrência de Caras em n Lançamentos de uma Moeda.

probabilidade do evento C. No exemplo, P (C) = P (cara) = 0,5.

2.3 - CONCEITOS BÁSICOS

2.3.1 - Experimento Aleatório

Segundo Parzen [36], é um "fenômeno empírico, caracterizado pelo fato de que, sob certas circunstâncias, nem sempre provoca o mesmo resultado, mas os resultados obtidos, ainda que diferentes, conduzem a uma certa regularidade estatística".

Um experimento aleatório é aquele cujos resultados não podem ser previstos com certeza mas é possível enumerar todos os seus resultados possíveis. Além disso, um experimento aleatório pode, conceitualmente, ser repetido um grande número de vêzes sob condições similares.

2.3.2 - Espaço Amostral

A todo experimento aleatório E está associado um conjunto S de todos os resultados possíveis desse experimento. Tal conjunto é denominado o espaço amostral associado ao experimento.

Assim, quando lançamos uma moeda, os resultados possíveis são "cara" e "coroa". Portanto, o espaço amostral é:

$$S = \{ \text{cara, coroa} \}$$

2.3.3 - Evento

Qualquer subconjunto de um espaço amostral é um evento. Assim, no lançamento de uma moeda um resultado possível é "sai coroa". Assim, definimos o evento A, como:

$$A = \{ \text{sai coroa} \}$$

Deve-se observar que o espaço amostral S é também um evento. A probabilidade de um dado evento é a probabilidade de qualquer um dos seus elementos ocorrer.

Usando um diagrama de Venn, temos:

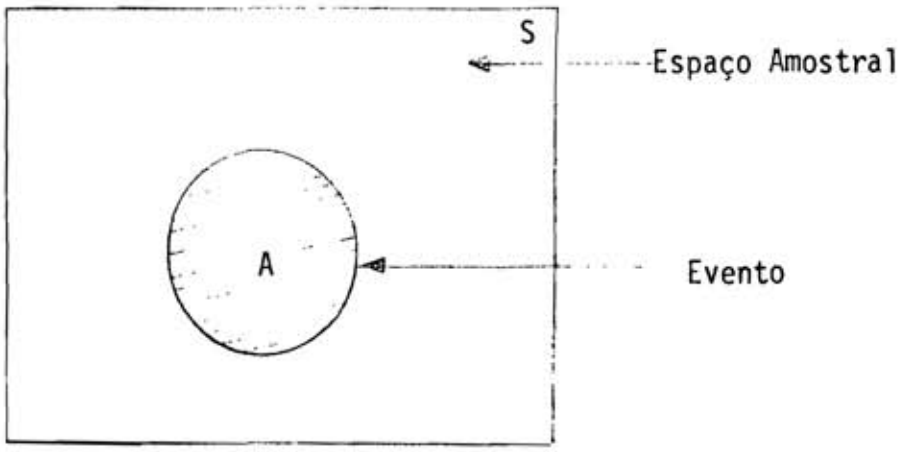


Figura 2.2 - Espaço Amostral e Evento.

2.3.4 - Combinação de Conjuntos

Quando tratamos com dois ou mais conjuntos, podemos realizar com eles as seguintes operações:

União

A união de dois conjuntos A e B, denotada por A U B, é o conjunto formado pelos elementos que pertencem a A ou a B, ou a ambos. Na figura 2.3, temos uma representação de A U B.

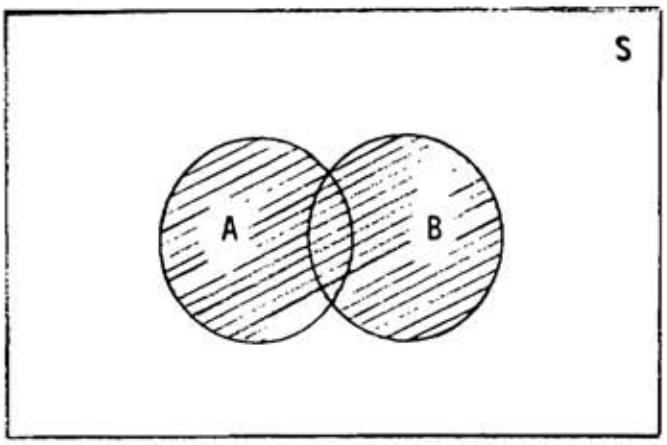


Figura 2.3 - Diagrama de Venn da União de Conjuntos.

Se C for uma coleção qualquer de conjuntos, a união de todos os conjuntos em C é definida como o conjunto dos elementos que pertencem a pelo menos um dos conjuntos de C e é denotada por

$$\bigcup_{A \in C} A$$

Intersecção

A intersecção de A e B, denotada por $A \cap B$ é o conjunto dos elementos que pertencem a A e a B. Se A e B não tiverem elementos em comum, $A \cap B$ é um conjunto vazio; dizemos, então, que A e B são conjuntos disjuntos. Na figura 2.4(a) temos a representação de uma intersecção não vazia de conjuntos enquanto que em 2.4(b)

$A \cap B = \emptyset$

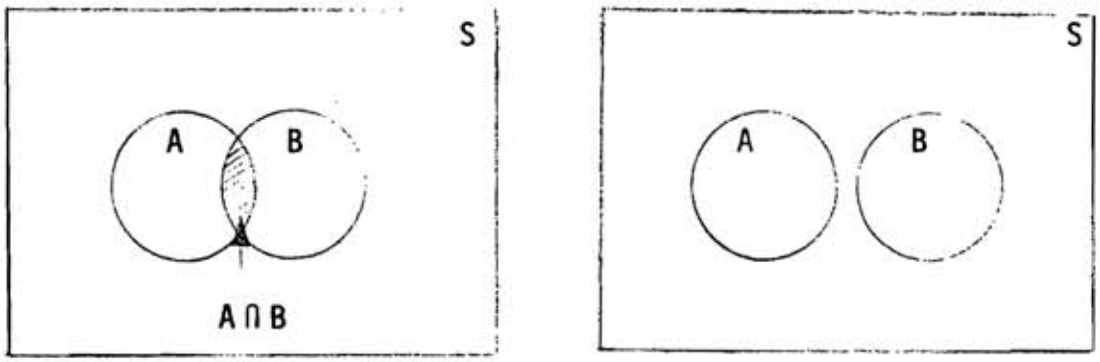


Figura 2.4 - Diagrama de Venn da Intersecção de Conjuntos.

No caso de C ser uma coleção qualquer de conjuntos, a intersecção de todos os conjuntos de C, denotada por $\bigcap_{A \in C} A$, é o conjunto dos elementos que pertencem a todos os conjuntos de C.

Conjuntos Complementares

O complemento de A em relação a B, denotado por B-A,

é definido como o conjunto dos elementos que pertencem a B e não pertencem a A. Na figura 2.5, a parte hachurada representa o complemento de A em relação a B.

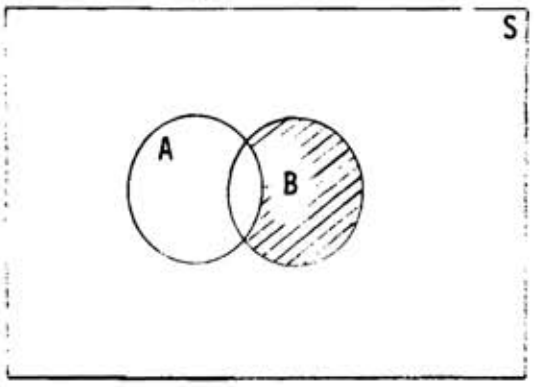


Figura 2.5 - Conjuntos Complementares.

Deve-se notar que $B - (B - A) = A$ sempre que $A \subset B$ e que $B - A = B$ se $A \cap B = \emptyset$. De uma forma geral, chamaremos "complemento de A" ao complemento de A em relação a S e o denotaremos simplesmente por $\overline{A}$.

2.4 - PROBABILIDADES ELEMENTARES E AXIOMAS DE PROBABILIDADES

Dado um evento A qualquer, a probabilidade da ocorrência de A é:

$$P(A) = \lim_{n \to \infty} f_A$$

Das propriedades de frequência relativa, obtemos os seguintes axiomas que devem ser satisfeitos por P (A):

Axioma 1:

$$0 \leq P(A) \leq 1$$

Axioma 2:

$P(S) = 1$, onde S representa o evento certo, ou seja, S é o espaço amostral.

Axioma 3:

$P(A \cup B) = P(A) + P(B)$, se A e B são disjuntos.

Dos axiomas acima, decorre um conjunto de propriedades, quais sejam:

Propriedade 1:

$P(\overline{A}) = 1 - P(A)$, onde $\overline{A}$ é o complemento de A. Como A e $\overline{A}$ são disjuntos e $A \cup \overline{A} = S$, vem:

$$P(S) = 1 = P(A \cup \overline{A}) = P(A) + P(\overline{A}) \Longrightarrow P(\overline{A}) = 1 - P(A)$$

Propriedade 2:

Se $\emptyset$ é um evento impossível de ocorrer, então $P(\emptyset) = 0$

Como $A = A \cup \emptyset$, temos:

$$P(A) = P(A) + P(\emptyset) \Longrightarrow P(\emptyset) = 0$$

Propriedade 3:

Se A e B são dois conjuntos quaisquer, então:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Sabemos que $A \cup B = A \cup (B \cap \overline{A})$ e $B = (A \cap B) \cup (B \cap \overline{A})$

Pelo axioma 3, temos:

$$P(A \cup B) = P\left[A \cup (B \cap \overline{A})\right] = P(A) + P(B \cap \overline{A}) \qquad (2.1)$$

e $$P(B) = P\left[(A \cup B) \cup (B \cap \overline{A})\right] = P(A \cap B) + P(B \cap \overline{A}) \qquad (2.2)$$

Subtraindo 2.2 de 2.1, vem:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

2.5 - EVENTOS MUTUAMENTE EXCLUSIVOS

Sejam os eventos $A_1, A_2, \ldots, A_n$ os resultados possíveis do experimento aleatório E. Dizemos que eles são mutuamente exclusivos se apenas um dos A_i, $i = 1, 2, \ldots, n$, pode ocorrer, em qualquer tentativa. Assim, dois eventos A_1 e A_2 são mutuamente exclusivos se sua intersecção é vazia, isto é, $A_1 \cap A_2 = \emptyset$; $A_1, A_2, \ldots, A_n$ são mutuamente exclusivos se suas intersecções dois a dois forem vazias.

Para obtermos a probabilidade de pelo menos um evento da coleção ocorrer, adicionamos as probabilidades desses eventos, ou seja.

$$P(A_1 \cup A_2 \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

Exemplo: Suponhamos que nas eleições para um clube, duas chapas, A e B, concorram e cada uma apresenta 2 candidatos para o cargo de presidente do clube. Pelo regulamento das eleições, não pode haver empate. A probabilidade de cada candidato vencer é:

Candidato 1 da chapa A : 0,20

Candidato 2 da chapa A : 0,12

Candidato 1 da chapa B : 0,54

Candidato 2 da chapa B : 0,14

Como somente um candidato pode vencer, a probabilidade de

vitória da chapa A é:

$$P(\text{candidato 1 de A}) + P(\text{candidato 2 de A}) = 0{,}20 + 0{,}12 = 0{,}32$$

A probabilidade da chapa B vencer é: 0,54 + 0,14 = 0,68, enquanto que a de vencer o candidato 2 da chapa A ou o candidato 1 da chapa B é 0,12 + 0,54 = 0,66.

2.6 - PROBABILIDADES CONDICIONAIS

Sejam A e B dois eventos pertencentes ao espaço amostral associado ao experimento E. A probabilidade de B ocorrer quando A tiver ocorrido é dada por:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \qquad P(A) > 0$$

Assim $P(A \cap B) = P(B|A) \cdot P(A)$, o que é conhecido na literatura como "Teorema da Multiplicação de Probabilidades".

Suponhamos que um escritório possua 100 funcionários. Alguns deles fumam (F), outros não (N); e, dentre eles, alguns são homens (H), outros são mulheres (M). A tabela 2.2 dá o número de funcionários de acordo com a classificação em fumantes, não fumantes e separando-os de acordo com o sexo.

TABELA 2.2

Classificação dos Funcionários de um Escritório.

Classif.	F	N	Total
H	30	20	50
M	10	40	50
Total	40	60	100

Uma pessoa entra no escritório, seleciona um dos empregados ao acaso e verifica que é homem. Qual a probabilidade de que seja fumante ?

Em termos da notação introduzida, desejamos $P(F|H)$. Considerando somente o espaço amostral reduzido H (isto é, os 50 homens), temos:

$$P(F|H) = \frac{30}{50} = \frac{3}{5} = 0,6$$

enquanto que, empregando a definição de probabilidade condicional, esta será

$$P(F|H) = \frac{P(F \cap H)}{P(H)}$$

e $$P(F \cap H) = \frac{30}{100} = 0,3 , \qquad P(H) = \frac{50}{100} = 0,5$$

portanto, $P(F|H) = \frac{0,3}{0,5} = 0,6$

2.7 - EVENTOS INDEPENDENTES

Dois eventos A e B, pertencentes a um espaço amostral S, são independentes entre si quando:

$$P(A \cap B) = P(A) \cdot P(B)$$

Isto significa que a ocorrência de A não afeta a ocorrência de B, ou seja:

$$P(B|A) = P(B) \quad \text{e} \quad P(A|B) = P(A)$$

Seja A o evento "sai cara no primeiro lançamento de uma moeda" e B o evento "sai cara em um segundo lançamento".

Então:

$$P(A) = 0,5$$

$$P(B) = 0,5$$

e $P(A \cap B) = (0,5)\ (0,5) = 0,25 \Longrightarrow P(B|A) = \frac{0,25}{0,5} = 0,5 = P(B)$

pois a ocorrência de cara no segundo lançamento não é influenciada pelo conhecimento de que no primeiro ocorreu cara; os eventos A e B são independentes.

2.8 - PROBABILIDADES REVISADAS

Dizemos que os eventos $B_1, B_2, \ldots, B_n$ constituem uma partição do espaço amostral S se $\bigcup_t B_t = S$ e $B_i \cap B_j = \emptyset$, $\forall$ i, j.

Seja $B_1, B_2, \ldots, B_n$ uma partição de S e seja A um evento pertencente a S, como está mostrado na figura 2.6.

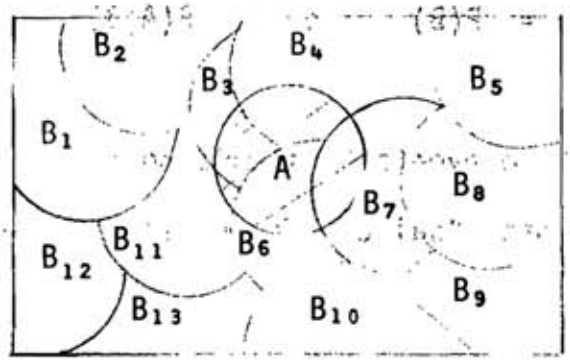


Figura 2.6 - Partição de um Espaço Amostral.

Podemos escrever A como:

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_n)$$

Vemos que A é formado pela união de n eventos mutua

mente exclusivos. Portanto,

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_n)$$

ou $$P(A) = P(A|B_1).P(B_1) + P(A|B_2).P(B_2) + \ldots + P(A|B_n).P(B_n)$$

Se soubermos que A ocorreu e quisermos determinar a probabilidade de qualquer B_j, $j = 1, 2, \ldots, n$, esta é dada por:

$$P(B_j|A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A|B_j).P(B_j)}{\sum_{j=1}^{n} P(A|B_j).P(B_j)}$$

Este resultado é conhecido como "Teorema de Bayes" e é de grande utilidade na determinação das probabilidades "a posteriori", ou seja, para revisar probabilidades face a novas informações.

Cada um dos B_j de S tem a ele associada uma probabilidade, que é a probabilidade "a priori", à qual já nos referimos no Capítulo 1.

Consideremos o seguinte exemplo:

Três transportadoras A, B e C realizam o transporte de peças do mesmo tipo da cidade M para a cidade N. A transportadora A tem capacidade de transporte duas vêzes maior que a de B e C, as

quais têm capacidades iguais. No entanto, a transportadora C adota um sistema de acondicionamento das peças diferente daquele usado por A e B, o que danifica 4% das peças. Das peças acondicionadas por A e B, apenas 2% se danificam.

Ao chegar em N, as peças são colocadas juntas num armazém e depois retiradas para serem utilizadas.

Qual a probabilidade da peça retirada estar danificada e qual a probabilidade desta peça danificada ter sido transportada por A?

Definamos os seguintes eventos:

$$D \equiv \{ \text{a peça está danificada} \}$$

$$E_A \equiv \{ \text{a peça foi transportada por A} \}$$

$$E_B \equiv \{ \text{a peça foi transportada por B} \}$$

$$E_C \equiv \{ \text{a peça foi transportada por C} \}$$

Sabemos que:

$$P(D) = P(D|E_A).P(E_A) + P(D|E_B).P(E_B) + P(D|E_C).P(E_C)$$

e que $P(E_A) = 0,5$

$P(E_B) = P(E_C) = 0,25$

$P(D|E_A) = P(D|E_B) = 0,02$

$P(D|E_C) = 0,04$

Então:

$$P(D) = 0,02 \cdot 0,5 + 0,02 \cdot 0,25 + 0,04 \cdot 0,25 = 0,025$$

A probabilidade da peça danificada ter sido transportada pela transportadora A é dada por:

$$P(E_A|D) = \frac{P(D|E_A).P(E_A)}{P(D)} = \frac{0,02 \times 0,5}{0,025} = 0,4$$

2.9 - VARIÁVEIS ALEATÓRIAS

Seja E um experimento aleatório e S o espaço amostral a ele associado. A função X que a cada elemento $s \in S$ associa um número real é denominado variável aleatória.

Quando temos duas funções X e W cada uma associando um número real a todo $s \in S$, o par (X, W) constitui uma variável

aleatória bidimensional.

No caso geral, se houver n funções, teremos uma variável aleatória n-dimensional.

Para alguns autores, o termo "variável aleatória" é confusa, sendo mais apropriada a denominação "função aleatória".

Suponhamos que o experimento E consiste em se lançar uma moeda duas vêzes. Podemos ter as configurações mostradas na figura 2.7.

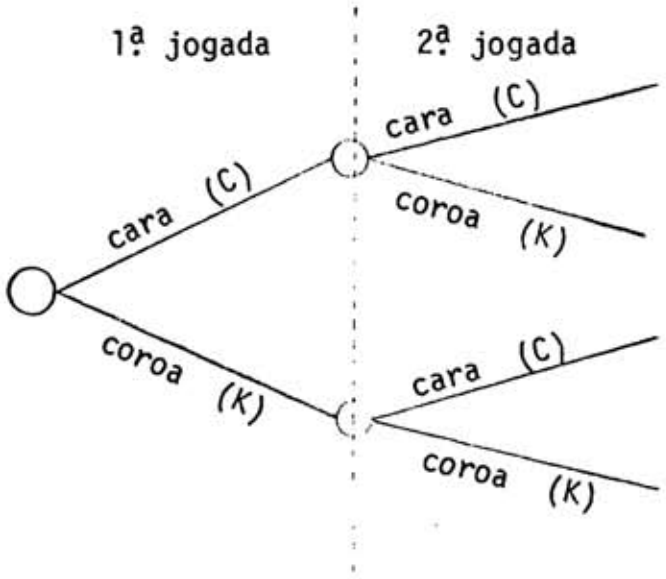


Figura 2.7 - Resultados de dois lançamentos sucessivos de uma moeda.

O espaço amostral S associado a este experimento é:

$$S = \{ CC,\ CK,\ KC,\ KK \}$$

Definamos a variável aleatória X como o "número de coroas (K) obtidas nos dois lançamentos", isto é:

$$X(CC) = 0;\ X(CK) = X(KC) = 1 \quad e \quad X(KK) = 2$$

Notemos que uma variável aleatória, sendo uma função, não impede que dois ou mais $s \in S$ correspondam a um mesmo valor $X(s) \in R$.

Se o número de valores possíveis de X for finito ou infinito numerável, X é dita ser uma variável aleatória discreta, ou seja, os valores possíveis de X podem ser postos em correspondência com os inteiros.

Se a variável X puder assumir qualquer valor num determinado intervalo, X é dita ser uma variável aleatória contínua.

2.10 - FUNÇÃO DENSIDADE DE PROBABILIDADE

Se X for uma variável aleatória discreta, a cada resultado possível x_i poderemos associar um valor $p(x_i) = P(X = x_i)$ que deve satisfazer às seguintes condições:

1) $p(x_i) = P(X = x_i) \geq 0, \qquad \forall_i$

2) $\sum_{i=1}^{\infty} p(x_i) = 1$

A função $p(.)$ é denominada função de probabilidade no ponto ou função massa de probabilidade.

Se X for uma variável aleatória contínua, definimos a "função densidade de probabilidade" de X, $f_X(.)$, como uma função que satisfaz às seguintes condições:

1) $f_X(x) \geq 0, \qquad \forall\, x$

2) $\int_{-\infty}^{+\infty} f_X(x)\, dx = 1$

Exemplo: O departamento de controle de qualidade de uma indústria de lâmpadas recebe uma grande partida das mesmas, retirada da linha de produção. Estima-se que a probabilidade de uma lâmpada estar perfeita é $\frac{4}{5}$, sendo $\frac{1}{5}$ a probabilidade da mesma apresentar defeito.

As lâmpadas são ensaiadas uma a uma até que apareça a primeira defeituosa.

Definamos X (uma variável aleatória) como o número de testes necessários até que a primeira defeituosa apareça. O espaço amostral associado ao experimento é:

$$S = \{ D,\ ND,\ NND,\ NNND,\ \ldots \}$$

onde D corresponde a uma lâmpada defeituosa e N a uma não defeituosa.

Portanto, os valores possíveis de X são: 1, 2, 3, ... n, ... Verificamos que $X = i$ somente se a i-ésima lâmpada testada for defeituosa e as $i - 1$ anteriores forem perfeitas. Pelo fato de uma lâmpada ensaiada não afetar a condição da próxima, poderemos escrever:

$$p(i) = P(X = i) = \left(\frac{4}{5}\right)^{i-1} \cdot \left(\frac{1}{5}\right), \qquad i = 1, 2, \ldots$$

e $$\sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} \left(\frac{4}{5}\right)^{i-1} \left(\frac{1}{5}\right) = \frac{1}{5}\left[1 + \frac{4}{5} + \frac{16}{25} + \ldots\right] = 1$$

No caso de uma variável aleatória bidimensional discreta (X, Y), a função de probabilidade $p(.,.)$ deverá satisfazer às seguintes condições:

1) $p(x_i, y_j) = P(X = x_i, Y = y_j) \geq 0, \qquad \forall_{i,j}$

2) $$\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} p(x_i, y_j) = 1$$

No caso contínuo, a função densidade de probabilidade conjunta $f_{XY}(.,.)$ deverá satisfazer a:

1) $f_{XY}(x, y) \geq 0, \qquad \forall (x, y)$

2) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{XY}(x, y)\, dx\, dy = 1$

Em geral, para uma variável aleatória n-dimensional $(X_1, X_2, \ldots, X_n)$, devemos ter:

a) caso discreto:

1) $p(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) \geq 0$

2) $\sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} \cdots \sum_{j_n=1}^{\infty} p(x_{j_1}, x_{j_2}, \ldots, x_{j_n}) = 1$

b) caso contínuo:

1) $f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) \geq 0$

2) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n = 1$

2.11 - FUNÇÃO DISTRIBUIÇÃO ACUMULADA

Seja X uma variável aleatória e x um número. A função $F_X(x) = P(X \leq x)$ é denominada função distribuição acumulada de X ou, simplesmente, função distribuição.

Quando X for uma variável aleatória discreta, temos:

$$F_X(x) = \sum_i p(x_i)$$

onde o somatório é feito sobre todos os i's para os quais $x_i \leq x$.

Se X for uma variável aleatória contínua:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$$

Neste caso, podemos calcular a função densidade de probabilidade, fazendo:

$$f_X(x) = \frac{d\, F_X(x)}{dx}$$

quando $F_X(.)$ for diferenciável.

A função distribuição $F_X(.)$, tanto para o caso contínuo quanto para o caso discreto, satisfaz às seguintes propriedades:

1) $F_X(-\infty) = 0$

2) $F_X(+\infty) = 1$

3) $F_X(.)$ é não decrescente, isto é, $F_X(x) \leq F_X(y)$ quando $x < y$.

Exemplos de gráficos da função distribuição para os casos discreto e contínuo, são apresentados nas figuras 2.8(a) e 2.8(b), respectivamente:

Se tivermos uma variável aleatória bidimensional (X, Y), sua função distribuição acumulada conjunta será definida por:

$$F_{XY}(x, y) = P(X \leq x, \quad Y \leq y)$$

Para obter a função densidade de probabilidade conjunta de X e Y a partir da função distribuição, fazemos:

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \, \partial y}$$

quando $F_{XY}(., .)$ for diferenciável.

2.12 - DISTRIBUIÇÃO DE PROBABILIDADE MARGINAL

Seja (X, Y) uma variável aleatória bidimensional, com X e Y discretas. Definimos a distribuição de probabilidade marginal de X, como:

$$p(x_i) = P(X = x_i) = \sum_{j=1}^{\infty} p(x_i, y_j), \quad \text{para todo} \quad x_i, \quad i = 1, 2, \ldots$$

De maneira análoga, a distribuição de probabilidade mar

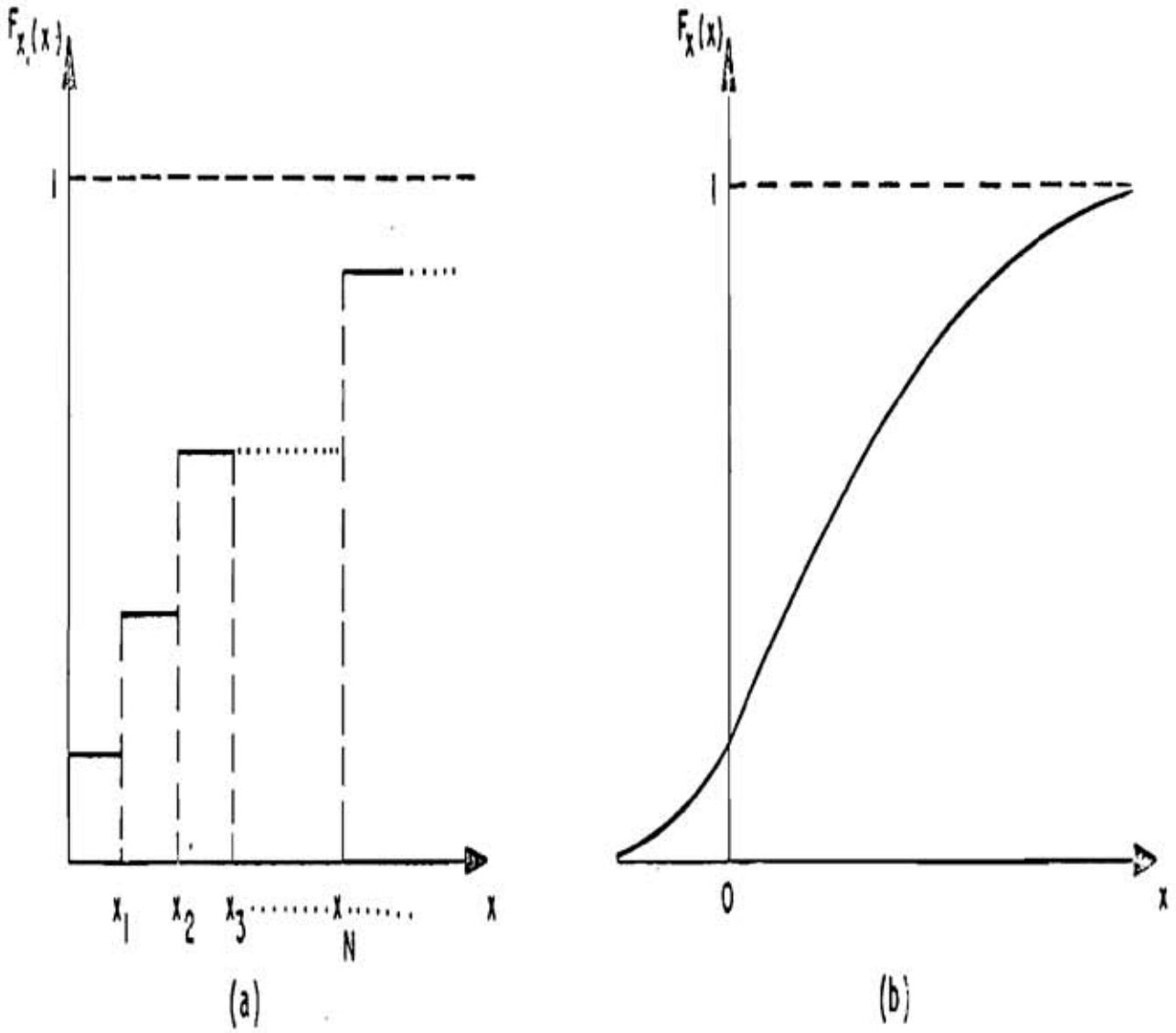


Figura 2.8 - Função Distribuição Acumulada: Casos Discreto e Contínuo.

ginal de Y será:

$$q(y_j) = \sum_{i=1}^{\infty} p(x_i, y_j), \quad \text{para todo } y_j, \ j = 1, 2, \ldots$$

No caso de (X, Y) ser contínua,

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y)\, dy$$

e

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y)\, dx$$

definem, respectivamente, as distribuições de probabilidade marginais de X e Y.

Como exemplo, seja a função distribuição acumulada de (X, Y) dado por:

$$F_{XY}(x, y) = \begin{cases} 1 - e^{-2x} - e^{-2y} + e^{-2(x+y)}, & 0 \leq x < \infty \\ & 0 \leq y < \infty \\ 0, & x < 0 \quad y < 0 \end{cases}$$

Para determinar $f_X(x)$, ou seja, a distribuição de probabilidade marginal de X, necessitamos da distribuição conjunta $f_{XY}(x, y)$.

Esta é dada por:

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \ \partial y} = 4\, e^{-2(x+y)} , \quad \begin{matrix} 0 \leq x < \infty \\ 0 \leq y < \infty \end{matrix}$$

$$f_{XY}(x, y) = 0 , \qquad x < 0 \quad e \quad y < 0$$

Como

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y)\, dy$$

obtemos $$f_X(x) = \begin{cases} \int_0^{\infty} 4\, e^{-2(x+y)} dy = 2\, e^{-2x} , & x \geq 0 \\ 0 , & x < 0 \end{cases}$$

2.13 - DISTRIBUIÇÃO DE PROBABILIDADE CONDICIONAL

Para variáveis aleatórias bidimensionais, (X, Y),

a) $p(x_i|y_j) = P(X = x_i \mid Y = y_j) = \frac{p(x_i \cap y_j)}{q(y_j)} , \quad q(y_j) > 0 ,$

define a distribuição de probabilidade condicional de X dado que $Y = y_j$, quando X e Y são discretas (de maneira análoga, definimos $p(y_j|x_i)$) e

b) $f_{X|Y}(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)} , \quad f_Y(y) > 0,$ define a distribuição de probabilidade condicional de X dado que Y = y, quando X e Y

são contínuas (de maneira análoga, definimos $f_{Y|X}(y|x)$).

No caso de X e Y serem variáveis aleatórias independentes, teremos:

$$f_{X|Y}(x \mid y) = f_X(x) \qquad e \qquad f_{Y|X}(y \mid x) = f_Y(y) \tag{2.3}$$

quando X e Y forem contínuas.

Se X e Y forem discretas, teremos, se elas forem independentes:

$$\begin{aligned} p(x_i|y_j) &= p(x_i) , \quad \forall i \\ q(y_j|x_i) &= q(y_j) , \quad \forall j \end{aligned} \tag{2.4}$$

A demonstração de 2.3 e 2.4 é direta, se lembrarmos que duas variáveis aleatórias são independentes se e somente se:

a) $p(x_i, y_j) = p(x_i).q(y_j)$, $\forall i,j$, para o caso discreto e

b) $f_{XY}(x, y) = f_X(x).f_Y(y)$, no caso contínuo.

Para a variável (X, Y) definida no exemplo da seção anterior, temos:

$$f_{Y|X}(y|x) = \frac{4\, e^{-2(x+y)}}{2\, e^{-2x}} = 2\, e^{-2y}$$

2.14 - VALOR ESPERADO DE UMA VARIÁVEL ALEATÓRIA

Seja X uma variável aleatória discreta com valores possíveis x_i, $i = 1, 2, \ldots$. O valor esperado da variável X (ou esperança matemática de X), denotado por E(X), é definido como:

$$E(X) = \sum_{i=1}^{\infty} x_i\, p(x_i)$$

Se X for uma variável aleatória contínua, com função densidade de probabilidade $f_X(x)$, seu valor esperado é definido como:

$$E(X) = \int_{-\infty}^{+\infty} x\, f_X(x)\, dx$$

Se, neste caso, a função distribuição acumulada de X for $F_X(x)$, teremos, que a área A na figura 2.9, será o valor esperado de X, ou seja:

$$E(X) = \int_0^{\infty} \left[1 - F_X(x)\right] dx$$

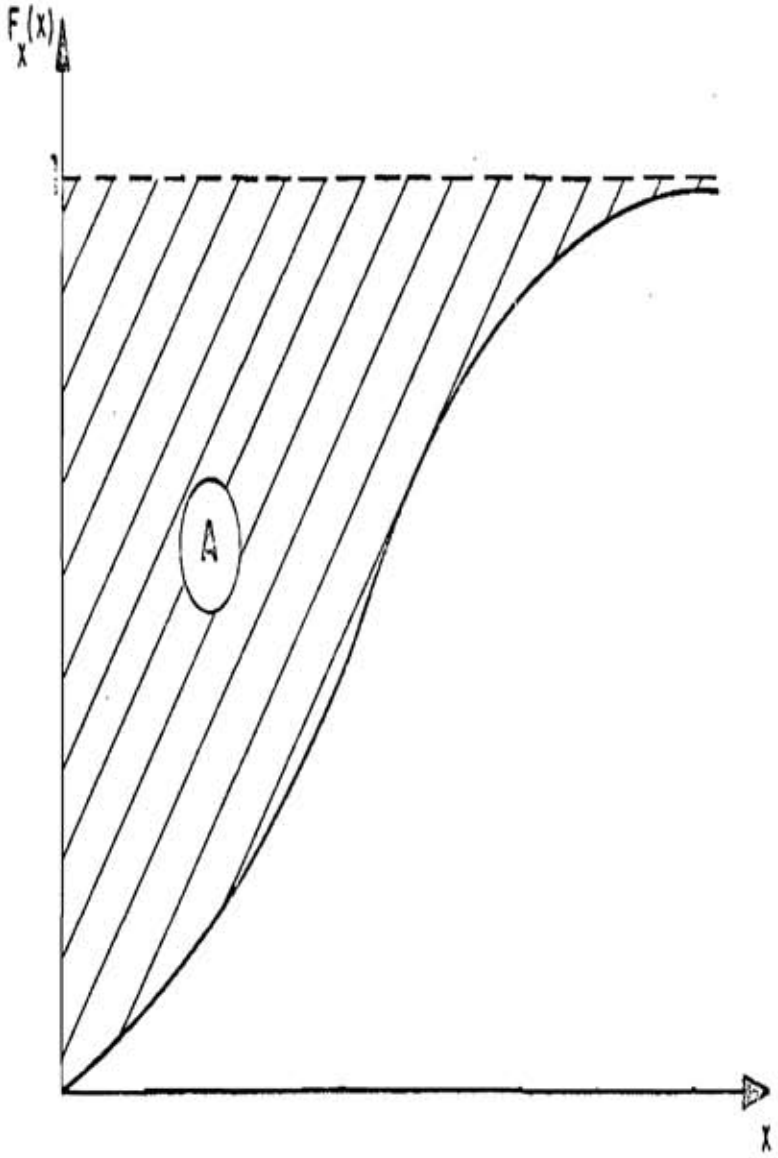


Figura 2.9 - Valor Esperado de uma Variável Aleatória.

2.14.1 - Propriedades do Valor Esperado

a) $E(X + Y) = E(X) + E(Y)$

b) $E(X) = C$, para $X = C$, onde C é uma constante.

c) $E(CX) = C.E(X)$, para C constante

d) $E(XY) = E(X).E(Y)$, quando X e Y forem independentes.

A demonstração dessas propriedades são encontradas em [36].

2.15 - PROBABILIDADES SUBJETIVAS

Chamamos probabilidade subjetiva atribuida por um indivíduo a um evento, a um valor numérico entre 0 e 1 que reflete o grau de confiança desse indivíduo na ocorrência de tal evento, levando em conta seu estado de informação.

Considera-se que tais valores são consistentes com os princípios básicos de probabilidades e estende-se para as probabilidades subjetivas todo o tratamento das probabilidades clássicas.

Sob o ponto de vista de probabilidade subjetiva, faz sen

tido falar-se em probabilidade de uma hipótese.

A probabilidade que um indivíduo associa a uma hipótese é um valor que leva em conta as informações que possui e reflete o crédito que ele dá à veracidade da hipótese. Uma hipótese, em princípio, é verdadeira ou falsa. Sob o ponto de vista clássico, não faz sentido falar-se em probabilidade de uma hipótese pois, simplesmente, não existe frequência histórica. Esta é uma restrição séria pois, em problemas de decisão, frequentemente os chamados estados da natureza são, na verdade, hipóteses. O uso das probabilidades subjetivas supera esta barreira. O conceito de probabilidade subjetiva se mostra também de grande valia quando o cálculo de probabilidades é possível mas os dados necessários não são disponíveis ou suficientes. Essa situação é também muito frequente na prática. Um indivíduo que reuna boas informações sob o evento em questão pode expressar um valor subjetivo de probabilidade que, posteriormente, poderá ser revisado pelos dados disponíveis. Como codificar as informações existentes na mente de um indivíduo e como revisar as probabilidades resultantes serão assunto do próximo capítulo.

CAPÍTULO III

PROBABILIDADES: ASPECTOS PRÁTICOS

3.1 - CONSIDERAÇÕES INICIAIS.

Neste capítulo nos ocuparemos particularmente com a discussão de técnicas para a obtenção de valores subjetivos de probabilidades. Apresentaremos inicialmente alguns conceitos que podem ser úteis na prática, definiremos nossa posição em relação às probabilidades subjetivas e passaremos à apresentação das técnicas mencionadas.

Conceito de "odds".

Muitas vezes, é mais cômodo para quem faz estimativas pensar em termos da razão entre as probabilidades de eventos alternativos

É comum nos meios esportivos ouvirmos frases como: "O Palmeiras está cotado na razão de 3/1 para o jogo de amanhã".

Analisando o significado de tal assertiva vemos que ela é equivalente a dizer que na opinião dos entendidos (ou apostadores)

"$\frac{\text{P (Palmeiras ganhar o jogo de amanhã)}}{\text{P (Palmeiras não ganhar o jogo)}} = 3$"

ou a "a probabilidade do Palmeiras ganhar o jogo de amanhã é 0,75".

"Chances", "cotação", "razão", "índice" são termos comumente usados em português para indicar essa "razão de probabilidades" a que nos referimos. Por julgarmos que nenhuma das palavras mencionadas é bastante boa para expressar a idéia geral do conceito e ignorarmos um vocábulo que melhor expresse tal idéia, adotaremos a palavra inglesa "odds".

Formulação Matemática.

Consideremos o evento A e seu evento complementar $\bar{A}$. Sabemos que

$$P(A) + P(\bar{A}) = 1$$

ou

$$P(\bar{A}) = 1 - P(A)$$

Chamaremos "odds" de A ao valor $\Omega(A)$ definido como

$$\Omega(A) = \frac{P(A)}{P(\bar{A})} = \frac{P(A)}{1-P(A)}$$

Observando a expressão acima podemos concluir que:

1. $\Omega(A) \geq 0, \forall A$

2. Se $\Omega(A) = \frac{a}{b}$, onde a e b são dois valores reais (não necessariamente inteiros)

então $P(A) = \frac{a}{a+b}$

Avaliação Objetiva de Probabilidades

Segundo entendemos a probabilidade objetiva, ou clássica, poderia ser vista como uma probabilidade estimada à luz de evidências bastante fortes que conduzissem a opiniões convergentes. Essas "evidências" a que nos referimos são basicamente históricas e geométricas. Apoiadas em hipóteses de aceitação mais ou menos fácil, tais "evidências" conduzem a valores de probabilidade consagrados como "objetivos".

Assim, quando se diz que "a probabilidade de se obter a face '2' quando se atira um dado é 1/6" estamos baseados em considerações como:

- o dado é simétrico: um cubo de seis faces iguais,

e hipóteses adicionais como:

- o dado é "honesto": tudo faz crer que sua massa seja uniformemente distribuída,

ou, ainda, considerações de "frequência-relativa":

- "em 10.000 vezes que se atirou o dado anteriormente, 1672 vezes obtive

mos a face 2"; apoiados na observação de simetria, concluiríamos pelo valor 1/6.

Num contexto mais geral, a única maneira objetiva de se estimar a probabilidade de um evento é a frequência relativa: número de casos favoráveis observados dividido pelo número total de casos observados.

A avaliação ou estimativa de probabilidades em face de observações já foi exaustivamente estudada pela Estatística Clássica e a literatura a respeito é incontável. Não cuidaremos disso neste trabalho. Nossa preocupação básica será estudar métodos que permitam estimativas pessoais de probabilidade quando não se dispõe de dados mais concretos que possibilitem uma abordagem clássica.

Necessidade de Estimativas Subjetivas da Probabilidade.

Imaginemos que nós somos os donos de uma indústria e, na tomada de uma decisão importante, constatamos a necessidade de avaliar a probabilidade de que uma determinada máquina dure 5 ou mais anos. Não dispomos de quaisquer dados históricos, nem existe a possibilidade de acesso em tempo hábil a informações do fabricante. Há contudo um velho empregado da firma que vem trabalhando há muitos anos com tal máquina. Não há dúvida que uma estimativa de tal indivíduo, convenientemente extraída, é bastante valiosa. Seria válida uma decisão baseada em tal estimativa, ainda que "subjetiva"? Acreditamos que a resposta sensata é "sim".

Consideremos uma outra situação hipotética. Um indivíduo planeja um passeio ao campo neste domingo. Em seu processo de tomada de decisão, julga necessário conhecer a probabilidade de fazer tempo bom até o fim do dia. Esta é uma situação típica em que dados históricos não são disponíveis. Existe porém seu velho tio, homem criado no campo, em cujas "previsões meteorológicas" ele aprendeu a confiar. É sensato admitir que esse homem poderia fornecer uma probabilidade bastante confiável, desde que lhe fosse fornecido um método de traduzir seus conhecimentos em um valor numérico conveniente ? Poderia o indivíduo basear sua decisão nessa probabilidade subjetiva ? Admita agora que, ouvindo o rádio, ele tem acesso à previsão do serviço meteorológico local. Esta previsão é mais objetiva, talvez, que a de seu velho tio. Envolve, talvez, o processamento de algumas informações mais "concretas". Mas, notemos, é também uma "estimativa". Admitamos ainda uma terceira possibilidade: o indivíduo consulta vários jornais velhos e constata uma estatística que fornece o número de "domingos em fase de Lua Nova" em que choveu, nos últimos dois anos, e o número total de domingos de Lua Nova desses dois últimos anos. O quociente lhe dá uma probabilidade. Podemos dizer que esta probabilidade é objetiva ?

Em qual das três estimativas o indivíduo basearia sua decisão ? Não seria surpresa se ele optasse pela "subjetiva" probabilidade de seu velho tio.

Nosso propósito com essas ilustrações é ressaltar a idéia

de que são frequentes as situações em que dados para avaliação de probabilidades são de difícil obtenção, inexistem, ou, mesmo, são de pouca validade, enquanto que "experts" (pessoas com vivência ou conhecimento do assunto em questão), capazes de nos fornecer um valor "subjetivo" para a probabilidade requerida, estão à nossa disposição.

Nosso ponto é que devemos fazer uso dessas evidências não "concretas" e que é válido e muito útil contar com "juízos de valor" traduzidos em probabilidades. Nas seções seguintes, tentaremos formalizar alguns métodos de obter tais probabilidades.

3.2 - PROBABILIDADES SUBJETIVAS

3.2.1 - Como Obter uma Estimativa da Probabilidade

Para a obtenção de estimativas de probabilidade existem duas linhas básicas de procedimentos:

1) Pergunta direta ao estimador: "Qual é, para você, a probabilidade de ocorrer A?" ou "Quais são, em sua opinião, os "odds" de A? ou "O que você considera mais provável: A ou B?".

Embora cômoda, esta linha é bastante criticada, especialmente pelos estatísticos teóricos *. As objeções mais comumente levantadas são:

---

* *Ver SAVAGE* [45]

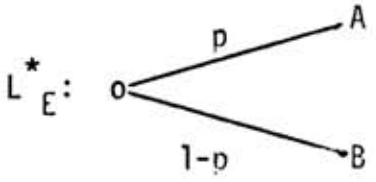
a. Nada garante que o conceito de "mais provável" seja intuitivo e não dê margem a interpretações equívocas.

b. A resposta a uma perqunta direta não reflete, necessariamente, o comportamento de uma pessoa em face da incerteza. Refletiria, tão somente, o "comportamento verbal" de uma pessoa em face de uma interrogação.

2) Método indireto ou comportamental: a pessoa seria colocada diante de uma situação de fato e sua escolha entre duas alternativas refletiria seu juízo. Uma adaptação mais praticável dessa idéia é colocar a pessoa diante de uma situação hipotética e perguntar-lhe qual seria sua atitude em tais circunstâncias.

Nesta linha se enquadra a idéia de definir probabilidade "subjetiva" através de uma loteria.

Seja E um evento aleatório e L_E uma loteria cujo prêmio é A, se E ocorre e B, caso contrário. Seja L^*_E uma outra loteria que dá o prêmio A com probabilidade p e B com probabilidade 1-p. Representadas abaixo estão L_E e L^*_E:



Se formos indiferentes entre essas duas loterias, dizemos que nossa probabilidade subjetiva de E é p e denotamos essa probabilidade por P*(E). Então, P*(E)= p .

Uma crítica comum a esta abordagem é considerá-la pouco pragmática. Para alguns é irreal colocar o indivíduo diante de loterias hipotéticas em busca de pontos de indiferença. Mais fácil e intuitivo seria a estimativa direta. Contudo, os teóricos da moderna estatística consideram tal conduta como a única normativamente correta.

O dispositivo que descreveremos a seguir busca tornar mais simples e concreta a utilização dos conceitos acima expostos.

O "Jogo de Dardos" como Medida de Probabilidade

Consideremos um disco rodando e um dardo que é largado em um instante de tempo qualquer, sobre o disco.

Qual a probabilidade de que esse dardo atinja uma determinada faixa do disco delimitada por duas linhas radiais?
Ou referindo-nos à figura 3.1, qual a probabilidade de que o dardo atinja a região B?

Parece-nos ser de aceitacão geral que a probabilidade bus

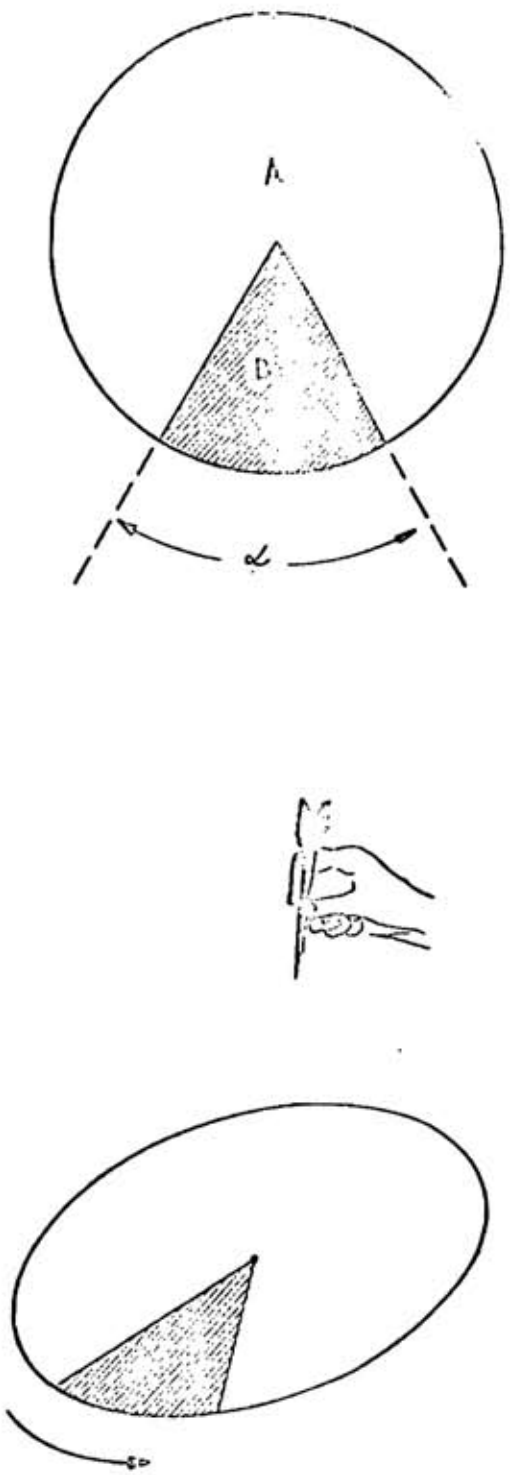


Figura 3.1 - Jogo de Dardos: Dispositivo para Determinação de Estimativas Subjetivas de Probabilidades.

cada é a razão entre a área hachurada e a área total do disco. Ou, como a área B é diretamente proporcional ao ângulo α , podemos graduar nosso disco em uma escala de probabilidades. Assim, para $\alpha = 360º$ (a área B corresponde ao disco todo), teríamos uma probabilidade igual a 1; para $\alpha = 180º$, uma probabilidade 1/2, etc.

Explicando melhor, nosso disco terá duas cores: amarelo e azul. Deslocando um ponteiro, aumentamos a área azul de 0 (disco totalmente amarelo) a 1 (disco totalmente azul) e vice-versa. A cada fração de área azul, poderemos ler diretamente no verso do disco a fração de intervalo [0, 1] correspondente (escala de "probabilidades").

Suponhamos agora que desejamos avaliar nossa probabilidade subjetiva da ocorrência do evento incerto E. Poderia ser a"probabilidade de que a próxima pessoa a atravessar a porta seja uma mulher" ou a "probabilidade de que nos próximos seis meses entre no mercado um produto concorrente", etc.

Propomo-nos a seguinte situação:

LOTERIA 1 - ocorrendo E receberemos um milhão de cruzeiros; caso contrário, não receberemos nada.

LOTERIA 2 - se o dardo, nas condições anteriormente expostas, atingir a área azul receberemos 1 milhão de cruzeiros; caso contrário , nada receberemos. Em qual das duas loterias preferiríamos apos

tar? Vamos então variando a área azul, até que as loterias se tornem indiferentes. Nesse ponto, poderemos ler diretamente no verso do disco nossa "probabilidade" da ocorrência de E.

3.2.2 - Alguns Critérios de Consistência

O caráter subjetivo das probabilidades que estamos tratando não implica, de forma alguma, que não possam ou devam ser revistas e confrontadas com outros valores que permitam testar sua "razoabilidade". Na verdade é bastante desejável que encontremos formas de avaliar a consistência de nossas estimativas.

Do mesmo modo, muitas vezes se torna mais fácil a estimativa indireta: a probabilidade desejada sai como resultado de outras probabilidades mais fáceis de estimar ou calcular que se ligam a ela por princípios básicos da Teoria de Probabilidades.

Consistência com as Propriedades Fundamentais

Usamos frequentemente as propriedades fundamentais das pro

babilidades para testar a consistência de atribuições subjetivas. Podemos testar o valor atribuído por nós a um evento E, estimando também a probabilidade do evento complementar $\bar{E}$ e verificando se as duas somam um. Caso isso não ocorra devemos modificar um dos dois valores. Isso é especialmente útil quando, por razões inerentes ao problema, nos sentimos mais confiantes em relação à estimativa do evento complementar.

O que se faz, também, habitualmente, é subdividir um evento em eventos mais simples e mutuamente exclusivos e confrontar a probabilidade atribuída ao evento composto com a soma das probabilidades dos eventos componentes.

EXEMPLO: Nossa vitrola está apresentando distorção. Analisando as possíveis causas de defeito, um técnico nosso amigo estabelece a seguinte árvore, mostrada na figura 3.2.

Pedimos que ele atribua probabilidades, e ele, então, nos fornece as seguintes estimativas:

P (defeito interno) = 0,20

P (defeito no alto-falante) = 0,25

P (defeito na cápsula) = 0,50

Neste ponto interrompemos nosso amigo e o fazemos ver que

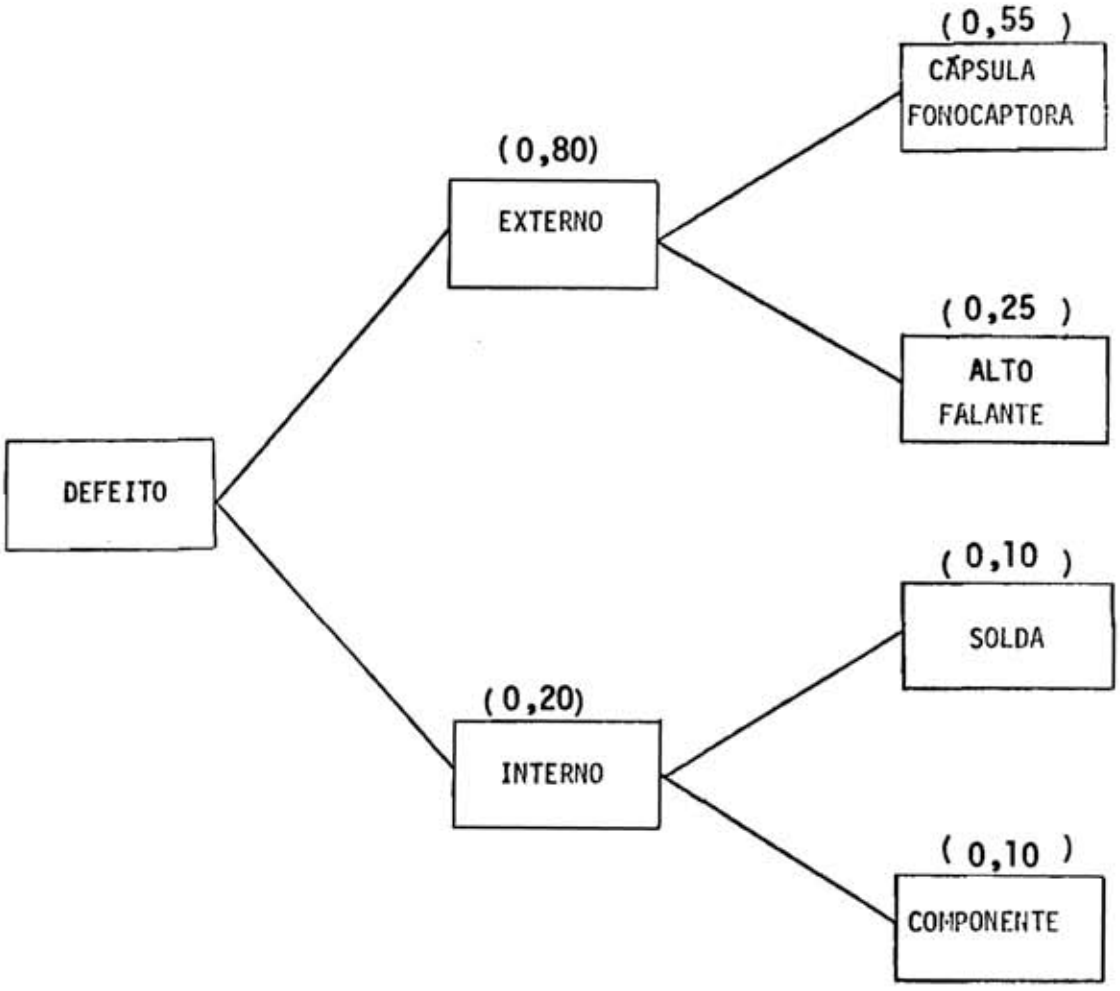


Figura 3.2 - Árvore dos possíveis defeitos na vitrola e as probabilidades finais atribuidas pelo técnico.

hã uma inconsistência. Com efeito, diante das estimativas feitas, teremos:

P(defeito externo) = 1- P(defeito interno) = 0,80

P(defeito externo) = P(defeito cápsula) + P(defeito alto-falante) = 0,50 + 0,25 = 0,75

Diante de nossa argumentação, ele reanalisa os valores atribuídos e, sentindo-se confiante no valor associado a "defeito interno", decide alterar para 0,55 a probabilidade de defeito na cápsula.

Interrogado sobre a probabilidade de defeito de componente, nosso amigo se mostra inseguro. Perguntamos então pela probabilidade de defeito de solda e ele fornece o valor 0,10.

Mostramos então a nosso amigo que a probabilidade de defeito de componente está agora determinada e deve valer:

P(defeito componente) = P(defeito interno) - P(defeito solda) = 0,20 - 0,10 = 0,10

Nosso amigo aceita sem relutância o valor obtido e nos damos por satisfeitos.

Uso de julgamentos condicionais

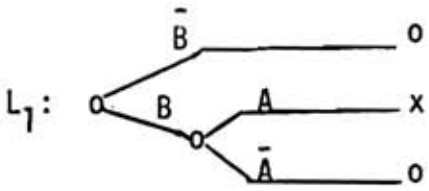
A consistência entre julgamentos a priori e julgamentos condicionais pode ser uma ferramenta bastante útil na avaliação de probabilidades. Sabemos, da teoria, que se A e B são dois eventos quaisquer e $P(B) \neq 0$ então

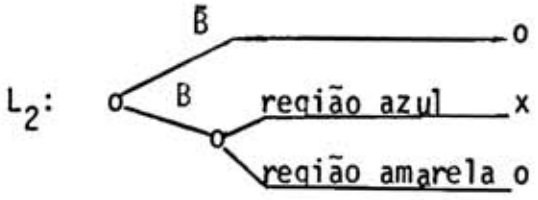
$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad (3.1)$$

A interpretação das probabilidades condicionais em termos de loterias, pode ser entendida como se segue. Apresentamos ao Estimador duas loterias compostas:

L_1 - Se B não ocorrer ele nada recebe. Se B ocorrer ele concorre a uma elevada importância X em dinheiro, caso A ocorra, e nada, em caso contrário.

L_2 - Não ocorrendo B o prêmio é nulo. Ocorrendo B, joga-se o dardo sobre o disco: sendo atingida a região azul o prêmio é X (o mesmo da loteria anterior). Se a região atingida for a amarela o prêmio é nulo.





O avaliador considera as duas loterias e se decide por uma delas. Varia-se então a área da região azul do disco: o Estimador reconsidera as loterias e faz nova escolha. O processo continua até que, para uma dada área azul, o "Estimador" se sente indiferente entre as duas loterias. Nesse ponto a leitura no verso do disco indicará $P(A|B)$.

Consideremos que nosso Estimador atribui probabilidades aos eventos $A \cap B$ e B. Usando a fórmula (3.1) ele determina, então, um valor para $P(A|B)$. Através do método indicado acima, por exemplo, ele obtém um outro valor para $P(A|B)$. Para ser consistente, os dois valores devem se igualar. Em face de desigualdade o Estimador deve reconsiderar as probabilidades envolvidas. Sentindo-se fortemente favorável ao resultado dado pela estimativa direta, os valores $P(A \cap B)$ e $P(B)$ precisarão ser revistos. Em caso de dúvida, contudo, aconselhamos a adotar o valor fornecido pela fórmula (3.1). A justificativa para tal procedimento é o "conservadorismo"* natural do homem.

O exemplo que se segue se deve a Raiffa, Schlaifer e Pratt [46] e ilustra como julgamentos condicionais podem ser usados no processo de estimativa de probabilidades.

---

* *Estudos realizados por psicólogos dedicados à Análise de Decisões, como o Dr. Ward Edwards, revelam que o elemento humano se comporta de uma maneira sensivelmente "conservadora" ao avaliar a influência de novas informações. Assim, tende sempre a subestimar a informação "B ocorreu" ao avaliar a "probabilidade de ocorrer A dado que B ocorreu". Dessa forma os valores atribuídos a $P(A|B)$ em uma estimativa direta, são sempre inferiores aos obtidos através do teorema de Bayes.*

Consideremos que, por razões acadêmicas, selecionamos um dentre os 1200 estudantes de uma Escola de Pós-Graduação em Administracão e desejamos estimar a probabilidade de que este estudante saiba calcular corretamente $\int_0^{2\pi} \text{sen}x\, dx$. Vamos supor que saibamos que aproximadamente 40% dos estudantes possuem formacão em Engenharia. Ao estimar a possibilidade de uma resposta correta (S) se torna mais fácil estimar primeiro as probabilidades p de S dado que o estudante é um engenheiro (E) e dado que ele não é engenheiro ($\bar{E}$). Se, por exemplo, determinamos que:

$$P(S|E) = 0,70 \quad e \quad P(S|\bar{E}) = 0,20$$

então a probabilidade desejada deve* ser

$$P(S) = 0,70 \times 0,40 + 0,20 \times 0,60 = 0,40$$

Ao estimar a probabilidade $P(S|\bar{E})$ poderia, por sua vez, ser mais fácil estimar primeiro as probabilidades de que um não-engenheiro tenha tido curso de cálculo (C) e que não tenha ($\bar{C}$) e então estimar as probabilidades condicionais de S condicionado em ($\bar{E}$ e C) e em ($\bar{E}$ e $\bar{C}$).

3.2.3 - Revisão de Probabilidades em Face de Novas Informacões (Sob o Ponto de Vista Bayesiano).

O teorema de Bayes representa, como dissemos no capítulo 2, uma forma de revisar nossa opinião acerca da ocorrência de um evento ou

---

* *Para sermos consistentes com a fórmula (3.1) devemos ter:*
$P(S) = P(S \cap E) + P(S \cap \bar{E}) = P(S|E).P(E) + P(S|\bar{E}).P(\bar{E})$

da veracidade de uma hipótese* em face de novas evidências (ou dados). Como vimos, podemos exprimí-lo como:

$$P(H|D) = \frac{P(D|H)\ P(H)}{P(D)} \qquad (3.2)$$

onde P(D) e P(H) são diferentes de zero.

Em (3.2), P(H) representa a probabilidade "a priori" de alguma hipótese H. Observemos que, embora não esteja representada como tal, essa é uma probabilidade condicional. Na verdade, segundo entendemos, todas as probabilidades são realmente condicionais: toda estimativa de probabilidade é condicionada ao conhecimento do Estimador ou aos dados disponíveis por ocasião da estimativa. P(H) é então a probabilidade de H condicionada a toda informação disponível sobre H antes que se tome conhecimento de D. De uma maneira análoga P(H|D) representa a probabilidade de H condicionada ao conhecimento de D e mais todo o conhecimento anterior a D. P(D|H) representa formalmente a probabilidade de que um dado D seja observado se a hipótese H for verdadeira.

Para um conjunto de hipóteses mutuamente exclusivas e exaustivas H_i, as $P(D|H_i)$ representam o impacto do dado D sobre cada uma das hipóteses. A chave de qualquer processamento de informações utilizando o teorema de Bayes é obter P(D|H) para todo D e todo H. Essa probabilidade pode

---

* *Ver Edwards et allii* [10].

ser obtida através de cálculos, usando modelos estatísticos, ou através de estimativas humanas diretas.

A probabilidade P(D) é comumente de pouco interesse. Normalmente é calculada ou eliminada, como veremos a seguir. A hipótese H como vimos é uma de uma partição de hipóteses H_i mutuamente exclusivas e exaustivas. Portanto devemos ter $\sum_i P(H_i|D) = \sum_i P(H_i) = 1$ e a fórmula (3.2) implica então que

$$P(D) = \sum_i P(D|H_i).P(H_i)$$

Desde que disponhamos de $P(D|H_i)$ e $P(H_i)$ para todas as hipóteses H_i, disporemos também do valor P(D).

Uma outra maneira de contornar o problema é, como dissemos, eliminando P(D). Isto pode ser obtido utilizando-se o teorema de Bayes na forma de "odds".

O teorema de Bayes na forma de "odds".

Consideremos a hipótese H, sua negação $\overline{H}$ e o dado D. Podemos escrever

$$P(H|D) = \frac{P(D|H)\ P(H)}{P(D)}$$

$$P(\bar{H}|D) = \frac{P(D|\bar{H})\ P(\bar{H})}{P(D)}$$

Dividindo membro a membro as duas expressões anteriores, obtemos:

$$\frac{P(H|D)}{P(\bar{H}|D)} = \frac{P(D|H)}{P(D|\bar{H})} \cdot \frac{P(H)}{P(\bar{H})}$$

$$\Omega = R\,\Omega_0 \qquad (3.3)$$

onde

Ω_o - são os "odds" a priori ou a razão entre as probabilidades a priori de H e $\bar{H}$.

R - ou a razão $P(D|H)/P(D|\bar{H})$ é chamado "razão de verossimilhança" (likelihood ratio);

Ω - são os "odds" revisados ou, "a posteriori". A noção de "razão de verossimilhança" aqui usada é a mesma da estatística clássica.

É importante observar que se multiplicarmos P(D|H) e $P(D|\bar{H})$ por uma mesma constante o resultado em (3.3) não se altera. Esse fato é de extrema importância prática e é consequência do "princípio da verossimilhança"*. Tão somente "likelihood ratios" e não valores de P(D|H), propria

---

* *Ver Raiffa* [40] *ou Edwards et allii* [11]

mente, são necessários para aplicações do Teorema de Bayes na forma de "odds". Observemos que, se um dado é igualmente provável em face de uma hipótese ou de sua negação, isso significa que tal dado é irrelevante. Isto se traduz em nossa fórmula por R = 1 o que implica que os odds"a posteriori" se igualam aos odds "a priori". Portanto, um dado é tão mais significativo para uma hipótese quanto mais a razão de verossimilhança se afasta de 1.

Independência condicional entre dados.

Quando analisamos a influência de uma série de dados $D_1, D_2, \ldots, D_n$ sobre uma hipótese H, é particularmente interessante que tais dados sejam condicionalmente independentes. Dizemos que $D_1, D_2, \ldots, D_n$ são condicionalmente independentes se

$$\begin{cases} P(D_1|H, D_2, \ldots, D_n) = P(D_1|H) \\ P(D_2|H, D_1, D_3, \ldots, D_n) = P(D_2|H) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ P(D_n|H, D_1, D_2, \ldots, D_{n-1}) = P(D_n|H) \end{cases}$$

Dada uma hipótese H e uma série $D_1, D_2, \ldots, D_n$ de dados, podemos escrever

$$\Omega(H|D_1, D_2, \ldots, D_n) = \frac{P(H, D_1, D_2, \ldots, D_n)}{P(\bar{H}, D_1, D_2, \ldots, D_n)}$$

ou

$$\Omega(H|D_1, \ldots, D_n) = \frac{P(D_n|H, D_1, \ldots, D_{n-1})\ P(H, D_1, D_2, \ldots, D_{n-1})}{P(D_n|\bar{H}, D_1, \ldots, D_{n-1})\ P(\bar{H}, D_1, D_2, \ldots, D_{n-1})}$$

$$= \frac{P(D_n|H, D_1, \ldots, D_{n-1})}{P(D_n|\bar{H}, D_1, \ldots, D_{n-1})}\ \Omega(H|D_1, D_2, \ldots, D_{n-1})$$

Havendo independência condicional entre os dados, teremos

$$\Omega(H|D_1, \ldots, D_n) = \frac{P(D_n|H)}{P(D_n|\bar{H})}\ \Omega(H|D_1, D_2, \ldots, D_{n-1})$$

$$= R(D_n, H)\ \Omega(H|D_1, D_2, \ldots, D_{n-1})$$

Prosseguindo em manipulações algébricas análogas, chegamos a

$$\Omega(H|D_1, \ldots, D_n) = \prod_i^n R(D_i, H)\ \Omega(H) \qquad (3.4)$$

Esta expressão é especialmente cômoda: as "razões de verossimilhança" (que indicam a relevância de cada dado em particular para a

hipótese considerada) podem ser avaliados independentemente de considerações sobre os demais dados. Os "odds finais" saem, então, como o resultado do produto das diversas razões e dos "odds a priori".

Embora, na prática, a maioria dos dados possam ser tratados como condicionalmente independentes, é preciso que se proceda com cuidado em certos casos.

Consideremos, por exemplo*, as informações "o suspeito de assassinato é canhoto" e "o golpe fatal foi desferido por trás e pela esquerda". Tais dados considerados juntos tem muito mais influência sobre a probabilidade de que "o suspeito é o assassino" do que teriam se analisados separadamente.

Uso da Função Logaritmo.

A forma de produtório da expressão (3.4) sugere-nos que alguma simplificação pode ser obtida pelo uso da função logaritmo. Com efeito, tomando-se o logaritmo dessa expressão teremos:

$$\text{Log}\ \Omega(H|D_i, \ldots, D_n) = \sum_{i=i}^{n} \text{Log}\ R(D_i, H) + \text{Log}\ \Omega(H)$$

Usando-se uma escala logarítmica calibrada em "odds", o

---

* *Este exemplo foi sugerido por W. Edwards* [10].

processamento bayesiano se resumirá em uma adição de segmentos; o valor final de $\Omega(H|D_i, \ldots, D_n)$ será lido diretamente na escala.

O grupo do professor Ward Edwards, no Engineering Psychology Laboratory, de Michigan, utiliza habitualmente uma escala desse tipo.

3.3 - PARÂMETROS CONTÍNUOS

3.3.1 - Introdução.

Comumente, na prática, trabalhamos com parâmetros que podem assumir quaisquer valores dentro de um intervalo. Um caso de grande interesse prático é o da "proporção de elementos com um dado atributo em uma determinada população". Essa proporção é uma variável típica que, em princípio, pode assumir qualquer valor entre 0 e 1. Podemos estar interessados na proporção de componentes defeituosos em um determinado lote, ou na proporção de escolares com condições mínimas de alimentação em uma determinada escola, ou na proporção de médicos em uma dada população.

Nossa variável pode ser ainda a nota de um aluno que pode assumir qualquer valor entre 0 e 100. Podemos estar interessados, igualmente, no volume de vendas de um determinado produto que pode assumir, em princípio, qualquer valor entre 0 e a capacidade máxima de produção da fábrica.

Notemos que, em grande número de casos, nossa variável não é propriamente contínua. A proporção de indivíduos com uma determinada característica em uma dada população de 1000 indivíduos, na realidade, pode assumir apenas um dos 1001 valores: (0), (0,001), (0,002), ..., (0,999), (1,000). Contudo, para efeitos de codificação de incerteza, nosso trabalho será extremamente simplificado se tratarmos essa variável aleatória como estritamente contínua.

O comportamento incerto dessas variáveis é retratado na teoria de probabilidades pela distribuição acumulada ou pela função densidade de probabilidade. O que se faz, na prática, ao se codificar a incerteza associada a uma variável aleatória contínua, é obter uma curva que represente a distribuição acumulada de probabilidades dessa variável. Se estivermos interessados na função densidade de probabilidades podemos obtê-la diretamente da primeira, como veremos posteriormente.

3.3.2 - Como obter uma curva de distribuição subjetiva de probabilidades.

Um indivíduo com vivência de um determinado problema possui em sua mente conhecimentos e informações que qualificam as previsões que possa fazer sobre o comportamento de uma variável incerta, diretamente envolvida no contexto do problema. Em falta de melhores dados, essas previsões podem ser extremamente úteis. Cabe ao Analista de Decisões, analisando o problema, retirar do indivíduo essas opiniões quantificadas na

forma de distribuição de probabilidades.

A pessoa cujas informações devem ser quantificadas é o perito ou "expert", isto é, o indivíduo com maior conhecimento específico sobre a variável em questão. Chamaremos tal indivíduo de Estimador.

Existem dois procedimentos básicos para se tentar codificar essas informações na forma de uma curva de distribuição de probabilidades. Ambos se baseiam em respostas fornecidas diretamente pelo Estimador a questões propostas pelo Analista.

Os dois métodos possuem uma primeira fase comum. Nessa primeira fase, o Analista procura pôr o Estimador à vontade, caracterizar bem a variável a ser investigada e assumir, se necessário, algumas condições de contorno*. É ainda nesta etapa que o Estimador fixa limites extremos para o intervalo no qual a variável em questão pode assumir valores. Em uma segunda fase cada método assume características distintas.

1) Método de Pontos.

O Estimador fornece ao Analista informações do tipo:

---

* *Ao se prever o comportamento das vendas de um determinado produto, por exemplo, pode ser conveniente que se fixe, como hipótese, que "o serviço de distribuição irá atender normalmente à demanda".*

(I) $P(X \geq a_1) = b_1$, $P(X \leq a_2) = b_2$, etc...

que dão diretamente pontos de uma curva de distribuição de probabilidades para a variável aleatória X. Com efeito, como $F(x) = P(X \leq x)$, teremos:

$$F(a_1) = 1 - b_1, \qquad F(a_2) = b_2, \text{ etc...}$$

Plotando em um gráfico os pares $(a_1, 1 - b_1)$, (a_2, b_2),etc. e interligando os pontos resultantes, teremos uma curva que representa a distribuição acumulada de X.

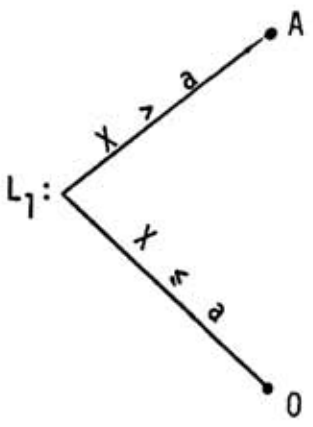
A questão é: como obter as respostas (I)?

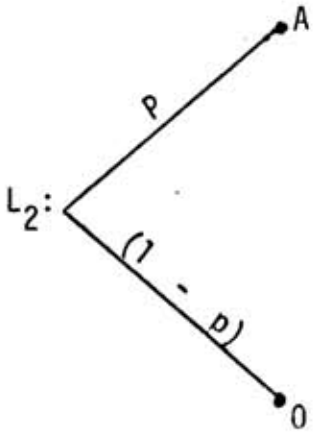
Essas podem ser as respostas a questões do tipo:

a) Em quanto você estimaria as chances de X ser maior (menor) que a?
b) Qual o valor p que o deixaria indiferente entre as loterias L_1, que lhe fornece o prêmio A* se X for maior que a e zero em caso contrário e L_2 que dá A com probabilidade p e zero com probabilidade (1 - p)?

---

* *A é um valor monetário positivo.*





Essa resposta pode ser mais facilmente obtida com o recurso do disco que descrevemos anteriormente. No caso, o disco faria o papel da loteria L_2.

A abordagem b é, em princípio, mais rigorosa. Exige, contudo, um esforço maior tanto por parte do Analista quanto do Estimador.

2) Método dos Intervalos

A idéia do método é conseguir do Estimador a subdivisão conveniente do intervalo de variação do parâmetro.

Suponhamos que X possa assumir valores entre 0 e 1. Através de um dos procedimentos que descrevemos a seguir, conseguimos que o Estimador subdivida o intervalo [0;1] em dois intervalos equiprováveis, digamos, [0;0,6] e [0,6; 1]. Isto significa que, para o Estimador é tão provável X cair entre 0 e 0,6 quanto entre 0,6 e 1,0. Em termos de lote

ria, ele é indiferente entre uma loteria que dá 10 milhões de cruzeiros caso X caia entre 0 e 0,6 (e nada em caso contrário) e outra loteria que dá o mesmo prêmio se X cair entre 0,6 e 1,0 (e nada em caso contrário). Em termos de probabilidade isto significa que $P(X \leq 0,6) = 0,5$. Temos portanto, um primeiro ponto de nossa curva.

Se conseguirmos agora que o Estimador subdivida o intervalo $[0; 0,6]$ em três intervalos equiprováveis, $[0; 0,38]$, $[0,38; 0,5]$, $[0,5; 0,6]$, por exemplo, teremos:

$$P(X \leq 0,38) = \frac{0,5}{3} = \frac{1}{6}$$

$$P(X \leq 0,50) = 2 \times \frac{1}{6} = \frac{1}{3}$$

o que nos fornece mais dois pontos para a função de distribuição acumulada:

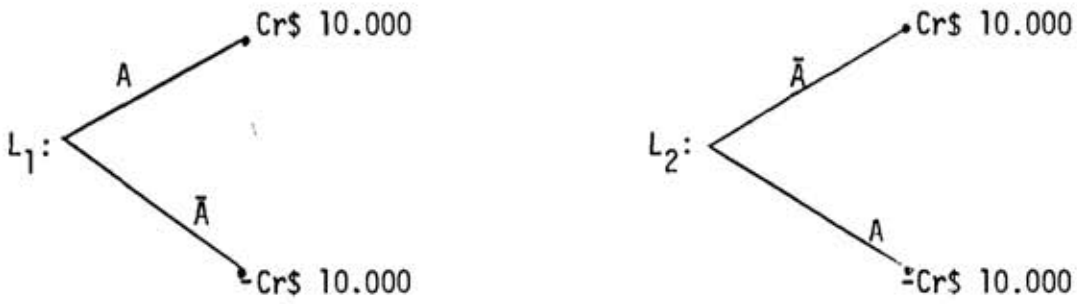
$$F(0,38) = \frac{1}{6} \quad \text{e} \quad F(0,50) = \frac{1}{3}.$$

Subdivisões subsequentes nos fornecerão novos pontos. Fazendo passar uma curva por esses pontos teremos uma representação gráfica da nossa função.

Há dois procedimentos básicos para se conseguir obter do Es

timador essa subdivisão em intervalos convenientes:

a) Pede-se inicialmente ao Estimador que subdivida o intervalo de variação do parâmetro em dois outros equiprováveis. Em outras palavras, pede-se ao Estimador um valor tal que ele sinta ser tão provável a variável cair acima quanto abaixo de tal valor. No exemplo considerado tal valor seria 0,6. A questão pode ser posta em termos de escolha entre loterias: o valor x_1 buscado é aquele que torna o Estimador indiferente entre as loterias L_1 e L_2 abaixo:



onde

A é o evento "o valor cai acima de x"

Ā é o evento "o valor cai abaixo de x"

Fixado esse primeiro valor, numa etapa seguinte tomamos, por exemplo, o intervalo limitado inferiormente por tal valor e superiormente pelo extremo absoluto do campo de variação do parâmetro (no caso mencionado , 1,0). A questão proposta ao Estimador tomaria uma forma análoga a: "Considere que a variável vai cair neste intervalo. Quero que vo

cê me forneça um valor x_2 que divida tal intervalo em dois equiprováveis".

Observemos que os valores x_1 e x_2, assim obtidos nos dariam as seguintes informações:

$$F(x_1) = 0,50$$

$$F(x_2) = 0,75$$

Prosseguindo no processo, iríamos considerando sempre novos intervalos e subdividindo cada um em dois outros equiprováveis, até dispormos de informação suficiente para traçar nossa curva.

b) Apresentamos ao Estimador alguns intervalos mutuamente exclusivos e pedimos que êle os ordene segundo as relações $\succ$ e $\equiv$,onde :

$1 \succ 2$ significa que é mais provável que nossa variável aleatória assuma um valor contido no intervalo 1 que no intervalo 2.

$2 \equiv 3$ significa que é tão provável que nossa variável assuma um valor contido em 2 quanto um valor contido em 3. Em outras palavras, 2 e 3 são equiprováveis.

Feita a ordenação, vamos alterando os intervalos até obter um conjunto de intervalos equiprováveis.

Para efeito de ilustração, vamos supor que estamos interessados na distribuição de uma variável que pode assumir valores entre 0 e 100.

Uma série de interações Analista-Estimador poderia ser ilustrada pela tabela abaixo:

TABELA 3.1

Exemplo do Processo Iterativo com Decisão entre Três Intervalos

INTERVALOS PROPOSTOS			OPINIÃO DO
1	2	3	ESTIMADOR
0 —— 50	50 —— 75	75 —— 100	$2 \succ 3 \succ 1$
0 —— 55	55 —— 70	70 —— 100	$2 \equiv 3 \succ 1$
0 —— 60	60 —— 70	70 —— 100	$3 \succ 1 \succ 2$
0 —— 60	60 —— 72	72 —— 100	$1 \equiv 2 \equiv 3$

Notemos, através do exemplo, que os novos intervalos a serem propostos em cada estágio dependem da resposta dada pelo Estimador no estágio anterior. O processo é, pois, iterativo.

No final desse primeiro ciclo de iterações, disporíamos, no caso do exemplo, dos seguintes resultados:

$$F(60) = P(X \leq 60) = \frac{1}{3}$$

$$F(72) = P(X \leq 72) = P(X \leq 60) + P(60 \leq X \leq 72)$$

$$= \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

Uma etapa seguinte seria, por exemplo, tomar o intervalo(60, 72) e subdividí-lo, digamos em dois outro equiprováveis. A tabela seguinte ilustra como isso poderia ser conseguido.

TABELA 3.2

Exemplo do Processo Iterativo com Decisão entre Dois Intervalos

INTERVALOS PROPOSTOS		OPINIÃO DO
1	2	ESTIMADOR
60 — 66	66 — 72	2 $\succ$ 1
60 — 70	70 — 72	1 $\succ$ 2
60 — 68	68 — 72	1 $\equiv$ 2

$$F(68) = P\ (X \leq 68) = P\ (X \leq 60) + P\ (60 \leq X \leq 68)$$

$$= P\ (X \leq 60) + \frac{1}{2}\ P\ (60 \leq X \leq 72)$$

$$= \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2}$$

Novos valores poderiam ser obtidos de maneira análoga.

Este método requer muita participação do analista, exigindo deste um maior treinamento. Pode, contudo, ser automatizado. Um programa interativo é usado pelo grupo de Análise de Decisões do Stanford Research Institute ("Probability Encoding Program"). O Estimador é posto diante de um terminal de computador e interage, diretamente com este. No final do processo a máquina imprime as curvas procuradas.

Consistência da Função Distribuição Obtida

Sabemos da teoria que uma função distribuição de probabilidades deve possuir algumas características. Assim, ela deve ser monotônica não decrescente, limitada superiormente por 1 e inferiormente por zero. No caso específico de uma variável aleatória contínua que assume valores em um intervalo limitado $[a, b]$, devemos ter $F(a) = 0$ e $F(b) = 1$. Portanto, para ser consistente com a teoria, a curva traçada pelos pontos obtidos através das técnicas apresentadas deve possuir tais características. Violações na forma da curva indicarão inconsistências nas opiniões subjetivas do Estimador. Detetada alguma inconsistência, o analista deve pedir ao Estimador que reveja as opiniões julgadas conflitantes.

Observemos, ainda, que inconsistências podem ser pressentidas através de observações de aspectos da curva que, embora teoricamente viáveis, não são de se esperar na prática. O bom senso do Analista deve funcionar em tais ocasiões. Na grande maioria dos casos reais, variações um tanto bruscas no comportamento das curvas de distribuição são altamente suspeitas.

Consideremos o exemplo da figura 3.3. Para evitar variação brusca na forma da curva, o Analista poderia simplesmente abandonar o ponto 4 ou adotar uma solução de compromisso passando a curva entre os pontos. As três possibilidades estão representadas na figura pelas curvas a, b, e c, respectivamente.

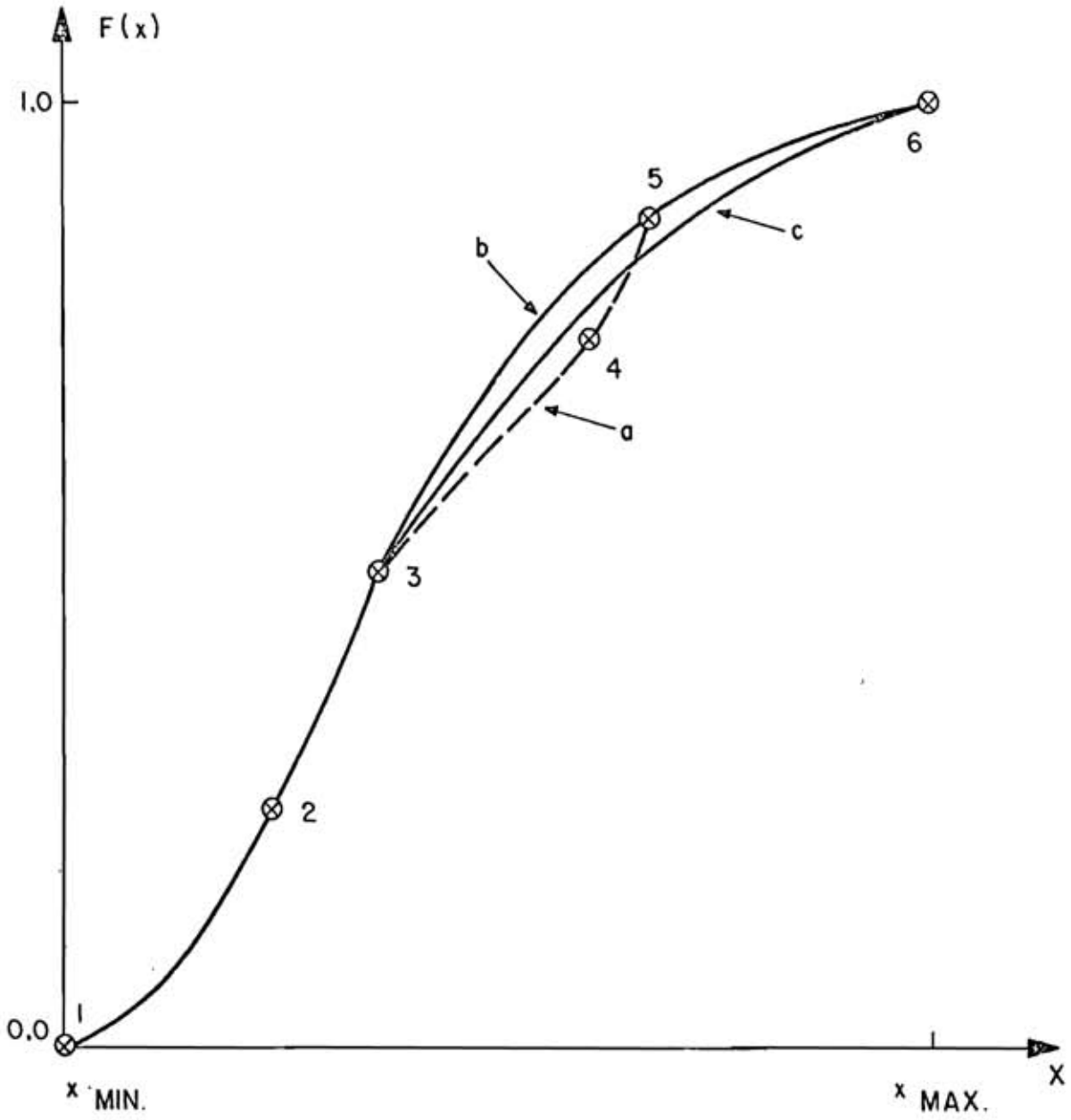


Figura 3.3 - Possíveis Variantes da Função Distribuição a Partir dos Pontos 1 a 6.

O procedimento mais rigoroso, porém, seria conduzir o Estimador a rever suas opiniões ou ainda tentar obter dêle mais alguns pontos para, em seguida, traçar uma curva"média".

3.3.3 - Obtenção da Função Densidade de Probabilidade Através da Função Distribuição

Sabemos da teoria que a função densidade de uma variável aleatória X se relaciona com a função distribuição dessa variável através da expressão:

$$f_X(x_o) = \left. \frac{d\, F_X(x)}{dx} \right|_{x = x_o}$$

A derivação gráfica, no entanto, pode levar a imprecisões muito grandes. O método que apresentaremos a seguir é sugerido por Raiffa [40] e acreditamos que, em face dos erros inerentes ao próprio processo de codificação de incertezas, apresenta suficiente precisão.

MÉTODO DO HISTOGRAMA

- Dividimos o intervalo de variação do parâmetro em um número n de subintervalos iguais.
- Para cada subintervalo (x_i, x_{i+1}) calculamos o valor θ_i dado por:

$$\theta_i = F(x_i + 1) - F(x_i)$$

- Construímos um histograma, associando a cada subintervalo o valor θ_i correspondente

- Aproximamos nosso histograma por uma curva suave passando por seus patamares. A curva obtida é proporcional à função desejada.
- Calculamos a área sob a curva e multiplicamos nossa escala vertical pelo fator κ, dado por:

$$\kappa = \frac{1}{\text{Área}}$$

A área sob a curva, considerando a nova escala, vale 1. A curva representa, agora, a função densidade de probabilidade subjetiva* da variável aleatória estudada. A precisão cresce, evidentemente, com o número n de subintervalos considerados (aconselhamos $n \geq 10$).

OBSERVAÇÕES:

1) Este método pressupõe que o parâmetro assuma valores em um intervalo finito. No caso pouco usual em que isto não ocorre devemos arbitrar dois valores limites L_1 e L_2 de tal forma que $F(L_2)-F(L_1) \geq 0{,}90$. Consideraremos então, para efeito do histograma, apenas o intervalo $[L_1, L_2]$ e extrapolaremos convenientemente a curva obtida para além dos limites fixados.

2)Caso a função distribuição apresente variação muito rápida dentro de certos intervalos ou muita lenta em outros, pode ser conveniente tomarmos subintervalos desiguais na construção do histograma. Nesse caso, deveremos associar a cada subintervalo $(x_i, x_i + 1)$ um valor θ_i^* dado

Subjetiva uma vez que estamos considerando que a função distribuição que a gerou é de natureza subjetiva.

por

$$\theta_i^* = \frac{F(x_i + 1) - F(x_i)}{x_{i+1} - x_i}$$

Nessas circunstâncias, o valor κ definido em (3.5) será sempre próximo de 1. Se tomarmos um número relativamente grande de subintervalos poderemos considerar κ praticamente igual a 1.

3.3.4 - Exemplo

Com o intuito de ilustrar as técnicas apresentadas, vamos construir um exemplo inteiramente hipotético. Admitamos que um certo partido político, cogitando da conveniência de lançar uma candidatura, deseja conhecer a opinião do povo de determinada cidade a respeito de certa personalidade local. Por razões políticas, contudo, está vetada a possibilidade de uma pesquisa de opiniões nos moldes tradicionais. Um analista, contratado para a tarefa, entra em contacto com um elemento da população local que parece conhecer bastante bem a opinião dos cidadãos da localidade e ser notoriamente apartidário. Usando a técnica de estimativa por intervalos, o Analista consegue que tal indivíduo forneça uma estimativa subjetiva da "proporção de cidadãos locais que apoiariam o candidato".

As informações fornecidas pelo Estimador estão sumarizadas na Tabela 3.3.

TABELA 3.3

Ilustração de um Possível Conjunto de Pontos de Indiferença Obtidos em um Processo de Estimativa por Intervalos

INTERVALO	PONTO DE INDIFERENÇA
0,0 - 1,0	0,60
0,60 - 1,0	0,68
0,68 - 1,0	0,75
0,75 - 1,0	0,80
0,0 - 0,60	0,50
0,0 - 0,50	0,42
0,0 - 0,42	0,36

Nessa tabela, "PONTO DE INDIFERENÇA" significa o ponto que divide o intervalo proposto em dois outros equiprováveis. Nossa variável aleatória, a "proporção de indivíduos favoráveis ao candidato", está evidentemente, contida no intervalo $[0, 1]$.

As informações da tabela acima, traduzidas em termos de probabilidades acumuladas, estão expressas na tabela 3.4.

TABELA 3.4

Pontos da Função Distribuição Obtidos a partir da Tabela 3.3

P	$F_p(P) = P(P \leq p)$
0,00	0,0000
0,36	0,0625
0,42	0,1250
0,50	0,2500
0,60	0,5000
0,68	0,7500
0,75	0,8750
0,80	0,9375
1,00	1,0000

O símbolo "P" indica a variável aleatória em questão e "p" representa valores entre 0 e 1.

Plotando os pontos da tabela em um gráfico e passando por êles uma curva suave, obtemos a função distribuição de P (figura 3.4).

Podemos, agora, obter a função densidade de probabilidade, $f_p(p)$, a partir da curva $F_p(p)$. Seguindo o método indicado no parágrafo anterior, construimos a tabela 3.5:

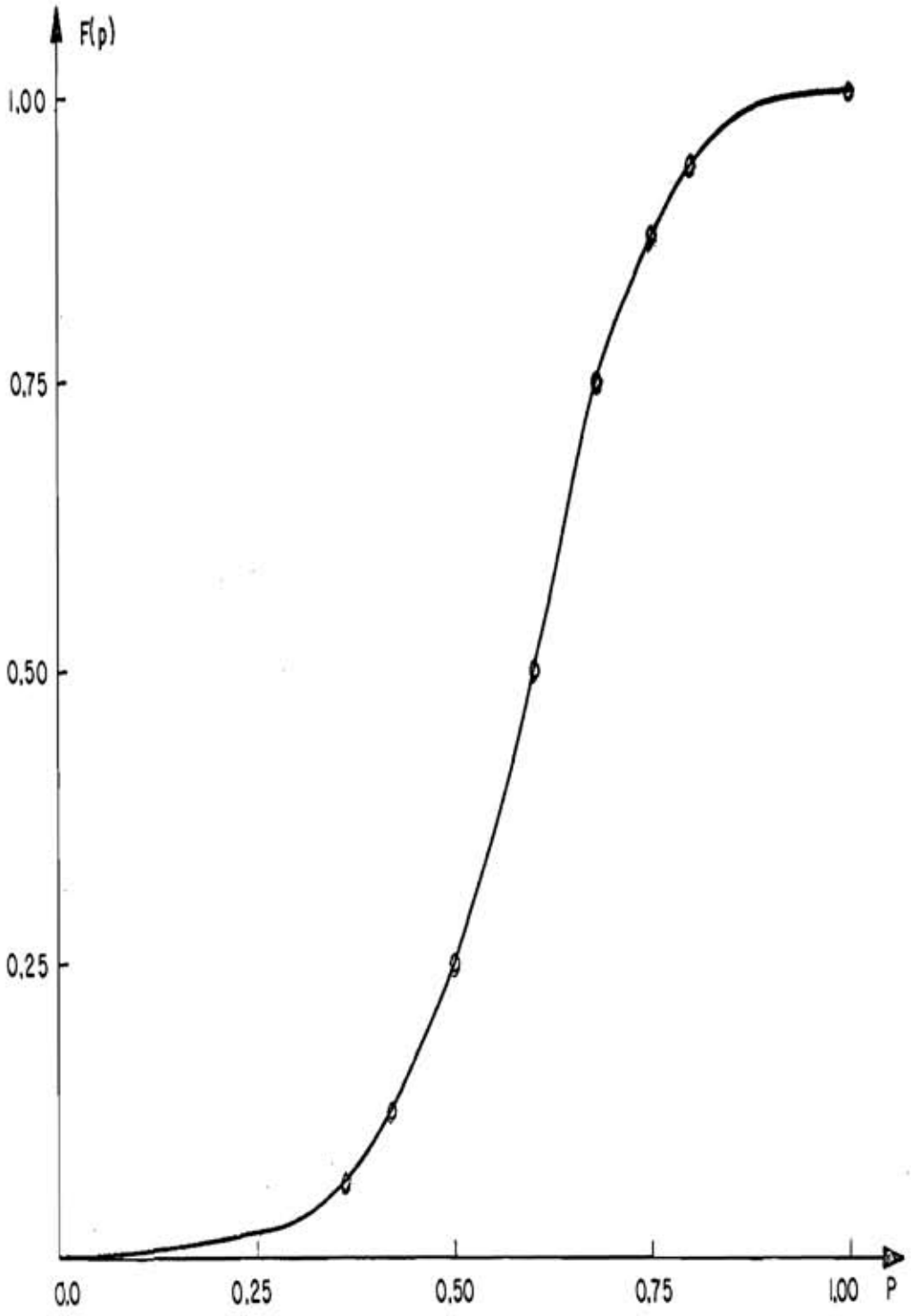


Figura 3.4 - Função Distribuição Subjetiva para a Fração de Eleitores Favoráveis ao Candidato Local.

TABELA 3.5

Pontos para Construção da Função $f_p(p)$, Obtidos da Distribuição Acumulada

INTERVALO i	θ_i
0 —— 0,1	0,005
0,1 —— 0,2	0,010
0,2 —— 0,3	0,015
0,3 —— 0,4	0,070
0,4 —— 0,5	0,150
0,5 —— 0,6	0,250
0,6 —— 0,7	0,280
0,7 —— 0,8	0,150
0,8 —— 0,9	0,050
0,9 —— 1,0	0,010

Construindo o histograma e aproximando-o por uma curva contínua, obtemos $g_p(p)$ que, a menos de uma constante de escala κ, representa a função densidade procurada.

Como foi visto, o fator de escala pode ser obtido calculando-se a área sob a curva $g_p(p)$. Uma primeira aproximação para κ é dada, contudo, pelo fator que torna unitária a área do histograma. No caso em que usamos intervalos iguais esse fator é dado por $\frac{1}{\Delta p}$, Δp sendo o comprimento de cada intervalo. Portanto

$$\kappa \cong \frac{1}{0,1} = 10.$$

A figura 3.5 representa a curva obtida.

3.3.5 - Revisão da Função Densidade de Probabilidade em Face de Informações Amostrais

No ítem 3.2.3 mostramos como o teorema de Bayes nos permite agregar novas informações às probabilidades "a priori" de um conjunto de eventos possíveis. Neste ítem, procuraremos mostrar, sem grande profundidade, como informações amostrais podem ser usadas para melhorar nossa opinião "a priori" sobre o comportamento de uma variável aleatória contínua. Em outras palavras, mostraremos como usar dados experimentais para revisar uma função densidade de probabilidade.

A Função de Verossimilhança

Consideremos a variável aleatória X e a informação amostral Z. Definimos a função de verossimilhança* de X e Z, L(z; x)** como:

a) $L(z; x) = P(Z = z|X = x)$ se a informação Z puder ser representada como uma variável aleatória discreta.

b) $L(z; x) = f(z|x)$, onde $f(z|x)$ é a função densidade de Z condicionada a X, se Z puder ser representada como uma variável alea

---

* *Em inglês, LIKELIHOOD FUNCTION.*

** *Para uma definição mais rigorosa, ver Raiffa & Schlaifer* [41].

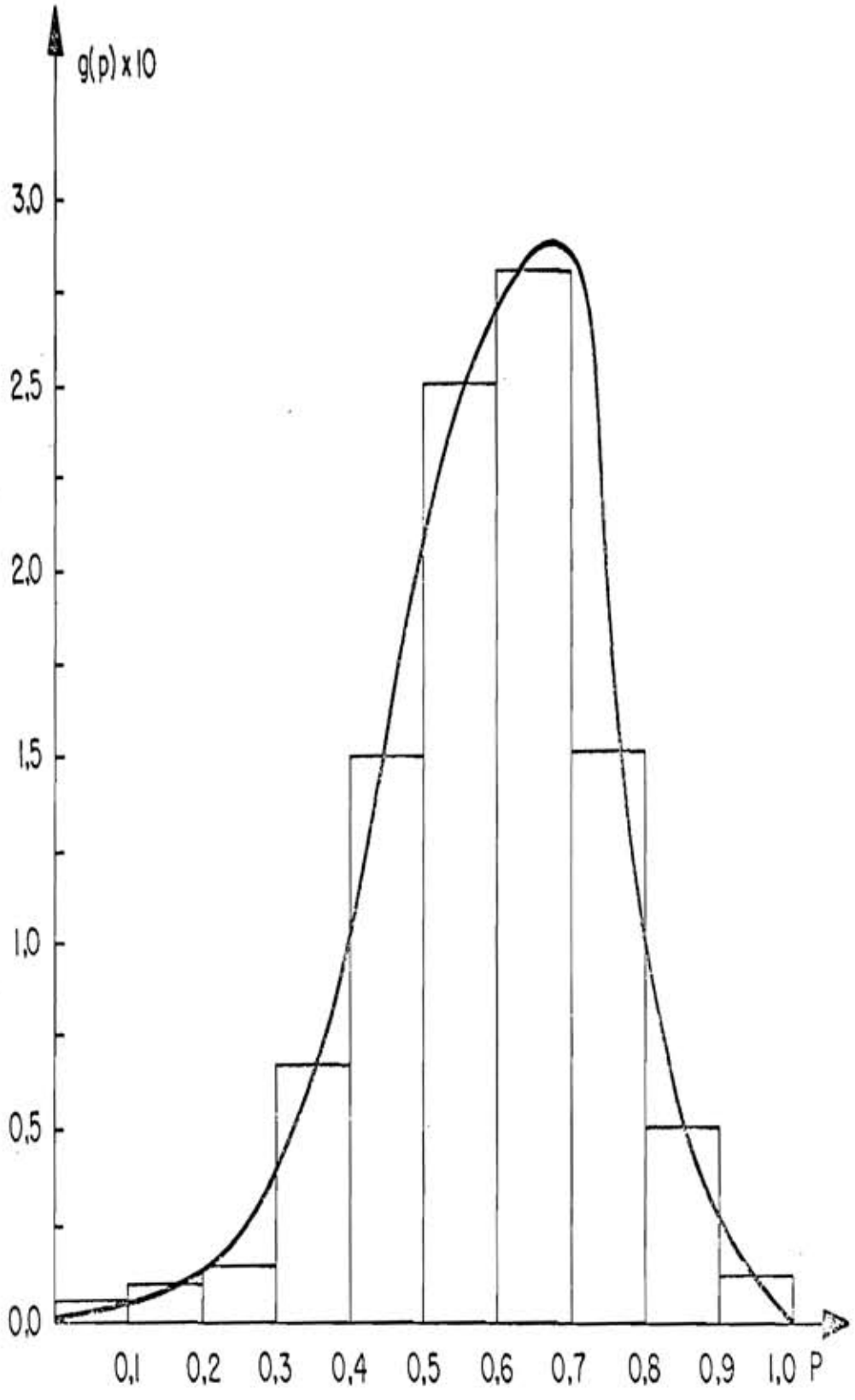


Figura 3.5 - Curva Proporcional à Função Densidade "a priori" Obtida a Partir da Figura 3.4.

tória contínua.

Exemplos

1) Seja P a variável aleatória que representa a proporção de elementos que apresentam um determinado atributo, dentro de uma população de M. A informação contida em um experimento que toma uma amostra aleatória de N elementos (N << M) e verifica o número R desses N que possuem o atributo A, pode ser representada por Z = (N, R).

A função de verossimilhança nesse caso será:

$L(z; p) = P(Z = z|P = p) = P(N = n, R = r \mid P = p) = P$ ("r elementos com o atributo A em uma amostra de n"|" a fração de elementos com o atributo A na população total é p").

Isto caracteriza perfeitamente um processo de Bernoulli e nos permite concluir:

$$L(z; p) = \binom{n}{r} p^r (1 - p)^{n-r} .$$

2) Consideramos uma variável aleatória X com distribuição normal, variância σ^2 conhecida e média μ desconhecida. A média μ pode ser tratada como uma variável aleatória e podemos associar a ela uma distribuição "a priori", fornecida por séries históricas ou julgamentos

subjetivos.

Consideremos a informação amostral Z associada ao experimento que leva em conta n observações independentes de X. Neste caso $Z = (X_1, X_2, \ldots, X_n)$, isto é, uma variável aleatória n-dimensional. Cada variável aleatória X_i representa os possíveis resultados da observação i e tem a mesma distribuição de X. Nessas condições

$$L(z; \mu) = g(x_1, x_2, \ldots, x_n|\mu) = f(x_1|\mu)\, f(x_2|\mu) \ldots f(x_n|\mu)$$

ou

$$L(z; \mu) = \prod_{i=1}^{n} \left\{ \frac{1}{\sqrt{2\pi}\,\sigma} \exp \left| -\frac{1}{2} \left[\frac{x_i - \mu}{\sigma} \right]^2 \right| \right\}$$

$$= \left(\frac{1}{\sqrt{2\pi}\,\sigma} \right)^n \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right]$$

Teorema de Bayes para Variáveis Contínuas

Seja $f_a(x)$ a função densidade "a priori" de uma v.a. X, Z a informação amostral e $f_d(x|z)$ a função densidade de X condicionada ao conhecimento de Z. O teorema de Bayes pode ser escrito na forma*:

$$f_d(x|z) = f_a(x) \cdot L(z; x)\, N(z) \qquad (3.6)$$

desde que

$$\int_{-\infty}^{+\infty} f_a(x)\, L(z; x)\, dx \neq 0$$

e onde N(z) é uma constante de normalização determinada pela condição:

---

* *Ver Raiffa e Schlaifer* [41].

$$\int_{-\infty}^{\infty} f_d(x|z)dx = N(z) \int_{-\infty}^{\infty} f_a(x)\, L(z;\, x)\, dx = 1$$

A expressão (3.6) pode ser generalizada se considerarmos duas funções $g_a(x)$ e $l(z;\, x)$ tais que:

$$f_a(x) = \kappa\, g_a(x)$$

$$L(z;\, x) = K(z)\, l(z;\, x)$$

onde κ e $K(z)$ são constantes em relação a x.

Diremos que $g_a(x)$ e $l(z;\, x)$ são proporcionais, respectivamente, a $f_a(x)$ e $L(z;\, x)$.

Podemos, agora, reescrever o teorema de Bayes como:

$$f_d(x|z) = M(z)\, l(z;\, x)\, g_a(x) \qquad (3.7)$$

onde a constante de normalização $M(z)$ pode ser calculada pela condição

$$M(z) \int_{-\infty}^{\infty} l(z;\, x)\, g_a(x)\, dx = 1 \qquad (3.8)$$

desde que:

$$\int_{-\infty}^{\infty} l(z;\, x)\, g_a(x)\, dx \neq 0$$

A expressão (3.7) é especialmente cômoda pois, ela nos diz que conhecendo-se duas funções proporcionais, respectivamente, à função densidade "a priori" e à função de verossimilhança, a função densidade revisada ("a posteriori") pode ser obtida fazendo-se o produto dessas duas

funções e procedendo-se à normalização, de modo a garantir que a área sob a curva resultante seja unitária.

Se $g_a(x)$ e $l(z; x)$ forem conhecidas analiticamente, fazendo uso das expressões (3.7) e (3.8) obteremos, normalmente, $f_d(x|z)$ em forma analítica.

Em particular, existem numerosos casos teóricos já bem estudados e de solução conhecida. O leitor interessado poderá encontrar muitas informações em Raiffa & Schlaifer [41].

A situação mais frequente, porém, em Análise de Decisões, é aquela em que conhecemos a função densidade "a priori" na forma de uma curva como a da figura 3.5. A função de verossimilhança, em geral, é mais fácil de se conhecer analiticamente através da escolha adequada de um modelo estatístico (como nos exemplos apresentados). Caso isso não ocorra, entretanto, podemos tentar levantar uma curva subjetiva para a função de verossimilhança. Se a informação Z for uma variável discreta, teremos que estimar $P(Z = z_0|X = x)$ para diferentes valores de x. Caso Z seja uma variável contínua, o problema se complica e aconselhamos, neste caso, que se busque diretamente $f_d(x|z)$, usando as técnicas apresentadas para levantamento de curvas de densidade. O Estimador deve, simplesmente, incorporar a informação amostral às informações já armazenadas em sua mente e proceder a novas estimativas.

Considerando-se que $g_a(x)$ ou $l(z; x)$ (ou ambas) são conhecidas apenas na forma gráfica, existem duas alternativas básicas:

1) APROXIMAR AS CURVAS POR FUNÇÕES ANALÍTICAS

Isto pode ser feito mais facilmente com o uso do computador. Obtidas expressões analíticas para as funções, a função densidade revisada pode ser calculada pelas expressões (3.7) e (3.8).

2) CONSTRUIR GRAFICAMENTE A FUNÇÃO REVISADA

– Para uma série de pontos convenientemente escolhidos, x_i, calculamos

$$g_d(x_i|z) = l(z; x_i)\, g_a(x_i)$$

– Plotamos em um gráfico os pontos $(x_i, g_d(x_i|z))$ e passamos por esses pontos uma curva suave.

– Calculamos a área sob a curva obtida e efetuamos uma mudança de escala, de modo a garantir que a área seja unitária. Em relação à nova escala, a curva representará, agora, a função densidade revisada.

EXEMPLO

No exemplo do item 3.3.4 obtivemos a função densidade de probabilidade subjetiva da "proporção de cidadãos favoráveis ao candidato".

Consideremos agora que o Analista, desejando melhorar a informação já obtida, escolhe aleatoriamente dez indivíduos da população e verifica que, desses, oito são favoráveis à candidatura da personalidade local.

A função de verossimilhança para este tipo de problema foi discutida no EXEMPLO 1 deste item. O resultado obtido nos permite escrever:

$$l(z = (8,10);\ p) = p^8(1 - p)^{10 - 8}$$

ou

$$l(r = 8,\ n = 10;\ p) = p^8(1 - p)^2$$

Podemos, agora, obter a função densidade revisada. Usaremos a notação $g_{10,8}(p)$ para indicar uma função proporcional a $f_d(p|n=10,\ r=8)$. Tomando dez valores de p, igualmente espaçados, construímos a tabela abaixo:

TABELA 3.6

Valores para Determinação da Distribuição "a posteriori" de P.

P_i	$(1 - p_i)$	p_i^8	$(1 - p_i)^2$	$g_a(p_i)$	$g_{10,8}(p_i)$
0,1	0,9	1.0×10^{-6}	0,81	$0,7 \times 10^{-2}$	$5,68 \times 10^{-8}$
0,2	0,8	$2,56 \times 10^{-4}$	0,64	$1,7 \times 10^{-2}$	$27,9 \times 10^{-8}$
0,3	0,7	$6,56 \times 10^{-5}$	0,49	$0,38 \times 10^{-1}$	122×10^{-7}
0,4	0,6	$6,61 \times 10^{-4}$	0,36	$1,07 \times 10^{-1}$	255×10^{-6}
0,5	0,5	$39,2 \times 10^{-4}$	0,25	2.10×10^{-1}	206×10^{-5}
0,6	0,4	$1,67 \times 10^{-2}$	0,16	$2,85 \times 10^{-1}$	760×10^{-5}
0,7	0,3	$5,57 \times 10^{-2}$	0,09	$2,00 \times 10^{-1}$	103×10^{-4}
0,8	0,2	$16,8 \times 10^{-2}$	0,04	$0,80 \times 10^{-1}$	538×10^{-3}
0,9	0,1	$43,0 \times 10^{-2}$	0,01	$0,25 \times 10^{-1}$	108×10^{-5}
1,0	0,0	1,0	0,00	0,00	0,0

Os valores $g_a(p_i)$ foram obtidos da figura (3.5) e $g_{10,8}(p_i)$ calculado como o produto dos valores das colunas 3, 4 e 5.

Plotando em um gráfico os valores p_i nas abcissas e $g_{10,8}(p_i)$ nas ordenadas, obtemos uma série de pontos e, passando por eles uma curva suave, teremos uma representação aproximada de $g_{10,8}(p)$. Observando os pontos obtidos verificamos que, para uma maior precisão, deveríamos incluir na tabela (3.6) mais alguns valores de p_i, situados entre 0,5 e 0,9. Calculando a área* sob a curva $g_{10,8}(p)$ obtivemos

$$A \cong 2{,}63 \times 10^{-3}$$

e, portanto, a constante de normalização deve ser

$$\kappa = \frac{1}{A} \cong 380$$

Na figura (3.6) representamos, simultaneamente, $f_a(p)$ e $f_d(p|n = 10, r = 8)$, para efeito de comparação. A curva f_d representa $f_d(p|n = 10, r = 8)$, se fizermos a leitura na escala da direita e $g_{10,8}(p)$, se fizermos a leitura na escala da esquerda. A curva f_a representa, na escala da direita, a função densidade "a priori" $f_a(p)$.

Observando os gráficos, constatamos que a informação amostral deslocou a curva para a direita e diminuiu a variância da distribuição (a nova curva é mais concentrada).

---

* *Utilizamos o método de aproximação por trapézios.*

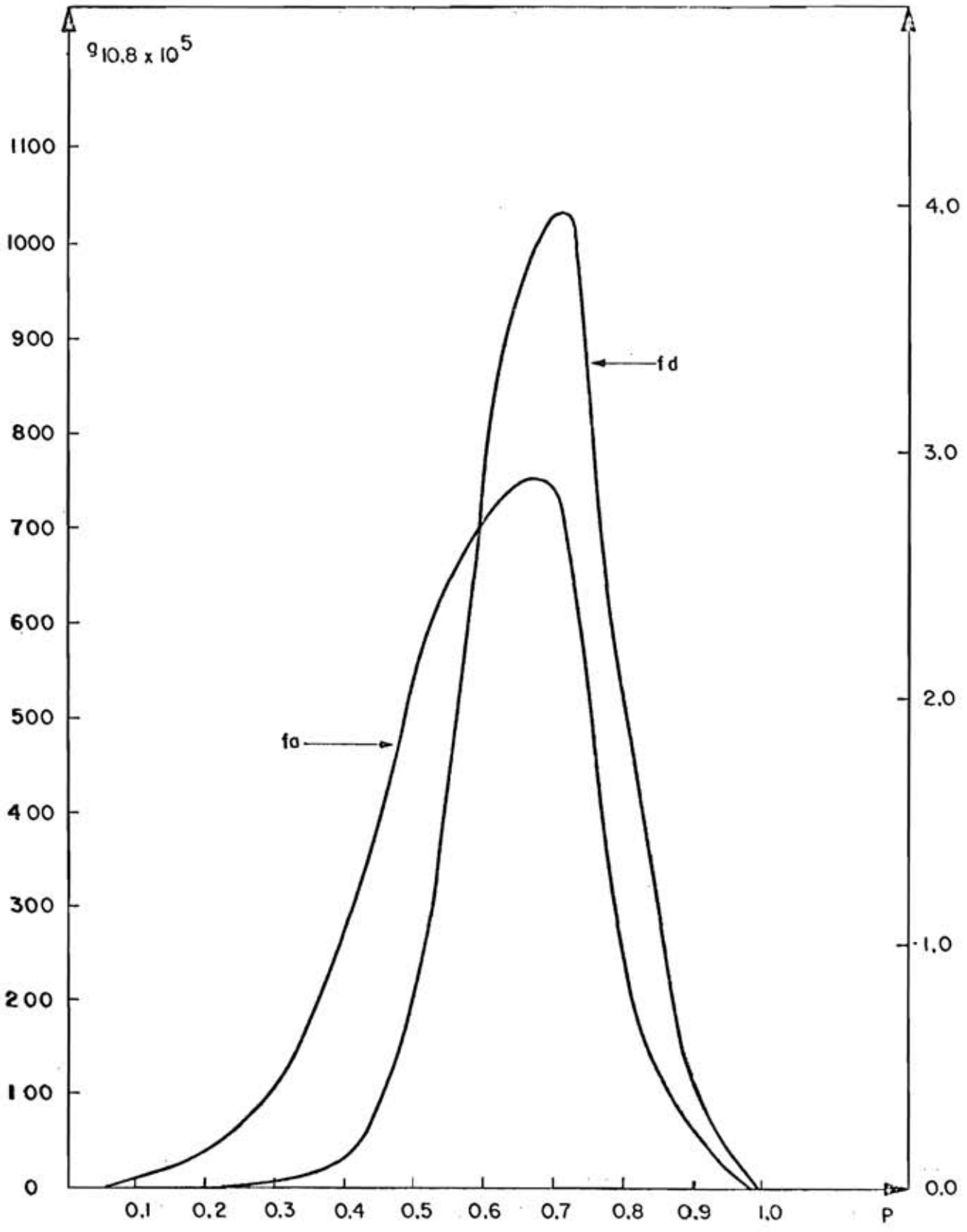


Figura 3.6 - Funções Densidade: "a priori" (f_a) e Revisada pela Informação de Oito Casos Favoráveis em Dez Observadas (f_d).

Esta última observação é perfeitamente lógica e de caráter geral: o acúmulo de informações tende a diminuir a incerteza e, consequentemente, a variância. Essa diminuição é tanto maior quanto maior for a informação amostral. No caso limite, se verificássemos todos os indivíduos da população, saberíamos ao certo qual a proporção de elementos com o atributo pesquisado. A função de massa se resumiria, então, em um único ponto com probabilidade 1 (variância nula). Na figura (3.7), vemos como a função densidade "a priori" (f_a) seria alterada se em uma amostra de 5 constatássemos 2 casos favoráveis ($f_{5,2}$) ou se em uma amostra de 50 constatássemos 28 favoráveis ($f_{50,28}$).

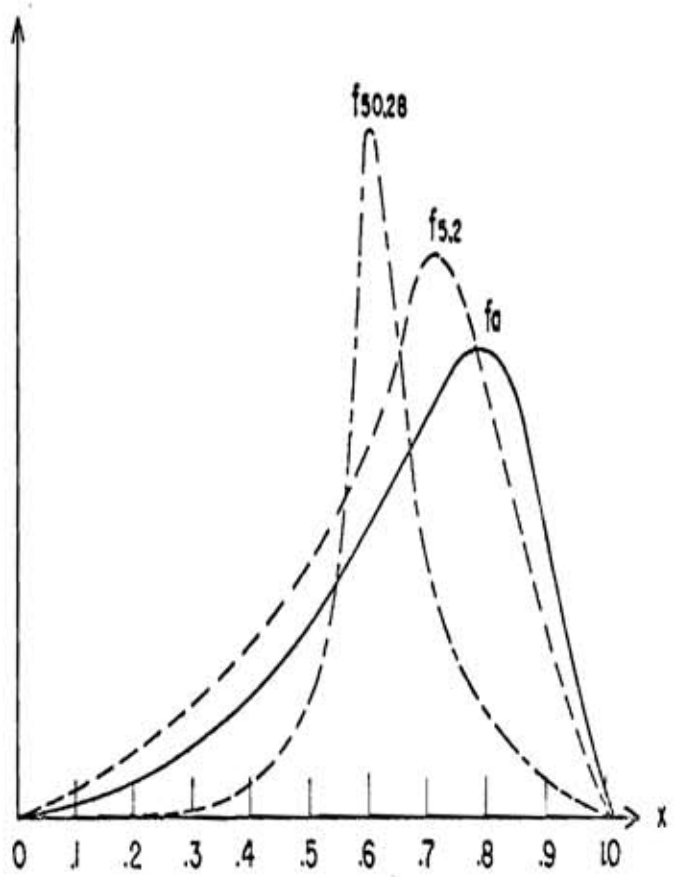


Figura 3.7 - Efeito de duas Informações Amostrais Diferentes sobre uma Função Densidade "a priori".

CAPÍTULO IV

TEORIA DA UTILIDADE

4.1 - INTRODUÇÃO

Quando o decisor seleciona uma alternativa entre as várias de que dispõe, dois fatores influenciaram sua decisão: suas preferências pelas consequências associadas a cada uma delas e as probabilidades de que estas consequências venham a se realizar.

Um dos problemas na análise de uma decisão é a quantificação das preferências já que, realmente, não é tarefa fácil atribuir valores numéricos a alternativas de ação. Na prática, o que se faz é confinar a atenção a aspectos menores de um problema, aspectos esses que são relevantes para o problema particular a ser resolvido.

Neste capítulo e no subsequente, veremos como quantificar as preferências do decisor em relação às alternativas apresentadas a ele por um dado contexto.

4.2 - ASPECTOS GERAIS

4.2.1 - Conceito de Utilidade

A tentativa de usar a utilidade como um critério de escolha

surgiu com um artigo escrito por Daniel Bernoulli, em 1783. No início, ela foi concebida como mensurável, ou seja, a cada entidade (objeto, situação, etc.) podia ser associado um número denominado utilidade daquela entidade. A utilidade seria,então, o valor que um indivíduo, o decisor, atribuísse a essa entidade. Críticas surgiram contra esta concepção inicial e, durante muito tempo, pensou-se em utilidade medida num sentido ordinal, isto é, apenas se falava em ordenação de preferências.

Assim a análise de curvas de indiferença, baseada na noção de utilidade ordinal, foi suficiente para sustentar a teoria da escolha determinística e quase toda a teoria econômica do consumidor tem por base esta concepção de utilidade.

Em anos recentes, voltou-se a falar em utilidade numericamente mensurável. Isto se deu a partir de 1944, com a publicação do trabalho de von Neumann e Morgenstern, "Theory of Games and Economic Behavior" [35]. O conceito de utilidade, no sentido de von Neumann e Morgenstern, é derivado a partir de determinadas hipóteses sobre o comportamento individual, as quais discutiremos posteriormente. Desde que estas hipóteses, que são as mesmas sobre as quais se baseia a análise de curvas de indiferença, sejam satisfeitas, pode-se falar em utilidade numérica.

A fim de salientar a importância da ideia de utilidade, consideremos as seguintes situações, que estão relacionadas ao lançamento de uma moeda:

Situação 1 - Há um prêmio de Cr$ 10,00 se cai cara e uma multa de Cr$2,00, se coroa.

Situação 2 - Tudo o que uma determinada pessoa possui são Cr$ 30 000,00. Ela recebe mais Cr$ 20 000,00 se cai cara e perde toda sua fortuna, se coroa.

Situação 3 - Um indivíduo dispõe de Cr$ 10,00 e seu maior desejo no momento é assitir a um jogo de futebol cujo preço é Cr$ 20,00. Se cai cara, ele recebe mais Cr$ 20,00; se coroa perde Cr$ 10,00.

Quando colocadas frente a estas situações e indagadas se estão dispostas a assumir o risco de cada uma delas, diferentes pessoas reagem diferentemente. Qual é a razão desse comportamento? O que parece claro é que a essência dessa diferença é o valor do dinheiro ganho ou perdido. Em outras palavras, o valor do dinheiro para seu proprietário não é proporcional à sua quantidade. Quando grandes quantias são envolvidas, nem as pessoas nem as empresas comportam-se como se todas as quantidades adicionais de dinheiro fossem de igual valor. Na maioria das vezes, a possibilidade de uma perda, mesmo com probabilidade pequena, faz com que, por exemplo, um empresário se arrisque somente se o ganho potencial for suficientemente grande para justificar o jogo.

Vemos assim, que os indivíduos diferem em suas atitudes frente ao risco e que essa diferença influencia suas escolhas. Podemos, então, dizer que as preferências por dinheiro são bastante desiguais, variando de

indivíduo para indivíduo e que, mesmo para uma determinada pessoa, a estrutura de preferências dependerá das quantidades envolvidas. Portanto, é extremamente importante ter uma medida de valor que não possua as deficiências do dinheiro e que seja, ao mesmo tempo, aplicável à avaliação de entidades que não podem ser traduzidas em termos monetários (benefícios sociais de determinado programa, efeito junto ao público de uma medida política,etc.)

A teoria da utilidade diz respeito justamente às preferências individuais e às suposições sobre estas preferências que possibilitam representá-las em uma maneira numericamente conveniente. Isto pode ser feito de várias maneiras, a mais comum das quais envolve uma tentativa de identificar uma função utilidade para um indivíduo, observando suas respostas a situações que envolvem probabilidades.

Bernoulli e Marshall* argumentaram que se a utilidade cresce linearmente com a quantidade de dinheiro, então as pessoas aceitarão um jogo que com 50% de chance lhe trará uma perda de Cr$ 1,00, se esta for compensada por uma chance igual de ganhar Cr$ 1,00. Se por outro lado, a utilidade adicional vai diminuindo à medida que as quantias envolvidas vão se tornando maiores, o risco de perder deverá ser compensado pela probabilidade de um ganho maior.

---

* *Citados por Samuelson* [43], *pp. 112*

4.2.2 - Diferentes Atitudes Frente ao Risco

Vimos que um problema típico de decisão é caracterizado por um conjunto de ações, a cada uma das quais é associado um conjunto de consequências. Qual consequência se concretiza depende do estado da natureza ou seja, de qual evento aleatório dentre um conjunto deles ocorre. A solução do problema é encontrada quando se determina a melhor ação a ser adotada. Admitindo que o decisor age sempre de maneira a maximizar alguma coisa, definimos como ação ótima aquela cujo valor para êle é máximo. E é claro que para determinar o melhor é preciso que avaliemos cada curso alternativo de ação. Se o problema envolve incerteza, usa-se o valor esperado para um ato, que é definido como a média ponderada de todas as consequências associadas àquele ato, onde os pesos são as probabilidades de que cada consequência venha a se realizar.

Assim temos:

$$V(a_j) = \sum_i x_{ij} p_i$$

onde $V(a_j)$ - Valor esperado de a_j

x_{ij} - i-ésima consequência associada ao ato j

p_i - probabilidade de que S_i seja o estado real, neste caso sendo x_{ij} a consequência obtida.

A definição acima independe da unidade de valor pela qual as consequências são traduzidas.Quando estes valores são em termos monetá

rios, falamos em "valor monetário esperado". Apesar de ser bastante frequente representar-se o valor de uma consequência por uma soma em dinheiro, nem sempre o valor monetário esperado é um guia válido para ação. Consideremos, por exemplo, o caso de dois gerentes que devem decidir se aceitam ou não um projeto que, com probabilidade 0,5, dará um retorno de Cr$ 40 000,00 (ou um prejuizo de Cr$ 20 000,00 com probabilidade o,5). Na tabela 4.1 temos os retornos e respectivas probabilidades, correspondentes às decisões (atos) de aceitar o projeto ou não aceitá-lo.

TABELA 4.1 - RECOMPENSAS E PROBABILIDADES

PROBABILIDADES	ATOS	
	ACEITA	NÃO ACEITA
0,5	40 000	0
0,5	-20 000	0
VALOR ESPERADO	10 000	0

O valor esperado do ato "aceitar o projeto" é Cr$ 10 000,00, enquanto que o de "não aceitar" é zero. Dependendo da situação específica de cada empresa, os dois indivíduos poderão chegar a decisões opostas. No caso por exemplo, de um deles ter problemas de caixa, a possibilidade de uma perda de Cr$ 20 000,00 poderia fazer com que ele rejeitasse o projeto, ao mesmo tempo em que o outro poderia aceitá-lo. Vemos, assim, que, embora o valor esperado seja positivo, o padrão individual de julgamento do decisor é que, na maioria das vezes, determina a escolha. Isto porque o que

constitui "uma grande quantidade de dinheiro" varia de pessoa a pessoa. Assim, uma perda de Cr$ 100.000,00 para muitas pessoas seria uma catástrofe, forçando-as a mudanças radicais no estilo de vida, padrão de gastos, etc., enquanto que para outras, uma companhia de seguros por exemplo, uma saída de Cr$ 100 000,00 é quase que uma rotina do dia-a-dia.

Segundo a atitude que adotam face ao risco, de uma maneira geral, podemos agrupar os indivíduos em três categorias: os chamados "médios",ou indiferentes ao risco, os aversos ao risco e aqueles que sentem-se bem assumindo o risco.

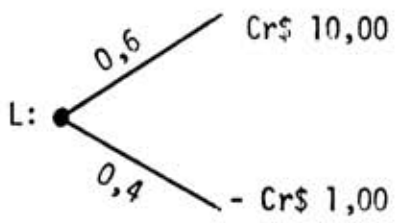
Os primeiros são aqueles que agem com base no valor monetário esperado. Um "médio" é indiferente entre:

a) um ato que resulta em x cruzeiros com probabilidade p e y cruzeiros com probabilidade 1-p e

b) a quantia (px + (1-p)y) cruzeiros, com certeza.

Isto quer dizer que as duas alternativas têm para êle o mesmo valor. Para o "médio" o Equivalente Monetário Certo é sempre igual ao valor esperado do jogo.

Assim, por exemplo, se temos a loteria L, onde



o valor de L para o médio será igual (Cr$ 10,00) x 0,6 + (-Cr$ 1,00) x 0,4 = Cr$ 5,60.

Para este tipo de pessoas, o valor monetário esperado pode ser usado como um indicador da ação ótima.

As pessoas que sentem aversão ao risco atribuem sempre a uma loteria um equivalente certo menor que o valor esperado dessa loteria. Para elas assumirem o risco, é preciso que os prêmios sejam altamente compensadores. É o caso de um banqueiro que está mais interessado em garantir o reembolso do que na concessão de empréstimos. Ele não concederia, por exemplo, um empréstimo a uma alta taxa de juros porque esta traduz a incerteza quanto à devolução do empréstimo. Ele é, então dito ser um "averso ao risco".

O terceiro tipo de comportamento é o das pessoas que se sentem bem assumindo o risco. O equivalente monetário certo de uma loteria, para estas pessoas, é sempre maior que o valor monetário esperado. Assim, no caso da loteria L do nosso exemplo, alguém poderia atribuir a ela um equivalente certo de Cr$ 8, 00, enquanto que um averso ao risco atribuiria, digamos Cr$ 2,00.

Podemos concluir que cada indivíduo tem uma curva de utilidade, ou seja, a magnitude do lucro ou perda que serão aceitos como equivalentes monetários e as preferências individuais pelo risco são únicas pa

ra cada pessoa. Logo, não há curvas de utilidade que servem como padrões gerais. É preciso uma para cada indivíduo e para cada instante de tempo, já que as preferências não são constantes ao longo do tempo.

4.2.3 - Utilidades Multidimensionais

Nem sempre as consequências associadas a um determinado curso de ação podem ser traduzidas em termos de uma única dimensão de valor tal como valores monetários. Com muita frequência, temos apenas "medidas" que devem traduzir o que é essencial sobre a consequência em consideração, e, deve-se salientar, isto não é fácil de ser obtido, principalmente em problemas complexos. E, como a maioria dos problemas em que se procura aplicar um processo formal de decisão são complexos, a análise desses problemas se torna difícil.

Consideremos, por exemplo, a compra de um carro. Dificilmente alguém baseia sua decisão apenas no preço do carro. Certamente, são também levados em consideração outros aspectos, como: consumo de gasolina, despesas de manutenção e preço de revenda, que podem ser medidos em termos monetários. No entanto, há ainda outros aspectos que influenciam na decisão (e muitas vêzes são os mais importantes) e que não podem ser traduzidos em termos monetários, tais como: desempenho, conforto, aparência, prestígio social que o carro confere ao seu proprietário, etc.

O exemplo acima, caracteriza-se pelo que é denominado "multidimensionalidade". Esta constitui uma das razões que dificultam o estabele

cimento de relações de preferência entre os possíveis resultados de uma decisão, sendo caracterizada por casos em que uma alternativa deve ser analisada com base em vários fatores ou atributos. Em muitos casos, nos quais é difícil traduzir uma consequência em termos de um simples quantidade numérica, podemos, no entanto, associar a cada atributo i um valor x_i que pode ser considerado como o "nível" da consequência para esse atributo. Assim, uma consequência x seria representada por $(x_1, x_2, \ldots x_n)$, onde n é o número de dimensões de valor que estão sendo consideradas.

Para determinar preferências entre alternativas complexas onde, para cada resultado há vários aspectos a serem considerados, o que se faz é tentar reduzir o problema a proporções menores. Usando técnicas já desenvolvidas para determinar utilidades de um simples atributo, chegamos à determinação de uma utilidade para cada consequência.

Assumindo que as várias dimensões são independentes, uma das abordagens mais comuns para avaliar alternativas multidimensionais é a função utilidade aditiva. Assim, em n dimensões temos:

$$U_j(x_{1j}, x_{2j}, \ldots, x_{nj}) = \sum_{i=1}^{n} u_i(x_{ij}) \qquad (4.1)$$

onde u_i é a função utilidade estabelecida para o i-ésimo atributo. Partindo desse tipo de função, o critério de decisão usado é maximizar a utilidade esperada, definida como:

$$E(U) = \sum_{j=1}^{m} U_j \, p_j$$

onde m é o número de consequências associadas a uma alternativa e U_j é a utilidade da j-ésima consequência, avaliada segundo (4.1) acima.

Além da representação aditiva, outras abordagens existem para a determinação de utilidades multidimensionais, das quais trataremos posteriormente.

4.3 - TRATAMENTO AXIOMÁTICO DA UTILIDADE

A teoria que apresentamos constitui uma teoria descritiva e, especificamente, ela parte da idéia de "comportamento racional". Em outras palavras, tenta obter uma medida do valor dos resultados associados a um dado curso de ação, a partir da definição de um conjunto de axiomas que, se satisfeitos, permitem "predizer" o comportamento do indivíduo racional, através de suas preferências. Dizemos que uma pessoa não é racional quando, conscientemente, toma uma decisão que é contra seus próprios objetivos estabelecidos. Isto é, se ela estabelece que adotará uma ação A e deliberadamente adota B, que ela sabe ser oposta a A, então ela está agindo irracionalmente.

A abordagem apresentada por Von Neumann e Morgenstern [35] para o problema de medida da utilidade baseou-se no uso de"jogos". Um jogo é uma escolha entre ações alternativas cujas consequências são incertas. O uso de jogos foi primeiramente sugerido por Frank Ramsey* mas, Von Neumann e Morgenstern elaboraram a estrutura básica da teoria que é constituída por um conjunto de axiomas e pelos teoremas deduzidos a partir dele.

Os elementos básicos com que a teoria trata são:

a) um conjunto C = {x, y, ..., z} , cujos elementos são as alternativas de ação.

---

* *Citado por Fishburn* [15]

b) As relações individuais de preferência-indiferença com relação aos elementos de C. Estas relações são denotadas por $\succsim$ que significa "é preferível ou indiferente a". Assim, sempre que dissermos $x \succsim y$, queremos dizer que "x é preferível ou indiferente a y" ou "y não é preferível a x".

c) O conjunto R dos números reais.

d) Uma operação em C e R, que pode ser interpretada como uma combinação de elementos de C. Assim, se x e y pertencem a C e p pertence a R,

$px + (1 - p)y$* também pertence a C

onde $0 \leq p \leq 1$

---

* *Esta expressão representa um jogo em que se obtém x com probabilidade p e y com probabilidade 1 − p.*

4.3.1 - Axiomas Básicos

Existem vários conjuntos de axiomas básicos que levam aos mesmos resultados. Dentre eles, o de mais fácil entendimento é o apresentado abaixo, formulado por Luce e Raiffa.

Axioma 1:

Um indivíduo, face a duas alternativas x e y, é sempre capaz de decidir se prefere x a y, y a x ou se é indiferente entre as duas. Isto significa que, dados x e y, pelo menos uma das alternativas abaixo se verifica:

$x \succsim y$

$y \succsim x$

Em outras palavras, as alternativas são sempre comparáveis.

Axioma 2:

Se $x \succsim y$ e $y \succsim w$ então $x \succsim w$

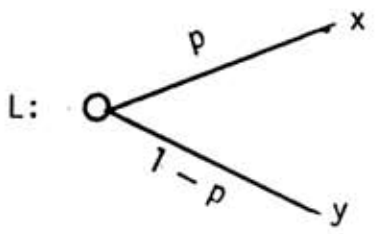
Os axiomas 1 e 2 constituem uma ordenação das preferências entre alternativas. O primeiro estabelece que, se duas alternativas não têm o mesmo valor, uma é preferível à outra. O segundo é o chamado "Axioma da Transitividade". Então, a relação de preferência-indiferença constitui uma ordem parcial* (fraca, quase-ordem) entre os elementos do conjunto C.

---

* *Para melhor compreensão da noção de ordem veja referências* [*14, 27 e 37*].

Axioma 3:

Se x é preferível a y, então x é preferível à loteria L, dada por:

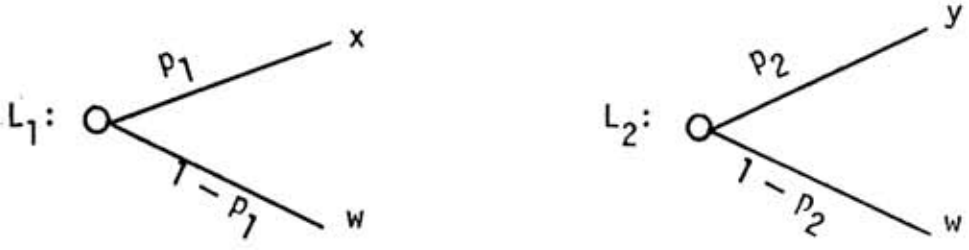


onde $0 < p < 1$, o que significa que x é preferível a $px + (1 - p)y$.

Este axioma é chamado por Savage [45] de "principle of sure-thing" e estabelece que se $x \succsim y$ e x pode ser obtido com certeza, então, o indivíduo não adota uma ação que arrisca x ao invés de simplesmente aceitar o resultado certo.

Axioma 4:

Sejam L_1 e L_2 as loterias seguintes:

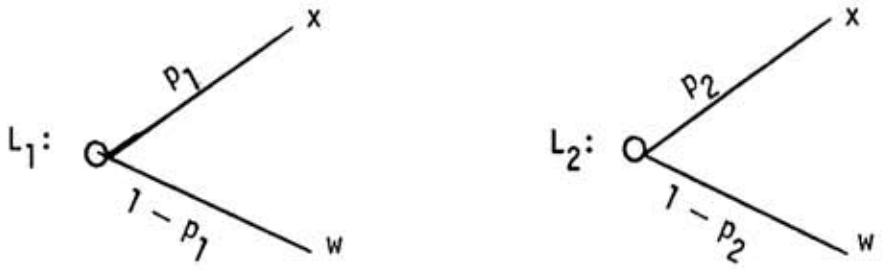


Se $x \succsim y$ e $p_1 = p_2$, então $L_1 \succsim L_2$

Isto quer dizer que, se dois resultados são igualmente prováveis, a possibilidade de ganhar o de maior valor é preferível.

Axioma 5:

Se $x \succ_r y \succ_r w$, então existem p_1 e p_2, $0 < p_1 < 1$, $i = 1,2$, tal que $L_1 \succ_r y$ e $y \succ_r L_2$ onde



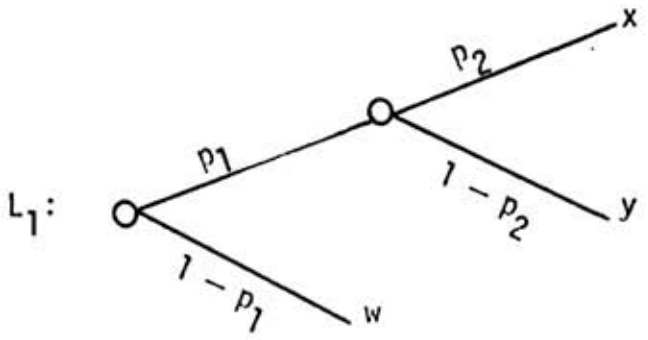
Isto significa que existe algum jogo, determinado por uma chance p_1 de ganhar x e $1 - p_1$ de ganhar w, que é preferível ao prêmio certo y; e existe também um jogo tal que y é preferível a ele. Segundo Von Neumann e Morgenstern [35,pp.630], este axioma expressa o que é chamado de Propriedade Arquimediana: não importa em que ponto y está entre x e w, este jogo sempre existe. O valor de p_1 dependerá da diferença em utilidade entre x e y e podemos dizer que este valor tem um papel crítico na atribuição de valores numéricos às consequências. Isto porque se $x \succ_r y \succ_r w$, é coerente que $L_1 \succ_r y$ se p_1 está próximo de um e, se p_1 está próximo de zero, a preferência é invertida. Portanto, para algum p, $0 < p < 1$, ou, mais especificamente, aquele p para o qual a preferência é invertida, verificar-se-á a indiferença. E é justamente esse ponto de indiferença que nos interessa.

Este aspecto será melhor compreendido quando, no capítulo seguinte, apresentarmos algumas técnicas para determinação de utilidades numéricas.

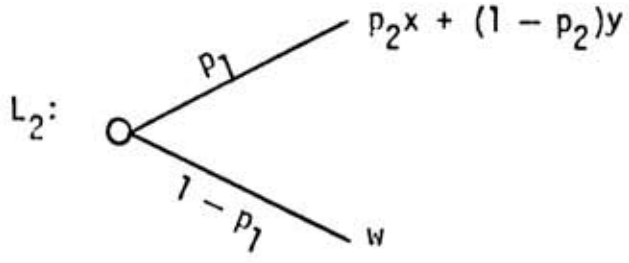
Axioma 6:

Se uma loteria tem como prêmio uma ou mais loterias, dizemos que ela é uma loteria composta. Toda loteria composta L_1 pode ser reduzida a uma loteria simples L_2 ou seja, L_1 é indiferente a L_2.

Assim, se L_1 é dada por:



e



então $L_1 \sim L_2$.

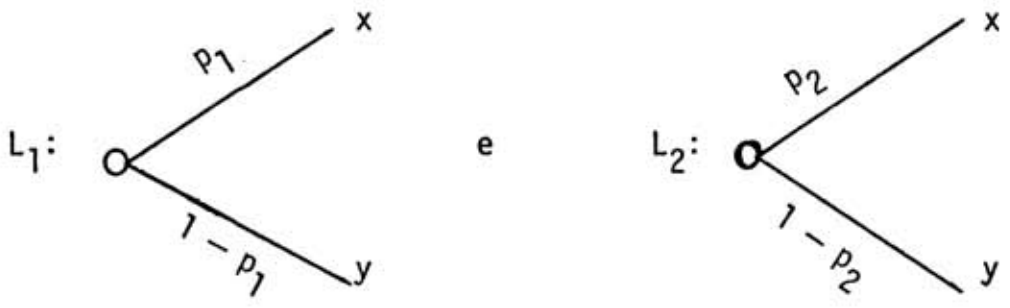
Este axioma estabelece que qualquer loteria complexa, do tipo L_1, pode ser decomposta nas alternativas básicas, através do uso de

probabilidades. Portanto, as utilidades dos jogos são reduzidas a formas mais simples.

A importância deste axioma está no fato de que ele estabelece uma ligação entre a teoria da utilidade e a de probabilidade, permitindo uma perfeita compatibilidade de operação nos dois domínios.

Axioma 7:

Dado um conjunto de alternativas, seja x a mais preferível e y a menos preferível e sejam p_1 e p_2 números tais que $0 < p_1, p_2 < 1$. Então $L_1 \succsim L_2$ se e somente se $p_1 \geq p_2$, onde:



Este é chamado "axioma da monotonicidade" e é equivalente a um princípio de dominância. Isto parece plausível já que, se o indivíduo estritamente prefere x a y, então ele tem um motivo forte para preferir L_1, onde x pode ser obtido com probabilidade maior.

Teorema 1:

Se o conjunto de axiomas é satisfeito, existe uma função u que a cada elemento x ∈ C associa um número real u(x), denominado utilidade de x, e o indivíduo decidirá sempre de maneira a maximizar a utilidade esperada. A função u é tal que se x e y pertencem a C, então:

$$u(x) \geq u(y) \iff x \succsim y \qquad (4.2)$$

Portanto, se a relação de preferência-indiferença, para um indivíduo, satisfaz aos axiomas mencionados, existem* números u(x), u(y), etc., associados a cada um dos elementos de C, que refletem as preferências do indivíduo com relação a esses elementos.

Se $u\,|\,px + (1 - p)y| = p \cdot u(x) + (1 - p) \cdot u(y)$, a função u é dita ser linear. Por ela satisfazer à condição (4.2) acima, diz-se que u é uma função preservadora da ordem no conjunto C.

Assim, dadas duas loterias $L_1 = (p_1a_1, \ldots, p_na_n)$ e $L_2 = (q_1b_1, q_2b_2, \ldots, q_kb_k)$, por mais complexas que sejam, o conjunto de axiomas básicos permite-nos reduzí-las a formas mais simples e decidir entre elas. Isto é feito através da computação das utilidades esperadas, definidas como:

$$u(L_1) = \sum_i p_i u(a_i)$$

$$u(L_2) = \sum_j q_j u(b_j)$$

---

* *A prova da existência de u pode ser vista em* [35].

A loteria com maior utilidade é escolhida.

Os axiomas mencionados constituem a base da teoria da utilidade formulada por Von Neumann e Morgenstern. A pertir deles, vários teoremas são deduzidos, alguns dos quais enunciaremos sem, no entanto, apresentar provas. Estas podem ser vistas em Savage [45].

Teorema 2:

Se x e y pertencem a um conjunto finito C, $u : C \rightarrow R$ é uma função utilidade se e somente se $x \succsim y$ é equivalente a $u(x) \geq u(y)$.

Teorema 3:

Se u é uma função utilidade e a e b são números reais, $a > 0$, então $v = au + b$ é também uma função utilidade. Ou seja, a função utilidade é única, a menos de uma transformação linear.

Se u é uma função utilidade, satisfaz às condições de teorema 1; portanto, v também as satisfará desde que $u(x) \geq u(y)$ implica que $v(x) \geq v(y)$

Teorema 4:

Se u e v são funções utilidade, então existem a e b, $a > 0$ tais que $v = au + b$.

O teorema nos diz que quaisquer duas funções utilidades são transformações lineares crescentes uma da outra.

Vemos assim que se existe uma função utilidade existem infinitas, já que, dada uma, as outras podem ser obtidas através de transformações lineares positivas.

4.4 - UTILIDADES MULTIDIMENSIONAIS

Seja T = (x, y, ..., w) um conjunto de consequências associadas a um determinado ato.

x = (x_1, x_2, ..., x_n) é uma consequência, onde x_i é o "nível" de x com relação à i-ésima dimensão de valor.

Como já dissemos, para atribuir um valor a resultados que devem ser analisados com relação a múltiplos atributos o mais comum é usar-se uma função aditiva, que se baseia na proposição:

- Um número $u_i(x_i)$ pode ser associado a cada x_i de x, i = 1, 2, ..., n, tal que, se x = (x_1, x_2, ..., x_n) e y = (y_1, y_2, ..., y_n) pertencem a T, então:

$$x \succcurlyeq y \iff \sum_{i=1}^{n} u_i(x_i) \geq \sum_{i=1}^{n} u_i(y_i)$$

Se $u(x) = u_1(x_1) + u_2(x_2) + \ldots + u_n(x_n)$, então, a proposição acima significa que a utilidade de um vetor é igual à soma das utilidades de seus componentes.

A representação aditiva é bastante restritiva em virtude de uma hipótese básica, que é a de independência entre os vários atributos. Edwards [6] comenta esse fato num trabalho em que apresenta um modelo para determinar utilidades multidimensionais. Mas, segundo ele, mesmo que haja uma certa correlação entre os atributos, este fato não é muito grave, dada a pouca sensitividade do modelo aditivo a tal correlação. Este modelo será apresentado no próximo capítulo.

Uma outra abordagem do problema da multidimensionalidade é a sugerida por Fishburn [15], que também pressupõe independência entre as dimensões de valor.

Ela parte da noção de ordem "lexicográfica" que é definida como:

- Se $(a_1, a_2, \ldots, a_n)$ e $(b_1, b_2, \ldots, b_n)$ são vetores em R^n, dizemos que $\overset{L}{\gtrsim}$ é uma ordem lexicográfica, quando $(a_1, \ldots, a_n) \overset{L}{\gtrsim} (b_1, \ldots, b_n)$ se e somente se $a_{i*} \gtrsim b_{i*}$, onde i^* é o menor i para o qual $a_i \neq b_i$.

Utilidades lexicográficas se verificam quando um número

$u_i(x_i)$ pode ser atribuido a cada x_i de x, $i = 1, 2, \ldots, n$, tal que, se

$x = (x_1, x_2, \ldots, x_n)$ e $y = (y_1, y_2, \ldots, y_n)$ pertencem a T, então:

$$x \gtrsim y \iff \left[u_1(x_1), u_2(x_2), \ldots, u_n(x_n)\right] \overset{L}{\gtrsim} \left[u_1(y_1), u_2(y_2), \ldots, u_n(y_n)\right]$$

Esta abordagem foi sugerida para casos em que alguns atributos são predominantemente mais importantes que os outros.

Keeney [23], [24], [25], dá outro enfoque ao problema. Ele parte do conceito de independência em termos de utilidade que, para efeito de simplificação, chamamos de "independência em utilidade".

Seja $x = (x_1, x_2, \ldots, x_n)$ uma consequência associada a um ato e seja T um conjunto dessas consequências.

Definimos os vetores Y e Z, tais que: $y = (x_1, x_2, \ldots x_m)$ e $z = (x_m+1, x_m+2, \ldots x_n)$, onde y e z são valores particulares de Y e Z, respectivamente.

Portanto, podemos escrever que $x = (y, z)$. Para um valor fixo de z, que denominamos z_o, a utilidade condicional de y dado que $Z = z_o$ é denotada por $u(y, z_o)$.

Dizemos que Y é independente em utilidade de Z se as preferências do decisor com relação a Y, dado que $Z = z_0$, são as mesmas, independente de z_0.

Como a função utilidade é única, a menos de uma transformação linear positiva, então, se Y é independente em utilidade de Z, a função utilidade condicional de Y, dado Z, é uma transformação linear positiva da função utilidade condicional de Y, dado qualquer outro valor de Z.

Ou seja:

$$u(y, z) = C_1(z) + C_2(z)u(y, z_0), \quad \forall\, y, z$$

Um outro conceito usado é o de independência mútua em utilidade. Se Y é independente de Z em utilidade, isto não implica que Z seja independente de Y. Quando Z também é independente de Y, dizemos que Y e Z são mutuamente independentes em utilidade.

Assumindo Y e Z mutuamente independentes em utilidade, Keeney sugere várias representações para utilidades multidimensionais que, para melhor compreensão, serão apresentadas para duas dimensões.

a) Representação quase-aditiva.

u(y, z) é completamente especificada por uma função utilida

de condicional para Y, $u(y, z_0)$ e a utilidade de uma outra consequência $x_1 = (y_1, z_1)$, isto é:

$$u(y, z) = u(y_0, z) + u(y, z_0) + k\, u(y_0, z)u(y, z_0)$$

onde

$$k = \frac{u(y_1, z_1) - u(y_1, z_0) - u(y_0, z_1)}{u(y_1, z_0)\, u(y_0, z_1)} \qquad (4.3)$$

b) Representação aditiva.

Se em (4.3) acima $k = 0$, temos uma representação aditiva, ou seja:

$$u(y, z) = u(y_0, z) + u(y, z_0)$$

c) Representação por produto.

$$u'(y, z) = u(y, z) + \frac{1}{k} \quad , \qquad k \neq 0$$

ou

$$u'(y, z) = u_z(y_0, z).\, u_y(y, z_0)$$

Uma outra forma multiplicativa, também sugerida por Keeney [26], é

$$u(x_1, x_2, \ldots, x_n) = \frac{1}{k}\left\{\prod_i \left[kk_i\, u_i(x_i) + 1\right] - 1\right\}$$

onde k_i é um fator de escala para $u_i(x_i)$, $0 < k_i < 1$, $\forall_i$, e k é determinado a partir dos valores individuais dos k_i's.

d) Curvas de "Isopreferências".

Na maioria das vezes, é difícil para o decisor determinar funções utilidade condicionais. Um procedimento mais fácil consiste em obter uma curva em que para diferentes pares (y, z) a preferência é a mesma, isto é, obter um conjunto de consequências igualmente desejáveis. Keeney mostra que, desde que elas cubram o mesmo domínio, uma utilidade condicional pode ser substituída por uma curva de isopreferência.

No capítulo seguinte, apresentamos um modelo para determinação de utilidades multidimensionais que usa uma função aditiva. A razão de nos determos nesta forma é que ela possui a vantagem da simplicidade. Ao invés de determinarmos uma função utilidade n-dimensional, passamos a trabalhar com n funções unidimensionais, usando para isto os procedimentos já existentes para determinar utilidades de um único atributo.

4.5 - CONCLUSÕES

Os axiomas que constituem a base da teoria apresentada permitem-nos, desde que nos comportemos como eles prescrevem, usar observações de nossas preferências simples entre jogos para derivar uma medida da uti

lidade dos resultados associados aos vários cursos alternativos de ação. Esta medida é tal que agiremos sempre de maneira a maximizar a utilidade esperada, definida como a soma das probabilidades de cada resultado vêzes a sua utilidade.

$$U(X) = \sum_{i=1}^{n} u(x_i)p(x_i)$$

onde X é uma variável aleatória com n valores possíveis. Se os resultados são descritos em uma escala contínua, temos:

$$U(X) = E\,|\,u(x)\,| = \int_{-\infty}^{+\infty} u(x)\ f(x)\ dx$$

onde f(x) é uma função densidade de probabilidade e u(x) é a utilidade de um resultado no intervalo x + dx, satisfazendo aos critérios de racionalidade.

Uma forma específica para u(x) foi proposta por Daniel Bernoulli*. Segundo ele, a utilidade do dinheiro seria uma função linear do logaritmo do seu valor, ou:

$$u(x) = \log x$$

Esta forma foi proposta para resolver o chamado "Paradoxo de São Petersburgo", o qual descrevemos a seguir.

Suponhamos que uma pessoa jogue uma moeda honesta até que apareça cara. Se k for o número de lançamentos necessários, ela recebe

---

* *Citado por Von Neumann e Morgenstern* [35].

$C_k = 2^k$ cruzeiros. Qual é o valor esperado desse jogo? Quanto alguém pagaria pelo privilégio de jogá-lo?

Seja C = recompesa

N = número de tentativas até que apareça cara pela primeira vez.

Sabemos que $p_k = P(C=2^k) = P(N=k) = \frac{1}{2^k}$ $k = 1, 2, \ldots$

ou $$p_k = \underbrace{\frac{1}{2} \cdot \frac{1}{2}, \ldots \frac{1}{2}}_{\text{probabilidade de (k-1) coroas}} \cdot \underbrace{\frac{1}{2}}_{\text{probabilidade de cara}} = \frac{1}{2^k}$$

O valor monetário esperado do jogo será:

$$E(C) = \sum_{k=1}^{\infty} p_k C_k = \sum_{k=1}^{\infty} \frac{1}{2k} \cdot 2^k = 1 + 1 + \ldots + 1 + \ldots = \infty,$$

o que significa que este jogo é, para um médio, preferível a qualquer quantia finita. Ou melhor, um médio estaria disposto a pagar qualquer quantia pelo direito ao jogo.

Paradoxalmente, ninguém parece disposto a pagar mais do que, digamos, Cr$ 10,00, por este jogo.

Bernoulli sugeriu que se usasse não a esperança matemática, mas o que ele chamou de "esperança moral", definindo a utilidade como o logaritmo dos bens possuídos por uma pessoa, expressos em termos monetários. Bernoulli partiu da idéia de que a utilidade de um ganho x, para um indivíduo, seria não somente proporcional a x mas também inversamente proporcional à riqueza do indivíduo.

O valor do jogo seria um certo número finito y, onde

$\log y = x$ e

$x = \sum_{k=1}^{\infty} (\log 2^k) \frac{1}{2^k}$, ou seja, x é o limite para o qual a série acima converge.

A objeção mais frequente ao critério de esperança de utilidades é que ele não leva em conta o prazer de jogar, desde que considera o indivíduo indiferente entre uma loteria composta e a loteria simples obtida pelo cálculo de probabilidades. Em outras palavras, ele negligencia a utilidade do risco em si. Argumentam os críticos da teoria que é preciso uma utilidade que leve em conta a utilidade intrínseca do risco.

CAPÍTULO V

UTILIDADE: ASPECTOS PRÁTICOS

5.1 - UTILIDADE UNIDIMENSIONAL

Como já foi visto, não é sempre verdade que as pessoas maximizem o valor monetário esperado. Nas situações que envolvem risco ou incerteza a maioria das pessoas demonstra aversão ao risco.

Por isso é preciso a introdução do conceito de "equivalente certo" em substituição ao valor esperado. Isto nos leva ao conceito de utilidade e o que se espera maximizar numa decisão envolvendo incerteza é a utilidade esperada.

Vimos também que, satisfeitos certos postulados, é possível levantar a curva de utilidade do decisor para um determinado bem ou situação, em particular para o dinheiro.

Veremos inicialmente, como levantar a curva de utilidade do dinheiro para um decisor por dois métodos diferentes: primeiro, através de uma comparação entre loterias e equivalentes certos e, segundo, admitindo que a curva de utilidade é uma função exponencial

$$U(x) = C(1 - e^{-kx})$$

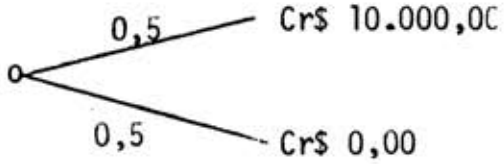
e tentando estimar o parâmetro k.

Veremos, a seguir, um processo mais geral, embora menos formal, para determinar a função utilidade de um bem ou situação qualquer que não seja exprimível em termos monetários.

5.1.1 - Função Utilidade do Dinheiro

Veremos neste item como levantar a função utilidade do dinheiro para um decisor, através de comparações entre loterias e equivalentes certos.

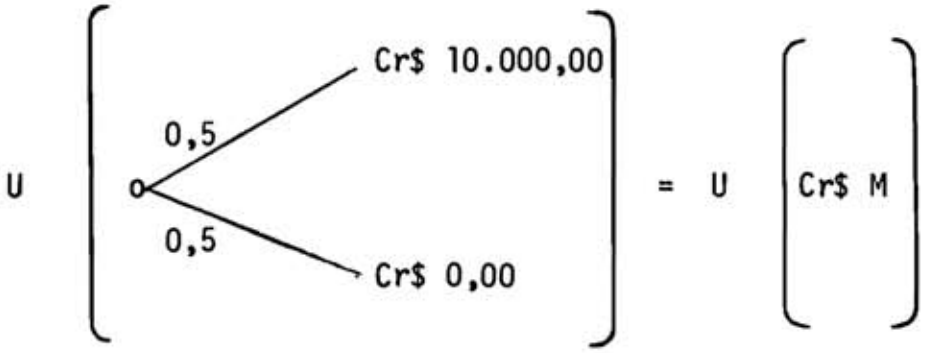
Antes, porém, relembremos alguns conceitos fundamentais. Suponha que o decisor é indiferente entre participar da loteria



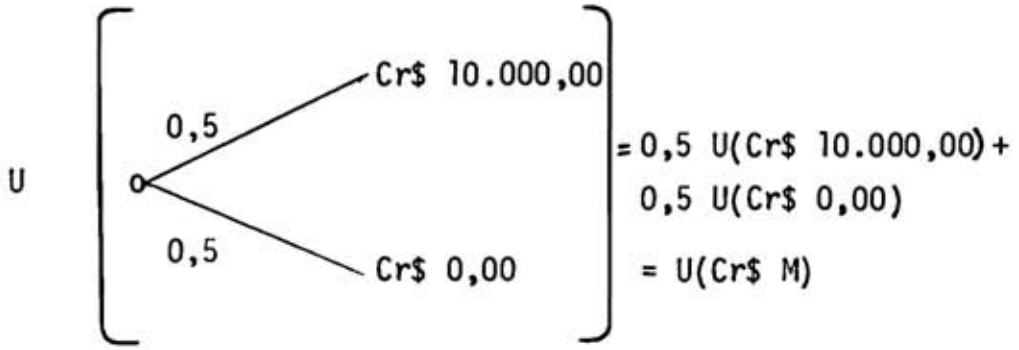
(ou seja uma loteria que, com 50 % de chance, lhe dá dez mil cruzeiros e com 50 % de chance não lhe dá nada) e receber (com certeza) uma quantia fixa: M cruzeiros.

Neste caso, M, como vimos, é chamado o equivalente certo da loteria, pois é aquela quantia fixa pela qual o decisor está disposto a trocá-la.

A indiferença sentida pelo decisor entre a loteria e o equivalente certo implica em que, para ele, as utilidades são iguais, ou seja:



Além disso, tomamos como pressuposto que uma propriedade de valor esperado é válida para utilidades, ou seja:



Resumindo, temos que, se as preferências do decisor satisfazem certos postulados, é possível construir sua função utilidade. Esta função possui duas propriedades importantes:

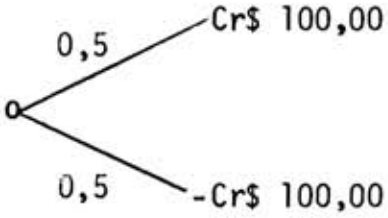
- A utilidade de uma loteria é a utilidade esperada de suas recompensas.
- A loteria preferida é a que tiver maior utilidade.

Suponhamos que somos analistas de decisões e estamos assessorando um empresário (o decisor). Durante o processo de análise, é preciso levantar sua função utilidade. Vejamos como fazê-lo.

Inicialmente, escolhemos dois valores monetários (digamos -Cr$ 100,00 e+Cr$ 100,00) bastante separados, de modo a englobar todos os

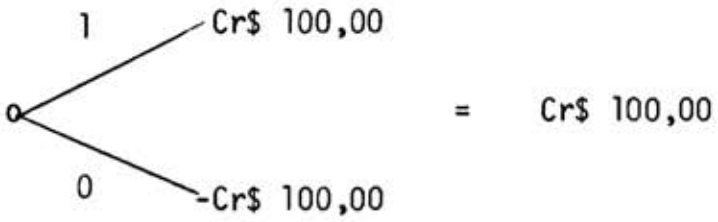
valores de lucros ou perdas que possam vir a ocorrer no problema que estiver sendo analisado.

Com estes dois valores podemos formular uma loteria que com probabilidades iguais, tenha como recompensa uma das quantias citadas.



Mas podemos variar a probabilidade de cada uma das recompensas, gerando uma infinidade de loterias.

A melhor possível (e que qualquer pessoa certamente gostaria de participar) seria aquela que desse + $ 100,00 de recompensa com certeza (ou seja com probabilidade 1).



Lembrando uma das propriedades fundamentais da função utilidade: "a loteria preferida é a que tiver maior utilidade", vemos que devemos atribuir a esta loteria o valor máximo da nossa escala de utilidades.

Se arbitrarmos a variação da nossa função utilidade entre

zero e um, teremos:

$$U\left[\begin{array}{ll} 1 & \text{Cr\$ 100,00} \\ 0 & \text{-Cr\$ 100,00} \end{array}\right] = U(\text{Cr\$ 100,00}) = 1$$

Com os mesmos valores de recompensa, a pior loteria possível é aquela em que se perde Cr$ 100,00 com certeza, ou seja

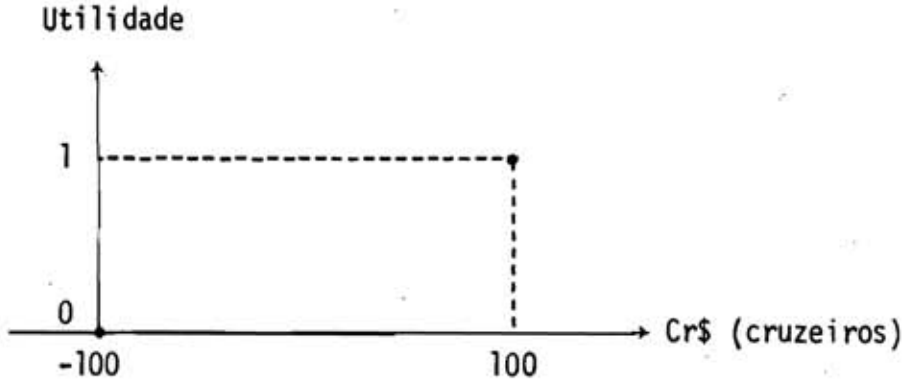
$$\begin{array}{ll} 0 & \text{Cr\$ 100,00} \\ 1 & \text{-Cr\$ 100,00} \end{array} = \text{- Cr\$ 100,00}$$

Portanto, a esta loteria deve ser atribuída a menor utilidade possível, no nosso caso zero:

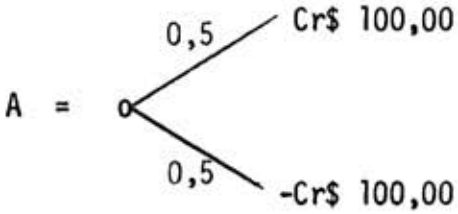
$$U\left[\begin{array}{ll} 0 & \text{Cr\$ 100,00} \\ 1 & \text{-Cr\$ 100,00} \end{array}\right] = U(\text{-Cr\$ 100,00}) = 0$$

Notemos que embora tenhamos feito U(-Cr$ 100,00) = 0, isto é arbitrário; nós não fixamos a origem da escala de utilidades pois, como já foi dito, qualquer transformação monotônica crescente de uma função utilidade é também uma função utilidade.

Se quisermos, podemos plotar a função utilidade num gráfico, colocando no eixo das abcissas a quantidade monetária e nas ordenadas a escala de utilidades. Para os dois pontos já obtidos, teríamos:



Tentemos então determinar alguns pontos intermediários da curva de utilidade: podemos calcular a utilidade esperada da loteria formulada inicialmente.



$$U(A) = \frac{1}{2} U(Cr\$\ 100,00) + \frac{1}{2} U(-Cr\$\ 100,00) =$$

$$\frac{1}{2}(1) + \frac{1}{2}(0) = \frac{1}{2}$$

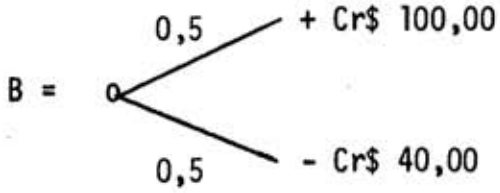
A pergunta a ser formulada ao decisor, neste ponto, é simples: por qual quantidade de cruzeiros ele estará disposto a trocar esta loteria, ou seja, qual é para ele o equivalente monetário certo desta loteria.

Sua resposta, seja ela qual for, cria uma correspondência entre equivalente certos e utilidades de loterias $\$(x) \longleftrightarrow U(A)$

Estes dois pontos podem ser colocados no gráfico, dando-nos um ponto intermediário da curva de utilidade. Outros pontos intermediários podem ser obtidos da mesma maneira.

Suponhamos que o decisor diz ser indiferente entre a loteria (A) e -Cr$ 40,00, ou seja, ele prefere perder quarenta cruzeiros com certeza a entrar na loteria, ou ainda, $U(-Cr\$\ 40{,}00) = \frac{1}{2}$.

Então lhe propomos a seguinte loteria



e a pergunta a ser feita é a mesma: qual o equivalente monetário certo dessa loteria, ou seja, por que quantia o decisor está disposto a trocar esta loteria.

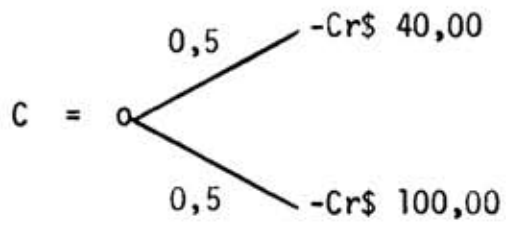
Seja qual for a resposta (por exemplo: x = Cr$ 15,00), temos

$$U(Cr\$\ x) = U(B) = \frac{1}{2}\ U(Cr\$\ 100{,}00) + \frac{1}{2}\ U(-Cr\$\ 40{,}00)$$

$$= \frac{1}{2}(1) + \frac{1}{2}(1/2) = 0{,}75$$

E temos, assim, mais um ponto intermediário da curva de utilidade.

Outra loteria que poderia ser proposta é a seguinte:



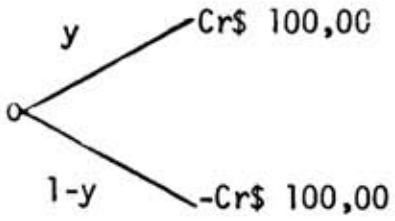
e seja qual for o equivalente certo atribuído pelo decisor (por exemplo: x = -Cr$ 75,00), teremos

$$U(Cr\$\ x) = U(C) = \frac{1}{2}\ U(-Cr\$\ 40{,}00) + \frac{1}{2}\ U(-Cr\$\ 100{,}00)$$

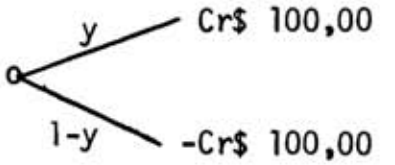
$$\frac{1}{2}\ (1/2) + \frac{1}{2}\ (0) = 1/4$$

Prosseguindo com uma série de loterias e perguntas deste tipo, podemos levantar mais alguns pontos e estimar a curva de utilidade do decisor (figura 5.1).

Dado, então, que o ponto (x, y) está sobre a curva de utilidade, isto significa que o decisor é indiferente entre x cruzeiros com certeza e a loteria



Pensemos, por um momento qual a forma da curva de utilidade, para um decisor médio (sem aversão ao risco); ele trocaria uma loteria qualquer



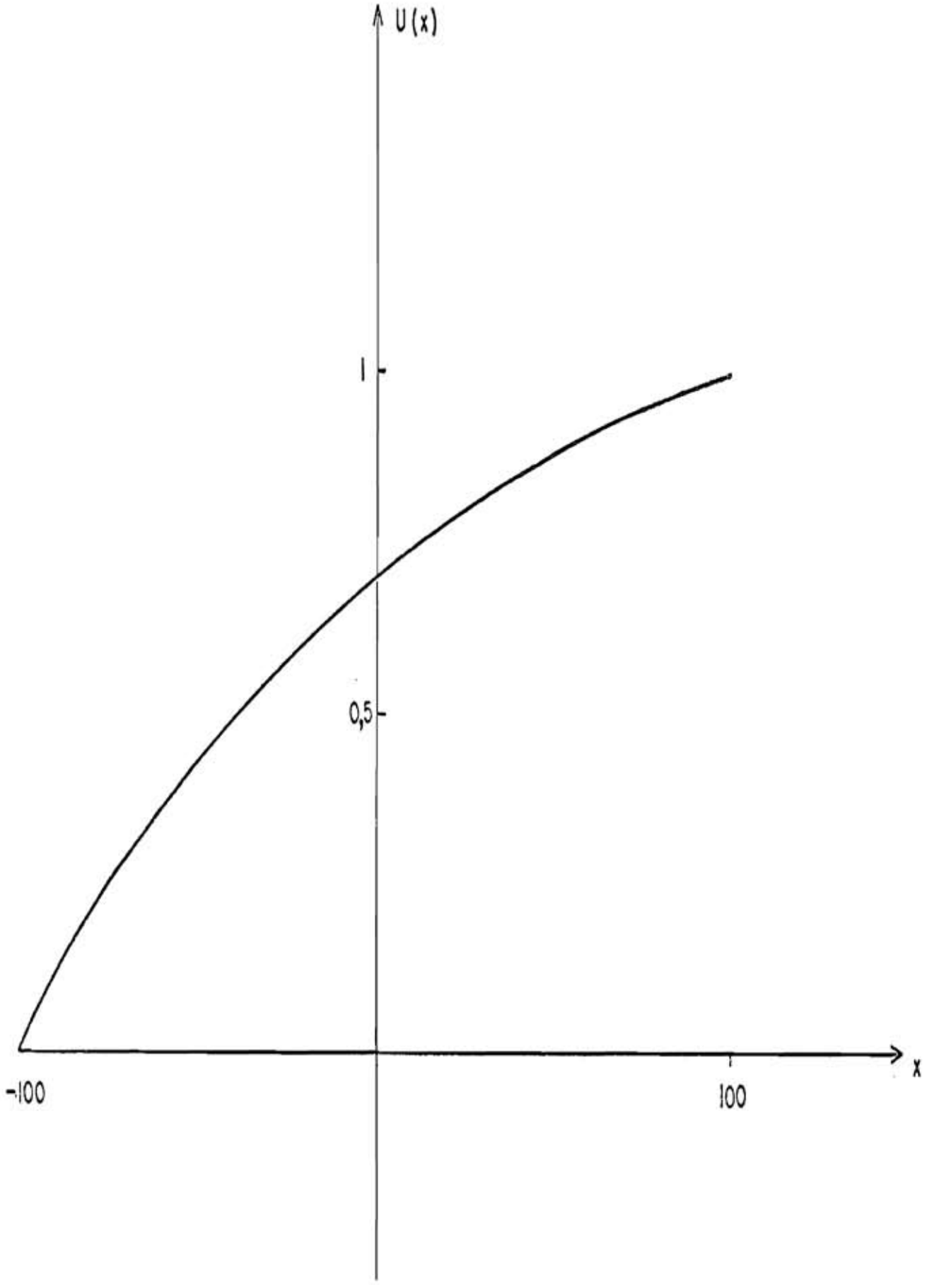


Figura 5.1 - Curva de Utilidade.

por seu valor esperado, ou seja, por uma quantia x, tal que

$$x = y(100) + (1 - y)(-100) \qquad (5.1)$$

ou seja:

$$U(x) = y\,U(100) + (1 - y)U(-100)$$

mas temos que:

$$U(x) = y$$

e, substituindo na equação 5.1, temos

$$x = U(x)(100) + (1 - U(x))(-100)$$

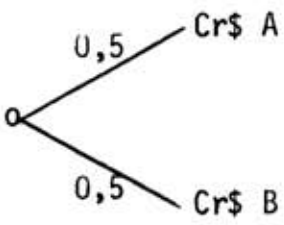
ou seja

$$U(x) = \frac{x}{200} + \frac{1}{2}$$

que é a equação de uma reta; portanto, neste caso, a curva de utilidade do decisor médio seria uma reta ligando os pontos (-100, 0) e (100, 1).

Mas, consideremos o decisor mais comum que é aquele que tem aversão ao risco.

Dada uma loteria qualquer



ele a trocaria por uma quantia certa C e, como ele possui aversão ao risco, esta quantia seria menor que o valor esperado da loteria, ou seja;

$$C < \frac{1}{2} A + \frac{1}{2} B$$

mas por definição

$$U(C) = \frac{1}{2} U(A) + \frac{1}{2} U(B)$$

vemos claramente na figura 5.2, que sua curva de utilidade será sempre côn cava.

5.1.2 - Função Utilidade Exponencial

Apresentamos, a seguir, um método mais simples e rápido pa ra avaliar a função utilidade do dinheiro, para o decisor.

Duas considerações são importantes neste método: o decisor possui aversão ao risco e sua função utilidade é exponencial. A primeira se justifica pelo fato de ser comprovado experimentalmente que a grande maioria das pessoas apresenta aversão ao risco quando em situações de in certeza. A segunda hipótese também se justifica experimentalmente, pois a maioria das funções utilidade encontradas pelo método anterior tem formato semelhante a uma exponencial.

Vamos, então, considerar funções utilidade da forma

$$U(x) = C(1 - e^{-\gamma x})$$

A constante C é apenas um fator de escala, ou seja, depende apenas da nossa escala de utilidades e das recompensas, não dependendo, portanto, do decisor.

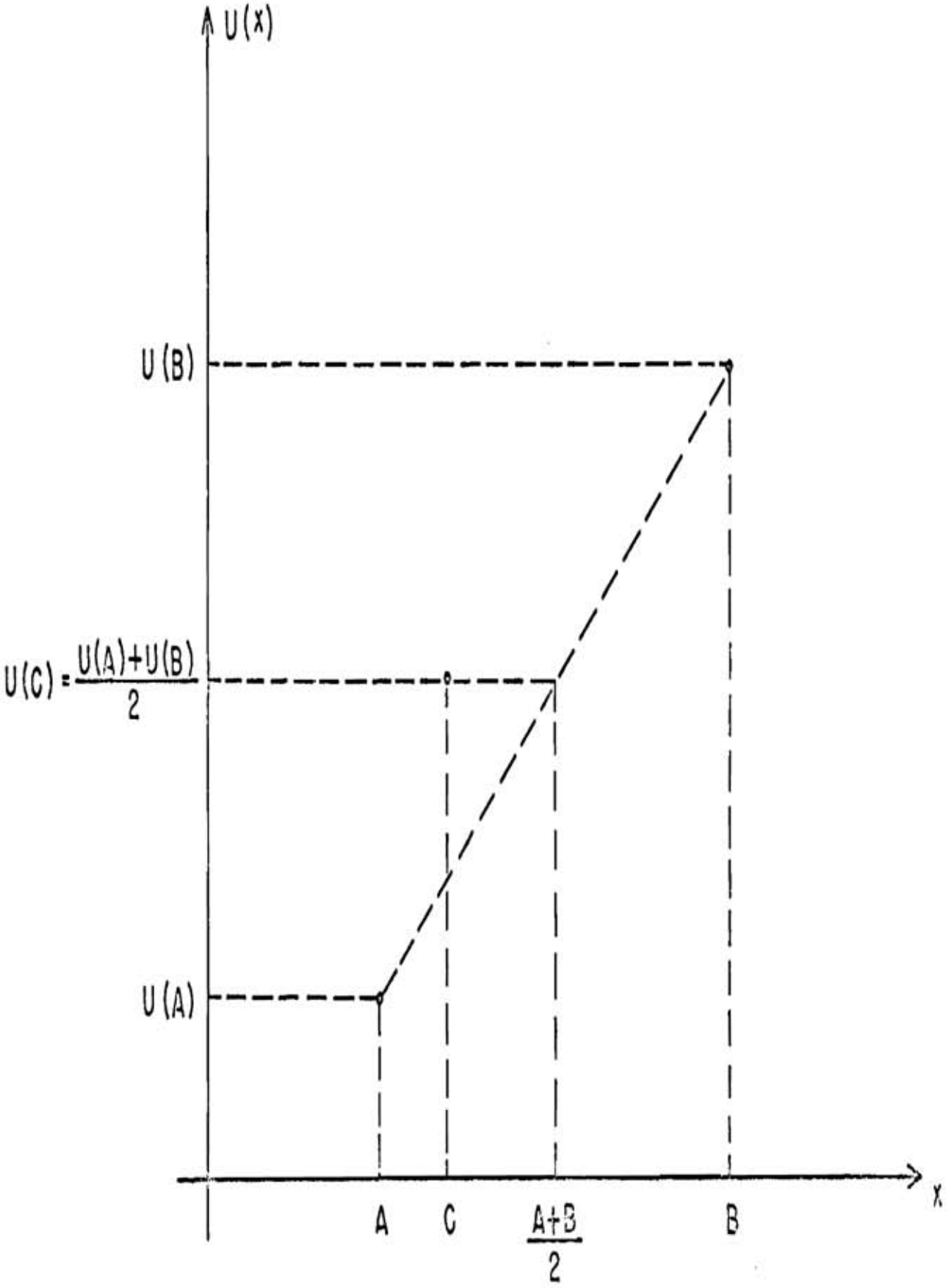


Figura 5.2 - Curva de Utilidade de um Indivíduo Averso ao Risco.

Assim, se a recompensa máxima que se espera ocorrer for Cr$ 1.000,00 e quisermos que a este valor corresponda a utilidade +1, ou seja U(Cr$ 1 000,00) = 1, devemos ter:

$$1 = C(1 - e^{-\gamma x 1.000})$$

$$C = 1/(1 - e^{-1.000\,\gamma})$$

Vemos então que a nossa função utilidade depende de apenas um parâmetro γ denominado "coeficiente de aversão ao risco".

Na realidade, este parâmetro exprime a aversão ao risco apresentada pelo decisor. Assim, quanto maior for γ, mais averso ao risco será o decisor. Naturalmente $\gamma = 0$ para o decisor médio.

Na figura 5.3 é apresentada uma família de funções utilidade exponenciais para que o leitor possa ter uma idéia da influência do coeficiente de aversão ao risco no formato da função utilidade.

O método consiste em avaliar o coeficiente de aversão ao risco do decisor. Isto é feito facilmente com auxílio das figuras 5.4, 5.5 e 5.6. Vemos, em cada uma destas figuras, uma loteria (onde uma das recompensas é variável) e um gráfico que nos dá o coeficiente de aversão ao risco em função do valor monetário da recompensa não fixada.

O que o decisor deverá fazer é, para uma das loterias, determinar o valor da recompensa variável que lhe parecer mais conveniente

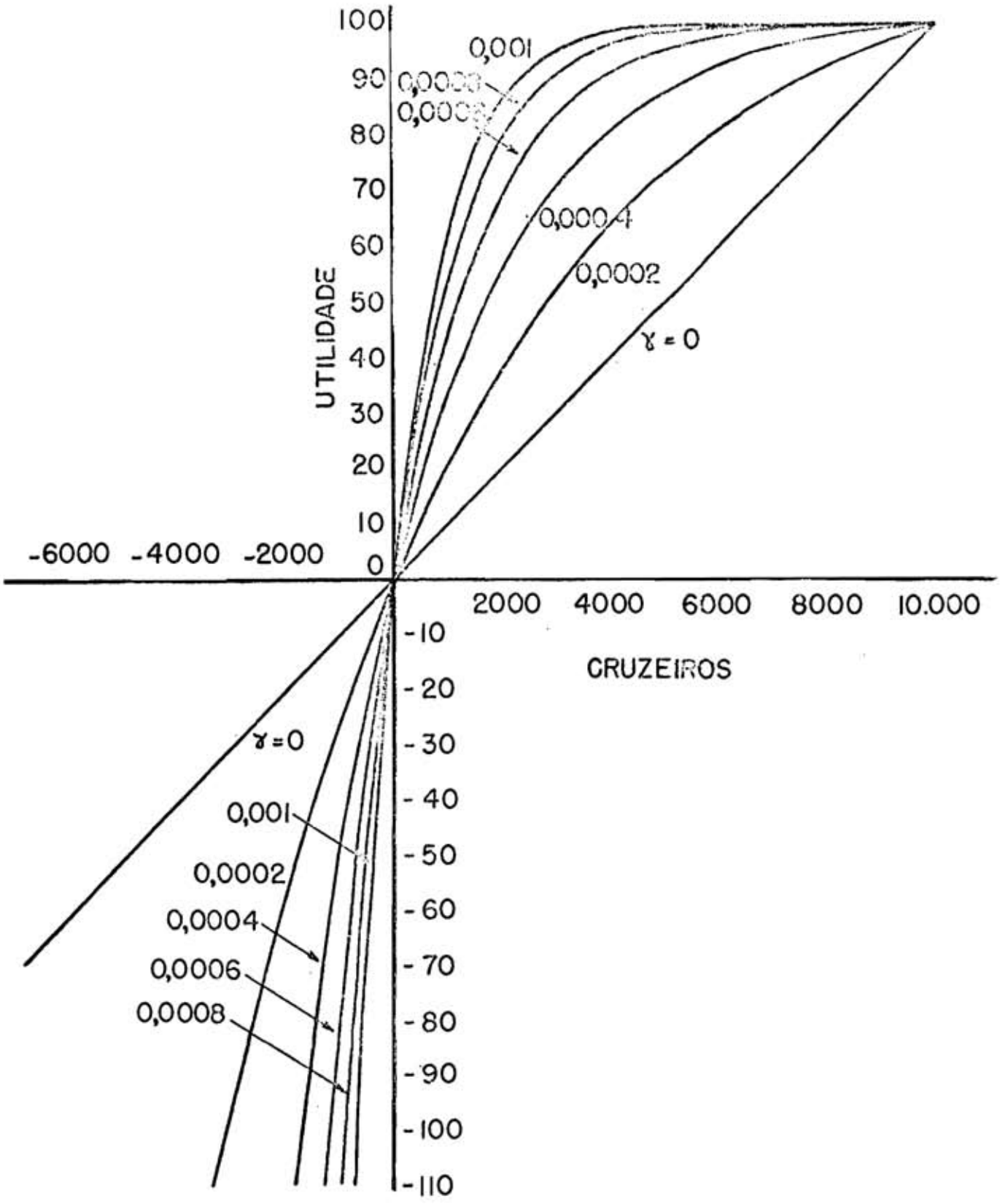


Figura 5.3 - Família de Exponenciais.

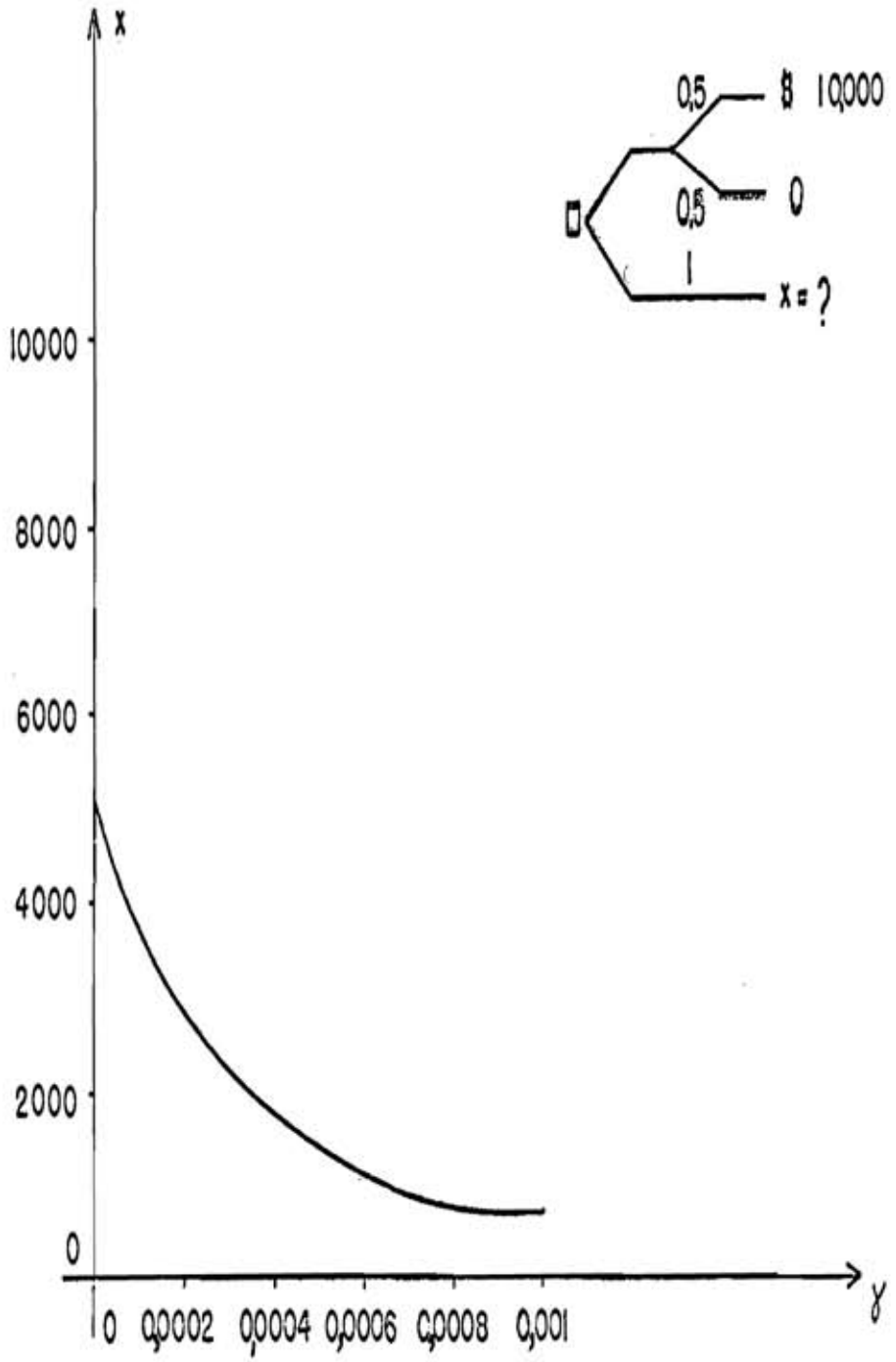


Figura 5.4 - Determinação do Coeficiente de Aversão ao Risco através de Loterias (1).

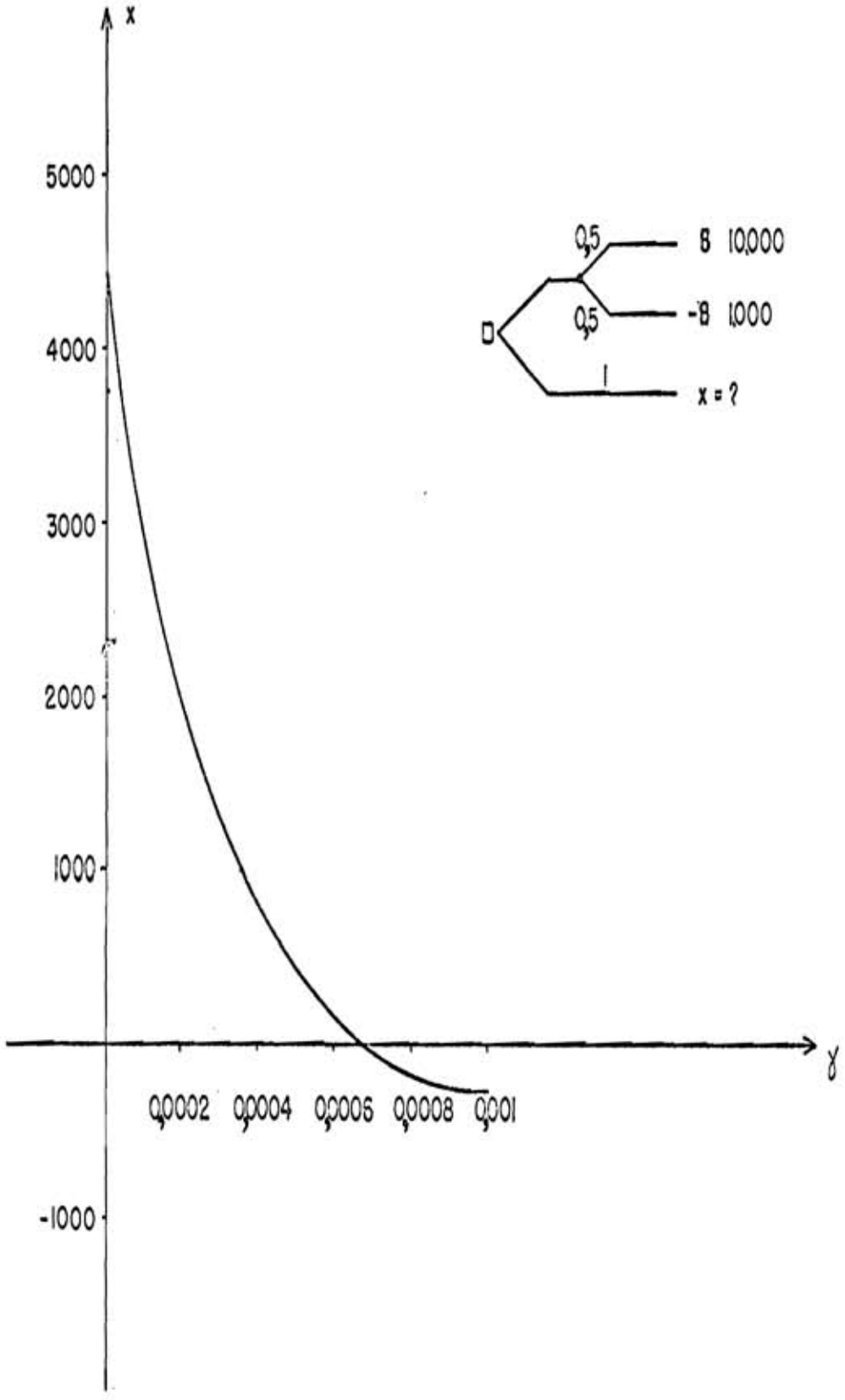


Figura 5.5 - Determinação do Coeficiente de Aversão ao Risco através de Loterias (2).

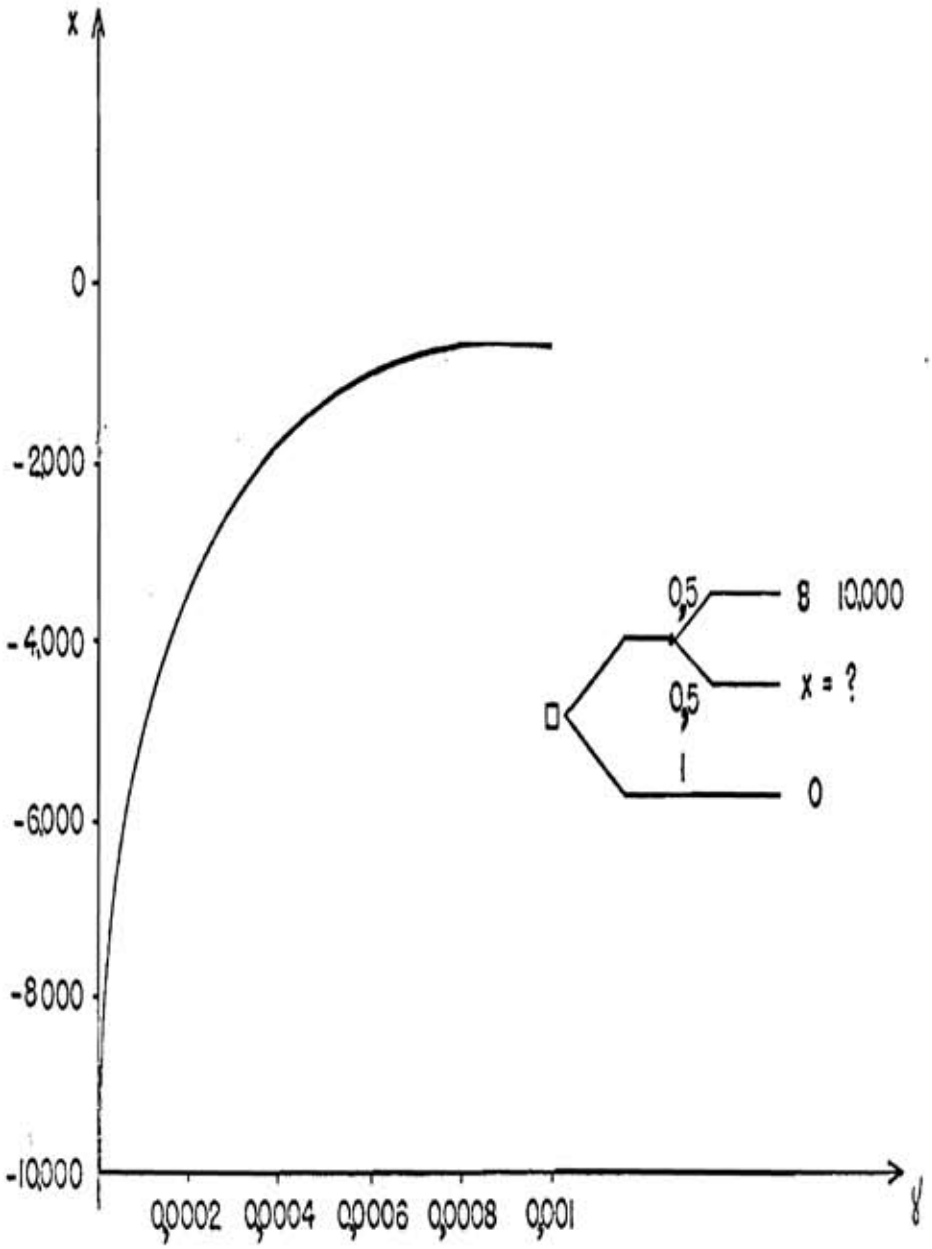


Figura 5.6 - Determinação do Coeficiente de Aversão ao Risco através de Loterias (3).

e, em seguida, obter no gráfico o valor correspondente do coeficiente de aversão ao risco.

O decisor poderá tomar como seu coeficiente de aversão ao risco uma média para as três loterias ou, então, somente o valor determinado pela loteria em que ele se sentir mais à vontade.

Naturalmente, é possível formular muitas outras loterias semelhantes e construir os gráficos correspondentes, simplesmente variando o valor das probabilidades e das recompensas fixas. Na prática, é importante que os valores das recompensas fixas sejam da mesma ordem de grandeza que os valores que ocorrem no problema em análise.

Quando se utiliza o método apresentado neste parágrafo, é importante que se faça o estudo da sensibilidade do problema em análise face a variações no coeficiente de aversão ao risco.

Este estudo é facilmente realizado com o auxílio do programa de processamento de árvores de decisão apresentado em anexo.

Na figura 5.7, temos um exemplo da sensibilidade do equivalente certo de duas loterias face a variações no coeficiente de aversão ao risco. Dependendo do valor deste coeficiente, deve-se preferir uma ou outra loteria.

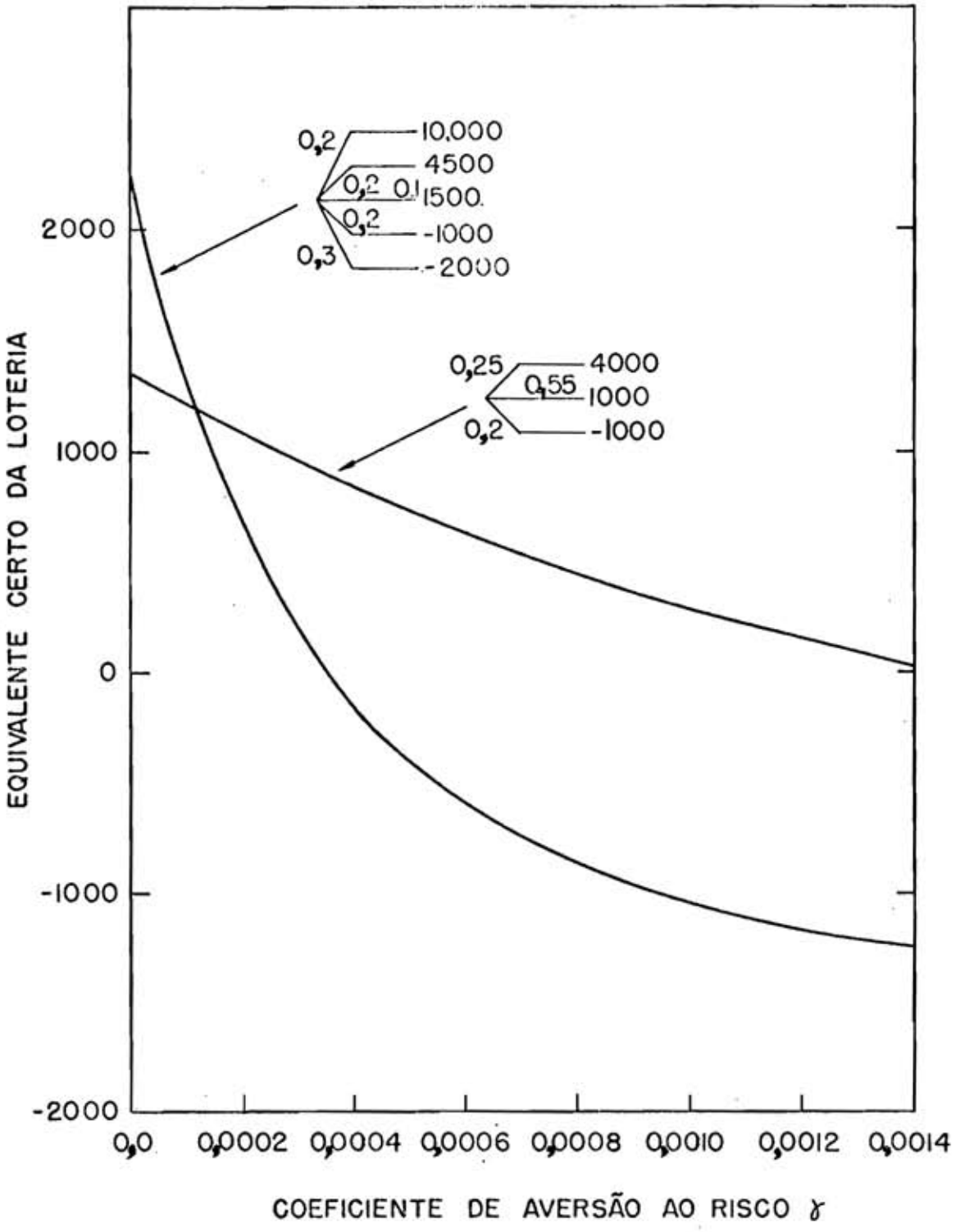


Figura 5.7 - Sensibilidade do Equivalente Certo a Variações no Coeficiente de Aversão ao Risco.

Uma vez realizado o estudo de sensibilidade, se for constatado que o coeficiente de aversão ao risco do decisor é próximo a um valor crítico no problema em análise, a curva de utilidade do decisor deverá ser levantada com maior cuidado, eventualmente usando-se o método do item 5.1.1.

5.1.3 - Curvas de Utilidade para Aspectos não Monetários

Pode ocorrer que a entidade cuja utilidade deva ser avaliada não seja diretamente mensurável em termos monetários. Neste caso, será preciso um procedimento algo diferente dos que foram vistos anteriormente.

Continuaremos a falar em utilidade, embora isto não seja correto. Seria melhor falar em "medidas de valor".

Neste item não estaremos falando em utilidade no sentido usado em economia e muito menos no sentido específico de Von Neumann - Morgenstern. A principal diferença é que a entidade a ser avaliada não estará geralmente ordenada naturalmente (ao contrário de dinheiro para o qual é sempre válido que quanto mais melhor). Às vezes, será mesmo impossível criar uma ordenação para as entidades que estão sendo avaliadas. Por isso não teremos necessariamente funções monotônicas crescentes. Em particular, não teremos às vezes, nem menos funções no sentido usual do termo. Mesmo o nome do parágrafo é algo impróprio pois não teremos necessariamente um gráfico ou curva que represente a utilidade da entidade avaliada.

Estamos, no fundo, preocupados apenas em estabelecer regras de correspondência entre quantidades ou qualidades de uma certa entidade e valores numa escala totalmente arbitrária (aos quais, por extensão de linguagem, continuaremos a chamar de utilidade).

As regras de correspondência citadas deverão fazer corresponder um único valor (utilidade) a cada possível valor da entidade que está sendo medida, sendo que estas regras deverão ser formuladas de maneira simples, uniforme e facilmente reproduzível. As regras poderão ser geralmente representadas por fórmulas matemáticas ou gráficos; contudo às vezes isto não será possível sendo, então, válidas quaisquer formas de tornar explícitas as relações de preferência existentes na mente do decisor.

Um cuidado que sempre deve ser tomado é com o campo de variação da entidade que estiver sendo avaliada. Este deverá cobrir todas as alternativas logicamente possíveis, não se restringindo aos valores "razoáveis" ou "esperados".

A escala de valores deverá ser definida seguindo-se as seguintes convenções:

- números positivos serão associados a situações com valor(utilidade) positivo, ou seja, situações pelas quais o decisor possua preferência ou que lhe gerem sentimentos positivos.
- números negativos serão associados a situações pelas quais o decisor possua aversão, ou seja, situações que tenham utilidade negativa.

A escala de utilidade será limitada acima por +1 e abaixo por -1.

O valor +1 será associado à situação que produza no decisor a maior satisfação, ou seja, que tenha a maior utilidade imaginável no contexto em análise. Analogamente, -1 será associado à pior situação possível, ou seja, àquela de menor utilidade.

A escala de utilidades será contínua entre +1 e -1 sendo que qualquer valor real é passível de ser associado a um valor da entidade avaliada.

Utilidade zero só será associada às situações pelas quais o decisor sinta total indiferença, ou seja, nem satisfação nem aversão.

Duas situações terão a mesma utilidade se o decisor for totalmente indiferente entre elas.

Vejamos, então, um procedimento para tornar explícitas as relações de preferência do decisor:

O primeiro passo é determinar qual o tipo de variação da entidade que está sendo avaliada. Ela poderá variar numa escala contínua (como por exemplo: o peso de um objeto) ou discreta (por exemplo: número de ocorrências de um determinado evento).

Caso se verifique que a variação é discreta, será preciso analisar em quantas categorias ela pode ser representada: duas (por exemplo: a existência ou não de um plano de ação), várias (como por exemplo o número de alunos de uma classe que acerta um teste: podemos ter nenhum, um, etc. até o total de alunos na classe), uma infinidade (neste caso podendo ser posta em correspondência com o conjunto dos inteiros não-negativos).

É importante verificar se as classes em que se divide a entidade possuem ou não alguma ordem natural: no exemplo acima sobre número de acertos, existe uma ordem natural que á a dos números inteiros; se por outro lado quisermos a utilidade de um dado número de cores para uma certa pessoa não haverá uma ordem natural. Se houver uma ordem natural usamos esta ordem; caso contrário fixamos uma ordem qualquer.

Caso a entidade a ser avaliada possua uma variação contínua, devemos verificar se esta é limitada ou não. Por ex.: o peso de um objeto possui um limite inferior (zero) mas não possui um limite superior lógico. O mesmo ocorre com frequências absolutas, como o número de habitantes de uma cidade. Já frequências relativas (como por ex.: a razão dos alunos de uma classe presentes à aula pelo total de alunos) possuem limite inferior zero e superior um.

Se não houver limites, nós os definimos usando bom senso.Se a entidade for limitada, os limites devem ser fixados. Lembremo-nos sempre que falamos de "limites lógicos", não de valores cuja ocorrência seja es

perada. Neste ponto, o bom senso é indispensável.

Após estas considerações de ordem geral, passemos à determinação específica de alguns pontos da curva de utilidade.

Os mais fáceis de determinar (ou seja, aqueles que o decisor geralmente se sente mais à vontade para avaliar) são os extremos da curva: pontos que correspondem à máxima satisfação (utilidade) e à máxima aversão. Associamos a estes valores utilidades +1 e -1, respectivamente.

A seguir verificamos qual o sentido de variação das relações de preferência (ou aversão). A satisfação aumenta (ou diminui) sempre entre os extremos? Existe algum ponto de máximo (mínimo) interno, ou seja, algum ponto ao qual o decisor associa a máxima (mínima) utilidade?

Através de considerações desta ordem é possível levantar a forma geral da curva de utilidade.

Uma vez determinada a forma geral da curva de utilidade é preciso fazê-la realmente representar as preferências do decisor.

Isto pode ser feito por um processo iterativo: são feitas perguntas ao decisor e, se as respostas contrariam a curva, esta é ajustada. Novas perguntas e correções são feitas até que a curva represente a função utilidade do decisor. Estas perguntas devem ser feitas no sentido de comparar dois a dois diferentes valores da entidade e saber qual deles é mais útil

para o decisor.

O fato de a utilidade de um dado ponto ser positiva ou negativa pode ser avaliado pondo-se o decisor frente à situação específica e perguntando-lhe quais os sentimentos (satisfação, aversão, indiferença)que ela lhe provoca.

O analista não deve se preocupar muito com a determinação do zero da curva de utilidade; este deve resultar naturalmente da própria curva mas pode, evidentemente, ser verificado.

5.2 - UTILIDADE MULTIDIMENSIONAL

5.2.1 - Situação do problema e necessidade de análise multidimensional

Na tentativa de sistematizar o processo de análise de decisões temos discutido seus principais aspectos: a árvore de decisões, utilidades e probabilidades. Incorporamos estes três conceitos num modelo normativo para os atos do decisor e concluimos que ele deveria maximizar sua utilidade esperada.

Temos, então, condições para analisar um grande número de problemas, no entanto, todos eles sujeitos a uma restrição fundamental: estaremos sempre procurando um único objetivo, ou seja, haverá uma única entidade (por ex.: dinheiro) cuja utilidade esperada deverá ser maximizada.

Embora esta restrição seja básica para o uso do modelo estudado, ela não diminui sua importância ou aplicação porque é muito grande o número de situações em que há uma única dimensão significante para a decisão a ser tomada. A maioria dos problemas de análise financeira recaem neste caso pois, geralmente, estaremos interessados em maximizar alguma quantidade mensurável em termos monetários (por ex.: lucro líquido).

No entanto, em muitos problemas envolvendo decisões nós nos vemos frente a mais de um objetivo. Isso nos leva a considerar mais de uma dimensão de valor, sem que possamos, sem prejuízo da análise, escolher uma dentre elas em que basear nossa decisão.

Vemo-nos então, frente a problemas de utilidade multidimensional, ou seja, para a tomada de decisão desejamos maximizar não a utilidade de uma entidade, mas maximizar a utilidade de várias entidades simultaneamente.

Apresentaremos a seguir um método heurístico para a análise de problemas deste tipo.

De um modo geral, o método envolve quatro etapas principais. Em primeiro lugar, é preciso listar todas as dimensões de valor em que o decisor baseará sua decisão. Em seguida, é preciso dar pesos relativos a cada uma destas dimensões. Em terceiro lugar, é preciso criar para cada dimensão de valor uma regra de correspondência entre os valores de cada alternativa e uma escala fixada pelo decisor, ou seja, levantar a função uti

lidade de cada dimensão de valor e, por fim, é preciso uma regra aritmética que combine as diversas dimensões num único índice, por exemplo, uma simples média ponderada.

Dissemos que o processo é heurístico porque não é possível provar que ele conduza à decisão ótima. No entanto, ele se justifica por várias considerações teóricas e práticas.

Quanto ao aspecto teórico existem vários estudos sobre utilidade multidimensional que, no entanto, não serão aqui tratados por fugirem aos objetivos deste manual. Utilidade multidimensional foi estudada, entre outros, por Fishburn [14], [15], [16], Keeney [23] , Raiffa [40].

O método, tal qual será apresentado, foi desenvolvido por Edwards [6] e estudado e validado na prática por Miller [32] e Fischer [13].

Outros métodos utilizados para a análise multidimensional e que não serão aqui apresentados podem ser encontrados em Mac Crimmon [30]

Antes de passar à descrição do método, é necessário que se façam algumas considerações; de um modo geral o método compreende duas fases distintas: a geração de um algoritmo e sua aplicação ao problema de decisão.

O método de análise a ser apresentado é normativo (ou seja, diz como deve agir o decisor) e não descritivo (ou seja, não pretende ser uma descrição de como são, na prática, tomadas as decisões).

5.2.2 - Geração do Algoritmo

Apresentaremos agora uma sequência de passos lógicos ao fim dos quais teremos um algoritmo através do qual poderemos tomar uma decisão.

1º passo: Identificar as dimensões de valor importantes para a decisão.

Inicialmente, procura-se gerar o maior número possível de dimensões de valor ou critérios de avaliação que possam vir a influir na decisão, sem nos preocuparmos com suas interrelações ou importância. De posse desta lista inicial, procuramos selecionar as não triviais e sempre que possível, dimensões de valor independentes, ou seja, que representem aspectos diversos da decisão a ser tomada.

Não deve haver nesta etapa preocupações com as unidades em que serão medidas as entidades: poderemos ter misturadas dimensões mensuráveis em termos objetivos, como dinheiro e tempo, com outras totalmente subjetivas e que não podem ser medidas, como importância e aspecto físico.

Naturalmente, nesta fase, a participação do analista é mínima, devendo apenas motivar o decisor que deverá gerar o maior número possível de dimensões e depois selecionar as mais importantes.

2º passo: Classificar as dimensões de valor em ordem de importância.

De posse da lista de dimensões de valor relevantes para a decisão, é preciso ordená-la, ou seja, o decisor deverá escolher dentre os critérios qual o mais importante, qual o seguinte e assim por diante, até esgotar a lista.

3º passo: Atribuir pesos relativos às dimensões de valor.

Se no passo anterior o decisor concluiu que todas as dimensões de valor têm a mesma importância, atribuimos a cada uma delas um mesmo número (10 por ex.), que será o seu peso relativo.

Se o decisor conseguiu ordenar as várias dimensões, escolhemos a menos importante e atribuimos a ela peso 10 (ou qualquer outro número; a escolha é totalmente arbitrária). Consideremos a dimensão seguinte na ordem de importância: o decisor terá que decidir quanto ela é mais importante que a primeira. Atribuimos à segunda dimensão um número que reflita esta razão de importâncias, continuamos a comparar todas as outras dimensões com a menos importante e a atribuir-lhes pesos relativos. Esgotada a lista, deve-se verificar a consistência dos julgamentos. Desta forma, caso se atribua peso 30 a uma certa dimensão enquanto se atribui peso 90 a uma outra, isto quer dizer que a segunda dimensão é três vezes mais importante que a primeira. Caso o decisor não concorde com isto, ele deverá alterar um dos pesos (ou ambos).

Se, ao comparar duas dimensões quaisquer, o decisor concluir que elas são igualmente importantes, atribuimos a ambas o mesmo valor relativo. Se, ao comparar duas dimensões, o decisor concluir por pesos relativos que contrariem a ordem da lista original, não há problemas: altera-se a ordenação na lista de critérios e refaz-se a atribuição de pesos relativos.

Nesta etapa, o analista deve participar ativamente, orientando o decisor na atribuição dos pesos relativos e incentivando as reformulações que contribuam para a consistência dos pesos atribuídos.

Para encerrar este passo, um detalhe de normalização: somamos os pesos atribuídos a cada dimensão de valor, dividimos cada um pelo total e multiplicamos por 100. Assim, todos os pesos serão representados por números entre 0 e 100 e cuja soma é 100. Deste modo cada peso pode ser encarado como a proporção de importância da dimensão de valor dentro do sistema.

Notemos aqui que,se houver muitas dimensões de valor,algumas delas terão provavelmente pesos insignificantes frente a outras. Neste caso, desprezamos estas dimensões e refazemos o passo 3 apenas com as dimensões realmente importantes.

4º passo: Levantar curvas de utilidade para cada dimensão de valor.

Talvez este seja o passo mais difícil para o analista. Ele

deverá tornar explícitas as preferências e aversões do decisor para todas as dimensões de valor.

Qualquer dos métodos citados no item 1 (utilidade unidimensional) pode ser usado. É importante lembrar que embora a escala de utilidades seja arbitrária ela deverá ser a mesma para todas as curvas de utilidade.

5º passo: Agregar as várias dimensões de valor num único índice.

Até agora, as várias dimensões de valor foram estudadas isoladamente. Neste passo final da geração do algoritmo deveremos agregá-las num único índice.

A maneira mais simples e usual de implementar esta agregação é através de uma combinação linear das dimensões de valor, usando como coeficientes os pesos relativos atribuídos no passo 3.

Assim,se chamarmos de u_j o índice de utilidade da j-ésima dimensão de valor e de C_j o peso relativo da mesma, teremos

$$K = \sum_j C_j \, u_j$$

onde K é a utilidade agregada de todas as dimensões de valor. Veremos, em itens seguintes, as diferentes interpretações do índice K e como utilizá-lo no processo de decisão entre diferentes alternativas e na avaliação de projetos.

Para encerrar esta seção, apresentaremos um método alternativo para a implementação do 1º passo, ou seja, a geração das dimensões de valor. Este método pode ser útil quando não temos uma visão muito clara do sistema em análise e encontramos dificuldade em escolher critérios eficientes para avaliá-lo.

O método consiste em gerar uma hierarquia de representação do sistema através de uma estrutura em forma de árvore onde serão representados, em níveis sucessivos, o sistema, seus subsistemas e os componentes destes últimos. A cada um dos componentes finais (de mais baixo nível na hierarquia) deverá ser associado um critério de avaliação, ou seja, uma dimensão de valor para cada componente. O propósito deste processo é criar uma estrutura pictórica das interrelações de valor que existem na mente do decisor.

Obviamente, os dois métodos não são incompatíveis. Pode-se, inicialmente, gerar uma lista de dimensões de valor e, depois, depurá-la com auxílio da hierarquia de representação do sistema, escolhendo uma dimensão de valor para cada componente deste.

Se for usado o método alternativo proposto para o passo um, a estrutura criada poderá ser usada para auxiliar os passos dois e três. A metologia é a mesma do processo original, apenas a ordenação e a atribuição de pesos poderão ser feitas a cada nível de representação do sistema, tornando menores e mais fáceis os julgamentos a serem feitos pelo decisor.

5.2.3 - Aplicação do Algoritmo ao Processo de Decisão

Temos até agora desenvolvido um modelo de análise multidimensional. Vejamos como aplicá-lo ao processo de decisão.

Colocados frente a várias alternativas, não pudemos decidir entre elas utilizando a utilidade esperada, simplesmente porque havia mais de um aspecto relevante à decisão. Através do modelo de análise multidimensional procuramos contornar este problema.

Temos agora um processo de agregar as utilidades das várias dimensões relevantes à nossa decisão, num único índice.

O que se deve fazer é avaliar o índice de utilidade agregada para cada alternativa, ou seja, entrar com o valor que uma alternativa oferece para cada dimensão de valor na função utilidade correspondente e, com as utilidades encontradas, calcular o valor do índice agregado para aquela alternativa.

Assim, sendo U_{ij} o índice de utilidade da alternativa i para a dimensão de valor j e C_j o peso relativo da j-ésima dimensão de valor, temos

$$K_i = \sum_j C_j U_{ij}$$

onde K_i pode ser considerado como a utilidade agregada da alternativa i.

No caso em que temos de selecionar uma única alternativa, a regra de decisão é simples: escolhe-se a alternativa que maximiza a utilidade agregada.

No caso de termos que selecionar mais de uma alternativa, devemos ter um outro critério, por exemplo, um orçamento máximo. Neste caso, escolheríamos as alternativas de maior utilidade até esgotarmos o orçamento.

Neste caso, poderíamos também utilizar a análise custo-benefício que será vista no próximo item.

5.2.4 - Análise Custo-Benefício

O método de análise multidimensional nos sugere uma maneira de sistematizar a análise custo-benefício.

Até agora, temos tratado aspectos monetários da decisão como simples dimensões de valor.

Para realizar uma análise custo-benefício deve-se proceder de maneira ligeiramente diferente; todos os aspectos diretamente mensuráveis em termos monetários devem ser isolados e analisados separadamente dos aspectos (dimensões de valor) subjetivos, ou seja, não representáveis em termos monetários.

Os vários tipos de custos, despesas e mesmo lucros monetários dos projetos, devem ser submetidos a uma análise financeira convencional e, através de um fluxo de caixa, pode-se chegar ao custo total do projeto.

Os aspectos subjetivos são encarados como benefícios e analisados pela técnica apresentada anteriormente.

Chegamos então, para cada projeto, a dois índices. Seu custo (medido em cruzeiros ou na utilidade equivalente à esta quantia) e seu benefício (representado pelo índice de utilidade agregada).

Com estes dois índices obtem-se a razão custo-benefício.

Se o problema for escolher um projeto entre vários, a regra de decisão é escolher o de menor razão custo-benefício.

Geralmente, o problema não se apresenta de maneira tão simples. Por exemplo, é comum termos de selecionar entre um grande número de projetos apenas um sub-conjunto que poderá ser implementado, devido a restrições orçamentárias. Neste caso, um procedimento válido é o seguinte: calcula-se para cada projeto em análise sua razão custo-benefício. Ordena-se os projetos pela ordem crescente desta razão e seleciona-se os primeiros projetos, até que a restrição orçamentária se esgote. (Figura 5.8).

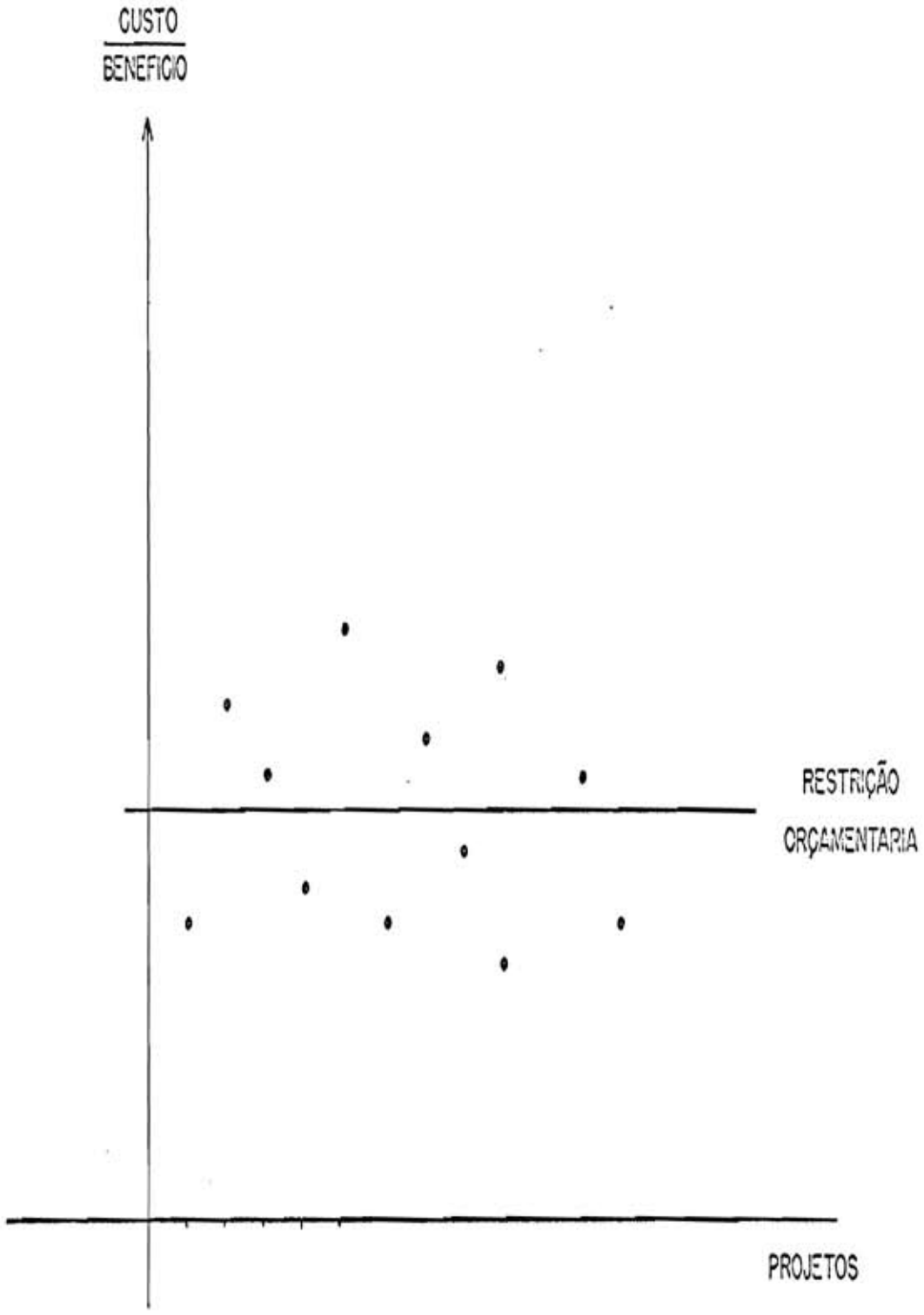


Figura 5.8 - Seleção de Projetos através da Análise Custo-Benefício.

Um ponto que deve ser notado é que o procedimento citado anteriormente é uma simplificação que só se aplica quando os vários projetos apresentam relações custo-benefício da mesma ordem de grandeza.

O conceito em questão é o de curvas de isopreferência. No caso em que as dimensões de valor são o custo e o benefício, elas assumem a forma da figura 5.9.

No entanto, na análise feita no início desta seção foram consideradas como sendo retas. (Figura 5.10). Isto só é válido numa região limitada do mapa de curvas de isopreferência (figura 5.11). Daí concluirmos que a análise só se aplica a diferentes projetos, com razões custo-benefício da mesma ordem de grandeza.

Um exemplo de aplicação da análise custo-benefício pode ser encontrado num trabalho realizado por uma equipe do INPE que sugere o emprego desta metologia na avaliação de projetos a serem incluidos no Plano Básico de Desenvolvimento Científico e Tecnológico (PBDCT) [20].

5.2.5 - Aplicação à Avaliação de Projetos

Até agora, falamos apenas em seleção de alternativas para projetos a serem implementados. No entanto, o método apresentado também se aplica (com ligeiras modificações) à avaliação de projetos em andamento.

Naturalmente, esta avaliação é no sentido de atingimento,ou

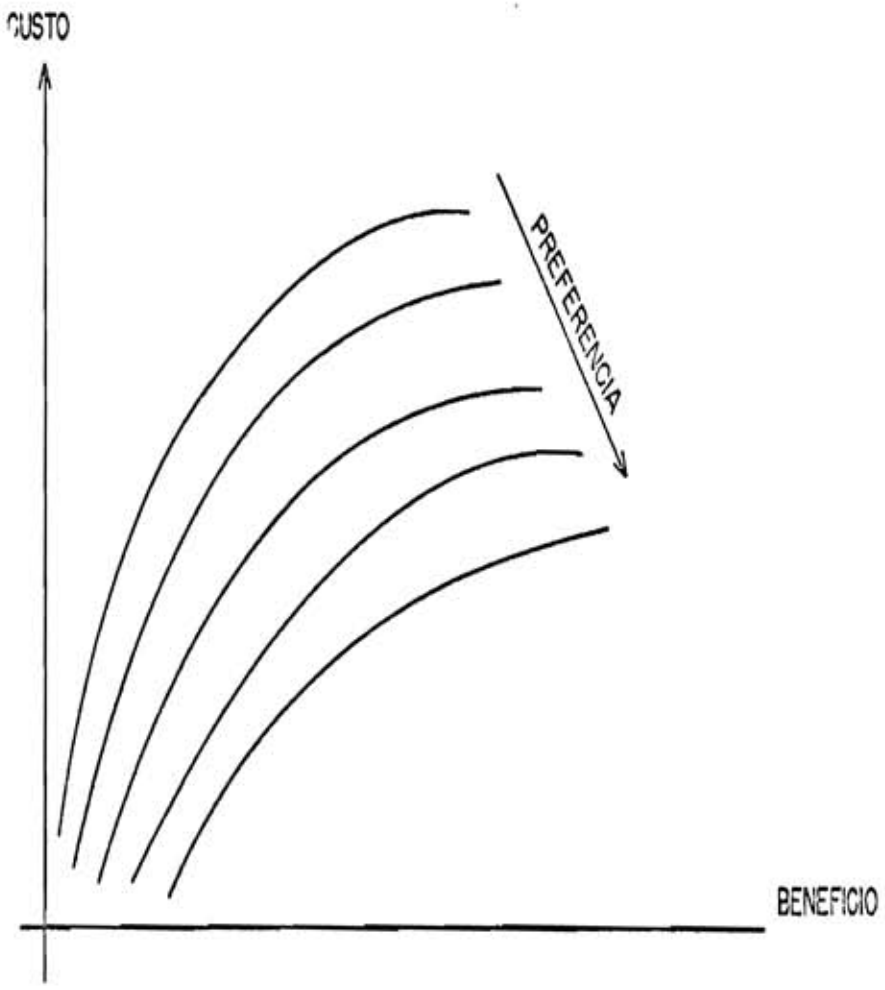


Figura 5.9 - Curvas de Isopreferência entre Custo e Benefício.

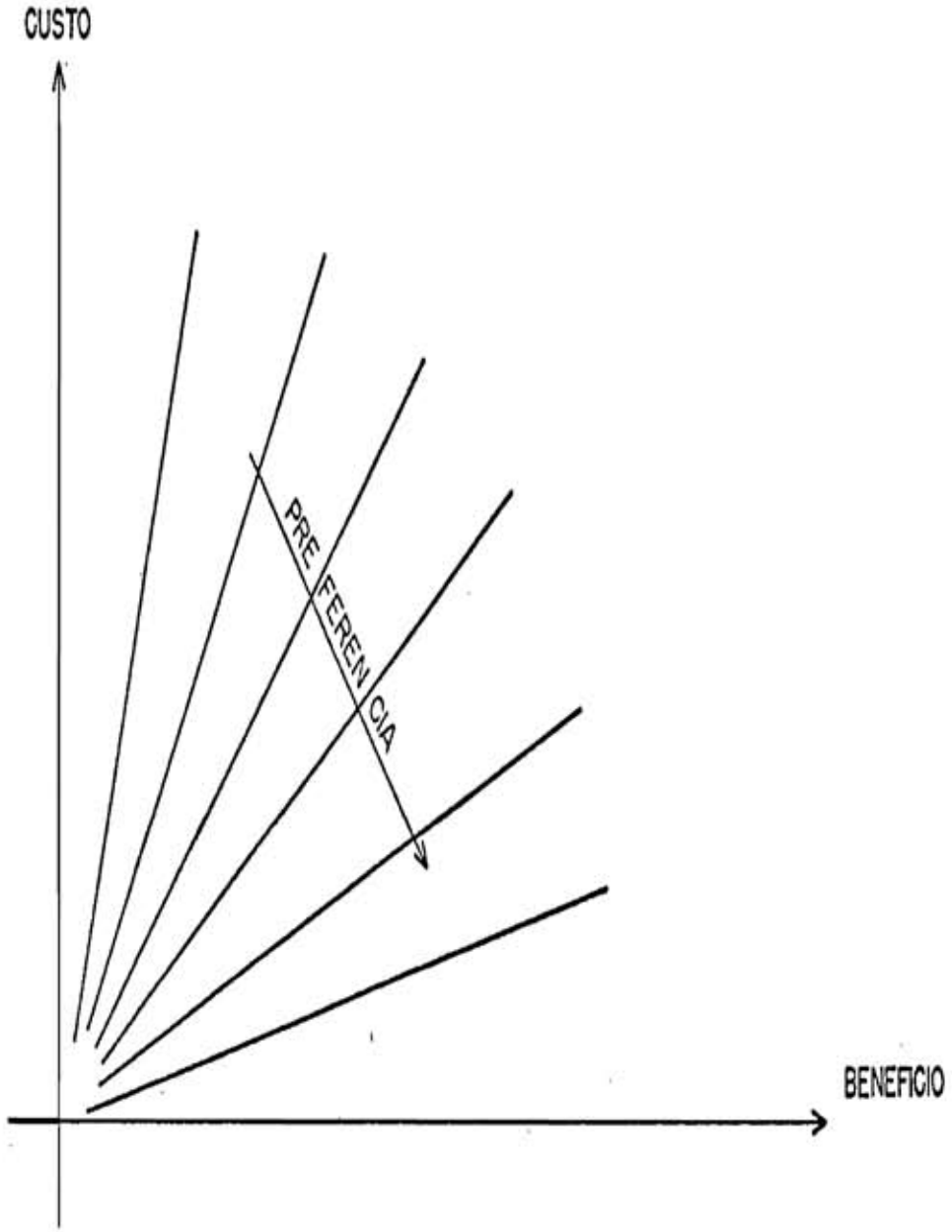


Figura 5.10 - Retas de Custo Proporcional ao Benefício.

Figura 5.11 - Aproximação das Curvas de Isopreferência por Retas de Custo Proporcional ao Benefício.

não dos objetivos propostos. Para tanto, é recomendável que se use o método alternativo proposto no item 2.2.

Para a criação da hierarquia de representação do sistema ou projeto deve ser usada a árvore de objetivos criada na fase inicial do planejamento. Os passos restantes de geração do algoritmo se aplicam integralmente.

Na fase de aplicação do algoritmo, devem ser consideradas três alternativas: o estado inicial da natureza, anterior à implantação do projeto, o estado atual já com a realidade modificada com a implantação em maior ou menor grau do projeto e o estado ideal que se espera obter se o projeto atingir integralmente seus objetivos.

Naturalmente, este estado ótimo trará máxima satisfação para o decisor, devendo, portanto, ser atribuído a ele o máximo valor da nossa escala de utilidades. Como estamos usando um modelo linear para agregação de nossas variáveis, vemos que, em particular, cada dimensão de valor terá para essa alternativa o valor máximo da escala de utilidades.

A realidade anterior ao projeto não terá necessariamente utilidade negativa ou mesmo nula. As diversas dimensões de valor deverão ser avaliadas normalmente e a utilidade calculada.

O acompanhamento do projeto deverá ser feito periodicamente através da observação da realidade modificada e introdução dos parâmetros

mensurados na função de utilidade agregada. Neste caso, o índice K pode ser considerado como a efetividade do projeto.

A variação deste índice ao longo do tempo pode, então, ser facilmente observada e espera-se que o índice parta de um valor inicial (realidade anterior à do projeto) e aumente até o valor máximo.

Se tal não ocorre, ou ocorre com velocidade menor que a esperada, medidas corretivas devem ser tomadas. A própria estrutura de valores nos fornece indicações de onde alocar mais recursos ou dedicar maiores esforços.

PARTE III

EXEMPLOS DE APLICAÇÃO

CAPÍTULO VI

EXEMPLO DE CODIFICAÇÃO DE INCERTEZA

6.1 - COLOCAÇÃO DO PROBLEMA

Uma agência de publicidade é um ambiente típico onde, habitualmente, são tomadas decisões envolvendo variáveis de comportamentos praticamente imprevisíveis por métodos tradicionais.

Com o propósito de testar a aplicabilidade das técnicas apresentadas neste manual, no tocante à codificação da incerteza, entramos em contato com elementos de renomada agência de publicidade de São Paulo.

Oficialmente, estivemos reunidos com representantes dessa agência em dois encontros de duração média de duas horas cada um. Extra-oficialmente, um de nosso grupo já tivera oportunidade de discutir idéias relacionadas à análise de decisões com o Sr. S, encarregado de uma das contas da agência e seu conhecido pessoal. Cumpre ressaltar que pudemos contar com toda a boa vontade por parte dos elementos contatados.

Em nosso primeiro encontro, o orientador do grupo e dois de nós estiveram reunidos com o Diretor Superintendente e com um dos en

carregados de conta da empresa. O objetivo dessa reunião foi a troca de informações de ambas as partes, na tentativa de fixar condições para o experimento pretendido.

Ficou então estabelecido o "aumento percentual de vendas de um produto", ao final de uma campanha publicitária, como sendo a variável a ser quantificada. Deliberou-se também que se escolheria um produto alimentício por estar menos sujeito a variáveis fora de controle como a moda, por exemplo. Essa última determinação partiu dos elementos da agência. A "estabilidade" do tipo de produto parecia deixá-los mais à vontade para fazer estimativas. Evidentemente, nada tínhamos a objetar. Em face do que foi deliberado, fomos encaminhados ao Sr. S que, na oportunidade, estava responsável pela campanha do produto J, com as características desejadas.

No segundo encontro, diretamente com o Sr. S e cerca de uma semana mais tarde, realizamos um experimento de codificação que será apresentado na sequência deste capítulo.

Vale a pena observar, e isto pode ser constatado ao longo das duas reuniões, a resistência compreensível do profissional em emitir julgamentos pessoais sobre o comportamento dos parâmetros envolvidos de incerteza.

O diálogo "fictício" que se segue é uma tentativa de i

lustrar uma forma de argumentação contra essa resistência inicial, que foi usada e parece surtir efeito:

Analista - Vamos lhe pedir que emita opiniões sobre o comportamento das vendas deste produto, imediatamente após o término dessa campanha que você preparou e está para ser lançada.

Profissional - Isso é impossível. Qualquer coisa pode ocorrer ! Os fatores que podem influenciar os resultados da campanha são incontáveis. Nós temos que confiar na sorte. Só podemos garantir que procuramos fazer a melhor campanha. Nada podemos afirmar quanto às vendas futuras.

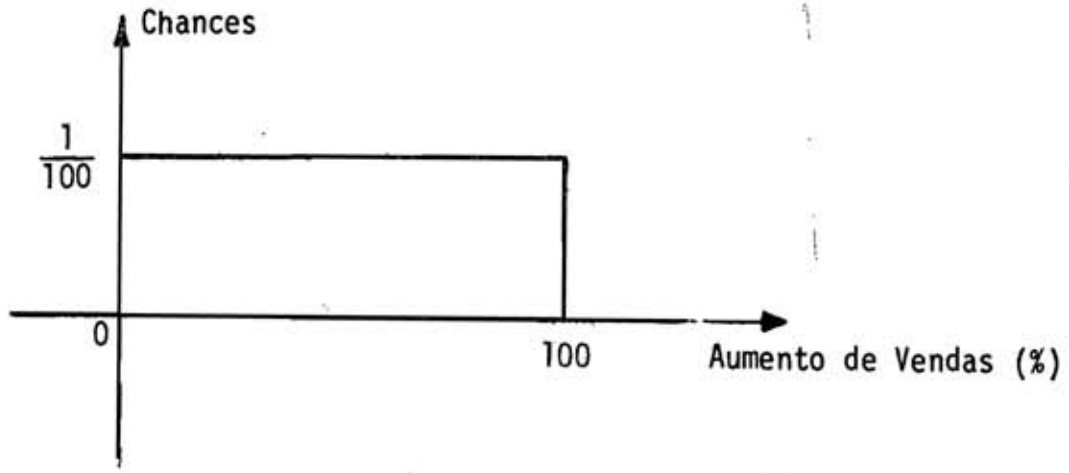
Analista - Concordo que seria impossível lhe pedir que sintetizasse em um número, todo o conhecimento da campanha e sua experiência com situações análogas. Não estamos lhe pedindo isso. Mas acreditamos que esse mesmo conhecimento e essa mesma experiência o habilitam a fornecer informações bem mais úteis do que "qualquer coisa pode ocorrer". Veja bem, talvez qualquer coisa possa ocorrer, mas "certas coisas" provavelmente podem ocorrer com maior chance. Você diria, por exemplo, que as chances de, no final da campanha, o aumento de vendas estar entre 100% e 150% são as mesmas de estar entre 30% e 80% ?

Profissional - Não, não diria. Para mim as chances do aumento de vendas ser maior que 100% são mínimas. Probabilidade quase zero eu diria.

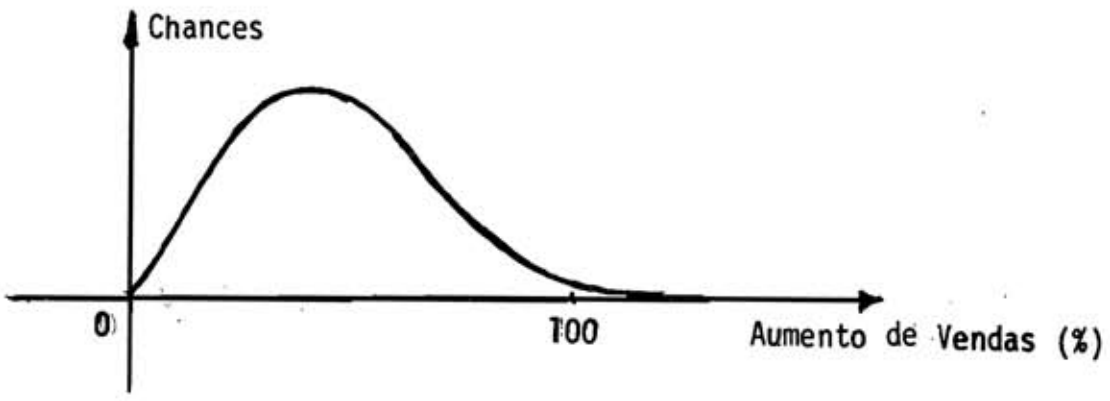
Analista - Muito bem. Já temos mais alguma informação agora. E quanto à possibilidade do aumento ser negativo, isto é, o volume de vendas após sua campanha ser menor que antes dela.

Profissional - Chance nula, para mim.

Analista - Ótimo, Você já conseguiu limitar bastante a "qualquer coisa". Se agora você julgasse que entre 0% e 100% não há condições de se afirmar nada, seriamos forçados a concluir que a distribuição do aumento de vendas, para você, é uniforme nesse intervalo. Isto pode ser representado pelo gráfico abaixo:



Mas eu acredito que você tem condições de fornecer informações menos genéricas. Acredito que possamos chegar a uma curva como a da figura abaixo:



Profissional - Muito bem. Podemos tentar.

Vencida, de certo modo, a resistência inicial do estimador, a etapa seguinte consistiu da aplicação das técnicas apresentadas na seção 3.3 para levantamento da função distribuição de variáveis aleatórias contínuas.

6.2 - O EXPERIMENTO: RESULTADOS E COMENTÁRIOS

Caracterização da Variável:

Aumento das vendas semanais do produto J (em relação às vendas antes de ser iniciada a campanha), na primeira semana após o término da campanha publicitária.

Época do Experimento:

Nossa entrevista foi realizada na semana do lançamento da campanha. O produto J é um produto de estação e a campanha deveria durar quatro semanas apenas.

Condições de Contorno Fixadas:

Foi estabelecido que o estimador deveria considerar que o serviço de distribuição do produto iria atender normalmente à demanda

extra. Igualmente foram fixadas como "normais" as demais variáveis de estado tais como incentivos de preço oferecidos pelo produtor ao consumidor, vantagens oferecidas aos intermediários, etc.

O Estimador:

O elemento que funcionou como estimador foi o próprio Sr. S, encarregado da conta da companhia fabricante do produto J. O Sr. S foi o responsável direto pela campanha estando a par de todos os seus detalhes. É um indivíduo com formação em nível de pós-graduação na área de publicidade e com experiência profissional de vários anos nos Estados Unidos e no Brasil.

Intervalo de Variação do Parâmetro

O estimador decidiu atribuir valor zero à probabilidade do aumento de vendas ser negativo. Fixou também o limite superior em 100%, atribuindo uma chance de 1% ao evento "aumento de vendas superior a 100%".

1) Resultados Obtidos Utilizando-se o Método de Estimativas por Pontos

Inicialmente obtivemos uma série de estimativas utilizando o processo 1.a do item 3.3.2. Os resultados obtidos estão sintetizados na Tabela 6.1.

TABELA 6.1

Pontos da Curva de Distribuição Acumulada - Método de Pontos

Aumento de Vendas (%) x	$P(X \leq x)$
0	0,00
10	0,20
20	0,60
30	0,85
50	0,88
70	0,95
100	0,99

Cumpre observar que, inicialmente, o estimador havia associado uma probabilidade de 0,80 ao aumento de vendas inferior a 50%. Obtidas os demais valores, fize-mo-lo ver a inconsistência existente entre essa estimativa e aquela feita para o aumento de vendas inferior a 30% (0,85).

Diante dessa observação o estimador decidiu alterar para 0,88 a probabilidade de "$X \leq 50$".

De posse dos dados da Tabela 6.1, pudemos construir a curva da figura 6.1 que dá a distribuição subjetiva da variável pesquisada.

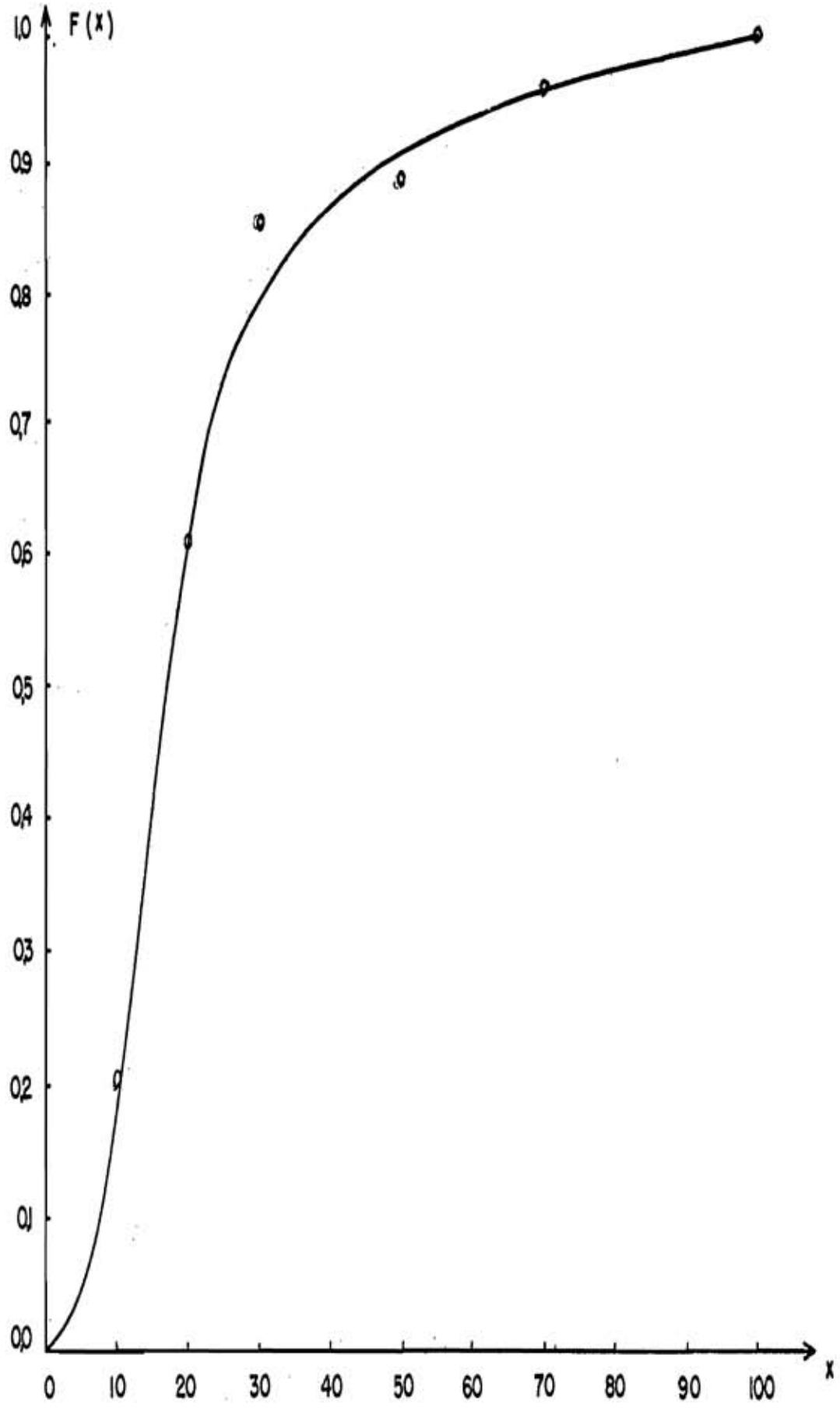


Figura 6.1 - Função Distribuição Subjetiva do Aumento de Vendas do Produto J Obtida pelo Método de Pontos.

A curva da figura 6.2 representa a função densidade e foi construida a partir da figura 6.1 pelo método do histograma apresentado no item 3.3.3.

2) Resultados Obtidos Utilizando-se o Método de Intervalos

Em seguida, usamos o método de intervalos descrito no item 3.3.2 para obter uma nova distribuição subjetiva.

O processo foi sempre conduzido com dois intervalos de cada vez. Feitas as explicações necessárias, o estimador não demonstrou dificuldades maiores em emitir os julgamentos que lhe foram solicitados. A tabela 6.2 ilustra, com a notação introduzida no item 3.3.2, uma das fases do processo.

TABELA 6.2

Processo Interativo com Decisão Entre Dois Intervalos

INTERVALOS PROPOSTOS		OPINIÃO DO ESTIMADOR
1	2	
(0, 50)	(50, 100)	1 $\succ$ 2
(0, 30)	(30, 100)	1 $\succ$ 2
(0, 20)	(20, 100)	2 $\succ$ 1
(0, 25)	(25, 100)	1 $\succ$ 2
(0, 22)	(22, 100)	1 $\equiv$ 2

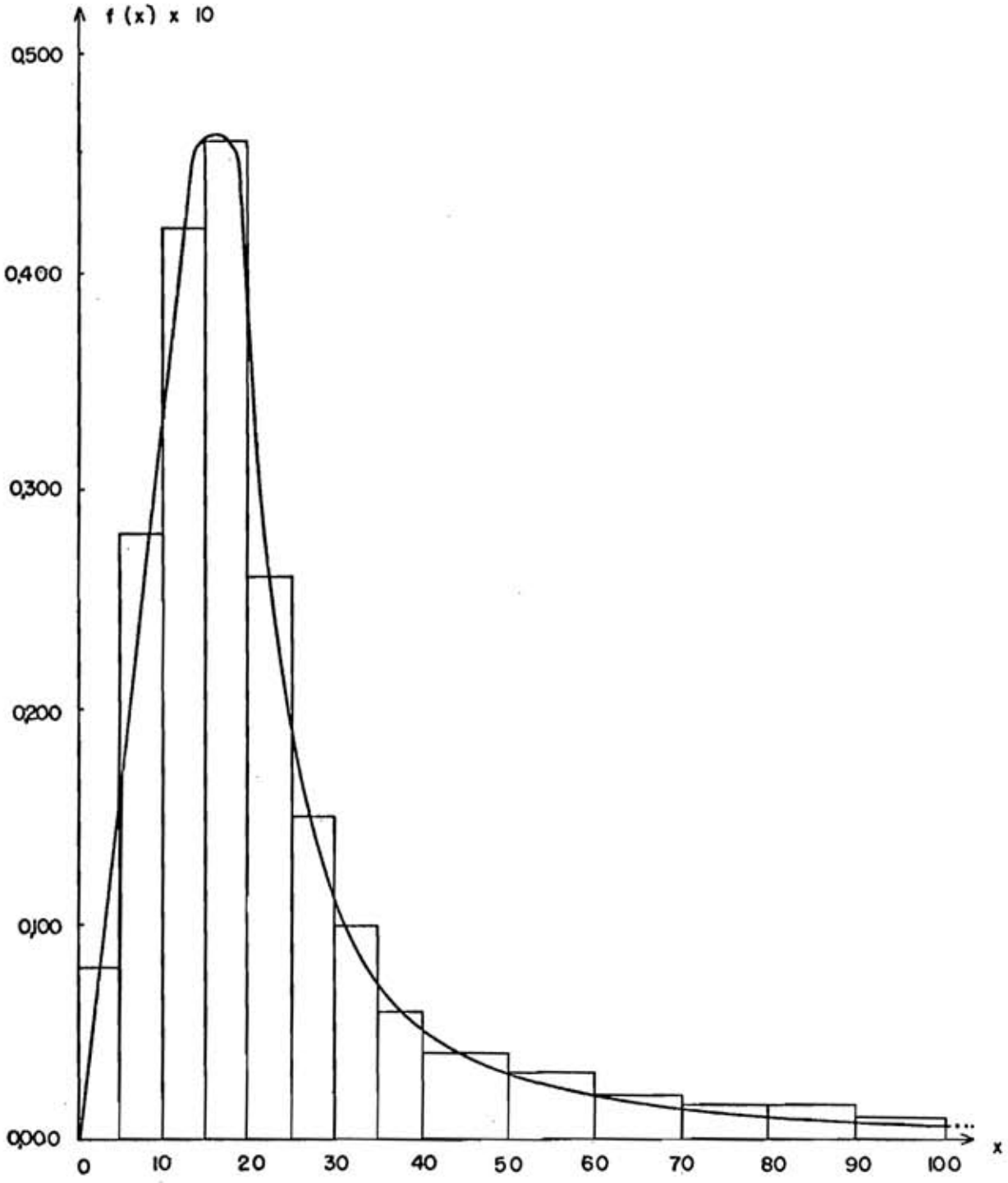


Figura 6.2 - Função Densidade de Probabilidade Correspondente à Curva da Figura 6.1.

Na tabela 6.3 sintetizamos os resultados finais obtidos.

TABELA 6.3

Pontos de Indiferença

INTERVALOS	PONTOS DE INDIFERENÇA
0 – 100	22
0 – 22	19
22 – 100	24
24 – 100	35
35 – 100	55
0 – 19	16

Como foi visto no item 3.3.4, cada ponto de indiferença divide o intervalo correspondente em dois outros equiprováveis.

Na tabela 6.4 apresentamos os dados da tabela anterior traduzidos em termos de probabilidades.

TABELA 6.4

Pontos da Curva de Distribuição Acumulada - Método de Intervalos

Aumento de Vendas (%) x	$P(X \leq x)$
0	0,000
16	0,125
19	0,250
22	0,500
24	0,750
35	0,875
55	0,938
100	1,000

Como se pode notar, assumimos que todos os valores possíveis para o aumento de vendas estão contidas no intervalo de 0 a 100. Em face da probabilidade negligível associada pelo estimador ao evento "aumento de vendas maior que 100%", essa simplificação não implica em erro significativo.

As figuras 6.3 e 6.4 representam as funções distribuição e densidade da variável em estudo.

Comentários

a) Como se pode observar pelas figuras 6.2 e 6.4, o estimador, subjeti

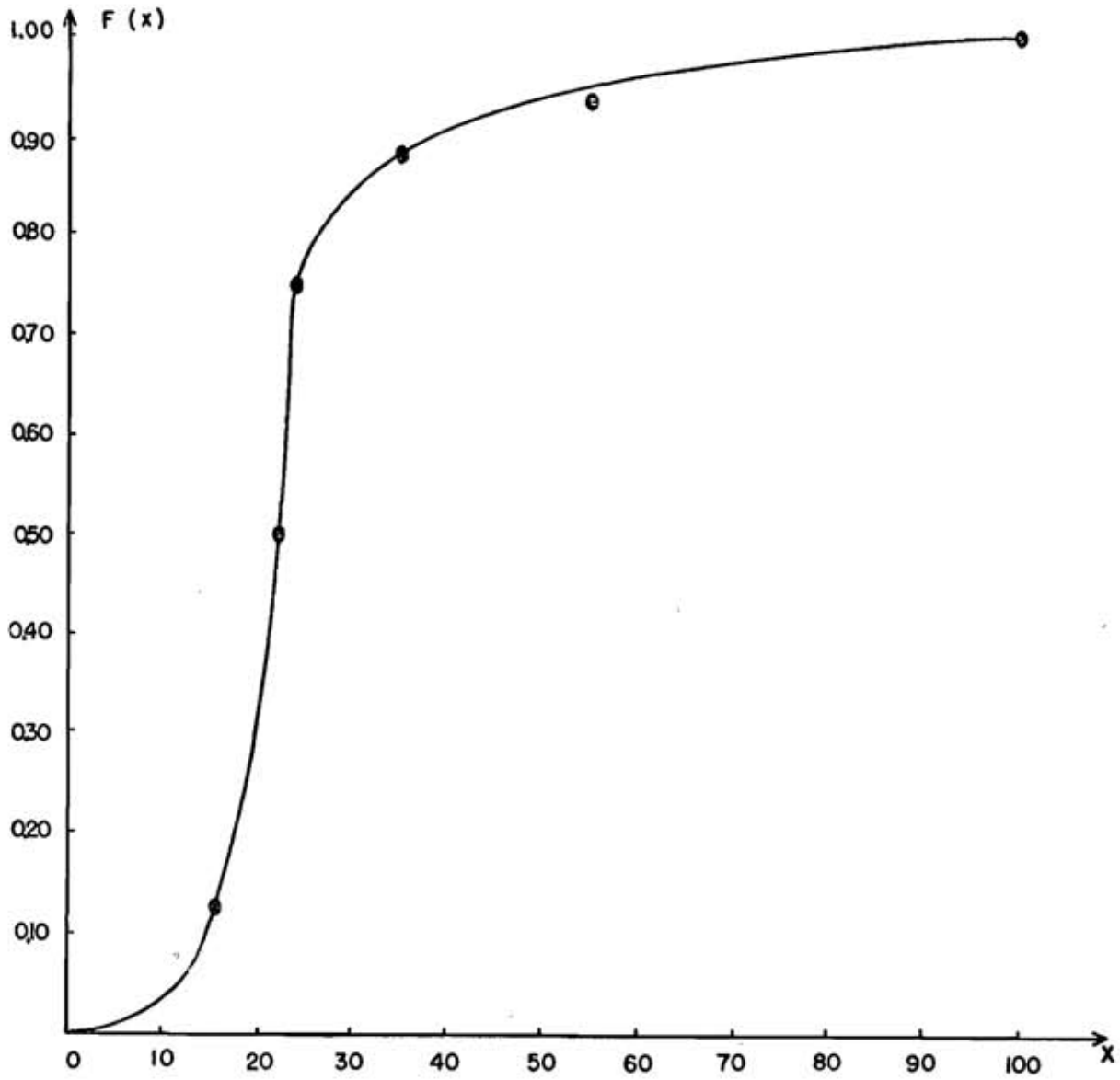


Figura 6.3 - Função Distribuição Subjetiva do Aumento de Vendas do Produto J Obtida pela Metodo de Intervalos.

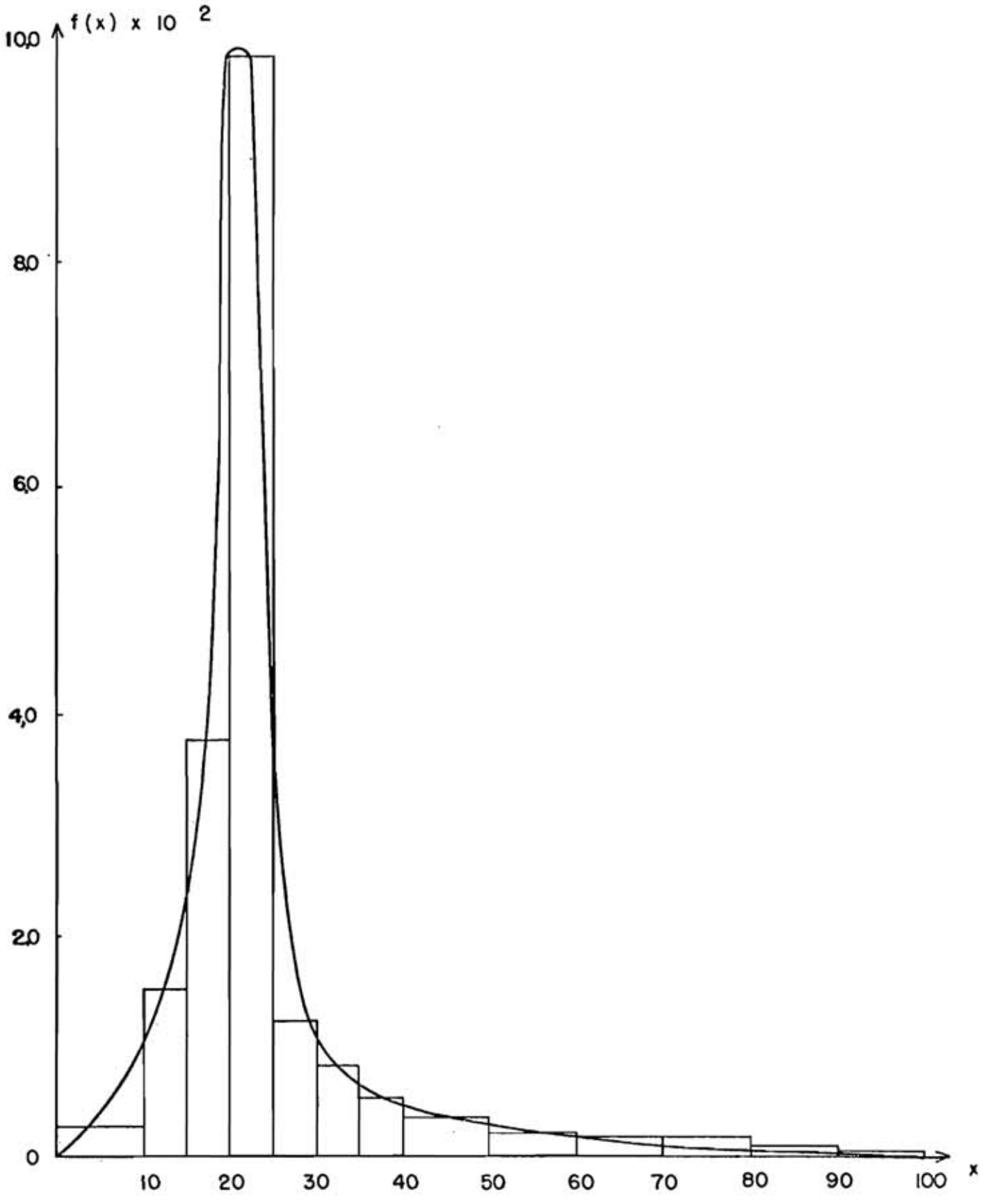


Figura 6.4 - Função Densidade de Probabilidade Correspondente à Curva da Figura 6.3.

vamente, atribuiu uma probabilidade bastante alta para um intervalo de aumento de vendas situado em torno de 20%. Se considerarmos que esse mesmo estimador de início relutava bastante em emitir qualquer opinião, esse resultado é de fato, surpreendente.

b) A densidade obtida pelo segundo método nos parece excessivamente concentrada. Isto é explicado por uma falha do Analista: antes de se proceder à aplicação do segundo método foram apresentados ao estimador os resultados mais imediatos obtidos na primeira abordagem. Sempre que isto ocorre, a experiência tem demonstrado que o estimador se fixa no valor mais provável, mascarando as estimativas subsequentes que se pretenda obter. A este fenômeno se costuma chamar "anchoring". A falha do Analista serviu, de qualquer modo, para ilustrar o fenômeno.

c) No final da reunião, apresentamos ao estimador um dispositivo análogo ao disco descrito no item 3.2.1. Com esse dispositivo obtivemos mais algumas estimativas por pontos (método 1.b). Contudo, o estimador demonstrou estar bastante influenciado pelas estimativas já feitas e os resultados obtidos concordaram de forma notável com os anteriores. A tentativa serviu, contudo, para observarmos que o processo pode ser aplicado: o estimador assimilou com relativa facilidade as "regras do jogo".

d) Devemos observar que foram obtidas relativamente poucas estimativas

por método. Em uma situação de fato onde se visasse algo além de uma experiência acadêmica, o Analista deveria, provavelmente, "forçar" um pouco mais o estimador.

e) Como último comentário queremos apresentar o resultado verificado "a posteriori". Segundo informações fornecidas pelo departamento de vendas do fabricante do produto J, as vendas na primeira semana após o término da campanha apresentaram um aumento (em relação ao nível anterior à campanha) de 18%.

Essa constatação parece uma evidência auspiciosa de que o procedimento usado pode ser aplicado na prática. Com efeito, a opinião subjetiva do estimador, como já frisamos, indicava uma alta probabilidade de que o aumento de vendas se situasse numa estreita faixa em torno de 20%.

CAPÍTULO VII

EXEMPLO DE ANÁLISE MULTIDIMENSIONAL

Como exemplo de aplicação das técnicas de Análise Multidimensional, apresentamos a seguir um modelo de avaliação de universidades.

Tomou-se como base para a análise o plano global de desenvolvimento (P.G.D.) preparado pelo CEPES (Comissão Especial para Execução do Plano de Melhoramento e Expansão do Ensino Superior), órgão do MEC.

Durante a elaboração do modelo, tomou-se o cuidado de considerar apenas variáveis cujos valores pudessem ser extraidos dos P.G.Ds. das universidades, ou que pudessem ser obtidos, facilmente, através de um questionário adicional.

1º passo - Identificação das Dimensões de Valor

Sem dúvida alguma, são inúmeras as dimensões de valor relevantes para a avaliação de uma universidade. No entanto, para tornar o modelo operacional devemos nos restringir apenas às mais importantes.

Entre todas as possíveis dimensões de valor foram sele

cionadas dez, consideradas as mais representativas.

São elas:

1. Relação entre o número de alunos e o número de professores.

   Neste caso, a nossa dimensão de valor seria representada por

$$x_1 = \frac{NA}{NP}$$

   onde NA = número de alunos matriculados na universidade

   NP = número equivalente de professores em tempo integral

   Este número pode ser calculado por

$$NP = \frac{\sum HS}{N}$$

   onde HS = número de horas semanais dedicadas por cada professor (sendo a somatória feita para todos os professores)

   N = número de horas por semana dedicadas por um professor em regime de tempo integral

2. Relação entre o número de professores com título e o número total de professores

Neste caso a nossa dimensão de valor seria representada por

$$x_2 = \frac{PT}{NP}$$

onde PT = número equivalente de professores com título, em tempo integral

NP = número equivalente de professores em tempo integral

O número de professores com título pode ser calculado através de uma relação aditiva do tipo:

$$PT = \frac{\sum HSM}{N} + a \times \frac{\sum HSD}{N}$$

onde HSM = número de horas semanais dedicadas por cada professor com nível de mestrado ou equivalente (sendo a somatória feita por todos os professores deste nível)

HSD = número de horas semanais dedicadas por cada professor com nível de doutoramento ou equivalente (sendo a somatória feita por todos os professores deste nível)

N = número de horas por semana dedicadas por um professor em regime de tempo integral

a = uma constante (da ordem de 2 ou 3)

3. Relação entre o número de publicações originadas na universidade e o número de professores

A nossa dimensão de valor seria

$$x_3 = \frac{PU}{NP}$$

onde NP = número equivalente de professores em tempo integral

PU = total de publicações originadas na universidade durante um certo período (por exemplo: últimos quatro anos)

Este número pode ser calculado por

$$PU = NPI + b\,NAR + c\,NLP$$

onde NPI = número total de publicações internas

NAR = número total de artigos publicados em revistas especializadas

NLP = número de livros publicados

b, c = constantes de ponderação

4. Relação entre a demanda e a oferta de vagas

A nossa dimensão de valor seria:

$$x_4 = \frac{DV}{OV}$$

onde DV = demanda de vagas, ou seja, número total de candidatos inscritos ao vestibular da universidade

OV = oferta de vagas, ou seja, número total de vagas oferecidas pela universidade nos seus vários cursos.

Em ambos os casos, poderia ser usada uma média dos últimos quatro anos.

5. Relação entre o número de formandos e o número de vagas

A dimensão de valor seria expressa por

$$x_5 = \frac{NF}{OF}$$

onde NF = número total de formandos nos vários cursos da universidade no ano T

OF = número total de vagas oferecidas pela universidade no ano T-X, onde X é a duração média dos cursos.

6. Relação entre o custo total do plano de trabalho e o número de alunos.

Esta dimensão de valor seria expressa por

$$x_6 = \frac{CT}{NA}$$

onde: NA = número total de alunos matriculados na universidade

CT = custo total do plano de trabalho, ou seja, despesa total prevista pela universidade.

7. Relação entre despesa com pessoal docente e despesa total

Esta dimensão de valor seria expressa por

$$x_7 = \frac{DPD}{DT}$$

onde DT = despesa total prevista pela universidade

DPD = despesa com pessoal docente

8. Relação entre despesa com pessoal técnico-administrativo e despesa com pessoal docente

Teríamos neste caso:

$$x_8 = \frac{DPT}{DPD}$$

onde DPT = despesa com pessoal técnico e administrativo

DPD = despesa com pessoal docente

9. Relação entre a dotação orçamentária e o custo total do plano de trabalho.

Neste caso: $x_9 = \frac{DO}{DT}$

onde DO = dotação de que disporá a universidade por orçamento, ou por recursos próprios

DT = despesa total prevista pela universidade

10. Relação entre custo dos projetos de obras e custo total do plano de trabalho.

E teríamos:

$$x_{10} = \frac{DPO}{DT}$$

onde DT = despesa total prevista pela universidade

DPO = despesa prevista com projetos de obras, ou seja, com a ampliação das atuais instalações.

Como se pode notar, as cinco primeiras dimensões de valor procuram refletir aspectos qualitativos do ensino oferecido pela universidade, enquanto as cinco últimas representam aspectos da alocação de verbas feita pela universidade.

2º passo - Levantamento das curvas de utilidade para cada dimensão de valor.

As curvas de utilidade para as várias dimensões de valor foram apenas esboçadas, não havendo em geral preocupação em determinar-se todos os valores numéricos necessários à sua perfeita descrição.

1. x_1 ≡ Relação entre o número de alunos e o número de professores. Se x_1 for muito grande, a qualidade de ensino será necessariamente baixa e a utilidade pequena. Se x_1 for pequeno haverá professores ociosos e portanto, a utilidade também será pequena. Portanto, haverá um ponto tal que, a utilidade será máxima (fig. 7.1).

2. x_2 ≡ Relação entre o número de professores com título e o número total de professores.
Naturalmente, deverá haver em cada universidade, um certo número de professores com nível de mestrado ou doutorado. Logo, pode-se considerar que abaixo de um certo valor de x_2 a utilidade será muito pequena. Pode-se também, considerar que quanto mais professores com títulos houver, maior será a utilidade. (Figura 7.2).

3. x_3 ≡ Relação entre o número de publicações e o número de professores. É importante que os professores publiquem, mas esta atividade não deverá prejudicar as atividades de docência; logo, haverá um valor ótimo para esta relação. (Figura 7.3).

4. x_4 ≡ Relação entre a demanda e a oferta de vagas.
Esta dimensão procura refletir a qualidade de ensino da universidade através da motivação criada nos alunos em potencial. Por isso, considera-se que quanto maior a demanda maior a utilidade.(Figura 7.4).

5. x_5 ≡ Relação entre o número de formandos e o número de vagas.
Espera-se que a maioria dos alunos chegue a se formar (ou seja, que haja uma baixa taxa de repetência e evasão). No entanto, se todos os alunos se formam (taxa de repetência nula) deve-se levantar dúvidas quanto à qualidade do ensino. (Figura 7.5).

6. x_6 ≡ Relação entre a despesa total e o número de alunos.
Um custo por aluno muito baixo se traduz em baixo padrão de ensino. Já um custo por aluno muito alto, pode representar desperdício de fundos (embora o padrão de ensino possa ser elevado). A ambos os casos, deve ser atribuída uma utilidade pequena, devendo haver, portanto, um valor ótimo intermediário. (Fig. 7.6).

7. x_7 ≡ Relação entre despesas com pessoal docente e despesa total.
Ao pessoal docente, deverá ser dedicado uma boa parcela do orçamento, mas não a sua totalidade, naturalmente. (Figura 7.7).

8. x_8 ≡ Relação entre despesas com pessoal técnico-administrativo e despesa com pessoal docente.
Também para esta dimensão de valor deve haver um compromisso ótimo entre excesso e falta de pessoal técnico-administrativo. (Figura 7.8)

9. x_9 ≡ Relação entre dotação orçamentária e despesa total.
Naturalmente, a universidade deverá contar com orçamento ou recursos próprios suficientes para cobrir a maior parte da despe

sa prevista. (Figura 7.9).

10. x_{10} ≡ Relação entre custo de projetos de obras e despesa total.
A universidade deverá ampliar suas instalações a uma taxa equilibrada e proporcional à sua despesa total. (Figura 7.10).

3º passo - Ordenação e atribuição de pesos relativos às dimensões de valor.

A ordem de importância e os pesos relativos das várias dimensões de valor foram atribuídos após discussão do problema em questão.

Após normalização, obtivemos os seguintes pesos relativos:

TABELA 7.1 - PESOS RELATIVOS

Dimensão de valor	Peso relativo
x_1	10
x_2	19
x_3	16
x_4	12
x_5	11
x_6	4
x_7	8
x_8	9
x_9	5
x_{10}	6
Total	100

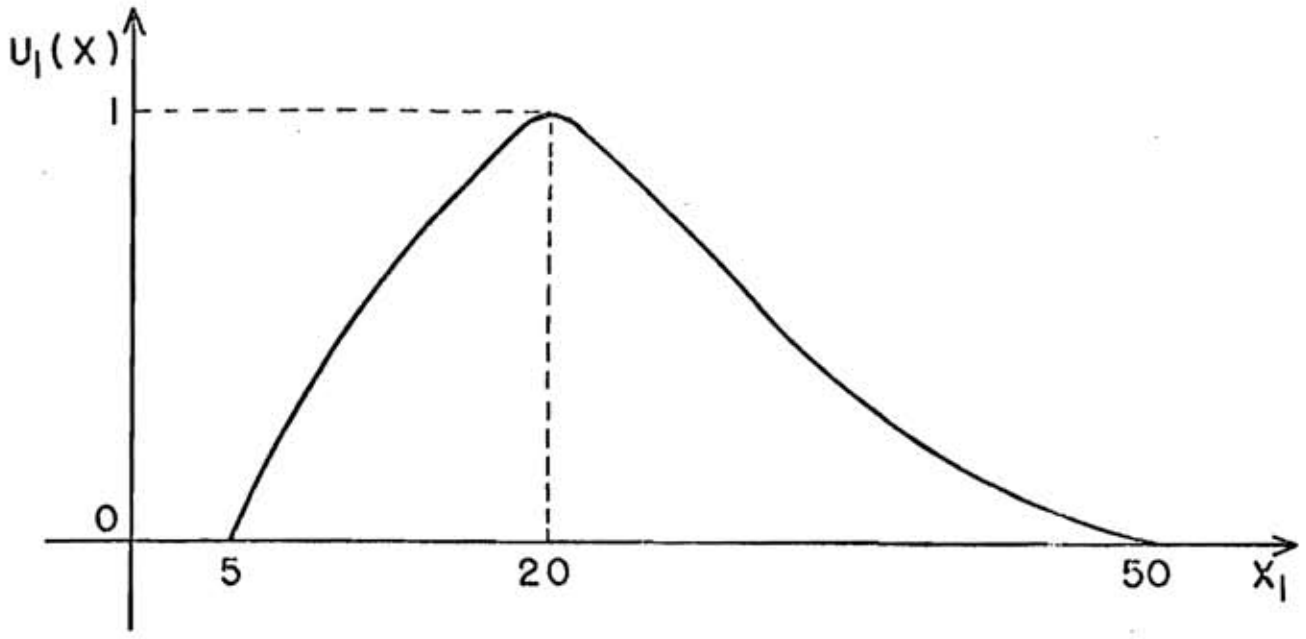


Figura 7.1 - Curva de Utilidade para x_1.

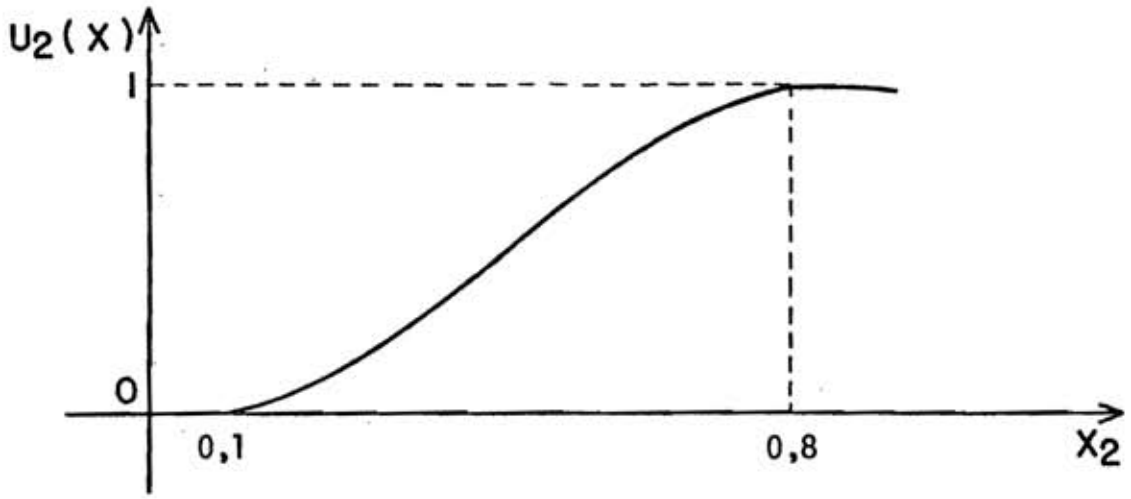


Figura 7.2 - Curva de Utilidade para x_2.

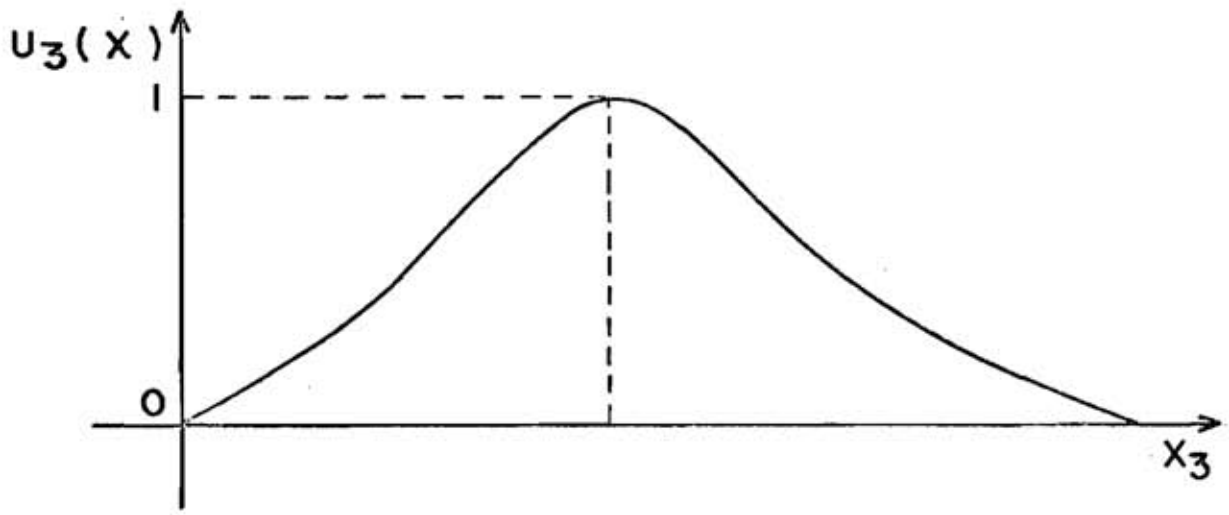


Figura 7.3 - Curva de Utilidade para x_3.

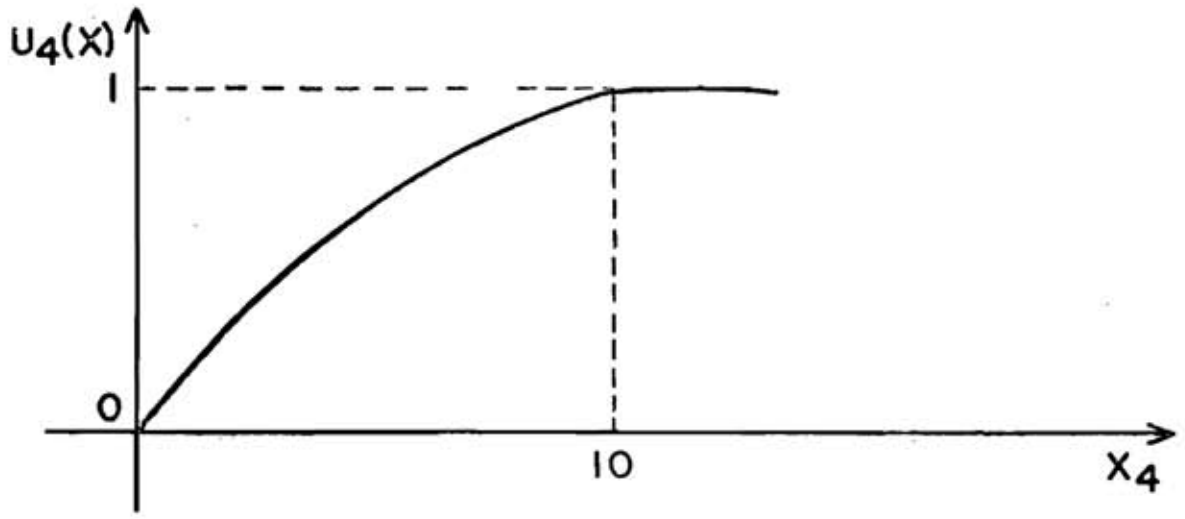


Figura 7.4 - Curva de Utilidade para x_4.

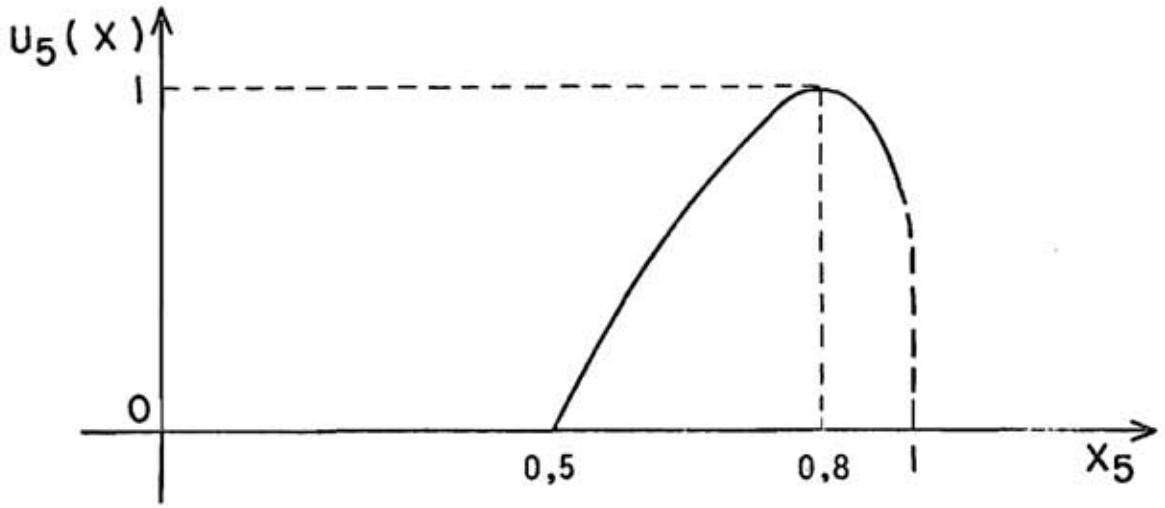


Figura 7.5 - Curva de Utilidade para x_5.

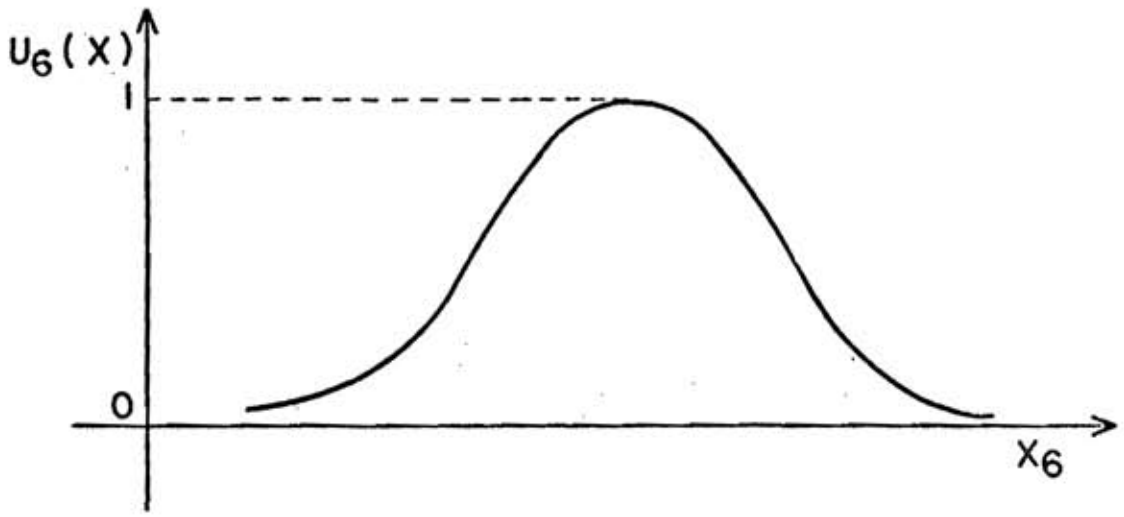


Figura 7.6 - Curva de Utilidade para x_6.

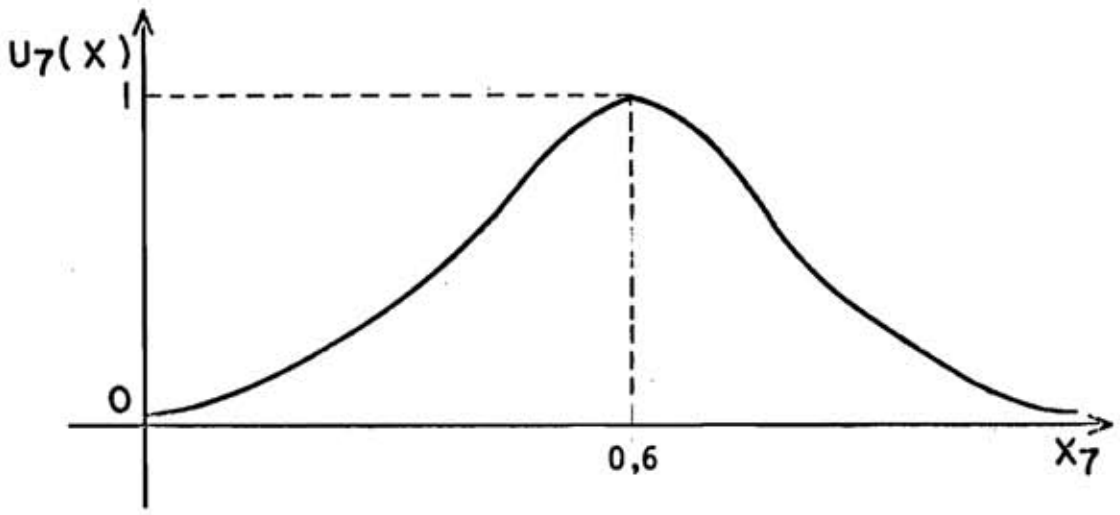


Figura 7.7 - Curva de Utilidade para x_7.

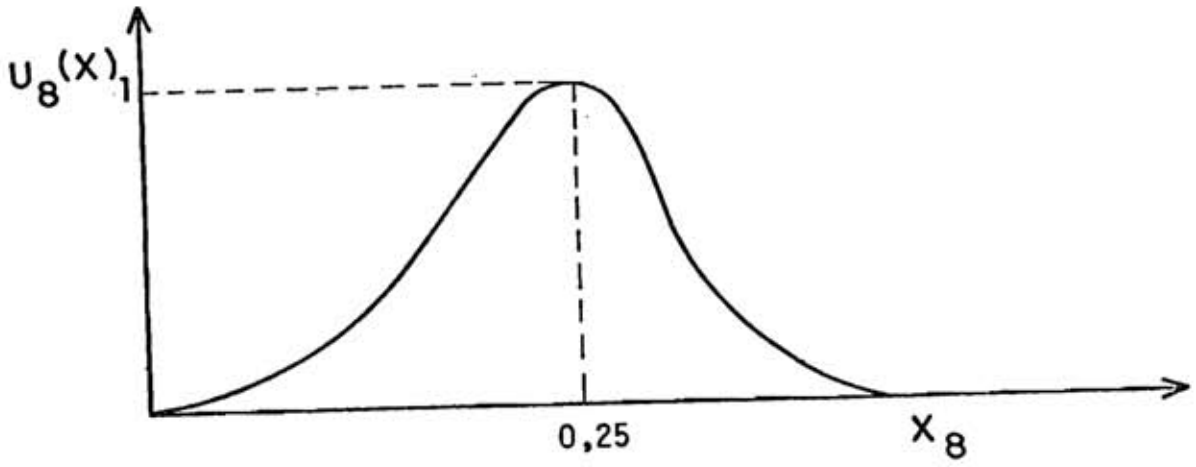


Figura 7.8 - Curva de Utilidade para x_8.

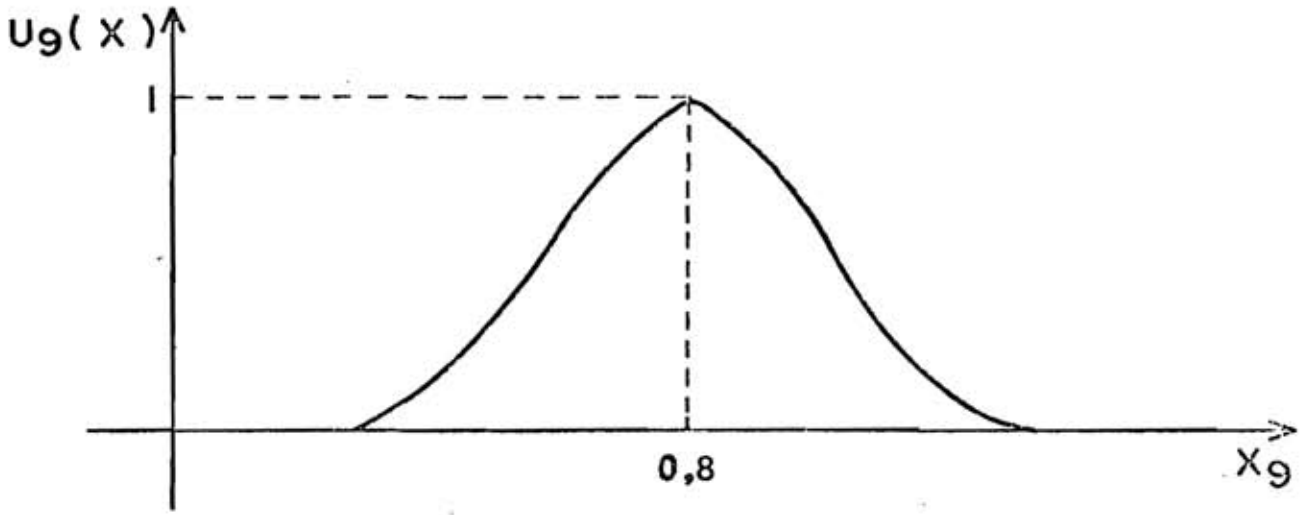


Figura 7.9 - Curva de Utilidade para x_9.

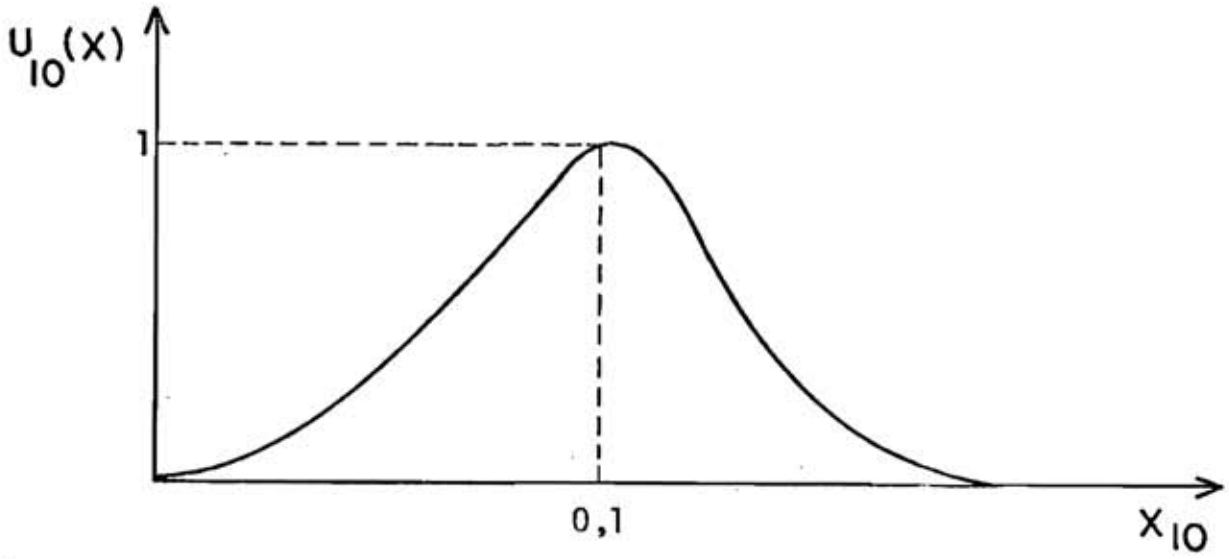


Figura 7.10 - Curva de Utilidade para x_{10}.

Após análise do problema, conclui-se que um modelo aditivo presta-se para a agregação das várias dimensões de valor num índice único.

Assim, sendo $U_i(x)$ a funçao utilidade da i-ésima dimensão de valor e P_i o peso relativo da mesma, temos:

$$K_j = \sum_{i=1}^{10} P_i \; U_i(x_{ij})$$

sendo K_j o índice de avaliação da universidade j

e x_{ij} o valor da dimensão i para a universidade j

ou seja

$$K_j = 10 \times U_1(x_{1j}) + 19 \times U_2(x_{2j}) + 16 \times U_3(x_{3j})$$

$$+ 12 \times U_4(x_{4j}) + 11 \times U_5(x_{5j}) + 4 \times U_6(x_{6j})$$

$$+ 8 \times U_7(x_{7j}) + 9 \times U_8(x_{8j}) + 5 \times U_9(x_{9j})$$

$$+ 6 \times U_{10}(x_{10j})$$

Para obter-se o valor de K para uma universidade, basta entrar com os valores correspondentes das dimensões de valor nas curvas de utilidade e substituir os valores obtidos na função acima.

CONCLUSÕES

Através dos estudos apresentados, conclui-se ser viável o emprego de técnicas de análise multidimensional para a construção de um modelo de avaliação de universidades.

A finalidade do presente trabalho é apenas exemplificar o estudo destas técnicas, apresentando a estrutura geral do modelo.

Naturalmente, no caso de sua implementação para uso na prática, caberia ao decisor: a seleção das dimensões de valor, a determinação de seus pesos relativos e o levantamento das curvas de utilidade.

APÊNDICE

DESCRIÇÃO DO PROGRAMA DE PROCESSAMENTO DE ÁRVORES

1.- INTRODUÇÃO

2.- DESCRIÇÃO DO PROGRAMA

3.- FORMATO DOS DADOS DE ENTRADA

4.- FORMATO DAS SAÍDAS

5.- PROBLEMA EXEMPLO

6.- LISTAGEM DO PROGRAMA

1 - INTRODUÇÃO

Apresentamos, a seguir, um programa para processamento de árvores de decisão.

A hipótese básica feita a respeito da função utilidade do decisor é que ela seja exponencial (o coeficiente de aversão ao risco é um dos parâmetros de entrada).

O programa foi escrito em FORTRAN para um computador B-6700, e usa aproximadamente 5.000 palavras de memória, sendo portanto facilmente adaptável a outros computadores menores. Sua elaboração deve-se a Eric Jan Roorda que estagiou com o grupo no período de julho a dezembro de 1973.

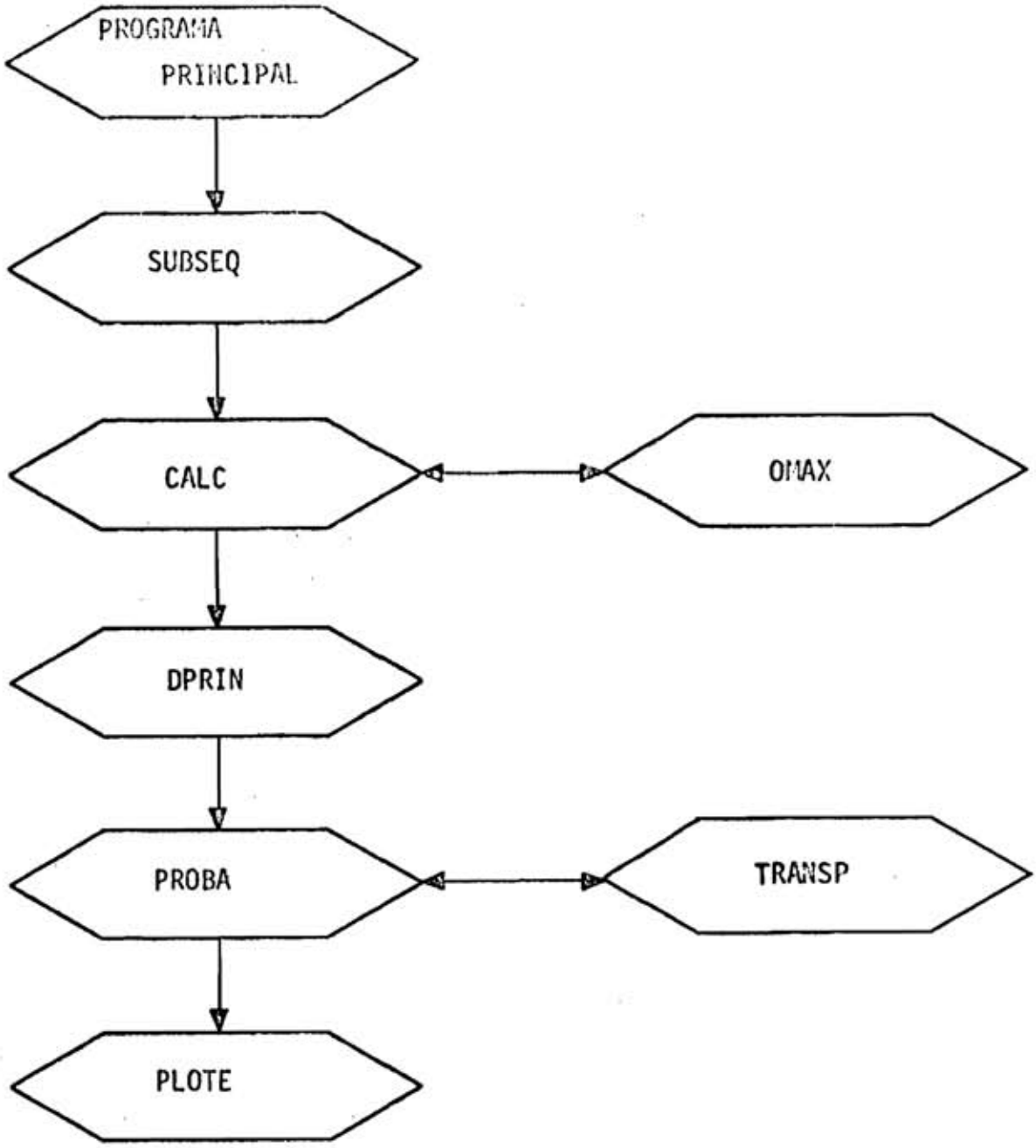


Figura A.1 - Diagrama de Blocos do Programa de Processamento de Árvores de Decisão.

2 - DESCRIÇÃO DO PROGRAMA

2.1 - PROGRAMA PRINCIPAL

No programa principal é feita a leitura dos dados, inicialização dos parâmetros e são chamadas as demais sub-rotinas.

2.2 - SUBROTINA SUBSEQ

Esta subrotina estabelece uma ligação entre os nós e seus subsequentes dentro das árvores. Ela enumera os nós desde o primeiro até os últimos (nós terminais).

2.3 - SUBROTINA CALC

Esta subrotina faz o cálculo do equivalente certo de um nó, dados os valores dos nós subsequentes e os parâmetros acarretados pelas decisões. O método de cálculo é função do tipo de nó que pode ser de decisão ou probabilístico.

2.4 - SUBROTINA DPRIN

É a subrotina de impressão dos resultados. Dá como saída uma listagem dos nós, o valor do coeficiente de aversão ao risco e uma tabela de decisões.

2.5 - SUBROTINA PROBA

Faz o cálculo da probabilidade de ocorrência de cada nó terminal (nós extremos da árvore de decisão).

2.6 - SUBROTINA PLOTE

Esta subrotina plota um gráfico das probabilidades dos possíveis resultados atingíveis.

2.7 - SUBROTINA TRANSP

Esta subrotina faz o transporte de todos os pagamentos resultantes de decisões, para os nós terminais da árvore.

2.8 - SUBROTINA OMAX

Calcula o máximo entre os valores dos equivalentes certos dos subsequentes de um nó.

NOTA: O programa foi escrito com uma limitação de 50 nós (incluindo os terminais) e no máximo dez nós subsequentes para cada nó. Se for necessário um programa para a resolução de uma árvore com mais nós ou mais ramificações por nó que o máximo previsto, basta mudar o dimensionamento nos cartões COMMON e ainda nos cartões FORMAT e os limites de alguns comandos DO.

3 - FORMATO DOS DADOS DE ENTRADA

No primeiro cartão consta o valor do coeficiente de aversão ao risco do decisor (lembremos que a função utilidade do decisor é considerada exponencial). Este valor deve ser declarado em formato real e pode ocupar as colunas 1 a 15.

Nos demais cartões consta a descrição da árvore de decisão, que é feita do seguinte modo: para cada nó (exceto os nós terminais) deve ser dado um nome e associado um subconjunto de cartões do seguinte formato:

1º cartão:

Colunas 1-6: Nome do nó, em formato alfanumérico.
(1ª letra na 1ª coluna).

Colunas 7-12: Tipo do nó, em formato alfanumérico.
(1ª letra na coluna 7).

Decisão : DEC

Probabilístico: PRØB.

Colunas 13-14: Número de nós em que se ramifica o nó que está sendo descrito (em formato inteiro).

Colunas 15-20: Nome do primeiro nó subsequente, em formato alfanumérico
(1ª letra na coluna 15).

Colunas 21-23: Probabilidade de ocorrência deste 1º nó subsequente (em formato real). Se o nó que está sendo descrito é do tipo decisão deixar estas colunas em branco para todos os subsequentes.

Colunas 24-33: Pagamento na decisão do 1º subsequente (ou valor de recompensa se for um nó terminal). Em formato real.

Os dados das colunas 15-20, 21-23 e 24-33 devem ser repetidos para todos os nós subsequentes ao que está sendo descrito, reservando-se do 2º subsequente em diante um cartão para cada subsequente, com os dados colocados respectivamente nas colunas 1-6, 7-9 e 10-19. Assim, completamos a descrição de um nó.

Esta descrição deve ser repetida para todos os nós da árvore (exceto os terminais).

4 - FORMATO DAS SAÍDAS

A primeira página de saída nos apresenta uma listagem de todos os nós e seus subsequentes.

Na segunda página temos para cada nó de decisão e seus subsequentes o valor do pagamento na decisão, o equivalente certo e a utilidade correspondente. Para cada nó probabilístico temos a utilidade a ele associada e para cada nó terminal o valor da recompensa correspondente.

As duas últimas páginas nos dão a distribuição dos resultados esperados ou seja a probabilidade de ganhar até uma certa quantia.

5 - PROBLEMA EXEMPLO

O engenheiro encarregado de pesquisas de uma fábrica de produtos químicos apresentou ao gerente da fábrica uma proposta de produção de um novo tipo de detergente. Analisando a situação, o gerente chegou às seguintes conclusões:

Para desenvolvimento do produto, serão gastos Cr$ 1 000,00. Como se trata de uma fórmula nova, não há certeza quanto aos resultados mas o engenheiro de pesquisas garante que com oitenta por cento de chance será obtido um produto competitivo com os já existentes no mercado.

O lançamento poderá ser de âmbito nacional mas poderá também ser feito um lançamento em escala regional e, dependendo dos resultados, sua extensão para uma escala nacional. Os custos e resultados de cada alternativa estão mostrados na árvore de decisão (Fig. A.2).

De posse destes dados, qual a estratégia que, uma vez adotada pelo gerente, lhe dá uma utilidade esperada máxima?

Na figura A.3 temos uma listagem dos cartões de dados necessários para a descrição da árvore de decisões.

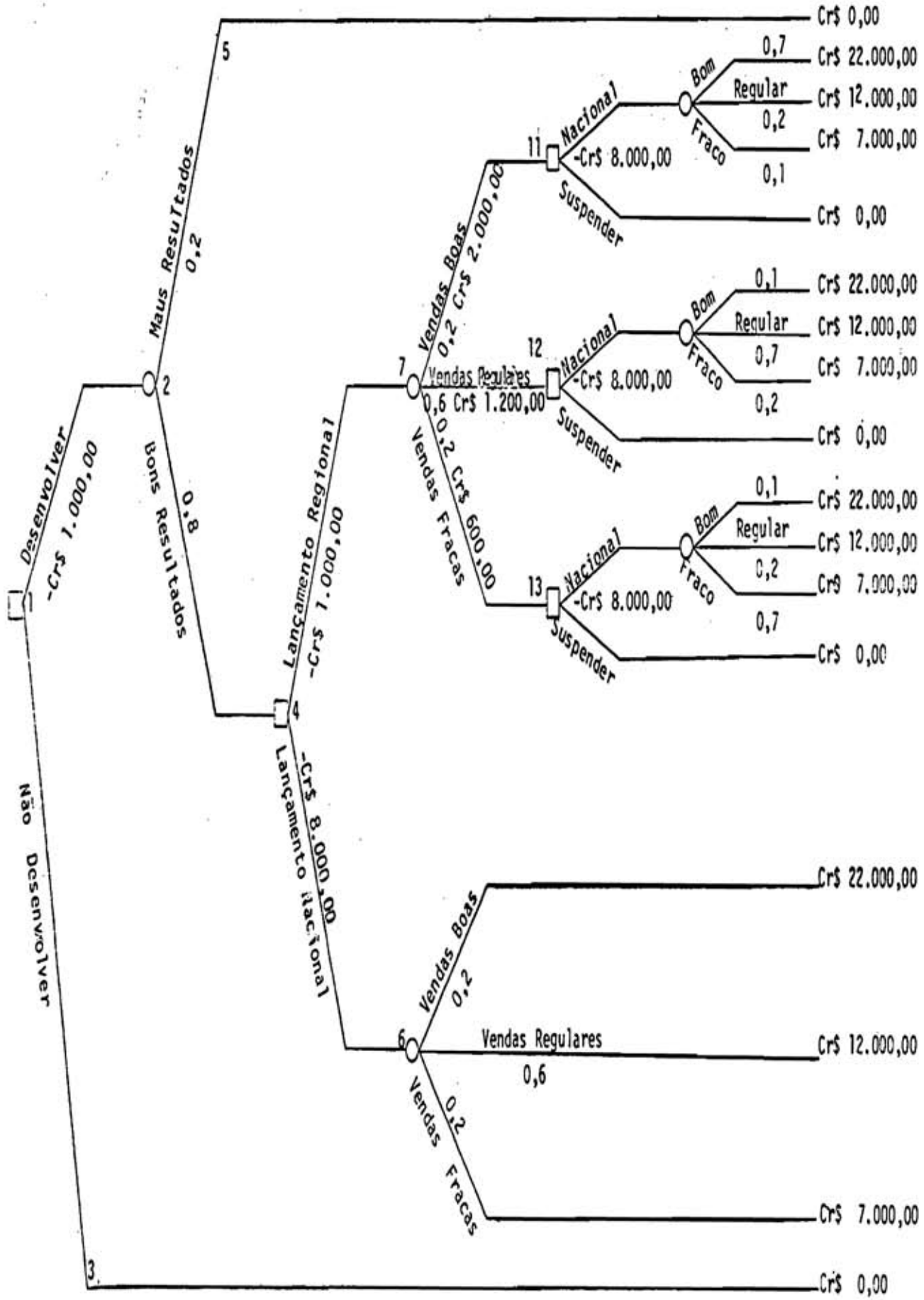


Figura A.2 - Árvore de Decisões para o Problema Exemplo.

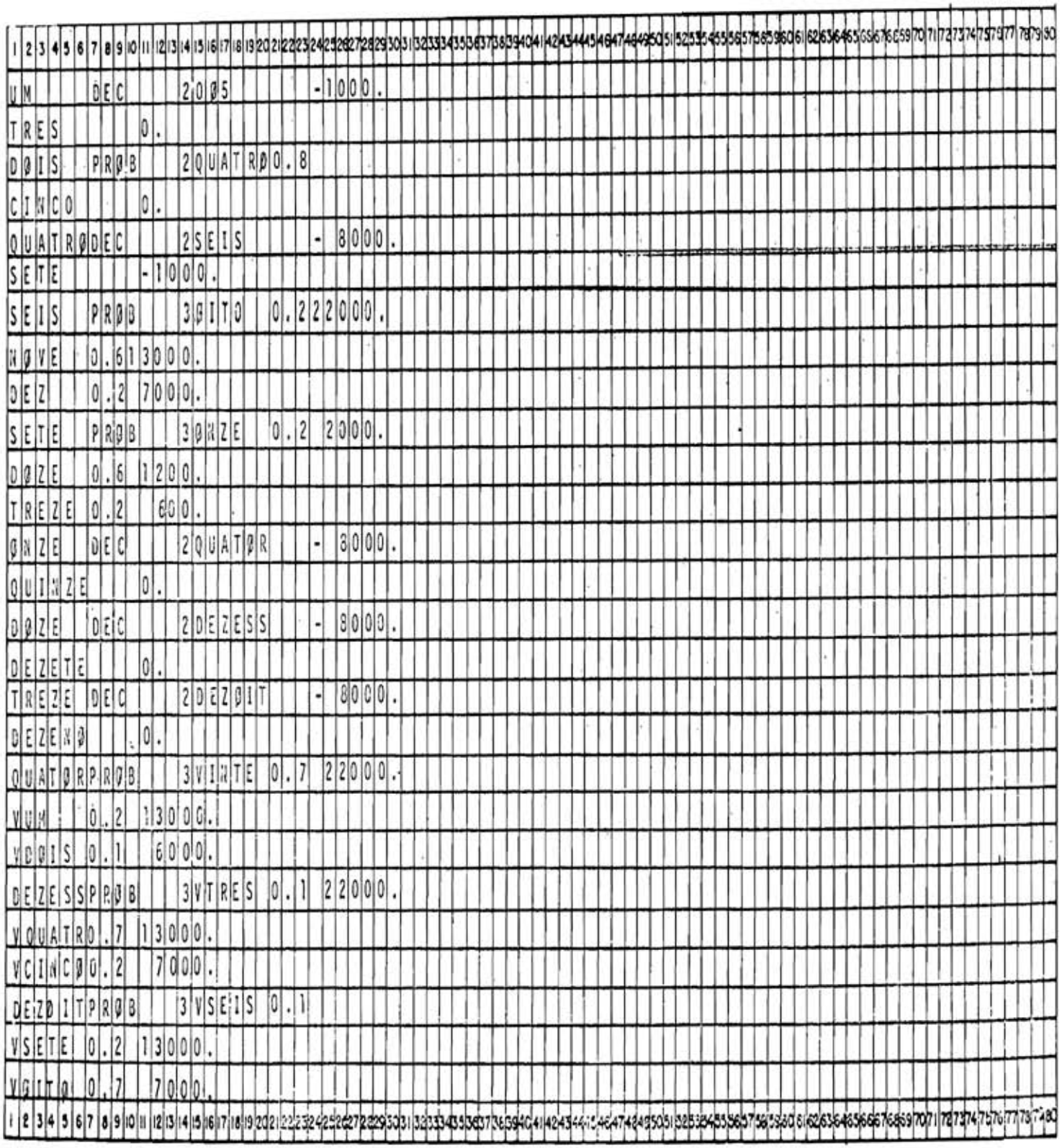


Figura A.3 - Dados de Entrada para o Problema Exemplo.

Utilizando o programa de processamento de árvores de decisão, obtemos uma curva do equivalente certo da decisão ótima em função do coeficiente de aversão ao risco do decisor (Figura A.4).

Nota-se que se o coeficiente de aversão ao risco é maior do que 0,00065 a decisão ótima é não desenvolver o novo produto. Já para um coeficiente de aversão ao risco menor que 0,00065, a estratégia ótima é desenvolver o produto e, em caso de bons resultados, fazer o lançamento em nível nacional.

Nas páginas seguintes, encontramos um exemplo das saídas do programa (para o problema exemplo, usando-se um coeficiente de aversão ao risco igual a 0,0005). E, por fim, uma listagem do programa de processamento de árvores de decisão.

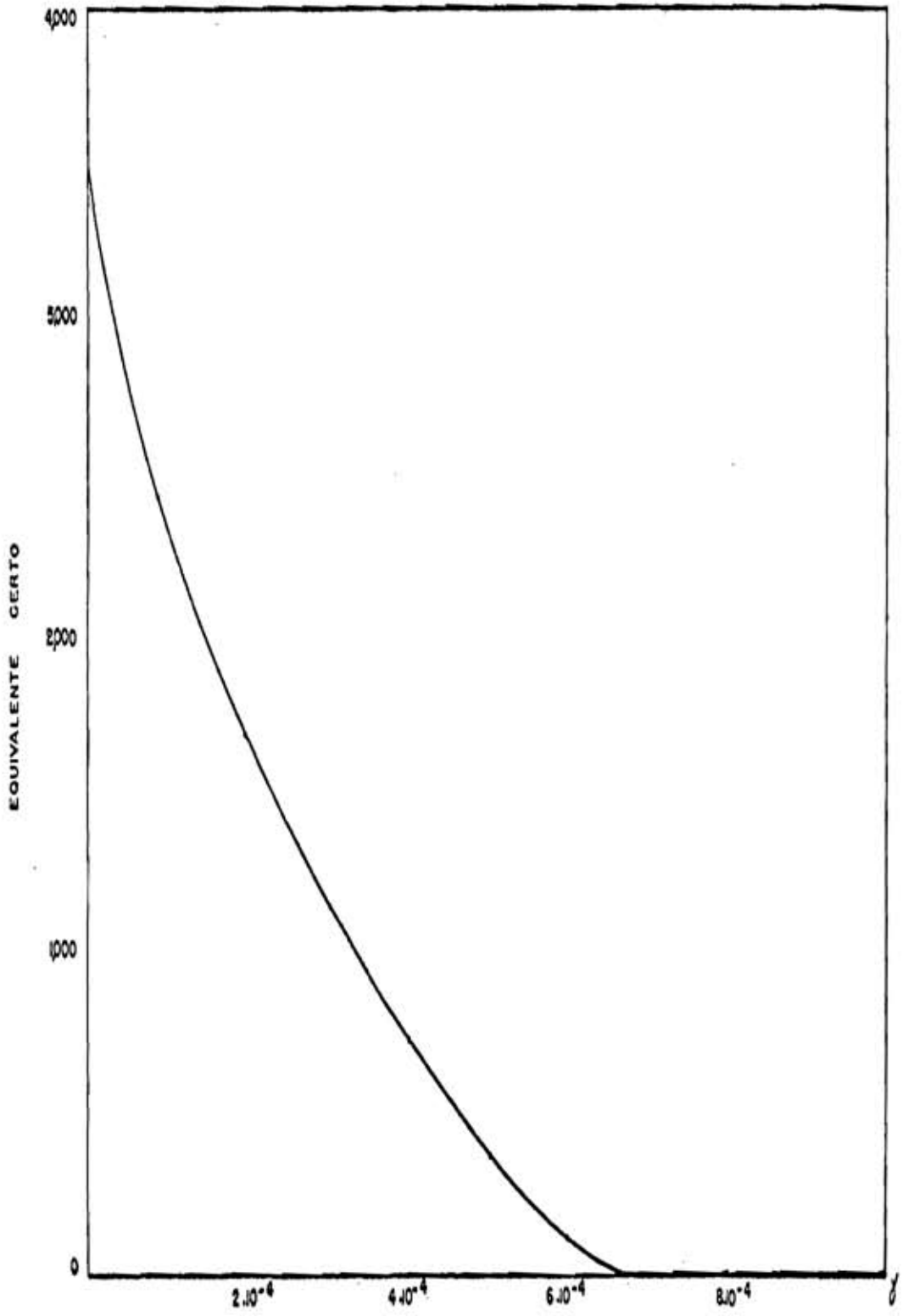


Figura A.4 - Equivalente Certo em Função do Coeficiente de Aversão ao Risco.

LISTAGEM DE TODOS OS NOS GAMA VALE ,SOOUDE*03

1	UM	DEC	1	2	DUIS
			2	12	THES
2	DUIS	PRIM	1	3	QUATRO
			2	13	CINCO
3	QUATRO	DEC	1	4	SEIS
			2	5	SETE
4	SEIS	PRIM	1	14	OITO
			2	15	NOVE
			3	16	DEZ
5	SETE	PRIM	1	6	ONZE
			2	7	DOZE
			3	8	TREZE
6	ONZE	DEC	1	9	QUATOR
			2	17	DEZESE
7	DOZE	DEC	1	10	DEZESS
			2	18	DEZOITO
8	TREZE	DEC	1	11	DEZOIT
			2	19	DEZENO
9	QUATOR	PRIM	1	20	VINTE
			2	21	VINT
			3	22	VINTS
10	DEZESS	PRIM	1	23	VINLS
			2	24	VINTAN
			3	25	VEINCO
11	DEZOIT	PRIM	1	26	VSEIS
			2	27	VSETE
			3	28	VOITO

NO	NOME DO NO	TIPO	DECISAO	ALTERNATIVA	NO SUCESSOR	NOME SUCESSOR	UTILIDADE	EQUIVALENTE CERTO	PAGAMENTO NA DECISAO
1	UM	DEC	1	1	2	DOIS	.13730E+04	.37297E+03	*.10000E+04
			0	2	12	TRES	0.	0.	0.
2	DOIS	PROB					.13730E+04		
3	QUATRO	DEC	1	1	4	SEIS	.99395E+04	.19395E+04	*.80000E+04
			0	2	5	SETE	.14460E+04	.44604E+03	*.10000E+04
4	SEIS	PROB					.99395E+04		
5	SETE	PROB					.14460E+04		
6	ONZE	DEC	1	1	9	QUATOR	.10483E+05	.24835E+04	*.80000E+04
			0	2	17	QUINZE	0.	0.	0.
7	DOZE	DEC	1	1	10	DEZESS	.99395E+04	.19395E+04	*.80000E+04
			0	2	18	DEZETE	0.	0.	0.
8	TREZE	DEC	0	1	11	DEZOIT	.76849E+04	*.31506E+03	*.80000E+04
			1	2	19	DEZENO	0.	0.	0.
9	QUATO	PROB					.10483E+05		
10	DEZES	PROB					.99395E+04		
11	DEZOIT	PROB					.76849E+04		
12	TRES	TERM					0.		
13	CINCO	TERM					0.		
14	OITO	TERM					.22000E+05		
15	NOVE	TERM					.13000E+05		
16	DEZ	TERM					.70000E+04		
17	QUINZE	TERM					0.		
18	DEZETE	TERM					0.		
19	DEZENO	TERM					0.		
20	VINTE	TERM					.22000E+05		
21	VUM	TERM					.13000E+05		
22	VDOIS	TERM					.70000E+04		
23	VTRES	TERM					.22000E+05		
24	VQUATR	TERM					.13000E+05		
25	VCINCO	TERM					.70000E+04		
26	VSEIS	TERM					.22000E+05		
27	VSETE	TERM					.13000E+05		
28	VOITO	TERM					.70000E+04		

```
LISTAGEM DAS PROBABILIDADES

     -2000.00000          0.16000

     -1000.00000          0.20000

      4000.00000          0.48000

     13000.00000          0.16000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000

         0.00000          0.00000
```

GRAFICO DAS PROBABILIDADES PLOTADO

PROB	RETORNO	RETORNO	,0	,1	,2	,3	,4	,5	,6	,7	,8	,9	1,
0,1600	*,2000E+04	*,2000E+04	*										
		*,1700E+04	*										
		*,1400E+04	*										
		*,1100E+04	*										
0,3600	*,1000E+04	*,8000E+03	*										
		*,5000E+03	*										
		*,2000E+03	*										
		,1000E+03	*										
		,4000E+03	*										
		,7000E+03	*										
		,1000E+04	*										
		,1300E+04	*										
		,1600E+04	*										
		,1900E+04	*										
		,2200E+04	*										
		,2500E+04	*										
		,2900E+04	*										
		,3100E+04	*										
		,3400E+04	*										
		,3700E+04	*										
0,8400	,4000E+04	,4000E+04	*										
		,4300E+04	*										
		,4600E+04	*										
		,4900E+04	*										
		,5200E+04	*										
		,5500E+04	*										
		,5800E+04	*										
		,6100E+04	*										
		,6400E+04	*										
		,6700E+04	*										
		,7000E+04	*										
		,7300E+04	*										
		,7600E+04	*										
		,7900E+04	*										
		,8200E+04	*										
		,8500E+04	*										
		,8800E+04	*										
		,9100E+04	*										
		,9400E+04	*										
		,9700E+04	*										
		,1000E+05	*										
		,1030E+05	*										
		,1060E+05	*										
		,1090E+05	*										
		,1120E+05	*										
		,1150E+05	*										
		,1180E+05	*										
		,1210E+05	*										
		,1240E+05	*										
		,1270E+05	*										
1,0000	,1300E+05	,1300E+05	*										*
1,0000	0,	,1330E+05	*										*
1,0000	0,	,1360E+05	*										*
1,0000	0,	,1390E+05	*										*
1,0000	0,	,1420E+05	*										*

```
B6700/B7700  F O R T R A N   C O M P I L A T I O N   M A R K   2.5,000

FILE  5=DADOS;UNIT READER
      COMMON N1,N2
      COMMON NS(50),NRESOL(50),NPOS(50),NDEC(50,5),NEG(50,10),NANT(50)
     1, NOPC(50)
      COMMON GAMA
      COMMON UNOME(50),SUBSQ(50,10),TIPO(50),PRBSQ(50,10),UTIL(50,10),
     1 U(50),UTI(50),UTILA(50,10),UA(50),PRB(50),U1(50),PRB1(50)
  100 FORMAT(F15.6)
  110 FORMAT(2A6,I2,  (A6,F3.1,F10.2))
      DO 10 I=1,50
      NRESOL(I)=0
      U(I)=.0
      NS(I)=0
      U1(I)=.0
      PRB1(I)=.0
      NANT(I)=0
      PRB(I)=.0
      NOPC(I)=0
      UTI(I)=.0
      DO 10 J=1,10
      UTIL(I,J)=.0
      PRBSQ(I,J)=.0
      UTILA(I,J)=.0
      U(I)=.0
      UA(I)=.0
   10 CONTINUE
      I=1
      READ(5,100)GAMA
   20 READ(5,110,END=30) UNOME(I),TIPO(I),NS(I),(SUBSQ(I,J),PRBSQ(I,J),
     1   UTIL(I,J),J=1,NS(I))
      I=I+1
      GO TO 20
   30 N1=I-1
      N2=I-1
      DO 35 K=1,N2
      DO 35 J=1,NS(K)
   35 UTILA(K,J)=UTIL(K,J)
      CALL SUBSEQ
      CALL CALC
      CALL UPRIN
      CALL PROBA
   40 CONTINUE
      STOP
      END
```

```
      SUBROUTINE SUBSEQ
      COMMON N1,N2
      COMMON NS(50),NRESOL(50),NPOS(50),NDEC(50,5),NEG(50,10),NANT(50)
     1, NOPC(50)
      COMMON GAMA
      COMMON ONOME(50),SUBSQ(50,10),TIPO(50),PRBSQ(50,10),UTIL(50,10),
     1 U(50),UTI(50),UTILA(50,10),UA(50),PRA(50),U1(50),PRB1(50)
      DATA DEC,PROB,TERM/6HDEC    ,6HPROB   ,6HTERM  /
      DO 20 K=1,N2
      DO 20 J=1,NS(K)
      DO 10 I=1,N2
      IF (ONOME(I).EQ.SUBSQ(K,J))GO TO 30
   10 CONTINUE
      N1=N1+1
      ONOME(N1)= SUBSQ(K,J)
      NANT(N1)=K
      PRB(N1)=PRBSQ(K,J)
      NEG(K,J)=N1
      U(N1)= UTIL(K,J)
      UTIL(K,J)=.0
      NRESOL(N1)=1
      TIPO(N1)=TERM
      GO TO 20
   30 NEG(K,J)=I
      NANT(I)=K
      PRB(I)=PRBSQ(K,J)
   20 CONTINUE
      RETURN
      END
```

S

```
      SUBROUTINE CALC
      COMMON N1,N2
      COMMON NS(50),NRESOL(50),NPOS(50),NDEC(50,5),NEG(50,10),NANT(50)
     1, NOPC(50)
      COMMON GAMA
      COMMON ONOME(50),SUBSO(50,10),TIPO(50),PRBSO(50,10),UTIL(50,10),
     1 U(50),UTI(50),UTILA(50,10),UA(50),PRB(50),U1(50),PRB1(50)
      DATA DEC,PROB,TERM/6HDEC   ,6HPROB  ,6HTERM  /
    5 DO 80 K=1,N2
      DO 10 J=1,NS(K)
      IF(NRESOL(NEG(K,J)))80,80,10
   10 CONTINUE
      IF(TIPO(K).EQ.DEC) GO TO 20
      IF(TIPO(K).EQ.PROB) GO TO 30
      GO TO 80
   20 CONTINUE
      CALL UMAX(K)
      GO TO 75
   30 IF(GAMA)60,40,60
   40 DO 50 J=1,NS(K)
   50 U(K)=U(K)+U(NEG(K,J))*PRBSO(K,J)
      GO TO 75
   60 DO 70 J=1,NS(K)
   70 UTI(K)=UTI(K)+PRBSO(K,J)*EXP(-GAMA*U(NEG(K,J)))
      U(K)=-ALOG(UTI(K))/GAMA
   75 NRESOL(K)=1
      DO 76 L=1,NS(K)
   76 NRESOL(NEG(K,L))=0
   80 CONTINUE
      IF(NRESOL(1))5 ,5.,90
   90 RETURN
      END
```

SEG

FORMAT

```
      SUBROUTINE UPRIN
      COMMON N1,N2
      COMMON NS(50),NRESOL(50),NPOS(50),NDEC(50,5),NEG(50,10),NANT(50)
     1, NOPC(50)
      COMMON GAMA
      COMMON ONOME(50),SUBSQ(50,10),TIPO(50),PRBSQ(50,10),UTIL(50,10),
     1 U(50),UTI(50),UTILA(50,10),UA(50),PRB(50),U1(50),PRB1(50)
  410 FORMAT(1H1," LISTAGEM DE TODOS OS NOS",5X,"GAMA VALE",E13.5)
  510 FORMAT(1H0,I3,5X,A6,5X,A6,10(T30,I3,5X,I4,5X,A6/))
  310 FORMAT(1H1,"NO  NOME      TIPO      DECI ALTER  NO  NOME        UTILI
     1          EQUIVALENTE   PAGAMENTO"
     1,    /,   "     DO NO                SAO  NATIVA     SUCESSOR   DADE
     3           CERTO          NA DECISAO")
  110 FORMAT(1H0,I3,2X,A6,4X,A6,5X,10(T25,I3,I4,I5,2X,A6,5X,3E13.5/))
  210 FORMAT(1H0,I3,2X,A6,4X,A6,27X,E13.5)
  610 FORMAT(1H0,10E13.5)
      DATA DEC,PROB,TERM/6HDEC    ,6HPROB  ,6HTERM  /
      PRINT 410 ,GAMA

      DO 20 I=1,N2
      PRINT 510, I,ONOME(I),TIPO(I),(J,NEG(I,J),SUBSQ(I,J),J=1,NS(I))
   20 CONTINUE
      PRINT 310
      DO 10 I=1,N1
      IF(TIPO(I).EQ.DEC)  GO TO  5
      PRINT 210, I,ONOME(I),TIPO(I),U(I)
      GO TO 10
    5 DO 6 J=1,NS(I)
    6 UA(NEG(I,J))=U(NEG(I,J))+UTILA(I,J)
      PRINT 110,I,ONOME(I),TIPO(I),(NOPC(NEG(I,J)),J,NEG(I,J),SUBSQ(I,J)
     1, U(NEG(I,J)),UA(NEG(I,J)),UTILA(I,J),J=1,NS(I))
   10 CONTINUE
      RETURN
      END
```

FORMAT
SEG

```
      SUBROUTINE PROBA
      COMMON N1,N2
      COMMON NS(50),NRESOL(50),NPOS(50),NDEC(50,5),NEG(50,10),NANT(50)
     1, NOPC(50)
      COMMON GAMA
      COMMON DNOME(50),SUBSO(50,10),TIPO(50),PRBSO(50,10),UTIL(50,10),
     1 U(50),UTI(50),UTILA(50,10),UA(50),PRB(50),U1(50),PRB1(50)
      DATA TERM/6HTERM  /
      DO 50 I=1,N1
      PRO=1.0
      IF (TIPO(I).EQ.TERM) GO TO 10
      GO TO 50
   10 IF(PRB(I))11,50,11
   11 LL=I
      LLL=I
   15 PRO=PRO*PRB(LLL)
      LLL=NANT(LLL)
      IF(LLL-1)40,40,15
   40 PRB(LL)=PRO
   50 CONTINUE
      CALL TRANSP
      DO 120 J=1,N1
      DO 120 I=1,N1
      IF (TIPO(I).EQ.TERM) GO TO 90
      GO TO 120
   90 IF (TIPO(J).EQ.TERM) GO TO 100
      GO TO 120
  100 IF (U(I)-U(J))120,120,110
  110 A=U(I)
      U(I)=U(J)
      U(J)=A
      B=PRB(I)
      PRB(I)=PRB(J)
      PRB(J)=B
  120 CONTINUE
      K=1
      DO 130 I=1,N1
      IF (TIPO(I).NE.TERM) GO TO 130
      IF(PRB(I))125,130,125
  125 U1(K)=U(I)
      PRB1(K)=PRB(I)
      K=K+1
  130 CONTINUE
      KKK=K-1
      CALL PLOTE(KKK)
      RETURN
      END
```

```
      SUBROUTINE PLOTE (KKK)
      COMMON N1,N2
      COMMON NS(50),NRESOL(50),NPOS(50),NDEC(50,5),NEG(50,10),NANT(50)
     1, NUPC(50)
      COMMON GAMA
      COMMON ORDNE(50),SUBSQ(50,10),TIPO(50),PRBSQ(50,10),UTIL(50,10),
     1 U(50),UTI(50),UTILA(50,10),UA(50),PRB(50),U1(50),PRB1(50)
      DIMENSION CHAR(80),CHARA(120)
  710 FORMAT(1H1,"LISTAGEM DAS PROBABILIDADES")
  810 FORMAT(1H0, 2F16.5)
  910 FORMAT(1H1,"GRAFICO DAS PROBABILIDADES PLOTADO")
 1010 FORMAT (1H ,T27,E11.5,T40,1H1,80A1)
 1110 FORMAT(1H ,T2,F6.4,T13,E11.5,T27,F11.5,T40,1H1,80A1)
 1210 FORMAT(1H0,T3,"PRO3",T15,"RETORNO",T29,"RETURNO",T39,".0",T47,".1
     *",T55,".2",T63,".3",T71,".4",T79,".5",T87,".6",T95,".7",T103,".8",
     *T111,".9",T119,"1.")
 1310 FORMAT(1H ,120A1)
      DATA BLANK ,PLOT,MENOS/1H ,1H*,1H-/
      DO 30 LL=1,120
   30 CHARA(LL)=MENOS
      PRINT 710
      PRINT 810 ,(U1(I),PRB1(I),I=1,N1)
      L=2
      KK=2
      ACRES=(U1(KKK)-U1(1))/50.
      X=U1(1)
      Y=PRB1(1)
      PRINT 910
      PRINT 1210
      PRINT 1310  ,CHARA
      DO 100 J=1,55
      DO 90 K=1,80
   90 CHAR(K)=BLANK
      I=Y*80.+.5
      DO 85 II=1,I-1
   85 CHAR(II)=MENOS
      CHAR(I)=PLOT
      GO TO (91,92)  , KK
   91 PRINT 1010 ,X,CHAR
      GO TO 93
   92 PRINT 1110 , Y,U1(L-1) ,X,CHAR
   93 X=X+ACRES
      IF(U1(L)-X)94,94,95
   94 Y=Y+PRB1(L)
      KK=2
   98 L=L+1
      GO TO 100
   95 KK=1
  100 CONTINUE
      RETURN
      END
```

FORMAT
SEG

```
      SUBROUTINE OMAX(K)
      COMMON N1,N2
      COMMON NS(50),NRESOL(50),NPOS(50),NDEC(50,5),NEG(50,10),NANT(50)
     1, NUPC(50)
      COMMON GAMA
      COMMON ONOME(50),SUBSQ(50,10),TIPO(50),PRBSQ(50,10),UTIL(50,10),
     1 U(50),UTI(50),UTILA(50,10),UA(50),PRA(50),UI(50),PRBI(50)
      U(K)=U(NEG(K,1))+UTIL(K,1)
      NPOS(K)=1
      NDEC(K,1)=1
      DO 30 J=2,NS(K)
      IF(U(K)-U(NEG(K,J))-UTIL(K,J))10,20,30
   10 U(K)=U(NEG(K,J))+UTIL(K,J)
      NPOS(K)=1
      NDEC(K,1)=J
      GO TO 30
   20 NPOS(K)=NPOS(K)+1
      NDEC(K,NPOS(K))=J
   30 CONTINUE
      NUPC(NEG(K,NDEC(K,NPOS(K))))=1
      PRB (NEG(K,NDEC(K,NPOS(K))))=1,0
      RETURN
      END
```

SEG

```
      SUBROUTINE TRANSP
      COMMON N1,N2
      COMMON NS(50),NRESOL(50),NPUS(50),NDEC(50,5),NEG(50,10),NANT(50)
     1, NOPC(50)
      COMMON GAMA
      COMMON ONOME(50),SUBSQ(50,10),TIPO(50),PRBSQ(50,10),UTIL(50,10),
     1 U(50),UTI(50),UTILA(50,10),UA(50),PRB(50),U1(50),PRB1(50)
      DATA TERM/6HTERM  /
    5 NNN=0
      DO 30 I=1,N2
      DO 30 J=1,NS(I)
      IF(TIPO(NEG(I,J)).EQ.TERM) GO TO 30
    6 IF (UTIL(I,J)) 10,30,10
   10 L=NEG(I,J)
      DO 20 K=1,NS(L)
      UTIL(L,K)=UTIL(L,K)+UTIL(I,J)
   20 UA(NEG(L,K))=-UTIL(L,K)
      UTIL(I,J)=.0
      NNN=1
   30 CONTINUE
      IF(NNN-1)40, 5,40
   40 CONTINUE
      DO 50 I=1,N2
      DO 50 J=1,NS(I)
      U(NEG(I,J))=UTIL(I,J) +U(NEG(I,J))
   50 UTIL(I,J)=.0
      RETURN
      END
```

SEG

BIBLIOGRAFIA

1. Beach, L.R. - *Accuracy and Consistency in The Revision of Subjective Probabilities*, IEEE Transactions on Human Factors in Eletronics, março de 1965, vol. HFE-7, nº 1, pp 29-36.

2. Bierman Jr., H., Charles P. Bonini e Warren H. Hausman - *Quantitative Analysis for Business Decisions*, Richard D. Irwin, 3ª ed., 1969.

3. Bohnert, H.G. - *The Logical Structure of the Utility Concept*, em Decision Processes, editado por R.M. Thrall, C.H. Coombs e R.L. Davis, John Wiley, 2ª ed., 1957.

4. Cost-Effectiveness - *The Economic Evaluation of Engineered Systems*, editado por J. Morley English, John Wiley, 1968

5. Debreu, Gerard - *Representation of a Preference Ordering by a Numerical Function*, em Decision Processes, editado por R.M. Thrall, C.H. Coombs e R.L. Davis, John Wiley, 2ª ed., 1957.

6. Edwards, Ward - *Social Utilities*, Symposium on Decision and Risk Analysis, A.S.E.E.-A.I.I.E., Annapolis, Maryland, Junho, 1971.

7. Edwards, Ward - *The Theory of Decision Making*, em Decision Making, editado por Ward Edwards e Amos Tversky, Penguin Modern Psychology, 1967.

8. Edwards, Ward - *Behavioral Decision Theory*, em Decision Making, editado por Ward Edwards e Amos Tversky, Penguin Modern Psychology, 1967.

9. Edwards, W., L.D. Philips e W.L.Hays - *Conservation in Complex Probabilistic Inference*, IEEE Transactions on Human Factors in Electronics, março, 1966, vol.HFE-7, nº 1, pp. 7-18.

10. Edwards, W., L.D.Philips, W.L.Hays e Barbara C. Goodman - *Probabilistic Information Processing Systems: Design and Evaluation,* IEEE Transactions on Systems Science and Cybernetics, vol. SSC-4, nº 3, setembro, 1968, pp. 248-264.

11. Edwards, E., Harold Lindman e Leonard J.Savage - *Bayesian Statistical Inference for Psychological Research,* Phychological Review, 1963, nº 70, pp. 193-242.

12. Feller, William - *An Introduction to Probability Theory and Its Applications,* John Wiley, 1966.

13. Fisher, Gregory W. - *Multi-dimensional Value Assessment for Decision Making,* Engineering Psychology Lab., University of Michigan, Technical Report, junho 1972.

14. Fishburn, Peter C. - *Decision and Value Theory,* John Wiley, 1964.

15. Fishburn, P.C. - *Utility Theory,* em Management Science, vol. 14, nº 5, janeiro, 1968, pp. 335-351.

16. Fishburn, P.C. - *Von Neumann-Morgenstern - Utility Functions on Two Attributes,* Operations Research, vol. 22, nº 1, janeiro/fevereiro, 1974, pp. 35-44.

17. Forester, John - *Statistical Selection of Business Strategies, Irwin Series in Quantitative Analysis for Business,* Richard D. Irwin, 1969.

18. Freeman, Harold - *Introduction to Statistical Inference,* Addison-Wesley, 1963.

19. Howard, R.A. - *The Foundations of Decision Analysis,* IEEE Transactions on Systems Science and Cybernetics, vol.SSC-4,nº 3, Set.,1968, pp.211-219.

20. INPE-PBDCT - *Proposta para Sua Revisão e Avaliação*, INPE-384/PR/013, Setembro de 1973.

21. Jedamus, Paul e Robert Frame - *Business Decision Theory*, McGraw-Hill, 1969.

22. Kaplan, R.J. e Newman, J.R. - *Studies in Probabilitic Information Processing*, IEEE Transactions on Human Factors in Electronics, março, 1969, vol. HFE-7, nº 1, pp. 49-62.

23. Keeney, Ralph L. - *Utility Independence and Preferences for Multialtributed Consequences*, em Operations Research, Julho/agosto, 1971, vol. 19, nº 4, pp. 875-892.

24. Keeney, R.L. - *Utility functions for Multiattributed Consequences*, em Management Science, vol. 18, nº 5, janeiro, 1972, pp. 276-287.

25. Keeney, R.L. - *Multiplicative Utility Functions*, em Operations Research, vol. 22, nº 1, janeiro/fevereiro, 1974, pp. 22-33.

26. Keeney, R. L. e Richard de Neufville - *Multiattributed Preference Analysis for Transportation Systems Evaluation*, em Transportation Research, pp. 63-75, vol. 7, nº 1, março, 1973.

27. Kelley, J.L. - *General Topology*, Von Nostrand Reinhold, 1955.

28. Laplace, P.S. - *A Philosophical Essays on Probability*, tradução de F.W. Fruscatt e F.C. Emory, Dover Publications, 1951.

29. Luce, R.D. e H. Raiffa, - *Games and decisions: Introduction and Critical Survey*, John Willey, 1967.

30. Mac Crimmon, K.R. - *Decision Making Among Multiple-Atribute Alternatives: A Survey and Consolidated Approach, Rand Memorandum AM-4823, ARPA,* Dezembro, 1968.

31. Meyer, P.L. - *Probabilidade, Aplicações à Estatística,* Tradução de Ruy de C. B. Lourenço Filho, Ao Livro Técnico, 1970.

32. Miller, James R. - *Professional Decision - Making: A Procedure for Evaluating Complex Alternatives,* Praeger Publishers, 1970.

33. Mood, Alexander M. e Franklin A. Graybill - *Introduction to the Theory of Statistics,* McGraw-Hill, 1963.

34. Mosteller F. e P. Nogee - *An Experimental Measurement of Utility,* em Decision Making, editado por Ward Edwards e Amos Tversky, Penguin Modern Psychology, 1967.

35. Neumann, J. Von e Oskar Morgenstern - *Theory of Games and Economic Behavior,* John Wiley, 1967.

36. Parzen, Emmanuel - *Modern Probability Theory and Its Applications,* John Wiley, 1960.

37. Peressini. Anthony L. - *Ordered Topological Vector Spaces,* Harper e Row, 1967.

38. Peters, William S. e George W. Summers - Statistical Analysis for Business Decisions, Prentice Hall, 1964.

39. Peterson, C.R. e Phillips, L.D. - *Revision of Continuous Subjective Probability Distributions,* IEEE Transactions on Human Factors in Eletronics, março, 1966, vol. HFE-7, nº 1, pp

40. Raiffa, Howard - *Decision Analysis: Introductory Lectures on Choices under Uncertainty*, Addison-Wesley, 1970.

41. Raiffa, H. e Robert Schlaifer - *Applied Statistical Decision Theory*, The M.I.T., Press, 1968.

42. *Readings in Mathematical Psychology*, editado por R. Duncan Luce, Robert R. Bush e Eugene Galanter, John Wiley, 1965.

43. Samuelson, Paul A. - *Probability and the Attempts to Measure Utility* em The Collected Scientific Papers of Paul A. Samuelson, The M.I.T. Press, 1970, pp. 117-126.

44. Samuelson, Paul A. - *Probability, Utility and the Independence Axiom*, em The Collected Scientific Papers of Paul A. Samuelson, The M.I.T. Press, 1970, pp. 137-145.

45. Savage, Leonard J. - *The Foundations of Statistics*, Dover Publications, 1972.

46. Schlaifer, R.O. - H. Raiffa e J. W. Pratt - *Introducation to Statistical Theory*, McGraw-Hill, 1965.

47. Schlaifer, R.O. - *Probability and Statistics for Business Decisions*, McGraw-Hill, 1959.

48. Slovic, Paul - *Value as a Determiner of Subjective Probability*, IEEE Transactions on Human Factors in Electronics, março, 1966, vol. HFE-7, nº 1, pp. 22-28.

49. Shuchman, Abe - *Scientific Decision Making in Business*, Holt, Rinehart and Winston, 1963.

50. Shum, D.A. - I.L. Goldstein e J.F. Southard - *Research on a Simulated Bayesian Information - Processing System*, IEEE Transactions on Human Factors in Eletronics, março, 1966; vol. HFE-7, nº 1, pp. 37-48.

51. Tribus, Myron - *Rational Descriptions, Decisions and Designs*, Pergamon Press, 1969.

52. Tversky, Amos - *Additivity, Utility and Subjective Probability*, em Decision Making, editado por Ward Edwards e Amos Tversky, Penguin Modern Psychology, 1967.

53. Weiss, Lionel - *Statistical Decision Theory*, McGraw-Hill, 1961.

54. Wilks, Samuel S. - *Mathematical Statistics*, John Wiley, 1962.

55. Yamane, Taro - *Statistics*, Harper e Row, 1964.