

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ANDRÉ LEITE WANDERLEY

**A EDUCAÇÃO DA FUNÇÃO UTILIDADE
MULTIATRIBUTO**

VIRTUS IMPAVIDA

RECIFE, DEZEMBRO DE 2008.

ANDRÉ LEITE WANDERLEY

**A EDUCAÇÃO DA FUNÇÃO UTILIDADE
MULTIATRIBUTO**

Tese submetida ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco como parte dos requisitos para obtenção do grau de **Doutor em Engenharia Elétrica**

ORIENTADOR: PROF. FERNANDO MENEZES CAMPELLO DE SOUZA, PH.D.

Recife, Dezembro de 2008.

©André Leite Wanderley, 2008

Wanderley, André Leite

**A educação da função utilidade multiatributo / André Leite Wanderley . – Recife : O Autor, 2008.
89 folhas : il., fig., tab.**

Tese (doutorado) – Universidade Federal de Pernambuco. CTG. Engenharia Elétrica, 2008.

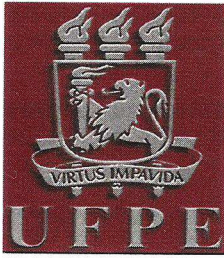
Inclui bibliografia

1. Teoria da utilidade – Educação multiatributo. 2. Preferências – Ordem linear – Educação das preferências. I. Título.

**621.3
621.3**

**CDU (2.ed.)
CDD (22.ed.)**

**UFPE
BC - 2008 - 146**



Universidade Federal de Pernambuco
Pós-Graduação em Engenharia Elétrica

PARECER DA COMISSÃO EXAMINADORA DE DEFESA DE TESE DE DOUTORADO

ANDRÉ LEITE WANDERLEY

TÍTULO

“A EDUCAÇÃO DA FUNÇÃO UTILIDADE MULTIATRIBUTO”

A comissão examinadora composta pelos professores: FERNANDO MENEZES CAMPELLO DE SOUZA, DES/UFPE, HÉLIO MAGALHÃES DE OLIVEIRA, DES/UFPE, VALDEMAR CARDOSO DA ROCHA JÚNIOR, DES/UFPE, FERNANDA MARIA RIBEIRO DE ALENCAR, DES/UFPE e, FRANCISCO DE SOUSA RAMOS, DCE/UFPE sob a presidência do primeiro, consideram o candidato **ANDRÉ LEITE WANDERLEY APROVADO.**

Recife, 29 de dezembro de 2008.

EDUARDO FONTANA
Coordenador do PPGE

FERNANDO MENEZES CAMPELLO DE SOUZA
Orientador e Membro Titular Interno

FERNANDA MARIA RIBEIRO DE ALENCAR
Membro Titular Externo

HÉLIO MAGALHÃES DE OLIVEIRA
Membro Titular Interno

FRANCISCO DE SOUSA RAMOS
Membro Titular Externo

VALDEMAR CARDOSO DA ROCHA JÚNIOR
Membro Titular Interno

Απειρον

AGRADECIMENTOS

Muitas pessoas foram deveras importante nesses anos de pós-graduação, de modo que gostaria de expressar meu sincero agradecimento a todos que participaram deste processo. Em particular:

A minha Família, pelo apoio e incentivo constante.

Ao Prof. Fernando Menezes Campello de Souza, orientador da dissertação, pela sua dedicação. O empenho e energia que ele coloca na universidade e na formação de seus alunos é exemplo a ser seguido. Muito obrigado professor.

Aos amigos Eric Arantes, Elton Bandeira, Eric Bouton, Márcio Lima, Alane Alves, Diogo de Carvalho, Diana Yomali, Katarina Santiago, Savana Cecília, Pedro Leon, Luis H. de Santana, Elton Bandeira, Bruno de Andrade, Peron Rios e Maria Fernanda, por tudo que me ensinaram nesses vários anos de convivência. Sem eles, nada disso teria sido possível, ou faria sentido.

Muito dos artigos usados neste trabalho foram obtidos graças a Eric Bouton, Marcus Muller, Luciano Brandão, Marcel Jar e Fernanda Maria, a eles minhas desculpas pelas demandas constantes e muito obrigado.

À Leda Clair, que muito ajudou na adequação desta tese às normas da ABNT.

Finalmente, gostaria de agradecer ao CNPQ e ao Programa de Pós-Graduação em Engenharia Elétrica.

Universidade Federal de Pernambuco

30 de Dezembro de 2008

A. L. W.

The effort of the economist is to see, to picture the interplay of economic elements. The more clearly cut these elements appear in his vision, the better; the more elements he can grasp and hold in mind at once, the better. The economic world is a misty region. The first explorers used unaided vision. Mathematics is a lantern by which what before was dimly visible now looms up in firm, bold outlines. The old phantasmagoria disappear. We see better. We also see further.

— **Irving Fisher**

Sans les mathématiques on ne pénètre point au fond de la philosophie. Sans la philosophie on ne pénètre point au fond des mathématiques. Sans les deux on ne pénètre au fond de rien.

— **Leibniz**

Resumo da Tese apresentada à UFPE como parte dos requisitos necessários para a obtenção do grau de Doutor em Engenharia Elétrica

A EDUCAÇÃO DA FUNÇÃO UTILIDADE MULTIATRIBUTO

André Leite Wanderley

Dezembro/2008

Orientador: Prof. Fernando Menezes Campello de Souza, Ph.D.

Área de Concentração: Sistemas

Palavras-chaves: Teoria da decisão, função utilidade, aversão ao risco, ordem.

Número de páginas: 89

SABE-SE que a Teoria da Decisão fornece uma abordagem prescritiva para a representação e manuseio de sistemas complexos de preferência. Contudo, a dificuldade em sua aplicação, reflexo essencialmente do alto custo da educação das preferências, tem limitado a sua aplicação. Estuda-se as implicações na estrutura da função utilidade de restrições nos *tradeoffs* geradas pela escolha de uma ordem linear particular. Mostra-se a relação deste conceito com a matriz de aversão ao risco de Duncan e explora-se esta ligação na escolha da forma funcional da utilidade. Apresenta-se um protocolo para a estimação da função utilidade multiatributo de von Neumann-Morgenstern, elemento fundamental para a Teoria da Decisão de A. Wald. Grande parte dos problemas de decisão é caracterizada pela existência de múltiplos atributos. Surge então o conceito de *tradeoffs*, de fundamental importância na caracterização das preferências do decisor, principalmente quando os atributos são contraditórios. Apresenta-se também um algoritmo para a educação de uma ordem linear sobre as conseqüências, neste caso multidimensionais, que representa as preferências do decisor e a qual será utilizada nos métodos clássicos de educação monoatributo de funções utilidades. Fornece-se uma ferramenta para a educação de funções utilidades multiatributo, parcimoniosa em relação ao tamanho do questionário e às hipóteses sobre a estrutura de preferências do decisor, facilitando assim o uso da teoria em epígrafe, tal ferramenta foi desenvolvida a partir de resultados na extensão de ordens parciais e ordenação topológica.

Abstract of Thesis presented to UFPE as a partial fulfillment of the requirements for the degree of Doctor in Electrical Engineering

A PROTOCOL FOR ELICITING THE MULTIATTRIBUTE UTILITY FUNCTION

André Leite Wanderley

December/2008

Supervisor: Prof. Fernando Menezes Campello de Souza, Ph.D.

Area of Concentration: Systems

Keywords:

Number of pages: 89

IT is known that Decision Theory provides a prescriptive approach for the representation and handling of complex preference systems. However, the difficulty in its application, essentially a reflex of the preferences elicitation high costs, has limited its application. One studies the implications the restrictions in the tradeoffs generated by the choice of a particular linear order have in the utility function structure. The relation is shown between this concept and the Duncan risk aversion matrix, and this relation is explored in choosing the utility function analytical expression. A protocol for the von Neumann-Morgenstern multiattribute utility function estimation is presented. This function is a fundamental element in the A. Wald Decision Theory. Most decision making problems are characterized by the existence of multiple attributes. Emerges then the concept of tradeoffs, of fundamental importance in the characterization of the decision maker preferences, mainly when the attributes are contradictory. An algorithm is also presented for the elicitation of a linear order over the consequences, in this case multidimensional, which represents the decision maker preferences and which will be used in the monoattribute utility functions classical elicitation methods. One provides a tool for the multiattribute utility functions elicitation which is parsimonious with respect to the protocol questionnaire size and the hypotheses concerning the decision maker preference structure, facilitating thus the use of the aforementioned theory. Such a tool has been developed using results from extensions of partial orders and topological sorting.

LISTA DE FIGURAS

2.1	Maneiras de apresentar uma loteria $[c_i, \alpha, c_j]$ para o decisor.	27
4.1	Faixas adjacentes com tons de cinza similares, contudo as faixas extremas com tons opostos.	52
4.2	Exemplo de partição do plano de dois atributos \mathcal{C}_1 e \mathcal{C}_2	55
4.3	Exemplo de partição do plano de dois atributos e relações entre c_{34} e demais consequências.	56
4.4	Diagrama de Hasse de $\mathcal{C} = (\mathcal{C}, \preceq)$ quando $n = 2, n_1 = 6$ e $n_2 = 5$	57
4.5	Exemplo de partição de um espaço de consequência com dois atributos. (a) Plano representando as consequências. (b) Diagrama de Hasse da ordem parcial induzida. (c) Extensões lineares da ordem parcial induzida.	58
4.6	Exemplo de partição de um espaço de consequências com dois atributos. (a) Plano representando as consequências. (b) Diagrama de Hasse da ordem parcial induzida.	62
4.7	Diagrama de Hasse do conjunto parcialmente ordenado $\mathfrak{J}(\mathcal{C})$	63
4.8	Cadeias maximais de $\mathfrak{J}(\mathcal{C})$ ou extensões lineares de \mathcal{C}	63
4.9	Exemplo de partição do plano de dois atributos mais simples possível. (a) Plano representando as consequências. (b) Diagrama de Hasse da ordem parcial induzida.	64
4.10	Exemplo de partição do plano de dois atributos mais simples possível.	64
4.11	Ordem induzida para $n = 2, n_1 = 3$ e $n_2 = 3$	65
4.12	Aplicação do algoritmo de ordenação topológica de \mathcal{C} com $n = 2, n_1 = 2$ e $n_2 = 3$. Visualização a partir do diagrama de Hasse. Os vértices em destaque são elementos minimais.	68
4.13	Matriz de incidência e diagrama de Hasse para ordem induzida. com $n = 2, n_1 = 3$ e $n_2 = 3$. Os vértices em destaque representam elementos minimais. Na matriz de incidência, isto significa que $\alpha(j) = 0$	69
4.14	Matriz de incidência e diagrama de Hasse para etapa 2 do processo de educação da ordem. Note-se que o decisor precisa escolher entre dois elementos minimais.	70
4.15	Matriz de incidência e diagrama de Hasse para etapa 3 do processo de educação da ordem. Novamente, o decisor precisa escolher entre dois elementos minimais.	70
4.16	Matriz de incidência e diagrama de Hasse para ordem induzida. Etapas 4, 5 e 6.	71
4.17	Matriz de incidência e diagrama de Hasse para ordem induzida. Etapas 7, 8 e 9.	72

4.18	Utilidade eduzida para o Decisor 1. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.	76
4.19	Utilidade eduzida para o Decisor 2. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.	76
4.20	Utilidade eduzida para o Decisor 3. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.	77
4.21	Utilidade eduzida para o Decisor 4. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.	77
4.22	Utilidade eduzida para o Decisor 5. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.	78
4.23	Utilidade eduzida para o Decisor 6. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.	78
4.24	Utilidade biatributo do decisor 1, com as dimensões separadas.	79
4.25	Utilidade biatributo do decisor 2, com as dimensões separadas.	80

LISTA DE TABELAS

2.1	Métodos para comparação de loterias.	26
4.1	Exemplos de limites para o número de perguntas em função de η	66
4.2	Aplicação do algoritmo de ordenação topológica de \mathcal{C} com $n = 2$, $n_1 = 2$ e $n_2 = 3$	68
4.3	Experimento trabalho <i>versus</i> tempo livre.	74

SUMÁRIO

I	INTRODUÇÃO	13
1.1	A Educação da Função Utilidade	14
1.2	Objetivos	16
1.2.1	Objetivos gerais	16
1.2.2	Objetivos específicos	16
1.3	Justificativa	17
1.4	Resultados	17
1.5	Estrutura da tese	17
2	DECISÕES E UTILIDADES	18
2.1	Teoria da utilidade	20
2.2	Construção da função utilidade	23
2.2.1	Métodos para a educação da função utilidade	25
2.3	Outras considerações	30
2.3.1	<i>Terminus technicus</i>	30
2.3.2	O “paradoxo” de Allais	31
2.4	A função de aversão ao risco de Arrow-Pratt	32
3	A AVERSÃO AO RISCO E A FUNÇÃO UTILIDADE	33
3.1	Aversão ao risco multidimensional	34
3.2	Prêmio pelo risco multidimensional	36
3.3	Comparando atitudes de aversão ao risco	38
3.4	Operações que preservam a aversão ao risco multidimensional	40
3.5	Extensão para o caso n -dimensional	42
3.6	Independência em utilidade e a aversão ao risco multidimensional	43
3.7	Prêmio em probabilidade multidimensional	45
3.8	Matriz de aversão ao risco	46
4	EDUCAÇÃO DE ORDENS LINEARES	49
4.1	Relações de preferência	51
4.2	Estrutura do conjunto de consequências	53
4.3	Ordem parcial	55

4.4	Extensão linear de uma ordem parcial	58
4.5	Ordenação topológica	67
4.6	Exemplo de educação: tempo livre <i>versus</i> salário	73
4.7	Considerações finais	81
5	CONCLUSÕES, COMENTÁRIOS E SUGESTÕES	82
	REFERÊNCIAS	84

CAPÍTULO I

INTRODUÇÃO

“Amicus Plato — amicus Aristoteles — magis amica veritas.”

— Isaac Newton

O PROBLEMA de como indivíduos formam suas preferências por alternativas (conseqüências incertas, ou mais precisamente, distribuições de probabilidade sobre possíveis conseqüências) é de interesse em áreas como psicologia, *marketing* e teoria da decisão, embora em cada uma delas o foco difira entre descritivo, investigativo e prescritivo [1].

O psicólogo almeja entender como o indivíduo processa informações; distinguir o processo que melhor representa o processamento cognitivo de informações. Sendo assim, a modelagem de preferências tem a tarefa de diferenciar um modelo do outro. O foco é **investigativo**.

O estudioso do *marketing* deseja entender como o indivíduo (consumidor) toma suas decisões de compra e como intervir nesse processo. A idéia, nesse contexto, é modificar a percepção da incerteza do consumidor sobre as conseqüências de suas compras, por meio de estratégias de propaganda ou mesmo modificações no produto. No caso de uma concessionária de automóveis usados, *e.g.*, o consumo de combustível, a disponibilidade, a confiabilidade, *inter alia*, são incertezas para o consumidor. Um modelo de comportamento para o consumidor de carros semi-novos é então necessário para o direcionamento inteligente das estratégias de *marketing* da empresa. O foco é **descritivo**.

Já no caso de teoria da decisão, o foco é **prescritivo** (ou **normativo**). O indivíduo ou agente decisor¹ está diante de um problema importante e complexo. O objetivo da modelagem da preferência é estimar a função utilidade do decisor sobre as possíveis conseqüências, de

¹De agora em diante denota-se **indivíduo decisor** ou **agente decisor** apenas por decisor. Há quem defenda o termo *decididor*.

modo a ajudá-lo a escolher uma alternativa no seu espaço de ações que seja coerente com um funcional objetivo.

A teoria da decisão, como estudada hoje, teve seu início na metade do século XX nos trabalhos de Abraham Wald [2] e usando resultados da teoria da utilidade de von Neumann e Morgenstern [3]. Desde então, teve contribuições das mais diversas áreas do conhecimento. Hoje considerada como assunto acadêmico *per se*, a teoria da decisão é foco de pesquisas em economia, estatística, ciências políticas e sociais e filosofia. Os cientistas políticos se interessam, por exemplo, em regras para eleições e outros aspectos de decisões em grupo. Os psicólogos se interessam em estudar o comportamento de decisores durante o processo de decisão. Filósofos focam principalmente a questão da racionalidade [4, 5].

1.1 A Educação da Função Utilidade

A teoria da decisão tem um construto central: a função utilidade do decisor. Sem esta função não se pode aplicar a teoria. Sem se medir as preferências dos indivíduos, não tem sentido falar em aplicação da teoria da decisão.

O processo de educação da função utilidade requer a elaboração de um protocolo. A maioria das críticas que são feitas à teoria da decisão recaem na apresentação de casos práticos nos quais indivíduos violam algum dos axiomas da teoria da utilidade, ou mesmo o resultado principal do valor esperado da função utilidade. Alguns experimentos são preparados como se fossem “armadilhas”, explorando mecanismos psicológicos (ancoragem, ilusão da certeza, superestimação ou subestimação de probabilidades, etc.), de forma a levar os indivíduos a cometerem erros relativos ao previsto pela teoria da decisão. Protocolos bem elaborados evitam esses problemas [6, 7].

Apesar de muitos estudos disponíveis na literatura, nota-se uma lacuna quanto a procedimentos bem estabelecidos e práticos. Os estudos dos mecanismos probabilísticos subjacentes às escolhas tem sido em várias direções. Duas delas se destacam: as chamadas escolhas probabilísticas (Campello de Souza [8, 9]) e os modelos estatísticos de erro para a estimação da função utilidade, *cf.* Hauser [1].

A situação torna-se complicada no caso em que as consequências ou *payoffs* têm de ser representadas por vetores, isto é, no caso multidimensional. Nesta situação, há uma imbricação entre dois fenômenos: o *tradeoff* entre as diversas dimensões dos *payoffs* e a matriz de aversão ao risco [10]. A curvatura geral da função utilidade de argumento multidimensional,

assim como as curvaturas das curvas de iso-utilidade (que caracterizam os *tradeoffs*), têm uma ligação íntima com a aversão ao risco. Esta é representada, por exemplo, pela chamada matriz de aversão ao risco [10]. Essas interações têm implicações diretas em qualquer que seja o protocolo de educação das preferências, e há que se lidar com isso.

Subjacente à essas questões está o conceito de ordem. Na teoria da utilidade de von Neumann e Morgenstern, o primeiro axioma exige a existência de uma ordem completa (ou total) sobre o conjunto de todas as distribuições de probabilidade sobre os bens (*payoffs*; os p_j 's). Ora, o conjunto \mathbb{R}^n , para $n \geq 2$, não é naturalmente ordenado. Há que se trabalhar pois no estabelecimento de uma ordem total, como parte do processo de educação. Isto não é feito, tipicamente, de maneira formal. Em geral, parte-se diretamente para uma expressão analítica geral e, então, estima-se os seus parâmetros (Keeney & Raiffa [11]).

O enfoque proposto nesta tese é o de considerar explicitamente o estabelecimento de uma ordem total e ao mesmo tempo explorar os mecanismos probabilísticos dos erros do processo de educação. Neste processo, explora-se a monotonicidade da função utilidade com relação a todos os seus argumentos.

A teoria da decisão clássica fornece uma abordagem normativa para a representação e o manuseio de preferências. Teorias, ferramentas e técnicas desenvolvidas nos últimos 60 anos fornecem uma estrutura adequada para a composição de sistemas de apoio à decisão. Neste contexto, pode-se conceber representações para preferências face à incerteza e múltiplos objetivos: estas representações capturam aspectos importantes da preferência tais como atitude diante do risco, *tradeoffs* entre objetivos conflitantes, *etc.* Em geral, a aplicação direta desta teoria tem sido difícil devido ao alto custo para a educação das preferências [11].

No núcleo da teoria da decisão encontra-se o paradigma da **máxima utilidade esperada**: assumindo-se uma função utilidade que quantifica a preferência em relação às distribuições de probabilidade sobre os possíveis resultados de um processo decisório. Diz-se que uma alternativa de ação é ótima se maximiza a utilidade esperada [5, 12, 13]. Note-se que se usa aqui o fato de que o resultado de uma ação é uma distribuição de probabilidade, *cf.* Dorfman, Samuelson e Solow² [14, p. 431]:

“In economic and military life it is probably quite exceptional to find a pay-off matrix that is free of probability distributions.”

²Esta afirmação, retirada do livro *Linear programming and economic analysis*, capítulo *Elements of Game Theory*, é uma rara exceção na literatura de teoria dos jogos, pois a vasta maioria da literatura se omite deste fato, assim como se omite do processo de educação da função utilidade.

Construída sobre este paradigma, a teoria da decisão iniciada por Wald [2] teve um impacto significativo em diversas disciplinas. Por exemplo, prover uma abordagem unificada para a inferência estatística [12, 13] e formalização matemática para a microeconomia [3, 15]. Nas últimas décadas tem sido utilizada para resolver problemas nas mais diversas áreas como finanças [16], medicina [17] e políticas públicas [18].

Para o caso de funções utilidade multiatributo (multidimensional, multicritério ou ainda multiobjetivo), a referência, ainda hoje, é o livro de Keeney e Raiffa [11]. Neste livro, os autores usaram o termo *Multiattribute Utility Theory* (MAUT) para este caso em particular. Embora esta não seja um teoria propriamente dita, tratando-se de métodos *ad hoc* para a obtenção de funções utilidade multilineares, obteve grande sucesso em aplicações onde existe independência entre os critérios.

1.2 Objetivos

No intuito de estabelecer protocolos de educação da função utilidade multiatributo fazendo-se menos hipóteses, tais como a independência entre os critérios, alguns objetivos foram estabelecidos para este trabalho.

1.2.1 Objetivos gerais

- ▷ Estudar os conceitos de aversão ao risco;
- ▷ Elaborar um método de estimação da função utilidade multiatributo.

1.2.2 Objetivos específicos

- ▷ Contribuir para o estudo da questão das preferências no contexto da teoria da utilidade de von Neumann e Morgenstern;
- ▷ Elaborar um protocolo para a educação de uma ordem linear sobre as preferências;
- ▷ Estudar a relação da matriz de aversão ao risco com os conceitos de independência entre os atributos; e implicações na estrutura da função utilidade.
- ▷ Realizar experimentos de educação das preferências.

1.3 Justificativa

A aplicação da “teoria” da utilidade multiatributo é bastante limitada, em parte pela falta de métodos gerais para educação dos parâmetros da função utilidade, em especial para os casos onde a função não é multilinear. Como consequência, a maioria dos estudos se restringem às especificações lineares, onde a estimação e interpretação dos parâmetros são mais diretas.

O aspecto mais difícil da aplicação da teoria da decisão é o alto custo da educação das preferências do decisor, *i.e.*, a estimação da função utilidade. Na prática, esta etapa consiste em entrevistas e aplicação de questionários, primeiro para a obtenção de características da função utilidade e, posteriormente, para obtenção de pontos necessários para a estimação dos parâmetros da função.

1.4 Resultados

Introduz-se um novo método para a educação da função utilidade multiatributo. Elabora-se um algoritmo para a educação de uma ordem linear sobre os *payoffs*, neste caso multidimensionais, que representa as preferências do decisor de modo a se poder usar os métodos clássicos de educação da função utilidade monodimensional. Além disso, apresenta-se as implicações na estrutura da função utilidade de restrições nos *tradeoffs* geradas pela escolha de uma ordem linear particular. Diversas representações do conceito de *tradeoff* usados na literatura foram estudadas. Estuda-se também a relação deste conceito com os conceitos de aversão ao risco e explora-se esta ligação na escolha da forma funcional da utilidade. Finalmente, esse estudo fornece uma ferramenta para educação de funções utilidades multiatributo parcimoniosa em relação ao tamanho do questionário e às hipóteses sobre a estrutura de preferências do decisor, facilitando assim o uso da teoria em epígrafe.

1.5 Estrutura da tese

No Capítulo 2 apresenta-se uma revisão dos conceitos da teoria da utilidade de von Neumann e Morgenstern. No Capítulo 3 estuda-se os conceitos de aversão ao risco e independência entre os atributos. No Capítulo 4 é desenvolvido um protocolo para a educação da função utilidade multiatributo. No Capítulo 5 finaliza este trabalho, com conclusões, comentários e sugestões para trabalhos futuros.

DECISÕES E UTILIDADES

“Your act was unwise,” I exclaimed “as you see by the outcome.” He solemnly eyed me. “When choosing the course of my action,” said he, “I had not the outcome to guide me.”

— Ambrose Bierce

A TEORIA DA DECISÃO surgiu de uma série de publicações de Abraham Wald que teve início em 1939, com o artigo *Contributions to the Theory of Statistical Estimation and Testing Hypotheses* [19], e culminou em 1950 com a publicação do livro *Statistical Decision Functions* [2]. Segundo Savage [20], Wald “*presents a new theory of the foundations of statistics.*”

Antes de Abraham Wald, o problema central da estatística era o de produzir inferências, ou seja, produzir afirmações razoavelmente seguras com base em informações incompletas. Esta nova teoria, entretando, muda o foco de inferência para a escolha de ações, isto é, o problema passa a ser o de como escolher uma ação “ótima” com base em informações incompletas [5, 12, 20]. A Teoria da Decisão de Wald generaliza e simplifica de maneira elegante a teoria de Neyman-Pearson, unificando problemas considerados distintos por Neyman e Pearson [13, 19, 21]. A teoria clássica se preocupa apenas em fazer afirmações a partir de uma amostra, sem levar em consideração o uso final dessas afirmações. Na Teoria da Decisão, ao contrário, combina-se as informações da amostra com outros aspectos relevantes do problema, de modo a escolher uma melhor ação (em um sentido pré-estabelecido) [2, 5, 13].

Além de unificar a teoria estatística de até então, em 1945 Wald mostrou que sua teoria é intimamente ligada com a Teoria do Jogos desenvolvida por von Neumann [3]. De fato, o problema de inferência estatística como formulado por Wald pode ser interpretado como um jogo de soma nula com dois jogadores, onde um é o “estatístico” e o outro é a “natureza” [22].

Esta estreita relação com o trabalho de von Neumann potencializou a abrangência da Teoria da Decisão. Usando-se a Teoria da Utilidade de von Neumann e Morgenstern [3] na teoria de Wald, aumentou-se sobremaneira o espectro desta última, de “*foundations of statistics*” para ferramenta fundamental em problemas de Economia, Engenharia, Psicologia, *etc.*

Os trabalhos de Campello de Souza [5], Berger [21], Ferguson [13] e Savage¹ [12] apresentam uma exposição detalhada da Teoria da Decisão.

Considera-se um problema de decisão uma situação com, no mínimo, os seguintes elementos:

1. Um conjunto \mathcal{A} de *ações* (ou *alternativas*). Uma ação a em \mathcal{A} disponível para o decisor capaz de realizar uma certa tarefa. Para que o problema de decisão seja bem definido, o conjunto \mathcal{A} deve ser exaustivo e mutualmente exclusivo.
2. Um conjunto Θ de *estados* ou *estados da natureza*. Um elemento θ em Θ representa uma influência nos possíveis resultados de uma ação. Não se tem controle sobre este elemento. Não se sabe qual estado irá se concretizar, é sob esta incerteza que se deve decidir por uma ação e aceitar as consequências.
3. Um conjunto de distribuições de probabilidade \mathcal{P} sobre possíveis consequências (resultados ou ainda *payoffs*) c de um conjunto \mathcal{C} . Se se decide pela ação a e o estado na natureza é θ , então tem-se a distribuição de probabilidade $p(c | \theta, a)$ sobre as possíveis consequências. Esta aplicação que leva um par ordenado $(\theta, a) \in \Theta \times \mathcal{A}$ em um elemento de \mathcal{P} é chamada de “Função consequência.”

Uma formulação completa do problema de decisão precisa ainda da hipótese de que é possível, para o decisor, construir uma relação binária² que capte sua preferência em relação aos elementos do conjunto \mathcal{P} , ou seja, sobre distribuições de probabilidade sobre as consequências. Nas palavras de von Neumann e Morgenstern [3, p. 17]:

¹Savage [12, p. 156] apresenta a teoria de Wald em seu “*The foundations of statistics*” como *Minimax theory*: “*In recent years there has been developed a theory of decision, here with due precedent to be called the minimax theory, that embraces so much of current statistical theory that the remaining chapters can largely be built around it. The minimax theory was originated and much developed by A. Wald [...]*”

²O conceito de relações foi primeiro introduzido por Gauss em seu *Disquisitiones Arithmeticae*, publicado em 1798.

“More precisely, we expect him, for any two alternative events which are put before him as possibilities, to be able to tell which of the two he prefers.

It is a very natural extension of this picture to permit such an individual to compare not only events, but even combinations of events with stated probabilities.”

Apresenta-se a seguir a formalização do conceito de preferência de von Neumann e Morgenstern e sua representação numérica, a função utilidade.

2.1 Teoria da utilidade

Considera-se que o decisor, quando avalia o peso das consequências das ações disponíveis, leva em consideração as probabilidades associadas as consequências (em \mathcal{C}). Apresenta-se nesta seção um modelo matemático pelo qual tal avaliação é feita. Este modelo se baseia na noção de que um decisor “racional” é capaz de expressar sua preferência entre consequências de maneira consistente com certos axiomas. Mostrar-se-á que a preferência de um decisor por uma consequência poderá ser expressa por uma função real, chamada *utilidade*, assim como a preferência sobre distribuições de probabilidade sobre consequências (\mathcal{P}) é baseada apenas no valor esperado da utilidade. Note-se que estas distribuições são consequências de determinadas ações, dessa forma esta utilidade induz uma preferência nas ações, jamais o contrário. Não faz sentido dizer que um indivíduo tem preferências por ações.

Define-se aqui racionalidade como a consistência para com os axiomas apresentados a seguir. Uma maneira equivalente porém mais abstrata, apresentada por Allais [23, p. 518], afirma que um decisor é racional se

- (a) Persegue objetivos coerentes ou consistentes entre eles;
- (b) Age de forma apropriada na busca destes objetivos.

Ainda segundo Allais [23, p. 518], estas duas condições têm como consequências:

- (1) Existe uma ordem completa sobre \mathcal{P} ;
- (2) Se dois elementos de \mathcal{P} são tais que o primeiro fornece uma probabilidade de ganho maior que o segundo, então o primeiro será preferido ao segundo;
- (3) Considera-se probabilidades objetivas.

Os pontos (1) e (2) são amplamente aceitos. O ponto (3) ainda gera muita controvérsia. Tanto A. Wald quando von Neumann eram “objetivistas.” Entretanto, Savage, como sugestão de Bruno de Finetti [20] e quatro anos depois do livro de Wald [2], publicou seu livro

“*The foundations of statistics*” em que apresenta uma exposição da teoria da decisão de Wald e da teoria da utilidade de von Neumann-Morgenstern, com axiomas de probabilidade subjetiva, além dos considerados anteriormente. Bem mais acessível que os trabalhos de Wald, principalmente do ponto de vista matemático, tornou-se a referência no assunto. Entretanto, estas mudanças (sobretudo Savage [12] e Anscombe-Aumann [24]) introduziram dificuldades — vide Ellsberg [25] e Machina [26], *cf.* Paradoxo de Ellsberg [25].

Segue-se a descrição clássica da teoria da utilidade. Define-se inicialmente um modelo de preferência sobre consequências aleatórias.

DEFINIÇÃO 2.1 Uma **relação de preferência** em \mathcal{P} é uma ordem linear, denotada \preceq , de \mathcal{P} ; *i.e.*, é uma relação binária que admite as seguintes propriedades:

(**linearidade**) se p_1 e p_2 são elementos de \mathcal{P} , então $p_1 \preceq p_2$ ou $p_2 \preceq p_1$;

(**transitividade**) se p_1, p_2 e p_3 são elementos de \mathcal{P} , e se vale $p_1 \preceq p_2$ e $p_2 \preceq p_3$, então também vale $p_1 \preceq p_3$.

Um elemento p em \mathcal{P} pode ser visto como uma **loteria** que, de acordo com a distribuição de probabilidade p , fornece uma consequência em \mathcal{C} . Supõe-se que o decisor seja “racional,” no sentido que ele possui uma relação de preferência entre loterias, como na Definição 2.1.

A afirmação $p_1 \preceq p_2$ significa que o decisor “racional” não prefere a distribuição p_1 à distribuição p_2 ; *i.e.*, ele prefere p_2 à p_1 ou é indiferente em relação às duas.

DEFINIÇÃO 2.2 Seja p_1 e p_2 elementos de \mathcal{P} . Diz-se que p_2 é **preferido** à p_1 e denota-se por $p_1 \prec p_2$, se vale $p_1 \preceq p_2$ e $p_2 \not\preceq p_1$. Além disso, diz-se que p_1 e p_2 são **equivalentes**, denota-se por $p_1 \sim p_2$, se vale $p_1 \preceq p_2$ e $p_2 \preceq p_1$.

Desta definição segue os seguintes lemas:

LEMA 2.1 A relação binária \sim é uma relação de equivalência.

LEMA 2.2 Se $p_1 \preceq p_2$ e $p_2 \prec p_3$, então $p_1 \prec p_3$.

A demonstração destes lemas é direta.

É importante notar que as seguintes propriedades são verdadeiras para o conjunto \mathcal{P} :

- (a) \mathcal{P} é convexo, *i.e.*, $p_1 \in \mathcal{P}$ e $p_2 \in \mathcal{P}$ implica que $\alpha p_1 + (1 - \alpha)p_2$ também pertence a \mathcal{P} para todo $0 \leq \alpha \leq 1$; e
- (b) Toda distribuição degenerada está em \mathcal{P} . Uma distribuição degenerada é aquela que concentra toda a massa probabilística em um único elemento de \mathcal{C} .

Tratar-se-á aqui apenas de distribuições que concentrem toda a massa probabilística em um número finito de elementos de \mathcal{C} . Note-se que se pode estender \mathcal{P} de modo a conter distribuições mais gerais (*e.g.*, distribuições contínuas) e as propriedades citadas continuarão válidas. Note-se que se pode entender a combinação linear $\lambda p_1 + (1 - \lambda)p_2$ como a combinação linear de dois vetores p_1 e p_2 , em que cada elemento do vetor indica a probabilidade associada a uma consequência em \mathcal{C} , embora \mathcal{P} não defina um espaço vetorial.

Pode-se construir uma função real, com domínio em \mathcal{P} , que represente a relação de preferência do decisor. As funções deste tipo dá-se o nome de *função utilidade*.

DEFINIÇÃO 2.3 Uma **utilidade** em \mathcal{P} é uma função real u , definida em \mathcal{P} , a qual é linear em \mathcal{P} ; *i.e.*, se p_1 e p_2 são elementos de \mathcal{P} e λ é um número real tal que $0 \leq \lambda \leq 1$, então

$$u(\lambda p_1 + (1 - \lambda)p_2) = \lambda u(p_1) + (1 - \lambda)u(p_2). \quad (2.1)$$

DEFINIÇÃO 2.4 Uma função utilidade u representa uma relação de preferência (\preceq) se, para todo p_1 e p_2 em \mathcal{P} , vale

$$p_1 \preceq p_2 \text{ se, e somente se, } u(p_1) \leq u(p_2).$$

Sejam c_1, c_2, \dots, c_k elementos de \mathcal{C} e $\alpha_1, \alpha_2, \dots, \alpha_k$ números reais não negativos tais que $\sum \alpha_i = 1$; denota-se $(\alpha_1 c_1, \dots, \alpha_k c_k)$ o elemento de \mathcal{P} que apresenta probabilidade α_i de se obter a consequência c_i . Se se define para cada $c \in \mathcal{P}$, $v(c) \equiv u(1c)$ para função utilidade u , então

$$u[(\alpha_1 c_1, \dots, \alpha_k c_k)] = \sum_{i=1}^k \alpha_i v(c_i). \quad (2.2)$$

é o valor esperado da função v usando a distribuição $(\alpha_1 c_1, \dots, \alpha_k c_k)$. Então, se existe uma função utilidade u que representa a preferência \preceq de um decisor, então este agirá de modo a maximizar o valor esperado da função v . A função v é chamada de **função valor** ou **utilidade de Bernoulli**.

Nem sempre uma relação de preferência pode ser representada por uma função utilidade. Para que isso seja sempre possível, faz-se necessário que a relação de preferência satisfaça a hipóteses adicionais [5, 13], a saber:

Hipótese 1: Se p_1, p_2 e q são elementos de \mathcal{P} e λ real tal que $0 < \lambda \leq 1$, então $p_1 \preceq p_2$ se, e somente se, $\lambda p_1 + (1 - \lambda)q \preceq \lambda p_2 + (1 - \lambda)q$.

Hipótese 2: Se p_1, p_2 e p_3 são elementos de \mathcal{P} tais que $p_1 \prec p_2 \prec p_3$, então existe números reais λ e μ com $0 < \lambda < 1$ e $0 < \mu < 1$, tais que

$$\lambda p_3 + (1 - \lambda)p_1 < p_2 < \mu p_3 + (1 - \mu)p_1.$$

Nesta condição, tem-se o seguinte teorema:

TEOREMA 2.1 Se a relação de preferência \preceq em \mathcal{P} satisfaz as hipóteses 1 e 2, então existe uma utilidade, u , em \mathcal{P} que representa \preceq . Além disso, u é única a menos de uma transformação linear.

A demonstração deste resultado é apresentada em detalhes em Campello de Souza [5] e Ferguson [13].

Note-se que se u representa uma relação de preferência \preceq , então $\hat{u} = \alpha u + \beta$, onde $\alpha > 0$ e β são números reais, é também uma função utilidade que representa \preceq .

Este resultado afirma que se um decisor possui uma relação de preferência que satisfaz certas restrições, ele irá se comportar como se tivesse associado uma utilidade numérica para cada consequência em \mathcal{C} , e sua preferência em \mathcal{P} coincidirá com uma preferência pelo elemento de \mathcal{P} com maior utilidade esperada.

2.2 Construção da função utilidade

Para construir a função utilidade u sobre um conjunto de distribuições \mathcal{P} , escolhe-se duas consequências, c_0 e c_n , induzidas pelas distribuições $p_0 = (1c_0)$ e $p_n = (1c_n)$, respectivamente, e de tal forma que $p_0 \prec p_n$. Define-se, por conveniência, as seguintes utilidades:

$$u(p_0) = v(c_0) \equiv 0 \quad \text{e} \quad u(p_n) = v(c_n) \equiv 1. \quad (2.3)$$

É usual escolher c_n e c_0 , isto é, p_n e p_0 , como os extremos da ordem linear, *i.e.*, o elemento mais preferido e o elemento menos preferido da relação de preferência do decisor, respectivamente. A escolha feita em (2.3) é arbitrária e serve para definir uma escala para função

utilidade u . Deste modo, valores de $u(p)$ para p no intervalo $[p_0, p_n]$ (não confundir com uma loteria) serão estabelecidos a partir de comparações às distribuições degeneradas p_0 (c_0 com probabilidade 1) e p_n (c_n com probabilidade 1).

Existem diversas maneiras de prosseguir com a educação da função utilidade. Por exemplo, pode-se comparar um elemento degenerado de \mathcal{P} (o que é equivalente a tomar-se um elemento de \mathcal{C}); e comparar com a loteria simples³ $\lambda p_n + (1 - \lambda)p_0$. Uma maneira de fazer essa comparação é solicitar ao decisor que escolha um número real λ_i ($0 < \lambda_i < 1$) de modo que o faça indiferente em receber qualquer uma das seguintes loterias: **(1)** c_i com probabilidade 1; **(2)** $[c_n, \lambda_i, c_0]$, onde receberá c_n (mais desejável) com probabilidade λ_i e c_0 (menos desejável) com probabilidade $1 - \lambda_i$. Simbolicamente, espera-se que o decisor seja capaz de encontrar um λ_i ($0 < \lambda_i < 1$) tal que seja válida a seguinte relação:

$$p_i \sim \lambda_i p_n + (1 - \lambda_i) p_0,$$

em que p_i é a distribuição de probabilidade degenerada que fornece c_i com probabilidade 1. Seja u uma função utilidade que representa esta relação de preferência, então

$$\begin{aligned} u(p_i) &= u(\lambda_i p_n + (1 - \lambda_i) p_0) \\ &= \lambda_i u(p_n) + (1 - \lambda_i) u(p_0) \\ &= \lambda_i. \end{aligned}$$

Como p_i , p_0 e p_n são loterias degeneradas, pode-se escrever:

$$v(c_i) = \lambda_i v(c_n) + (1 - \lambda_i) v(c_0) = \lambda_i.$$

Note-se que a idéia é construir primeiro a função valor, função utilidade quando restringe-se \mathcal{P} às loterias degeneradas, e com isso se valer da condição que u é linear em \mathcal{P} para encontrar a utilidade para uma distribuição qualquer (cf. Definição 2.3).

Por outro lado, se existe um elemento p_i que forneça c_1 com probabilidade 1, tal que $p_1 \prec p_0$, então eduz-se o λ_i ($0 < \lambda_i < 1$) tal que

$$p_0 \sim \lambda_i p_i + (1 - \lambda_i) p_n.$$

³Loterias simples são aquelas que contém apenas elementos degenerados de \mathcal{P} . Caso contrário são ditas compostas.

Tem-se então

$$\begin{aligned} u(p_0) &= \lambda_i u(p_i) + (1 - \lambda_i) u(p_n) \\ 0 &= \lambda_i u(p_i) + (1 - \lambda_i) \cdot 1 \Rightarrow \\ u(p_i) &= -\frac{1 - \lambda_i}{\lambda_i}. \end{aligned}$$

Considera-se agora que o elemento p_i é preferível ao elemento p_n de \mathcal{P} , *i.e.*, $p_n \prec p_i$. Procura-se então um valor λ_i ($0 < \lambda_i < 1$) tal que valha a relação

$$p_n \sim \lambda_0 p_0 + (1 - \lambda_i) p_i,$$

o que implica que a utilidade de p_i é dada por

$$u(p_i) = \frac{1}{1 - \lambda_i}.$$

É importante notar que, durante todo o processo de educação da função utilidade, é fundamental o conhecimento da relação de preferência do decisor. É preciso conhecer, no mínimo, uma ordem sobre as distribuições degeneradas de \mathcal{P} que, além de induzir naturalmente uma ordem sobre o conjunto \mathcal{C} , de consequências determinísticas, estende-se para todo o conjunto \mathcal{P} com o uso da função valor sobre \mathcal{C} e da linearidade sobre \mathcal{P} da função utilidade, *cf.* Equação (2.2).

De fato, se se fixa a relação de preferência \preceq onde $p_0 \prec \dots \prec p_n$, e define-se a função utilidade como

$$u(p) \equiv \sup \{ \lambda \mid p \succ \lambda p_n + (1 - \lambda) p_0 \}, \quad (2.4)$$

tem-se que esta representa a relação de preferência \preceq , vide a demonstração do Teorema 2.1 em Campello de Souza [5, p. 50] e Ferguson [13, p. 15]. Vê-se, portanto, que a definição inclui uma maneira de se medir a função utilidade.

2.2.1 Métodos para a educação da função utilidade

A maioria dos métodos de educação da função utilidade consideram loterias (jogos, distribuições de probabilidade) sobre duas consequências em \mathcal{C} , ou seja, consideram elementos de \mathcal{P} na forma $(\alpha c_1, (1 - \alpha) c_2)$ com $0 \leq \alpha \leq 1$, *i.e.*, tem-se o resultado c_1 com probabilidade α e o resultado c_2 com probabilidade $1 - \alpha$, e não loterias sobre loterias (combinações de distribuições de probabilidade). Se $c_1 = c_2$, $\alpha = 0$ ou $\alpha = 1$, a loteria é dita **degenerada**.

Outras maneiras de se denotar essa loteria são:

$$[c_1, \alpha, c_2] \quad \text{ou} \quad \begin{cases} c_1 \text{ com probabilidade } \alpha \\ c_2 \text{ com probabilidade } 1 - \alpha. \end{cases}$$

Em geral, estes métodos utilizam comparações de loterias da seguinte maneira:

$$[c_i, \alpha, c_j] \mathcal{R} [c_k, \beta, c_l] \tag{2.5}$$

em que um elemento é desconhecido e todos os outros são dados, de modo que o decisor especifique este elemento desconhecido para que a relação (2.5) seja válida. Note que em (2.5), c_i, c_j, c_l e c_k são elementos de \mathcal{C} , α e β são números reais tais que $0 \leq \alpha \leq 1$ e $0 \leq \beta \leq 1$; e \mathcal{R} é uma das relações \prec (menos preferível que), \sim (equivalente ou indiferente) e \succ (preferível a).

Os métodos básicos de comparação de loterias são apresentados na Tabela 2.1, separados em duas categorias:

1. *jogos padrões*: uma das loterias é degenerada;
2. *jogos pareados*: nenhuma das loterias é degenerada.

Note-se que, na Tabela 2.1, os elementos sublinhados devem ser especificados pelo decisor durante o processo de educação, conhecendo-se todos os demais.

Tabela 2.1: *Métodos para comparação de loterias.*

	Jogos padrões	Jogos pareados
1. Preferência	$[c_i, \alpha, c_j] \underline{\mathcal{R}} c_k$	$[c_i, \alpha, c_j] \underline{\mathcal{R}} [c_k, \beta, c_l]$
2. Equivalência em probabilidade	$[c_i, \underline{\alpha}, c_j] \sim c_k$	$[c_i, \underline{\alpha}, c_j] \sim [c_k, \beta, c_l]$
3. Equivalência em valor	$[\underline{c}_i, \alpha, c_j] \sim c_k$	$[\underline{c}_i, \alpha, c_j] \sim [c_k, \beta, c_l]$
4. Equivalente certo	$[c_i, \alpha, c_j] \sim \underline{c}_k$	—

Antes de continuar com a apresentação dos métodos, faz-se necessário apresentar algumas questões sobre a forma de representação das loterias. Estudos empíricos sugerem que se deve apresentar as loterias de maneira que todos seus elementos estejam explícitos, para facilitar a mediação cognitiva [11, 27, 28]. A Figura 2.1 mostra alguns exemplos de como se pode representar loterias em um questionário de educação de utilidade. Note que todos os elementos da loteria são representados.

Examina-se a seguir diferentes sequências de jogos padrões que podem ser usados para a educação da função utilidade. Os métodos serão apresentados segundo a divisão da Tabela 2.1:

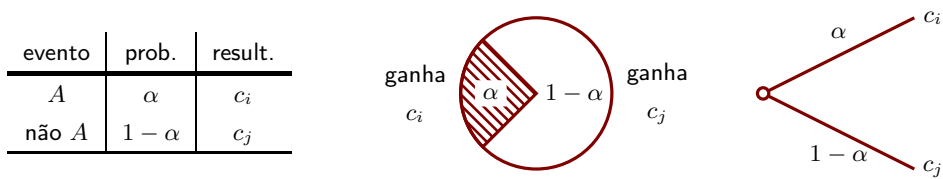


Figura 2.1: Maneiras de apresentar uma loteria $[c_i, \alpha, c_j]$ para o decisor.

métodos de comparação de preferência, métodos de equivalência em probabilidade, métodos de equivalência em valor e métodos de equivalente certo.

Métodos de comparação de preferência

No *método de comparação de preferência* entre uma loteria $[c_i, \alpha, c_j]$ e uma consequência certa c_k (ou seja, uma loteria degenerada) o decisor deve especificar a relação \mathcal{R} (\prec , \succ ou \sim) tal que a expressão $[c_i, \alpha, c_j] \mathcal{R} c_k$ seja satisfeita. Este método envolve uma sequência de tais comparações em que as probabilidades e consequências são escolhidas de uma forma particular.

Estes métodos são usualmente utilizados de duas formas durante o processo de educação da utilidade. A primeira investiga a atitude ao risco em análises preliminares e em testes de consistência de uma utilidade eduzida. Cada comparação de preferência fornece uma das seguintes restrições lineares que a função valor eduzida deve satisfazer:

$$\begin{aligned}
 [c_i, \alpha, c_j] \prec c_k &\Rightarrow \alpha v(c_i) + (1 - \alpha)v(c_j) < v(c_k) \\
 [c_i, \alpha, c_j] \succ c_k &\Rightarrow \alpha v(c_i) + (1 - \alpha)v(c_j) > v(c_k) \\
 [c_i, \alpha, c_j] \sim c_k &\Rightarrow \alpha v(c_i) + (1 - \alpha)v(c_j) = v(c_k).
 \end{aligned}$$

Com um conjunto destas restrições é possível construir uma família de funções valor admissíveis ou então para testes de consistência.

Métodos de equivalência em probabilidade

Requer que o decisor especifique uma probabilidade α para a qual $[c_i, \alpha, c_j] \sim c_k$, em que c_k está entre c_i e c_j . Este método se aplica para atributos (consequências) discretos ou contínuos. Inicialmente escolhe-se elementos de referência c_0 e c_n em \mathcal{C} , tais que $c_0 \prec c_n$. Em geral, estes elementos são escolhidos de forma que representem a pior consequência e a melhor consequência em \mathcal{C} . O problema agora é eduzir a utilidade em uma sequência $c_0 \prec c_1 \prec \dots \prec c_{n-1} \prec c_n$, utilizando-se um dos seguintes tipos de jogos.

Valores extremos: $[c_n, \underline{\alpha}_i, c_0] \sim c_i$. Este método usa os pontos de referência do conjunto de consequências \mathcal{C} em todas as comparações, foi proposto por von Neumann e Morgenstern [3]. Se $v(c_0) \equiv 0$ e $v(c_n) \equiv 1$, então

$$\begin{aligned} v(c_i) &= u([c_n, \alpha_i, c_0]) \\ &= \alpha_i v(c_n) + (1 - \alpha_i) v(c_0) \\ &= \alpha_i. \end{aligned}$$

Então a probabilidade eduzida para garantir a indiferença entre as loterias, é também a função valor do elemento c_i . Embora o método das loterias com valores extremos seja direto e intuitivo, ele é muito susceptível a uma dependência serial entre as respostas e a viés se os extremos forem muito “distantes” [5, 11, 27].

Valores adjacentes: $[c_{i+1}, \underline{\alpha}_i, c_{i-1}] \sim c_i$. Ao invés de utilizar valores extremos, este método usa consequências que são “localmente extremas” para uma determinada consequência c_i [28, 29]. Cada uma das respostas do decisor implica em uma equação da forma

$$v(c_i) = \alpha_i v(c_{i+1}) + (1 - \alpha_i) v(c_{i-1}), \quad i = 1, 2, \dots, n - 1.$$

Fazendo

$$\begin{aligned} v(c_0) &\equiv 0 \\ v(c_n) &\equiv 1 \\ f_0 &\equiv 1 \\ f_i &\equiv \frac{1 - \alpha_i}{\alpha_i}, \end{aligned}$$

pode-se resolver um sistema com $n - 1$ equações e $n - 1$ incógnitas, resultando na função valor

$$v(c_i) = \frac{\sum_{j=0}^{i-1} \prod_{k=0}^j f_k}{\sum_{j=0}^{n-1} \prod_{k=0}^j f_k}, \quad i = 1, \dots, n - 1.$$

Uma vantagem deste método é que o decisor não precisa responder a questões para as quais a probabilidade ficaria naturalmente próxima a zero ou a um, tornando esse processo de educação menos sensível a falta de precisão das probabilidade especificadas pelo decisor [28, 29].

Métodos de equivalência em probabilidade oferecem várias vantagens. Uma delas é que como não acontece encadeamento de respostas, *i.e.*, o uso de respostas em jogos subsequentes, a dependência serial entre as comparações de loterias pode ser atenuada permutando-se as sequências de perguntas [28]. Respostas expressas em probabilidade são menos susceptíveis a distorções devido ao risco [30], dependência de sequência [31] e outras fontes cognitivas de viés [27]. Embora muitos indivíduos achem difícil fazer afirmações sobre probabilidades, protocolos e treinamentos estão disponíveis para facilitar o processo de educação. Além disso, métodos de loterias adjacentes são robustos para probabilidades não muito perto de zero ou um, de modo que os julgamentos do decisor a respeito destas probabilidades não têm que ser precisos [29].

Métodos de equivalência em valor

Métodos de *equivalência em valor* solicitam ao decisor que especifique um valor $c \in \mathcal{C}$ tal que a relação de indiferença $[\underline{c}, \alpha, c_j] \sim c_k$ seja válida. Estes métodos assumem a existência de um contínuo de valores c em \mathcal{C} ; dessa forma pode-se garantir que exista um c que satisfaça a relação de indiferença [27, 30].

Métodos de equivalente certo

A comparação é feita de modo que o decisor especifique uma consequência c , chamada **equivalente certo**, para o qual a expressão $[c_i, \alpha, c_j] \sim \underline{c}$ é válida. Considere um contínuo de valores em \mathcal{C} de modo que c existe, e se a preferência é estritamente crescentes em \mathcal{C} , então c é único.

Considere fixo um conjunto de probabilidades $0 < \alpha_1 < \dots < \alpha_{n-1} < 1$ e dois pontos de referência c_0 e c_n em \mathcal{C} tal que $c_0 \prec c_n$ com $v(c_0) \equiv 0$ e $v(c_n) \equiv 1$.

Método Direto: $[c_n, \alpha_i, c_0] \sim c_i$. Este método é fácil de implementar e a educação da função valor é direta, $v(c_i) = \alpha_i$. Note-se sua similaridade com o método de equivalência de probabilidade utilizando-se valores extremos, posto isto, possuem as mesmas vantagens assim, como as mesmas desvantagens [3, 11].

Utiliza-se os pontos extremos do conjunto \mathcal{C} , c_0 e c_n , pois, caso contrário, seria necessário usar outros procedimentos para eduzir a função valor dos pontos menos preferíveis que c_0 e mais preferíveis que c_n . Isto pode provocar viés na estimação do equivalente certo quando os pontos extremos estão muito “distantes” [11].

Método de encadeamento: se pontos da função valor previamente eduzidos são utilizados na comparação de loterias subsequentes, então as respostas estão encadeadas. Simbolicamente, seja S_i o conjunto de valores eduzidos de um decisor antes da i -ésima resposta. Recursivamente S_i é definida como $S_i = S_{i-1} \cup \{c_{i-1}\}$ para $i = 1, \dots, n-1$ e $S_0 \equiv \{c_n\}$. Logo, para uma probabilidade dada $0 < \alpha_i < 1$ e elementos c_l e c_m em $\mathcal{C} \cap S_i$, a i -ésima comparação é $[c_l, \alpha_i, c_m] \sim c_i$.

Após cada comparação, o novo valor c_i é eduzido utilizando os valores previamente eduzidos e a igualdade a seguir:

$$\begin{aligned} v(c_i) &= u([c_l, \alpha_i, c_m]) \\ &= \alpha_i v(c_l) + (1 - \alpha_i) v(c_m). \end{aligned}$$

Para atenuar o viés introduzido para comparação usando valores extremos, c_0 e c_n , Krzysztofowicz e Duckstein [32] e Campello [5] propuseram um método onde primeiro particiona-se o espaço ordenado de c_0 a c_n em subintervalos, então aplica-se o método em cada uma das faixas para eduzir uma função valor para cada faixa. A função valor geral (para toda a faixa) será estimada a partir de jogos comuns nas faixas. Esta abordagem evita comparações entre valores muito “distante” que podem introduzir viés na educação.

2.3 Outras considerações

Antes de começar o estudo sobre a construção de uma relação de preferência, é mister dizer-se algo sobre a terminologia adotada, críticas clássicas à teoria da utilidade e aspectos normativos *versus* descritivos desta teoria.

2.3.1 *Terminus technicus*

A escolha de uma terminologia adequada é um problema de grande importância em todas as ciências [33]. Economia em particular, utiliza-se de conceitos e termos já parcialmente familiares ao grande público, sendo isto uma fonte constante de dificuldade.

O conceito de “utilidade” surgiu de várias formas diferentes em economia [33]: “grau final de utilidade” (Jevons), “utilidade efetiva”, “utilidade específica”, “eficiência marginal” (estes por Clark), “utilidade marginal” e “desejabilidade marginal” (Marshall), “rareté” (Walras), “Ophélimité” (Pareto) e “Wantability” (Fisher).

Neste trabalho, sempre que se usa o termo utilidade, refere-se ao construto de von Neumann e Morgenstern [3, p. 15].

2.3.2 O “paradoxo” de Allais

O que se chama hoje de **paradoxo de Allais** é o resultado de um experimento apresentado no artigo [23] “*Le comportement de l’homme rationnel devant le risque: critique des postulats et axiomes de l’école américaine.*” Nele se pede a decisores que respondam as seguintes questões:

1. Qual a relação de preferência entre as loterias A e B :

$$A : \text{ganhar 100 milhões com certeza.} \quad \text{e} \quad B : \begin{cases} 10\% \text{ de chance de ganhar 500 milhões.} \\ 89\% \text{ de chance de ganhar 100 milhões.} \\ 1\% \text{ de chance de nada ganhar.} \end{cases}$$

2. Qual a relação de preferência entre as loterias C e D :

$$C : \begin{cases} 11\% \text{ de chance de ganhar 100 milhões.} \\ 89\% \text{ de chance de nada ganhar.} \end{cases} \quad \text{e} \quad D : \begin{cases} 10\% \text{ de chance de ganhar 500 milhões.} \\ 90\% \text{ de chance de nada ganhar.} \end{cases}$$

Se um decisor é racional segundo os axiomas da teoria da utilidade, então se um decisor afirma que $B \prec A$, necessariamente também deve afirma que $D \prec C$. No entanto, segundo Allais, a maior parte das repostas foi $B \prec A$ e $C \prec D$, o que viola os axiomas.

No entanto, na sua amostra, Allais [23, p. 527] afirma que fazem parte indivíduos “[...] *que l’opinion commune considère comme très rationnels* [...].” Este sentido de racionalidade é diferente da hipótese de racionalidade de von Neumann e Morgenstern ou mesmo de Savage. Nestes últimos, um individuo racional é aquele que se comporta de acordo com certos axiomas, os quais o próprio Allais fornece uma definição mais abstrata [23, p. 518].

Neste sentido, o resultado de Allais é que os indivíduos “*que l’opinion commune considère comme très rationnels*” **não** são racionais no sentido de von Neumann e Morgensten. Posto isto, vê-se que o resultado de Allais não se trata de um paradoxo e sim da confirmação que a teoria da utilidade não é puramente descritiva. Ela é normativa, podendo ser considerada descritiva quando se trata de agentes racionais (no sentido axiomático).

Quanto ao “*que l’opinion commune considère comme très rationnels,*” subscrevo às idéias de Bertrand Russell [34, p. 71]:

“Man is a rational animal — so at least I have been told. Throughout a long life, I have looked diligently for evidence in favor of this statement, but so far I have not had the good fortune to come across it, though I have searched in many countries spread over three continents.”

2.4 A função de aversão ao risco de Arrow-Pratt

Desde Daniel Bernoulli (1938) que se coloca a questão da atitude de um decisor com relação ao risco. No seu “Paradoxo de São Petersburg” ficou claro que seria necessário ir-se além do valor médio de um jogo ou loteria, na hora de se tomar decisões envolvendo incerteza.

Intuitivamente, um decisor averso ao risco é aquele que é conservativo em relação a suas preferências. Considere um decisor frente a uma loteria que fornece uma consequência c ou uma consequência menos preferível c' , como igual probabilidade, *i.e.*, $[c, 1/2, c']$. Suponha-se que seja solicitado ao decisor que estabeleça sua preferência entre esta loteria e uma loteria que fornece com probabilidade 1 o valor esperado da loteria inicial, denota por \bar{c} . Se o decisor prefere a consequência certa \bar{c} à loteria $[c, 1/2, c']$ com as mesmas consequências esperadas, então o decisor prefere evitar os riscos associados à loteria. Um decisor deste tipo é dito *averso ao risco*.

Em tal situação, a utilidade do valor esperado de uma consequência de qualquer loteria deve ser maior que a utilidade esperada da loteria. Se se denota uma loteria por p , um decisor é averso ao risco se, para todas as loterias não degeneradas,

$$u[E(p)] > E[u(p)].$$

Isto é, um decisor é averso ao risco se a sua utilidade é côncava. Analogamente, um decisor é *propenso ao risco* se a sua utilidade é convexa e *neutro ao risco* se a sua utilidade é linear [35].

Existe pois uma relação entre a expressão analítica da função utilidade, e a psicologia do decisor com relação ao risco. Uma medida deste atitude com respeito ao risco, para o caso monodimensional, é a função de aversão ao risco de Pratt-Arrow [35]:

$$r(p) = -\frac{u''(p)}{u'(p)},$$

que pode ser vista como uma medida da curvatura da função utilidade. No caso multidimensional, várias extensões deste conceito foram apresentadas por diversos autores. Este é o tema do próximo capítulo.

A AVERSÃO AO RISCO E A FUNÇÃO UTILIDADE

All models are wrong, but some are useful.

— George Box

A MODELAGEM adequada para inúmeros problemas em economia e teoria da decisão envolve funções utilidades multidimensionais (ou multivariável, multiatributo, multicritério, *etc.*). Problemas de decisão nos quais as conseqüências são expressas em termos de mais de um atributo¹ têm se tornado cada vez mais importantes.

No caso em que as funções utilidades são unidimensionais (ou monodimensional, monoatributo), medidas de aversão ao risco, absoluto e relativo, foram apresentadas nos trabalhos seminiais de Pratt [35] e Arrow [36] (*apud* [37]), e desde então tornaram-se ferramentas indispensáveis para a análise de risco em situações envolvendo riscos unidimensionais.

A extensão do conceito de aversão ao risco para o caso multidimensional continua sem uma resposta concreta. Várias tentativas para a generalização para o caso multiatributo foram propostas, *cf.* de Finetti [38] e Richard [39], Stiglitz [40], Keeney [11, 41], Kihlstron e Mirman [42, 43], Paroush [44], Duncan [10], Karni [37], Levy e Levy [45], Ercolani [46] e Shah [47]. Entre estes, destacam-se dois, o conceito de Kihlstron e Mirman [42] e o conceito introduzido independentemente por de Finetti [38] e Richard [39]. O primeiro, mais freqüente na literatura, é compatível com o caso unidimensional. O segundo, embora tenha sido a primeira proposta em ordem cronológica, foi muito pouco estudado.

As primeiras abordagens do problema foram propostas por de Finetti [38] e Richard [39]. Seja $u(x, y)$ a função utilidade de um decisor para dois atributos. Para qualquer $x_0 < x_1$ e

¹Quando o argumento da função utilidade é uma distribuição de probabilidade sobre vetores.

qualquer $y_0 < y_1$, considere as loterias

$$L_1 = \begin{cases} (x_0, y_0) & \text{c.p. } 1/2 \\ (x_1, y_1) & \text{c.p. } 1/2 \end{cases} \quad L_2 = \begin{cases} (x_0, y_1) & \text{c.p. } 1/2 \\ (x_1, y_0) & \text{c.p. } 1/2 \end{cases}$$

Se o decisor prefere a loteria L_2 que tem uma chance² igual para (x_0, y_1) ou (x_1, y_0) à loteria L_1 onde ganha-se (x_0, y_0) ou (x_1, y_1) com igual probabilidade, *i.e.*, $u(L_2) > u(L_1)$, então o decisor é considerado *averso ao risco multidimensional* (abreviado *a.r.m.*). Este tipo de comportamento do decisor é considerado aversão ao risco, pois ele prefere uma combinação do “melhor” e do “pior” (L_2) a se arriscar entre todo o “melhor” ou todo o “pior” (L_1). Caso o decisor seja indiferente às loterias L_1 e L_2 , ele é dito *neutro ao risco multidimensional* (*n.r.m.*); e caso prefira L_1 à L_2 , *propenso ao risco multidimensional* (*p.r.m.*). Nota-se que este conceito de aversão ao risco não restringe o decisor em relação a um atributo em particular, *i.e.*, o decisor poder ser averso ao risco no sentido de de Finetti-Richard e propenso ao risco em relação a um atributo em particular segundo a definição de Arrow-Pratt.

Kihlstron e Mirman [42], propuseram uma abordagem compatível com o conceito de Arrow-Pratt. Um decisor é averso ao risco multidimensional se, diante da escolha entre um vetor aleatório n -dimensional \tilde{L} (uma loteria) e seu valor esperado $E[\tilde{L}]$, ele sempre prefere o segundo ao primeiro. Usando-se a desigualdade de Jensen, tem-se então que a matriz Hessiana da função utilidade tem que ser negativa semidefinida. Então, por definição, a função utilidade u é côncava em ambos os atributos. Portanto, aversão ao risco segundo Kihlstron e Mirman implica em aversão ao risco segundo Arrow-Pratt em cada atributo isoladamente.

3.1 Aversão ao risco multidimensional

Considera-se inicialmente o caso para dois atributos, denotados por x e y . Supõe-se que os conjuntos de todos os possíveis valores dos atributos x e y são intervalos da reta real, X e Y , respectivamente. Seja $P = X \times Y$ o produto cartesiano de X e Y , formado pelos pares ordenados (x, y) , onde $x \in X$ e $y \in Y$. Um elemento de P é chamado uma consequência $p = (x, y)$. Assume-se que existe uma função utilidade $u(x, y) = u(p)$ definida em P . Além disso, assume-se que esta função utilidade é suficientemente regular para justificar as demonstrações; geralmente isto vai significar que u é de classe $C^2[P]$ com primeira derivada positiva em todos os atributos. Deixa-se subentendido que $x_0, x_1 \in X$, $y_0, y_1 \in Y$, $x_0 < x_1$ e $y_0 < y_1$. Usa-se a

²“By chance I mean the same as probability.” Thomas Bayes.

seguinte notação no que diz respeito às derivadas parciais:

$$\begin{aligned}
 u_x(x, y) &\equiv \frac{\partial u(x, y)}{\partial x} \\
 u_y(x, y) &\equiv \frac{\partial u(x, y)}{\partial y} \\
 u_{xy}(x, y) &\equiv \frac{\partial^2 u(x, y)}{\partial x \partial y} \\
 &\dots
 \end{aligned}$$

Considere as seguintes loterias, para x_0, x_1, y_0 , e y_1 quaisquer:

- (i) Loteria L_1 dá ao decisor probabilidade 50% de receber (x_0, y_0) e 50% de probabilidade de receber (x_1, y_1) .
- (ii) Loteria L_2 dá ao decisor probabilidade 50% de receber (x_0, y_1) e 50% de probabilidade de receber (x_1, y_0) .

DEFINIÇÃO 3.1 O decisor é *avverso ao risco multidimensional (a.r.m)* se ele não prefere L_1 a L_2 , *i.e.*, $L_2 \succsim L_1$, para todo x_0, x_1, y_0 , e y_1 . Ele é *neutro ao risco multidimensional (n.r.m)* se e somente se fica indiferente entre L_1 e L_2 , *i.e.*, $L_1 \sim L_2$, para todo x_0, x_1, y_0 , e y_1 . Por último, o decisor é dito *propenso ao risco multidimensional (p.r.m)* se $L_1 \succ L_2$ para todo x_0, x_1, y_0 , e y_1 . Caso a preferência seja estrita (\succ), diz-se então que o decisor é *estritamente avverso ao risco multidimensional (s.a.r.m)* se $L_2 \succ L_1$ e *estritamente propenso ao risco multidimensional (s.p.r.m)* se $L_1 \succ L_2$ para todo x_0, x_1, y_0 , e y_1 .

A motivação para definição de que $L_2 \succsim L_1$ implica em *a.r.m* e $L_1 \succ L_2$ torna o decisor *p.r.m* é a seguinte. A loteria L_1 pode ser vista como um jogo “tudo ou nada” com 50% de chance. Por outro lado, a loteria L_2 garante que o decisor receberá pelo menos o “melhor” em um atributo.

Note-se que o comportamento diante do risco multidimensional não depende da aversão ao risco do decisor em relação a apenas um atributo, *i.e.*, o decisor pode ser avverso ao risco em relação ao atributo x , avverso ao risco em relação ao atributo y , e ainda propenso ao risco multidimensional.

TEOREMA 3.1 (RICHARD [39]) Um decisor é averso ao risco multidimensional se e somente se para todo x, y em P , $u_{xy}(x, y) \leq 0$; propenso ao risco multidimensional se e somente se para todo (x, y) em P , $u_{xy}(x, y) \geq 0$; e neutro ao risco multidimensional se e somente se $u_{x,y} = 0$ para todo (x, y) em P .

DEMONSTRAÇÃO 3.1 A demonstração é direta. Toma-se a diferença entre as utilidades esperadas de L_2 e L_1 :

$$\frac{1}{2} \{u(x_0, y_1) + u(x_1, y_0) - u(x_0, y_0) - u(x_1, y_1)\} = -\frac{1}{2} \int_{x_0}^{x_1} \int_{y_0}^{y_1} u_{xy}(x, y) dy dx. \quad (3.1)$$

A função utilidade será *a.r.m* se e somente se o lado esquerdo da Equação 3.1 é não negativa para todo $x_0 < x_1$ e $y_0 < y_1$, *i.e.*, a utilidade esperada de L_2 é maior ou igual a utilidade esperada de L_1 . De modo equivalente, o lado direito da Equação 3.1 deve ser não negativo para todo $x_0 < x_1$ e $y_0 < y_1$ o que acontece se e somente se para todo $(x, y) \in P$, $u_{xy}(x, y) \leq 0$. A demonstração é análoga para os casos *p.r.m* e *n.r.m*.

OEA

Dessa forma, tem-se um critério simples para a determinação do comportamento de aversão ao risco multidimensional quando se dispõe de uma função utilidade $u(x, y)$.

Note-se que o fato de $u(x, y)$ ser côncava não significa que $u_{xy} \leq 0$ e por isso $u(x, y)$ ser *a.r.m*. Pois para $u(x, y)$ ser côncava, sua matriz Hessiana deve ser negativa semi-definida, o que implica que os determinantes dos menores principais de sua matriz Hessiana devem alternar de sinal, *i.e.*, $u_{xx} \leq 0$ e $u_{xx}u_{yy} - u_{xy}^2 \geq 0$. A última equação implica que $u_{yy} \leq 0$ e

$$-(u_{xx}u_{yy})^{\frac{1}{2}} \leq u_{xy} \leq (u_{xx}u_{yy})^{\frac{1}{2}}.$$

Logo, é possível ter $u_{xy} \geq 0$ e ainda $u(x, y)$ ser côncava.

EXEMPLO 3.1 Considere a função utilidade $u(x, y)$, para $x > 0$ e $y > 0$, dada por:

$$u(x, y) = (xy)^{\frac{1}{4}}.$$

Tem-se então $u_{xy}(x, y) = 1/16(xy)^{-3/4} > 0$ e ainda assim u é estritamente côncava.

3.2 Prêmio pelo risco multidimensional

Introduz-se agora o conceito de prêmio pelo risco multidimensional. Para decisores que apresentam a preferência $L_2 \succ L_1$, pode-se modificar L_2 para um loteria L'_2 tal que $L'_2 \sim L_1$.

TEOREMA 3.2 (RICHARD [39]) $L_2 \succ L_1$ se e somente se existe (infinitos) um par (π_x, π_y) onde $0 \leq \pi_x \leq x_1 - x_0$ e $0 \leq \pi_y \leq y_1 - y_0$, com $\pi_x > 0$ ou $\pi_y > 0$ ou ambos, e a loteria L'_2 , que fornece $(x_1 - \pi_x, y_0)$ com 50% de chance e $(x_0, y_1 - \pi_y)$ com 50% de chance, é indiferente à loteria L_1 . Chama-se qualquer um desses pares que satisfazem essa condição de *prêmio pelo risco multivariado* (Π).

DEMONSTRAÇÃO 3.2 Se $L_2 \succ L_1$, tem-se então a seguinte relação entre as utilidades esperadas:

$$\frac{1}{2}u(x_0, y_1) + \frac{1}{2}u(x_1, y_0) > \frac{1}{2}u(x_0, y_0) + \frac{1}{2}u(x_1, y_1). \quad (3.2)$$

Seja agora L'_2 uma loteria com $\pi_x = x_1 - x_0$ e $\pi_y = 0$, então

$$\frac{1}{2}u(x_0, y_0) + \frac{1}{2}u(x_0, y_1) < \frac{1}{2}u(x_0, y_0) + \frac{1}{2}u(x_1, y_1), \quad (3.3)$$

donde a desigualdade levando-se em conta que $u_x(x, y) > 0$ para todo (x, y) em P . Das Equações 3.2 e 3.3 conclui-se que $L_2 \succ L_1 \succ L'_2$. A função utilidade $u(x, y)$ é contínua e então vale o teorema do valor médio. Logo, pode-se garantir que quando $\pi_y = 0$, existe um valor para π_x^* , tal que $0 < \pi_x^* < x_1 - x_0$, e $L'_2 \sim L_1$. Analogamente, quando $\pi_x = 0$ pode-se garantir que existe um π_y^* , tal que $0 < \pi_y^* < y_1 - y_0$, e $L'_2 \sim L_1$. Portanto, para qualquer $\pi_x \in [0, \pi_x^*]$, existe um $\pi_y \in [0, \pi_y^*]$ tal que $L'_2 \sim L_1$.

Para demonstrar a suficiência argumenta-se que se existe um par (π_x, π_y) que satisfaz a proposição $L'_2 \sim L_1$, então

$$\frac{1}{2}u(x_0, y_0) + \frac{1}{2}u(x_1, y_1) = \frac{1}{2}u(x_1 - \pi_x, y_0) + \frac{1}{2}u(x_0, y_1 - \pi_y) < \frac{1}{2}u(x_1, y_0) + \frac{1}{2}u(x_0, y_1) \quad (3.4)$$

donde a desigualdade segue do fato que $u_x(x, y) > 0$ e $u_y(x, y) > 0$ para todo $(x, y) \in P$. O lado direito da desigualdade na Equação 3.4 é a utilidade esperada de L_2 e portanto, $L_2 \sim L_1$.

OEA

Pode-se estender a definição de Π para os valores de x_0, x_1, y_0 e y_1 onde $L_2 \sim L_1$ e o decisor paga zero de Π , *i.e.*, $\pi_x = \pi_y = 0$.

Intuitivamente, o prêmio pelo risco multidimensional Π é a quantidade dos atributos X e Y que o decisor tem que desistir de x_1 e y_1 , na loteria L_2 , de modo a tornar L_2 equivalente à loteria L_1 . É um prêmio pelo risco no sentido que o decisor prefere desistir dessas quantidades em X e Y à se arriscar em uma loteria “tudo ou nada”.

Analogamente, se o decisor é propenso ao risco multidimensional, pode-se estabelecer um teorema com $L_1 \succ L_2$ e os prêmios pelo risco seriam adicionados a x_0 e y_0 , de modo a fazer com que a relação $L_1 \sim L_2''$ seja satisfeita.

3.3 Comparando atitudes de aversão ao risco

A partir da definição de prêmio pelo risco multivariado da Seção 3.2, pode-se comparar a propriedade de aversão ao risco multidimensional entre duas funções utilidades $u(x, y)$ e $v(x, y)$. Não serão discutidos em detalhes os casos onde a comparação é intuitiva e óbvia, tais como quando $u(x, y)$ é *p.r.m* e $v(x, y)$, sendo *a.r.m* ou *n.r.m*, terá uma aversão ao risco multidimensional maior que $u(x, y)$, *i.e.*, $u_{xy}(x, y) \geq 0$ e $v_{xy}(x, y) = 0$ ou $v_{xy}(x, y) \leq 0$ são condições suficientes para este caso. Os casos interessantes ocorrem quando $u(x, y)$ e $v(x, y)$ são ambos *a.r.m* ou *p.r.m*. Em particular, se $u(x, y)$ e $v(x, y)$ são ambos *a.r.m*, diz-se que $v(x, y)$ tem uma maior *a.r.m* que $u(x, y)$ se o Π pago por um decisor usando $v(x, y)$ é maior que o prêmio pelo risco usando $u(x, y)$ para a mesma loteria. Mais precisamente, seja $u(x, y)$ averso ao risco multidimensional, então para quaisquer x_0, x_1, y_0 e y_1 como definidos anteriormente, pode encontrar um par (π_x, π_y) de modo que $L_2' \sim L_1$, ou seja,

$$\frac{1}{2} \{u(x_0, y_1 - \pi_y) + u(x_1 - \pi_x, y_0)\} = \frac{1}{2} \{u(x_0, y_0) + u(x_1, y_1)\}. \quad (3.5)$$

Então, diz-se que $v(x, y)$ é mais *a.r.m* que $u(x, y)$ se e somente se $L_2' \succ L_1$ quando comparado usando $v(x, y)$, ou seja,

$$\frac{1}{2} \{v(x_0, y_1 - \pi_y) + v(x_1 - \pi_x, y_0)\} \geq \frac{1}{2} \{v(x_0, y_0) + v(x_1, y_1)\} \quad (3.6)$$

para todo x_0, x_1, y_0 e y_1 e todos os pares (π_x, π_y) tais que a Equação 3.5 é válida. Fica claro na Equação 3.6 que $v(x, y)$ deve ser *a.r.m*. Pode-se agora demonstrar o seguinte teorema, o qual estabelece condições suficientes para $v(x, y)$ ser mais *a.r.m* que $u(x, y)$.

TEOREMA 3.3 (RICHARD [39]) Seja $u(x, y)$ tal que $u_x(x, y) > 0$, $u_y(x, y) > 0$ e $u_{xy}(x, y) \leq 0$ para todo $(x, y) \in P$. Para todo $(x, y) \in P$ seja $v(x, y)$ tal que $v_x(x, y) > 0$ e $v_y(x, y) > 0$;

$$-v_{xx}/v_x \geq -u_{xx}/u_x; \quad (3.7a)$$

$$-v_{xy}/v_x \geq -u_{xy}/u_x; \quad (3.7b)$$

$$-v_{yy}/v_y \geq -u_{yy}/u_y; \quad (3.7c)$$

$$-v_{xy}/v_y \geq -u_{xy}/u_y. \quad (3.7d)$$

Para qualquer $x_0 < x_1$ e $y_0 < y_1$ dado, e seja o par (π_x, π_y) um Π tal que

$$u(x_1 - \pi_x, y_0) + u(x_0, y_1 - \pi_y) = u(x_0, y_0) + u(x_1, y_1) \quad (3.8)$$

Então

$$v(x_1 - \pi_x, y_0) + v(x_0, y_1 - \pi_y) \geq v(x_0, y_0) + v(x_1, y_1). \quad (3.9)$$

DEMONSTRAÇÃO 3.3 Seja $\psi(x, y) \equiv v_x(x, y)/u_x(x, y) > 0$ de modo que as condições 3.7a e 3.7b implicam em $\psi_x(x, y) \leq 0$ e $\psi_y(x, y) \leq 0$. Analogamente, seja

$$\phi(x, y) \equiv v_y(x, y)/u_y(x, y) > 0$$

, de modo que as condições 3.7c e 3.7d, respectivamente, implicam $\phi_x(x, y) \leq 0$ e $\phi_y(x, y) \leq 0$.

Além disso, note-se que a Equação 3.8 é equivalente a

$$\int_{y_0}^{y_1 - \pi_y} \int_{x_0}^{x_1 - \pi_x} -u_{xy}(x, y) dx dy = \int_{x_1 - \pi_x}^{x_1} u_x(x, y_1) dx + \int_{y_1 - \pi_y}^{y_1} u_y(x_1 - \pi_x, y) dy. \quad (3.10)$$

Inicialmente, assuma que $\psi(x_1 - \pi_x, y_1 - \pi_y) \geq \phi(x_1 - \pi_x, y_1 - \pi_y)$. Segue então de 3.7b:

$$\begin{aligned} \int_{y_0}^{y_1 - \pi_y} \int_{x_0}^{x_1 - \pi_x} -v_{xy}(x, y) dx dy &\geq \int_{y_0}^{y_1 - \pi_y} \int_{x_0}^{x_1 - \pi_x} -\psi(x, y) u_{xy}(x, y) dx dy \\ &\geq \psi(\mathbf{w}) \int_{y_0}^{y_1 - \pi_y} \int_{x_0}^{x_1 - \pi_x} -u_{xy}(x, y) dx dy, \end{aligned} \quad (3.11)$$

pois $\psi_x \leq 0$ e $\psi_y \leq 0$. Substituindo a Equação 3.10 na Equação 3.11:

$$\begin{aligned}
 \int_{y_0}^{y_1 - \pi_y} \int_{x_0}^{x_1 - \pi_x} -v_{xy}(x, y) dx dy &\geq \psi(x_1 - \pi_x, y_1 - \pi_y) \int_{x_1 - \pi_x}^{x_1} u_x(x, y_1) dx \\
 &\quad + \psi(x_1 - \pi_x, y_1 - \pi_y) \int_{y_1 - \pi_y}^{y_1} u_y(x_1 - \pi_x, y) dy \\
 &\geq \int_{x_1 - \pi_x}^{x_1} \psi(x, y_1) u_x(x, y_1) dx \\
 &\quad + \phi(x_1 - \pi_x, y_1 - \pi_y) \int_{y_1 - \pi_y}^{y_1} u_y(x_1 - \pi_x, y) dy \\
 &\geq \int_{x_1 - \pi_x}^{x_1} v_x(x, y_1) dx + \int_{y_1 - \pi_y}^{y_1} v_y(x_1 - \pi_x, y) dy. \quad (3.12)
 \end{aligned}$$

Integrando-se o lado direito da Equação 3.12 e cancelando-se os termos iguais chega-se a Equação 3.9.

Para o caso em que $\psi(x_1 - \pi_x, y_1 - \pi_y) < \phi(x_1 - \pi_x, y_1 - \pi_y)$, basta trocar ψ por ϕ e a demonstração é análoga.

OEA

Usando-se o Teorema 3.3 pode-se comparar a aversão ao risco multidimensional de dois decisores. As condições 3.7a e 3.7c impõem que o primeiro decisor, com função utilidade v , possua uma aversão ao risco (no sentido de Pratt [35]) maior em relação aos atributos X e Y do que o decisor com utilidade u . As condições 3.7b e 3.7d levam em conta as derivadas parciais cruzadas v_{xy} e u_{xy} , essenciais na definição de aversão ao risco multidimensional. Se as condições 3.7 são satisfeitas, então qualquer par (π_x, π_y) de risco pelo prêmio multidimensional para o segundo decisor não será, em geral, um risco pelo prêmio multidimensional para o primeiro decisor. Intuitivamente, o primeiro decisor é mais *a.r.m* que o segundo porque ele está disposto a pagar um valor “maior” de prêmio pelo risco multidimensional para qualquer situação de risco multidimensional.

Argumentando de forma similar ao Teorema 3.3, pode-se mostrar que as condições 3.7 com as desigualdades revertidas são suficientes para $v(x, y)$ ser mais *p.r.m* que $u(x, y)$ se $u(x, y)$ é *p.r.m*.

3.4 Operações que preservam a aversão ao risco multidimensional

Sabe-se da Seção 3.1 que uma função utilidade é dita *a.r.m* se e somente se $u_{xy} \leq 0$ para todo $(x, y) \in P$. Nesta seção mostra-se que algumas operações sobre tais funções utilidade preservam a propriedade *a.r.m*. Este conhecimento será útil quando for preciso caracterizar se

funções utilidade possuem a propriedade de *a.r.m* e no estabelecimento de fórmulas razoavelmente simples.

TEOREMA 3.4 (RICHARD [39]) Se u^1, u^2, \dots, u^n são *a.r.m* para todo (x, y) em P e se b_1, b_2, \dots, b_n são constantes positivas, então

$$u = a + \sum_{i=1}^n b_i u^i$$

é *a.r.m*.

DEMONSTRAÇÃO 3.4 Segue diretamente da diferenciação de u .

OEΔ

Teorema 3.4 no caso $n = 1$ mostra que transformações lineares com $b_1 > 0$ não modificam a propriedade *a.r.m*. Além disso, as transformações lineares positivas também não mudam a relação de aversão ao risco multidimensional entre duas funções utilidades $u(x, y)$ e $v(x, y)$, como mostrado no teorema a seguir:

TEOREMA 3.5 (RICHARD [39]) Se $u(x, y)$ é *a.r.m* então para $v(x, y)$ satisfazendo as condições 3.7, $v(x, y)$ mais *a.r.m* que $u(x, y)$, então $v' = a_1 + b_1 v(x, y)$, $b_1 > 0$, tem maior *a.r.m* que $u' = a_2 + b_2 u(x, y)$, $b_2 > 0$.

DEMONSTRAÇÃO 3.5 Tem-se do teorema anterior que $u'(x, y)$ é *a.r.m*. O resto da demonstração decorre do fato de que $-u'_{xy}/u'_x = -u_{xy}/u_x$.

OEΔ

TEOREMA 3.6 (RICHARD [39]) Se $u(x, y)$ é *a.r.m* e se $f(\cdot)$ é côncava, então $v(x, y) = f(u(x, y))$ é *a.r.m*.

DEMONSTRAÇÃO 3.6 Tem-se que $v_{xy} = f' u_{xy} + f'' u_x u_y$ e, por hipótese, $u_{xy} \leq 0$, $u_x > 0$ e $u_y > 0$. Logo $v_{x,y} \leq 0$ se $f(\cdot)$ é côncava.

OEΔ

3.5 Extensão para o caso n -dimensional

Nesta seção generaliza-se os Teoremas 3.2 a 3.6 para o caso em que o decisor leva em consideração n atributos, a saber p_1, p_2, \dots, p_n . Assume-se que o conjunto de todos os possíveis valores assumidos por p_i é um intervalo fechado da reta real denotado por P_i . Seja $P = P_1 \times P_2 \times \dots \times P_n$, produto cartesiano dos P_i , o conjunto das n -nuplas (p_1, p_2, \dots, p_n) com $p_i \in P_i$ para cada i . Um elemento de P é chamado uma conseqüência e denotado por $p = (p_1, p_2, \dots, p_n)$. Como a aversão ao risco multidimensional é uma propriedade binária, tem-se que focar nos atributos p_i e p_j para $i \neq j$. Chama-se então p_{ij} um elemento do conjunto $P = P_1 \times P_2 \times \dots \times P_{i-1} \times P_{i+1} \times \dots \times P_{j-1} \times P_{j+1} \times \dots \times P_n$, pode-se dessa forma escrever uma conseqüência como $p = (p_i, p_j, p_{ij})$.

Assume-se que existe uma função utilidade $u(p_1, p_2, \dots, p_n) = u(p)$ definida no conjunto das conseqüências³. Assim como na Seção 3.1, supõe-se que a função utilidade é suficientemente regular para justificar as demonstrações, em geral isto significa duas vezes continuamente diferenciável em todas as variáveis, com primeira derivada positiva.

A definição da *a.r.m* pode ser estendida, para qualquer $i, j = 1, \dots, n$ e $i \neq j$, substituindo-se x por p_i , y por p_j e adicionando p_{ij} nas loterias L_1 e L_2 . O decisor é então *a.r.m* em p_i e p_j se para qualquer $p_{ij} \in P_{ij}$, $L_2 \succsim L_1$ para todo p_i^0, p_i^1, p_j^0 , e p_j^1 (i.e., x_0, x_1, y_0 e y_1). A mesma extensão, *mutatis mutandis*, pode ser feita para os casos *n.r.m* e *p.r.m*.

EXEMPLO 3.2 Considere a função utilidade $u(p_1, p_2, p_3) = -(p_1 p_2)^{-1} + p_3^{\frac{1}{2}}$. Tem-se que

$$u_{p_1 p_2} = -\frac{1}{p_1^2 p_2^2}$$

$$u_{p_1 p_2} = 0$$

$$u_{p_1 p_2} = 0.$$

Portanto, é *a.r.m* em relação a p_1, p_2 ; e *n.r.m* para p_1, p_3 e p_2, p_3 .

Verifica-se que os Teoremas 3.2 a 3.6 continuam válidos, devendo-se apenas fazer as seguintes alterações nos enunciados e demonstrações: substituir x por p_i , y por p_j ; incluir p_{ij} em todos os pares (x, y) , e.g., (x_0, y_1) deve ser (p_i^0, p_j^1, p_{ij}) ; a expressão aversão ao risco multidimensional, *a.r.m*, deve ser substituída por *a.r.m* em p_i, p_j ; todas as funções em (x, y) agora são funções de (p_i, p_j, p_{ij}) ; e $P = P_1 \times P_2 \times \dots \times P_n$.

³Para ser mais preciso, a utilidade é definida sobre distribuições de probabilidade sobre vetores do espaço das conseqüências [5].

3.6 Independência em utilidade e a aversão ao risco multidimensional

Estuda-se as implicações da aversão ao risco multidimensional na estrutura da função utilidade. Denota-se por $p_{i\blacktriangleright}$ um elemento do conjunto $P_{i\blacktriangleright} = P_i \times P_{i+1} \times \cdots \times P_n$ e por $p_{\blacktriangleleft i}$ um elemento do conjunto $P_{\blacktriangleleft i} = P_1 \times P_2 \times \cdots \times P_{i-1}$. Pode-se agora escrever uma consequência como $(p_{\blacktriangleleft i}, p_{i\blacktriangleright})$ e a função utilidade $u(p) = u(p_{\blacktriangleleft i}, p_{i\blacktriangleright})$. Tomando-se $p_{\blacktriangleleft i}$ igual a um valor específico $p_{\blacktriangleleft i}^0$, pode-se falar em uma função utilidade condicional para $p_{i\blacktriangleright}$ dado $p_{\blacktriangleleft i} = p_{\blacktriangleleft i}^0$, sendo esta uma transformação linear positiva qualquer de $u(p_{\blacktriangleleft i}^0, p_{i\blacktriangleright})$.

DEFINIÇÃO 3.2 (INDEPENDÊNCIA EM UTILIDADE) Dado uma função utilidade $u(p_{\blacktriangleleft i}, p_{i\blacktriangleright})$, diz-se que $P_{i\blacktriangleright}$ é *independente em utilidade* de (i.u.) $P_{\blacktriangleleft i}$ se a preferência do decisor por loterias em $P_{i\blacktriangleright}$ para um valor fixo $p_{\blacktriangleleft i}^0$ em $P_{\blacktriangleleft i}$ é a mesma para qualquer valor $p_{\blacktriangleleft i}^0$.

A definição anterior implica na seguinte estrutura para funções utilidades se $P_{i\blacktriangleright}$ i.u. $P_{\blacktriangleleft i}$, fixado qualquer $p_{\blacktriangleleft i}^0$:

$$u(p) = a(p_{\blacktriangleleft i}) + b(p_{\blacktriangleleft i}) \cdot u(p_{\blacktriangleleft i}^0, p_{i\blacktriangleright}) \quad (3.13)$$

em que $b(\cdot) > 0$. Note-se que $P_{i\blacktriangleright}$ i.u. $P_{\blacktriangleleft i}$ não implica em $P_{\blacktriangleleft i}$ i.u. $P_{i\blacktriangleright}$.

Meyer [39, 48] obteve o seguinte resultado. Para alguma ordem dos atributos P_1, \dots, P_n ter-se-á

$$P_{k\blacktriangleright} \text{ i.u. } P_{\blacktriangleleft k}, \text{ para } k = 2, \dots, n \quad \text{e} \quad P_{n\blacktriangleright} \text{ i.u. } P_{\blacktriangleleft n} \quad (3.14)$$

se e somente se para algum a e $b > 0$, ou

$$u(p) = a + b \sum_{k=1}^n u_k(p_k), \quad (3.15)$$

em que $u_k(p_k)$ é a função utilidade para o atributo p_k ; ou

$$u(p) = a + b \prod_{k=1}^n u_k(p_k), \quad (3.16)$$

em que cada uma das $u_k(p_k)$ é tal que $u_k(p_k) > 0$ para todo $z_k, k = 1, \dots, n$; ou

$$u(p) = a - b \prod_{k=1}^n (-u_k(p_k)), \quad (3.17)$$

com $u_k(p_k) < 0$ para todo $z_k, k = 1, \dots, n$.

Em economia, especialmente quando se estuda o consumo ou renda em função do tempo, a forma aditiva 3.15 é usada para representar a função utilidade do decisor. Entretanto, esta

estrutura da utilidade possui uma estrutura (desejável?) que implica em *a.r.m* em P_i e P_j para todo $i \neq j$.

TEOREMA 3.7 (RICHARD [39]) Se as condições 3.14 são satisfeitas, tal que a função utilidade pode ser escrita em uma das formas 3.15, 3.16 ou 3.17, então o decisor é *n.r.m* para todo P_i e P_j se a utilidade é da forma 3.15; *p.r.m* para todo P_i e P_j se a utilidade é da forma 3.16; e *a.r.m* para todo P_i e P_j se a utilidade é da forma 3.17.

DEMONSTRAÇÃO 3.7 Derivando-se as Equações 3.15, 3.16 ou 3.17 com respeito a p_i e p_j , para i e j quaisquer e $i \neq j$, tem-se:

$$\frac{\partial^2 u}{\partial p_i \partial p_j} = 0, \quad (3.18)$$

$$\frac{\partial^2 u}{\partial p_i \partial p_j} = b \frac{u'_i u'_j}{u_i u_j} \prod_{k=1}^n u_k > 0, \text{ e} \quad (3.19)$$

$$\frac{\partial^2 u}{\partial p_i \partial p_j} = -b \frac{u'_i u'_j}{(-u_i)(-u_j)} \prod_{k=1}^n (-u_k) < 0, \quad (3.20)$$

respectivamente.

OEA

De fato, tem-se o seguinte teorema para utilidades *n.r.m*:

TEOREMA 3.8 (RICHARD [39]) Um decisor é *n.r.m* em P_i e P_j para $i, j = 1, 2, \dots, n$ e $i \neq j$ se e somente se a utilidade é da forma 3.15.

DEMONSTRAÇÃO 3.8 Se a utilidade é da forma 3.15, então mostrou-se no Teorema 3.7 que o decisor é *n.r.m* em P_i e P_j para $i, j = 1, 2, \dots, n$ e $i \neq j$. Se o decisor é *n.r.m* em P_i e P_j tem-se que $u_{p_i p_j} = 0$ para $i, j = 1, 2, \dots, n$ e $i \neq j$. A demonstração será por indução em n .

Para $n = 2$, $u_{p_1 p_2} = 0$ se e somente se $u(p_1, p_2) = u_1(p_1) + u_2(p_2)$. Assuma que a proposição é verdadeira para $n - 1$ atributos e considere $u(p)$ para n atributos. Por hipótese, a solução para $u_{p_i p_j} = 0$ para $i, j = 1, 2, \dots, n$ e $i \neq j$ é

$$u(p) = f_1(p_1, p_n) + f_2(p_2, p_n) + \dots + f_{n-1}(p_{n-1}, p_n). \quad (3.21)$$

Para todo $i = 1, \dots, n - 1$ tem-se

$$u_{p_i p_n} = \frac{\partial^2 f_i}{\partial p_i \partial p_n} = 0, \quad (3.22)$$

para todo p_i e p_n . Então

$$f_i(p_i, p_n) = u_i(p_i) + g_i(p_n). \quad (3.23)$$

Substituindo 3.23 em 3.21 e tomando-se

$$u_n(p_n) = \sum_{i=1}^n g_i(p_n), \quad (3.24)$$

completa-se a demonstração.

OEΔ

3.7 Prêmio em probabilidade multidimensional

Estende-se agora o conceito de prêmio em probabilidade de Arrow para o caso multiatributo. Esta generalização foi proposta por H. Levy e A. Levy [45].

Considere a loteria L_h que dá ao decisor h com probabilidade λ e $-h$ com probabilidade $1 - \lambda$, onde $h = (h_1, h_2, \dots, h_n)' \in \mathbb{R}^n$. Seja u uma função utilidade, com derivadas parciais positivas. O vetor h é tal que vale a seguinte afirmação:

$$u(p + h) > u(p) > u(p - h). \quad (3.25)$$

A Condição 3.25 assegura que receber o vetor h é preferível ao *status quo*, o qual é preferível a perder h . Esta condição é trivialmente satisfeita se $h_i > 0$ para $i = 1, \dots, n$, em que h_i representa os componentes do vetor h .

Seguindo a linha de raciocínio de Arrow para o caso monoatributo, H. Levy e A. Levy [45] estendem a idéia para o caso multiatributo. Assumindo que a condição 3.25 é satisfeita, então existe λ tal que o indivíduo fica indiferente entre entrar no jogo L_h ou permanecer com um valor certo p .

Tomando-se a utilidade esperada da loteria e igualando a utilidade de p , tem-se

$$u(p) = \lambda u(p + h) + (1 - \lambda)u(p - h). \quad (3.26)$$

Se o indivíduo é averso ao risco então

$$\lambda > \frac{1}{2} \quad (3.27)$$

Apresenta-se a seguir a matriz de aversão ao risco, extensão para o caso multidimensional da função de aversão ao risco de Arrow-Pratt, proposta por Duncan [10]. Esta matriz será importante no estudo das inter-relações entres os conceitos de aversão ao risco introduzidos anteriormente.

3.8 Matriz de aversão ao risco

A matriz de aversão ao risco, proposta por Duncan [10], é uma generalização da medida de aversão ao risco de Arrow [36] e Pratt [35]. Segundo Kihlstrom e Mirman [42], o principal problema da medida de aversão ao risco de Arrow e Pratt é a limitação de ser definida apenas para utilidades monoatributo:

$$r(x) \equiv -\frac{u''(x)}{u'(x)}$$

Duncan estendeu esta definição de modo a permitir o exame de inter-relações entre os atributos, construindo uma medida mais informativa.

Para introduzir a matriz de aversão ao risco de Duncan, considera-se inicialmente a definição geral de *prêmio pelo risco*.

Define-se como *função prêmio* $\pi(\mathbf{x}, Z)$ a função que associa um vetor $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_n) \in \mathbb{R}^n$ a certa loteria Z e consequências \mathbf{x} . Se $E(Z) = 0$, $\boldsymbol{\pi}$ deve satisfazer a equação:

$$u(\mathbf{x} - \boldsymbol{\pi}) = E[u(\mathbf{x} + Z)]. \quad (3.28)$$

Aqui u é uma função utilidade de von Neumann-Morgenstern que satisfaz as preferências do indivíduo. Note-se que, ao contrário do caso monodimensional, a solução deste problema não é única. Pode-se ter vários prêmios $\boldsymbol{\pi}$ como solução da Equação 3.28. Considere contínuas as derivadas de segunda ordem:

$$u_{ij}(\mathbf{x}) \equiv \frac{\partial^2 u(\mathbf{x})}{\partial x_i \partial x_j}.$$

Então

$$u(\mathbf{x} - \boldsymbol{\pi}) = u(\mathbf{x}) - \sum_{i=1}^n \pi_i u_i(\mathbf{x}) + \frac{1}{2} \sum_{i,j=1}^n \pi_i \pi_j u_{ij}(\mathbf{x} - \theta \boldsymbol{\pi}), \quad (3.29)$$

para algum $\theta, 0 \leq \theta < 1$. Por outro lado, se $\text{var}(Z) = \Sigma \equiv [\sigma_{ij}]$ existe, então

$$\begin{aligned} E[u(\mathbf{x} - \boldsymbol{\pi})] &= E \left[u(\mathbf{x}) - \sum_{i=1}^n Z_i u_i(\mathbf{x}) + \frac{1}{2} \sum_{i,j=1}^n Z_i Z_j u_{ij}(\mathbf{x}) + \sum_{i,j,k=1}^n \mathcal{O}(Z_j Z_j Z_k) \right] \\ &= u(\mathbf{x}) + \frac{1}{2} \sum_{i,j=1}^n \sigma_{ij} u_{ij}(\mathbf{x}) + \mathcal{O}(\text{tr } \Sigma), \end{aligned} \quad (3.30)$$

A Equação 3.28 exige a igualdade destas duas últimas quantidades. Segundo Duncan [10], uma solução aproximada para esta equação é:

$$\mathbf{u}'\boldsymbol{\pi} = -\sum_{i=1}^n \pi_i u_i(\mathbf{x}) = \frac{1}{2} \sum_{i,j=1}^n \sigma_{ij} u_{ij}(\mathbf{x}) = -\frac{1}{2} \text{tr}(U\Sigma), \quad (3.31)$$

sendo $U = [u_{ij}(\mathbf{x})]$ é a matriz Hessiana de u e o vetor $\mathbf{u} = (u_i(\mathbf{x}))$ formado pelas derivadas parciais de u . Note-se que esta equação restringe $\boldsymbol{\pi}$ a um hiperplano n -dimensional.

Denota-se-á por $\hat{\boldsymbol{\pi}}$ qualquer solução da Equação 3.31 e será dita *prêmio aproximado pelo risco*. As soluções $\hat{\boldsymbol{\pi}}$ têm a seguinte forma:

$$\hat{\boldsymbol{\pi}} = -\frac{1}{2}(\mathbf{u}')^{-} \text{tr}(U\Sigma), \quad (3.32)$$

para uma matriz qualquer A , $m \times n$, a matriz A^{-} , $n \times m$, é a inversa generalizada de A , tal que $AA^{-}A = A$. A solução natural é fazer

$$\hat{\pi}_1 = -\frac{1}{2} \sum_i u_i^{-1} u_{ij} \sigma_{ij}, \quad i = 1, \dots, n.$$

que pode ser reescrita da seguinte maneira

$$\hat{\boldsymbol{\pi}} = \frac{1}{2} \text{dg } R\Sigma, \quad (3.33)$$

em que

$$R = [r^{ij}] = \left[-\frac{u_{ij}}{u_i} \right] = [\text{diag } \mathbf{u}]^{-1} U,$$

e, se B é uma matriz $n \times n$, $\text{dg } B$ é um vetor n -dimensional com os elementos da diagonal principal de B (b_{ii}), enquanto que se \mathbf{c} é um vetor n -dimensional, $\text{diag } \mathbf{c}$ é uma matriz $n \times n$ com os elementos da diagonal principal iguais a c_i .

Note-se que quando $i = j$, tem-se o prêmio pelo risco definido por Arrow e Pratt

$$-\frac{1}{2} \frac{u_{ii}}{u_i} \sigma_i^2.$$

Portanto, é apropriado chamar R como *matriz de aversão ao risco*.

Por exemplo, considera a seguinte família de funções utilidades indexadas pelas constantes θ_1 e θ_2 :

$$u(\mathbf{x}) = -\theta_1 (e^{-x_1} + e^{-x_2}) - \theta_2 e^{-x_1 - x_2}$$

Tem-se então $r^{ii} = 1$, independente de θ_1 e θ_2 , enquanto que, para $i \neq j$, tem-se

$$r^{ij}(\mathbf{x}) = \frac{\theta_2}{\theta_1 e^{x_j} + \theta_2}$$

Keeney [41] propôs usar r^{ii} como medida de aversão ao risco. Propôs também a terminologia, para o caso em que r^{ii} é constante, que os atributos X_1 e X_2 são *mutualmente independentes em risco*. Contudo, o exemplo mostrado sugere mostra que isto seria uma simplificação exagerada.

Note-se que a independência em risco definida por Keeney pode ser válida e ainda R não ser diagonal. Por outro lado, Richard [39] aborda à aversão ao risco, como foi visto no início do capítulo, de forma qualitativa, sem definir uma medida. Segundo Richard, um indivíduo é averso ao risco multivariado se prefere loterias com “correlação negativa” às loterias com “correlação positiva.” Como foi visto, isto será o caso quando os elementos fora da diagonal de R forem não negativos. Isto é consistente com a abordagem de Duncan pois, da Equação 3.33, o prêmio aproximado pelo risco é uma função crescente dos coeficiente de correlação para variâncias fixas da loteria.

Os conceitos de aversão ao risco estudados aqui podem ser usados para o estudo do comportamento do decisor e sugerir expressões analíticas para os valores de utilidades eduzidos no próximo capítulo.

Embora este procedimento não seja necessário (e por vezes impossível) para a solução de um problema de decisão, é útil para comparação entre diferentes decisores do comportamento frente ao risco. Note-se que transformações afins de utilidades não alteram as medidas de aversão ao risco estudadas aqui.

A seguir tratar-se-á do problema da educação da função utilidade.

EDUCAÇÃO DE ORDENS LINEARES

Ordnung und Sichtung sind der Anfang der Beherrschung, und der eigentlich furchtbare Feind ist der unbekannte.

— Thomas Mann¹

A TEORIA DA UTILIDADE de von Neumann e Morgenstern, apresentada no capítulo anterior, não faz nenhuma referência a estrutura em si do conjunto das consequências, denotado aqui por \mathcal{C} . Por outro lado, tem como axioma a existência de uma relação de preferência \preceq que induz uma ordem nesse conjunto. Com efeito, a existência desta relação foi a única propriedade utilizada nos métodos de construção da função utilidade apresentados no Capítulo 2.

Os processos de educação da função utilidade supracitados são apresentados na literatura como métodos para educação da função utilidade monodimensional (*e.g.*, *cf.* Farquhar [28]). Neste trabalho, não foram apresentados desta forma, pois a dimensão de \mathcal{C} em nada muda aqueles processos. De fato, em todos os métodos, a única propriedade de \mathcal{C} utilizada foi a existência de uma relação de preferência \preceq que o ordenasse.

Em problemas nos quais o conjunto \mathcal{C} é monodimensional é natural assumir uma monotonicidade na relação de preferência sobre as consequências, por exemplo: dinheiro, participa-

¹Em tradução de H. T. Lowe-Porter: “*order and simplification are the first steps toward the mastery of a subject — the actual enemy is the unknown.*”

ção em mercado, unidades defeituosas, *etc.* Mesmo em atributos como temperatura corporal ou pressão arterial, onde existe um ponto “ideal,” é possível reestruturar ou particionar o conjunto \mathcal{C} de forma a assumir a monotonicidade. Sem perda de generalidade assume-se que as relações de preferências são monotônicas nos atributos.

Nos casos em que $\mathcal{C} = \mathcal{C}_1 \times \cdots \times \mathcal{C}_n$ é multidimensional ou multiatributo, mesmo considerando a existência de uma relação monotônica em cada atributo (\mathcal{C}_i), não se tem, *a priori*, uma ordem sobre todo o conjunto de consequências. Este é o único fato que leva à classificação (equivocada) dos métodos do Capítulo 2 como métodos unidimensionais.

Acredita-se que este problema pode ser contornado de maneira mais natural para uma classe extensa de problemas. Considera-se aqui uma nova engenharia da construção da função utilidade. Os fatos que levaram a esta nova engenharia são os seguintes:

1. A maior parte das aplicações considera atributos discretos e limitados, destes, grande parte numa escala ordinal. Quando não, a discretização é possível e condição *sine qua non* para a educação;
2. A simples hipótese de uma relação monotônica em cada um dos atributos implica uma ordem parcial sobre o conjunto \mathcal{C} , que pode sempre ser estendida para todo conjunto \mathcal{C} , a partir de um questionário;
3. De posse de uma ordem linear sobre \mathcal{C} , os processos de construção do capítulo anterior tornam-se aplicáveis.

Os métodos existentes na literatura constroem inicialmente uma função utilidade para cada atributo e a agregação para construção da utilidade multiatributo é feita por formas rígidas baseada em axiomas de independência entre os atributos. Apenas duas formas de utilidade são concebidas: aditiva e multiplicativa. Estes métodos têm sua origem, com raras exceções, nos trabalhos de Keeney e Raiffa [11, 41, 49]. Este axioma extra, além da forma de como as agregações das utilidades individuais são feitas, têm sido objeto de críticas na literatura.

Além disso, outros fatores aumentam a dificuldade destes métodos — por exemplo, é sabido que loterias com valores extremos introduzem erro na educação da probabilidade de indiferença entre jogos. Então, quando se faz a educação de utilidades monodimensionais, pode-se estar considerando loterias que, na ordem total, estariam distantes, apesar de vizinhas quando se considera apenas uma dimensão.

A seguir, considera-se os fatos apresentados acima para a elaboração de um protocolo de educação para uma relação de preferência que ordena completamente o conjunto \mathcal{C} .

4.1 Relações de preferência

Um princípio central em modelos de preferência, presente em todos os trabalhos, é a noção de **ordem**, formalizada pela transitividade e base para toda otimização. Mesmo em paradigmas de decisão que não envolvem maximização, como equilíbrio de Nash em jogos não-cooperativos [50], assume-se que os indivíduos possuem uma relação de ordem que representa suas preferências.

Como já introduzido (*cf.* Definição 2.1), a estrutura matemática que se relaciona com o estudo das relações de preferências são as relações binárias. Uma **relação binária** é uma ferramenta matemática usada para descrever associações entre elementos de conjuntos. Formalmente, uma relação binária de um conjunto A para um conjunto B , é um subconjunto R do produto cartesiano $A \times B$, *i.e.*, $R \subseteq A \times B$.

No caso de relações de preferência, o interesse recai sobre relações binárias sobre o espaço de consequências \mathcal{C} , (ou seja, relações que são subconjuntos do conjunto \mathcal{C}^2), que admitem certas propriedades: linearidade e transitividade.

A **linearidade** é satisfeita se a relação de preferência é completa, ou seja, se se denota a relação por “ \preceq ”, para qualquer c_i e c_j em \mathcal{C} , tem-se $c_1 \preceq c_2$ ou $c_2 \preceq c_1$. O decisor é capaz de comparar duas consequências quaisquer de \mathcal{C} . Esta propriedade é fundamental para a demonstração do Teorema 2.1, que garante a existência de uma função utilidade que represente esta relação.

O fato de não ser permitido a incomparabilidade entre duas possíveis consequências tem sido alvo de muitas críticas, sobretudo da escola francesa², que tem como principal expoente Bernard Roy [51]. Esta escola afirma que um decisor é, muitas vezes, incapaz de exprimir suas preferências devido principalmente à ambiguidade, à falta de informação e ao conflito entre atributos [51]. Contra estes argumentos, deve-se apenas lembrar que esta escola trata de preferência sobre ações. Este fato mascara a natureza aleatória do problema, introduzindo uma falsa sensação de certeza. Desta forma, incerteza em consequências das ações é interpretada como ambiguidade, falta de informação, *etc.*

Trata-se aqui de preferências sobre consequências, neste caso, a educação de uma ordem

²Allais sugeriu em seus artigos uma divisão entre escola americana e escola francesa, *cf.* [23].

completa sobre as consequências, acessíveis ao decisor, não requer muito esforço. É mais intuitivo e direto ordenar as consequências do que ordenar as ações (sendo equivalentes apenas no estranho mundo onde ações levam a consequências determinísticas). No caso da preferência por um atributo, a linearidade está associada diretamente ao fato das escalas de mensuração serem naturalmente ordenadas, *e.g.*, se o atributo for dinheiro, mais dinheiro é preferível a menos dinheiro. No caso de vários atributos, a ordem pode vir *a priori*, como no protocolo proposto neste trabalho, ou *a posteriori* como é o caso da literatura em geral. Dito isto, a linearidade não engendra nenhuma limitação, *cf.* [5–7, 28].

A **transitividade** é considerada como um axioma de racionalidade, de natureza normativo. A violação da transitividade pode levar ao processo decisório comportamentos cíclicos. Contra o axioma da transitividade existe o argumento de que o decisor possui um “poder limitado de discriminação” [52]. Por exemplo, considere as faixas de cinza na Figura 4.1, as faixas adjacentes de cinza apresentam graus próximos de cinza, embora seja clara a distinção entre as faixas extremas [5, p. 49]. O decisor, por possuir um poder limitado de discriminação, dirá que é indiferente entre as faixas adjacentes e não entre as faixas dos extremos, violando a transitividade. Isto acontece também com pequenas diferenças de peso.

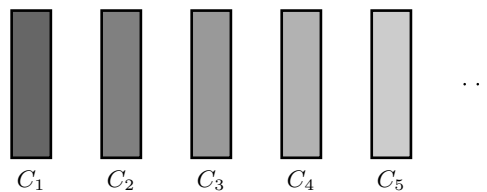


Figura 4.1: *Faixas adjacentes com tons de cinza similares, contudo as faixas extremas com tons opostos.*

Outro exemplo onde acontece a violação da transitividade é o paradoxo de Allais, apresentado na Seção 2.3.2. Nos dois exemplos, a transitividade é violada em condições extremas, em armadilhas psicológicas que não acontecem em casos práticos.

A estruturação do problema de decisão e a preparação para educação das preferências, com as devidas escolhas de escalas de mensuração para os atributos, contribuem sobremaneira para evitar comportamentos como os descritos anteriormente.

Estuda-se a seguir o problema da educação das preferências.

4.2 Estrutura do conjunto de consequências

Independente do contexto em que o problema de decisão se encontra, seja medicina, economia ou engenharia, *etc.* as variáveis que representam as consequências podem ser combinação dos atributos mais diversos: qualidade de vida, tempo de sobrevivência, valor monetário, tempo para lazer, probabilidade de falha de um determinado sistema, *etc.*

Qualquer que seja o problema em questão, os atributos são representados em escalas de medidas. Cada escala possui um certo número de operações possíveis sobre os valores dos atributos, que vai desde relações de pertinência às operações aritméticas [5, 53].

Uma escala de medida deve atender as seguintes propriedades:

Exaustividade: deve conseguir representar todos os fatos possíveis (abrangência);

Exclusividade: cada fato possui uma única representação (coerência).

Existem basicamente quatro escalas, propostas por Stevens [54] no seu artigo seminal “*On the Theory of Scales of Measurement.*” A saber:

Escala nominal — Nível mais elementar de mensuração. Utiliza números ou nomes para identificação ou classificação. Não permite nenhuma operação aritmética. Principal função é a distinção.

Escala ordinal — Além da distinção entre elementos, é enriquecida com uma relação de ordem entre os elementos.

Escala intervalar — Além de distinção e ordem, introduz uma regra de combinação entre elementos na forma de uma “distância.”

Escala de razão — A mais completa das escalas, acrescenta-se à escala intervalar uma origem. Todas as operações aritméticas são válidas.

Quando se pretende mensurar vários atributos de diferentes tipos, como é caso de interesse deste trabalho, várias escalas de mensuração podem ser necessárias. Esta “escala multidimensional” ocupa-se de maneiras de se combinar diferentes atributos para formar um único índice de mensuração. Note-se a semelhança entre escalas e utilidades. Uma utilidade é, portanto, uma escala de mensuração exclusiva (pois representa apenas um indivíduo) para “medir” as preferências do decisor. Além disso, como foi visto no Capítulo 2, a utilidade é representada por uma escala intervalar.

Sejam $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_n$ os conjuntos dos possíveis valores dos n atributos. Define-se o conjunto das consequências \mathcal{C} como o produto cartesiano dos atributos, *i.e.*, $\mathcal{C} \equiv \mathcal{C}_1 \times \mathcal{C}_2 \times \dots \times \mathcal{C}_n$. Uma consequência particular $c \in \mathcal{C}$ é dada por $c = (c^1, c^2, \dots, c^n)$, em que c^1, c^2, \dots, c^n são elementos correspondentes de $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_n$, respectivamente.

Considere a seguinte hipótese:

Hipótese de Monotonicidade: As preferências do decisor são monotônicas em relação a cada um dos atributos \mathcal{C}_i , $i = 1, 2, \dots, n$. Isto é, se c_1^i e c_2^i são elementos de \mathcal{C}_i tais que $c_1^i \leq c_2^i$ em sua escala de medida, então vale a relação de preferência $c_1^i \preceq c_2^i$, para todo $i = 1, 2, \dots, n$.

Esta é um hipótese corriqueira em economia [15, 55] (como sugere o nome “bens” ou “goods”). Note-se que esta hipótese é relativamente fraca, pois se um atributo não se comporta desta maneira, pode-se aplicar uma transformação ou partição de modo que se tenha um atributo para o qual a hipótese seja válida. Por exemplo, ao invés de se trabalhar com o valor do *deficit* (d), em percentual, de energia no setor elétrico, pode-se trabalhar com o percentual atendido, ou seja, $100 - d$.

Caso um atributo seja representado numa escala nominal, deve-se inicialmente ordenar este atributo. Considera-se neste trabalho que os atributos são naturalmente ordenados ou foram previamente ordenados.

Considera-se também os atributos \mathcal{C}_i , $i = 1, 2, \dots, n$, são discretos e finitos (logo, limitados). Tem-se, dessa forma, um “pior” e um “melhor” elemento para cada atributo. Denota-se esses elementos por c_1^i (pior) e $c_{n_i}^i$ (melhor), onde $n_i = \|\mathcal{C}_i\|$, e $i = 1, 2, \dots, n$. O número de possíveis consequência $c \in \mathcal{C}$, *i.e.*, a cardinalidade de \mathcal{C} , é dado por:

$$\|\mathcal{C}\| = \prod_{i=1}^n \|\mathcal{C}_i\| = \prod_{i=1}^n n_i.$$

Considere o caso no qual existem dois atributos, *e.g.*, saúde *versus* riqueza, custo *versus deficit* ou ainda salário *versus* tempo livre. Tem-se então o seguinte problema:

PROBLEMA 4.1 Dado o espaço de consequências $\mathcal{C}_1 \times \mathcal{C}_2 \subset \mathbb{R}^2$, sendo o eixo \mathcal{C}_1 particionado em n_1 intervalos e o eixo \mathcal{C}_2 em n_2 intervalos, como encontrar uma ordem linear sobre esses $n_1 \times n_2$ elementos, que satisfaça as preferências do decisor?

Resolvido este problema, *i.e.*, encontrada esta ordem linear, pode-se tratar do problema de educação da função utilidade multiatributo com os métodos monodimensionais apresentados no Capítulo 2. Em particular, pode-se utilizar o método das faixas sobrepostas, proposto em Campello de Souza [5].

DEFINIÇÃO 4.1 Denota-se por c_{jk} o par ordenado $(c_j^1, c_k^2) \in \mathcal{C}_1 \times \mathcal{C}_2$, um elemento do espaço das consequências. *cf.* Figura 4.2.

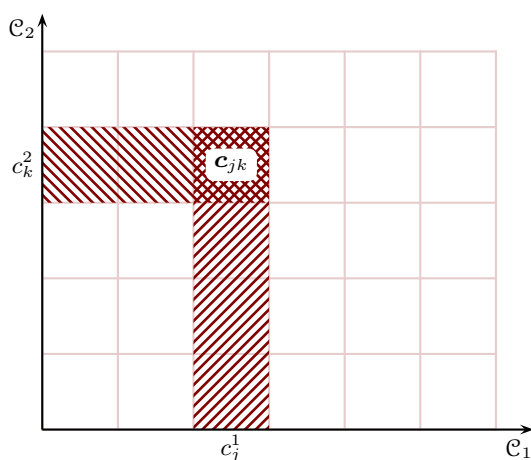


Figura 4.2: Exemplo de partição do plano de dois atributos \mathcal{C}_1 e \mathcal{C}_2 .

Note-se que algumas relação de preferência são induzidas pela monotonicidade dos atributos. Sabe-se que o elemento menos preferível é o c_{00} e o mais preferível é o elemento $c_{n_1 n_2}$. Para todo $c_{jk} \in \mathcal{C}$, tem-se $c_{00} \preceq c_{jk} \preceq c_{n_1 n_2}$.

A seguir estuda-se estas relações parciais induzidas pela monotonicidade.

4.3 Ordem parcial

Trata-se agora em detalhe do conceito de ordem. Referências clássicas são Alencar [56], Davey e Priestley [57] e Birkhoff [58].

Considere a seguinte definição:

DEFINIÇÃO 4.2 Chama-se **relação de ordem** em \mathcal{C} , ou apenas **ordem** em \mathcal{C} , toda relação \mathcal{R} em \mathcal{C} que é *reflexiva, anti-simétrica e transitiva*.

Em outros termos, uma relação \mathcal{R} em \mathcal{C} é uma *relação de ordem* em \mathcal{C} , ou simplesmente uma *ordem* em \mathcal{C} , se e somente se, possui as três seguintes propriedades:

(Reflexiva) $\forall c \in \mathcal{C} \Rightarrow c \mathcal{R} c$

(Antissimétrica) $\forall c, c' \in \mathcal{C}, c \mathcal{R} c' \text{ e } c' \mathcal{R} c \Rightarrow c = c'$

(Transitiva) $\forall c, c', c'' \in \mathcal{C}, c \mathcal{R} c' \text{ e } c' \mathcal{R} c'' \Rightarrow c \mathcal{R} c''$

Se \mathcal{R} é uma relação de ordem em \mathcal{C} , também se diz que \mathcal{C} é ordenado por \mathcal{R} ou que \mathcal{R} ordena \mathcal{C} e que o par $(\mathcal{C}, \mathcal{R})$ é um conjunto ordenado.

Diz-se que uma relação de ordem \mathcal{R} é **parcial** se existe pelo menos um par de elementos c' e c'' em \mathcal{C} tais que estes não são comparáveis na relação \mathcal{R} . Ou seja, não vale $c' \mathcal{R} c''$ nem $c'' \mathcal{R} c'$. Dito de outra forma, \mathcal{R} é uma **ordem parcial** se é reflexiva, assimétrica e transitiva, mas **não** é linear.

Considere a seguinte relação \preceq , construída no conjunto das consequências \mathcal{C} , juntos formam o conjunto parcialmente ordenado $\mathcal{C} = (\mathcal{C}, \preceq)$:

$$(r^1, r^2, \dots, r^n) \preceq (s^1, s^2, \dots, s^n) \Leftrightarrow r^i \leq s^i, \quad (4.1)$$

para $r^i, s^i \in \mathcal{C}_i$ e todo $i = 1, 2, \dots, n$. Além disso, é estritamente preferível se vale (4.1) e, para algum j ($1 \leq j \leq n$), a desigualdade estrita $r^j < s^j$ é satisfeita³. Na Figura 4.3 ilustra-se estas relações para o caso de dois atributos, com $n_1 = 6$ e $n_2 = 5$, usando-se como referência a consequência c_{34} (não incluindo as relações implicadas pela transitividade da relação).

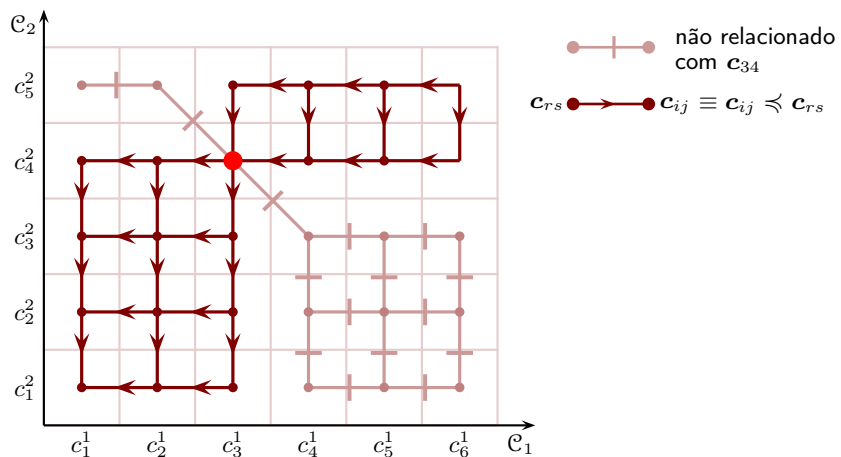


Figura 4.3: Exemplo de partição do plano de dois atributos e relações entre c_{34} e demais consequências.

³As relações “ \leq ” e “ $<$ ” representam as ordenações naturais nas devidas escalas de mensuração.

Uma maneira mais adequada para representação de relações de ordem é o diagrama de Hasse⁴ da relação. Neste diagrama, uma versão simplificada da ordem, representa-se cada elemento de \mathcal{C} como um vértice e desenha-se um segmento de linha entre dois elementos c e c' de \mathcal{C} , se $c \preceq c'$ e não existe c'' tal que $c \preceq c'' \preceq c'$. Diz-se então que c' **cobre** c , ou c' é o sucessor imediato de c , e desenha-se c' acima de c . Note-se que nem todos os elementos da relação de ordem são representados por segmentos de linhas ou qualquer outro símbolo. Por exemplo, reflexividade e transitividade ficam subentendidos.

O diagrama de Hasse para o exemplo anterior, com dois atributos $\mathcal{C}_1, \mathcal{C}_2$ com cardinalidades $n_1 = 6$ e $n_2 = 5$, é apresentado na Figura 4.4.

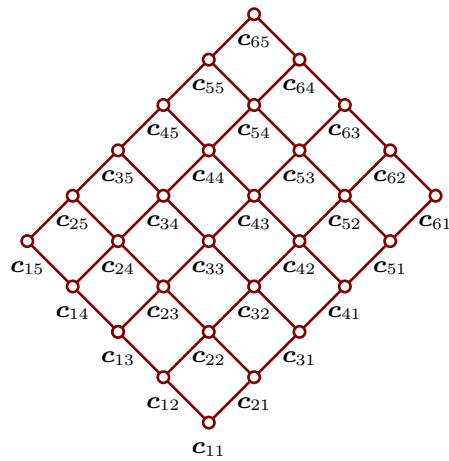


Figura 4.4: Diagrama de Hasse de $\mathcal{C} = (\mathcal{C}, \preceq)$ quando $n = 2$, $n_1 = 6$ e $n_2 = 5$.

Denote-se por $\mathcal{C} = (\mathcal{C}, \preceq)$ um conjunto parcialmente ordenado.

DEFINIÇÃO 4.3 (ELEMENTO MINIMAL) Um elemento c de \mathcal{C} é chamado **minimal** de \mathcal{C} se o único elemento c' tal que $c' \preceq c$ é $c' = c$.

Dito de uma outra maneira, c é minimal se, no diagrama de Hasse de \mathcal{C} , não existe segmento de linha abaixo de c .

Define-se um elemento **maximal** de um conjunto parcialmente ordenado \mathcal{C} de maneira análoga. Note-se também que um elemento pode ser, ao mesmo tempo, maximal e minimal. O elemento maximal e minimal do conjunto ordenado \mathcal{C} representado pelo diagrama de Hasse apresentado na Figura 4.4, é então c_{11} e c_{65} , respectivamente.

⁴Em homenagem a Helmut Hasse, matemático alemão que introduziu esta representação em seu livro *Höhere Algebra*, 1926, para facilitar a solução de equações polinomiais.

Neste modelo, o elemento minimal do conjunto \mathcal{C} corresponde à consequência menos desejada do espaço das consequências.

A estrutura do conjunto parcialmente ordenado \mathcal{C} foi completamente induzida pela hipótese de monotonicidade dos atributos. O objetivo é, contudo, construir uma ordem linear que representa as preferências do decisor. Até o momento tem-se uma ordem parcial \preceq sobre o conjunto das consequências, precisa-se estender esta ordem parcial, de modo que qualquer par de elementos em \mathcal{C} seja comparável. Tratar-se-á de extensões de conjuntos parcialmente ordenados na próxima seção.

4.4 Extensão linear de uma ordem parcial

Uma ordem parcial \preceq , em um conjunto \mathcal{C} , é uma **extensão** de uma ordem parcial \preceq , também em \mathcal{C} , se para todo par de elementos c e c' em \mathcal{C} que $c \preceq c'$, então vale também a relação $c \preceq c'$. Posto isto, uma **extensão linear** de \preceq é uma extensão que é uma ordem linear ou total.

Considere o caso em que se tem dois atributos e cada um deles duas possibilidades, como representado na Figura 4.5. Neste caso, há duas consequências não comparáveis na ordem parcial induzida pela monotonicidade, c_{12} e c_{21} . Existem duas possíveis extensões lineares para esta ordem parcial. Considerando a monotonicidade, reduziu-se o problema de uma busca entre $(n_1 \times n_2)! = 24$ ordens para apenas duas.

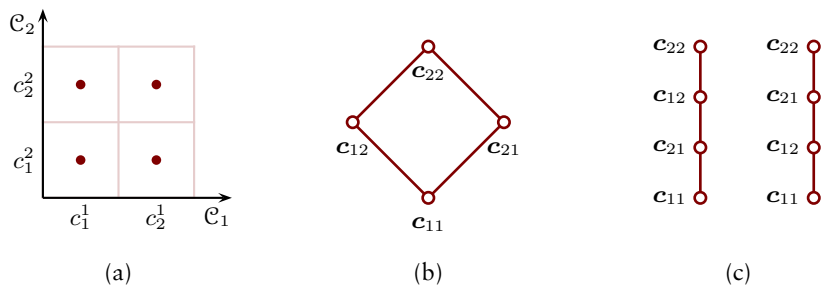


Figura 4.5: Exemplo de partição de um espaço de consequência com dois atributos. (a) Plano representando as consequências. (b) Diagrama de Hasse da ordem parcial induzida. (c) Extensões lineares da ordem parcial induzida.

Denota-se aqui por \preceq tanto a ordem parcial quanto sua extensão linear. A distinção se fará segundo o contexto, usando-se notações diferentes em casos limites.

Considere a seguinte definição:

DEFINIÇÃO 4.4 (ORDEM LEXICOGRÁFICA) A relação de preferência \preceq é dita **lexicográfica** se

$$(r^1, r^2, \dots, r^n) \preceq (s^1, s^2, \dots, s^n) \Leftrightarrow s^{\sigma(i)} \geq r^{\sigma(i)} \quad \text{e} \quad s^{\sigma(m)} > r^{\sigma(m)}$$

para uma dada permutação σ dos atributos e $i = 1, \dots, m; m \leq n$.

Preferências lexicográficas podem ser entendidas como “preferências” ou prioridades sobre atributos. Duas consequências em \mathcal{C} são comparadas a partir do atributo com prioridade mais alta. Caso haja diferença entre as consequências nesse atributo, escolhe de forma a maximiza o atributo prioritário, caso contrário segue-se para o próximo em ordem de prioridade e assim por diante.

Durante o processo de educação é interessante descobrir se o decisor é capaz de estabelecer prioridade entre os atributos. Embora não seja uma condição necessária, a existência de prioridades pode simplificar o processo de educação da ordem linear e da função utilidade, pois pode-se combinar conjuntos de atributos onde seja clara, para o decisor, suas prioridades em relação aos atributos.

Para o caso de dois atributos tem-se apenas duas permutação possíveis (e também duas ordens possíveis). Pode-se dar prioridade a \mathcal{C}_1 ou \mathcal{C}_2 . Estes são os casos apresentados na Figura 4.5 (c). Existem apenas duas extensões lineares e estas são exatamente as preferências lexicográficas.

A existência de extensões de ordens parciais é garantida por um teorema de existência devido a Szpilrajn⁵ [59], em 1930. Resolvida a existência, considera-se agora o número de soluções para este problema.

Denota-se por $\mathfrak{M}(\mathcal{C})$ o número de extensões lineares de um conjunto parcialmente ordenado $\mathcal{C} = (\mathcal{C}, \preceq)$. Uma maneira de calcular este número é a seguinte fórmula de recorrência devido a Sidorenko [60, 61]:

$$\mathfrak{M}(\mathcal{C}, \preceq) = \sum_{m \in M} \mathfrak{M}(\mathcal{C} - m, \preceq), \tag{4.2}$$

em que M denota o conjunto de elementos mínimos de \mathcal{C} e $\mathfrak{M}(\emptyset, \preceq) \equiv 1$. É usual a notação $\mathfrak{M}^m \equiv \mathfrak{M}(\mathcal{C} - m, \preceq)$.

⁵Segundo o Szpilrajn [59, p. 387], outra demonstração para o mesmo teorema de existência devida a Banach, Kuratowski e Tarski era conhecida na época, mas não publicada.

Para o caso mais simples, com dois atributos e duas possibilidades por atributo, *cf.* Figura 4.5, tem-se:

$$\begin{aligned}
 \mathfrak{M}(\mathcal{C}, \preccurlyeq) &= \mathfrak{M}^{c_{11}} \\
 &= \mathfrak{M}^{c_{11}, c_{12}} + \mathfrak{M}^{c_{11}, c_{21}} \\
 &= \mathfrak{M}^{c_{11}, c_{12}, c_{21}} + \mathfrak{M}^{c_{11}, c_{21}, c_{21}} \\
 &= \mathfrak{M}^{c_{11}, c_{12}, c_{21}, c_{22}} + \mathfrak{M}^{c_{11}, c_{21}, c_{12}, c_{22}} \\
 &= 2,
 \end{aligned}$$

em que os sobrescritos de \mathfrak{M} , na penúltima linha do resultado anterior, representam as extensões lineares de \mathcal{C} .

Cada vez que um elemento minimal m é retirado do conjunto \mathcal{C} , é necessário recalculer os elementos minimais de \mathcal{C} , portanto, o conjunto M muda a cada etapa do cálculo. Como, a cada etapa, retira-se um elemento de \mathcal{C} , e este processo sempre acaba, dado que \mathcal{C} é sempre finito. Portanto, a Equação 4.2 está bem definida.

Outra abordagem para construção de extensões lineares é a partir do conceito de *ideais de ordens*. Antes de introduzir-lo, seguem alguns conceitos básicos.

Dado um conjunto parcialmente ordenado \mathcal{C} , designa-se por **cadeia** (de \mathcal{C}) a todo subconjunto X de \mathcal{C} no qual todos os elementos são comparáveis. Por sua vez, designa-se por **anticadeia** (de \mathcal{C}) a todo subconjunto Y de \mathcal{C} no qual não existem dois elementos distintos comparáveis.

Denotando por $\text{cad}(\mathcal{C})$ e $\text{anticad}(\mathcal{C})$ os conjuntos de subconjuntos de \mathcal{C} que determinam, respectivamente, cadeias e anticadeias de \mathcal{C} ,

$$\max \{ \|X\| \mid X \in \text{cad}(\mathcal{C}) \} \quad \text{e} \quad \max \{ \|Y\| \mid Y \in \text{anticad}(\mathcal{C}) \}$$

designam, respectivamente, por **comprimento** (κ) e **largura** (λ), do conjunto parcialmente ordenado \mathcal{C} . Segue-se, *e.g.*, para as ordens parciais induzidas que:

$$\begin{aligned}
 \kappa(\mathcal{C}) &= \sum_{i=1}^n (n_i - 1) \\
 \lambda(\mathcal{C}) &= (n - 1) \min\{n_1 - 1, n_2 - 1, \dots, n_n - 1\}.
 \end{aligned}$$

Posto isto, considera-se a seguinte definição:

DEFINIÇÃO 4.5 Um **ideal de ordem** de um conjunto parcialmente ordenado \mathcal{C} é um subconjunto \mathcal{J} de \mathcal{C} que satisfaz, para todo $c \in \mathcal{J}$ e $c' \in \mathcal{C}$:

$$c' \preceq c \Rightarrow c' \in \mathcal{J}.$$

O conjunto de ideais de ordem de \mathcal{C} , ordenado pela relação de inclusão (\subset), forma também um conjunto parcialmente ordenado denotado por $\mathfrak{J}(\mathcal{C})$. Birkhoff [58] mostrou que $\mathfrak{J}(\mathcal{C})$ representa completamente \mathcal{C} .

Se \mathcal{C} tem cardinalidade n , então $\mathfrak{J}(\mathcal{C})$ possui comprimento n . Mais ainda, o **posto** (comprimento a partir do \emptyset , cf. Figura 4.7) $\nu(\mathcal{J})$ de qualquer ideal $\mathcal{J} \in \mathfrak{J}(\mathcal{C})$ é igual à cardinalidade do ideal de ordem \mathcal{J} .

Munido das estruturas supra-definidas pode-se agora apresentar um resultado que relaciona o número de cadeias maximais (i.e., cadeias com comprimento $\|\mathcal{C}\|$) do conjunto parcialmente ordenado $\mathfrak{J}(\mathcal{C})$, com o número de extensões lineares de \mathcal{C} [62, 63].

PROPOSIÇÃO 4.1 O número de extensões lineares $\mathfrak{M}(\mathcal{C})$ é igual ao número de cadeias maximais de $\mathfrak{J}(\mathcal{C})$.

DEMONSTRAÇÃO 4.1 De fato, a proposição pode ser vista como um corolário do teorema que mostra a existência de uma bijeção entre as extensões lineares de \mathcal{C} e as cadeias maximais de $\mathfrak{J}(\mathcal{C})$. Apresenta-se a demonstração devida a Stanley [62].

Seja a aplicação $\sigma : \mathcal{C} \rightarrow \{1, \dots, \|\mathcal{C}\|\}$ uma extensão linear de \mathcal{C} . Defina o ideal de ordem como $\mathcal{J}_i = \sigma^{-1}(\{1, 2, \dots, i\})$, $i = 1, 2, \dots, \|\mathcal{C}\|$. Então

$$\emptyset = \mathcal{J}_0 \subset \mathcal{J}_1 \subset \dots \subset \mathcal{J}_{\|\mathcal{C}\|} = \mathcal{C}$$

é uma cadeia maximal de ideais de ordem de \mathcal{C} . Por outro lado, dada uma cadeia maximal, a aplicação $\sigma : \mathcal{C} \rightarrow \{1, \dots, \|\mathcal{C}\|\}$ definida por $\sigma(i) \in \mathcal{J}_i - \mathcal{J}_{i-1}$ é uma extensão linear de \mathcal{C} .

OEΔ

Tome-se, e.g., um espaço com dois atributos ($n = 2$), onde o primeiro atributo possui duas possibilidades ($n_1 = 2$) e o segundo três ($n_2 = 3$), cf. Figura 4.6. O diagrama de Hasse do conjunto parcialmente ordenado de ideais de ordem de \mathcal{C} é apresentado na Figura 4.7.

A partir das cadeias maximais de $\mathfrak{J}(\mathcal{C})$ segue as extensões lineares de \mathcal{C} :

$$c_{11} \preceq c_{21} \preceq c_{12} \preceq c_{22} \preceq c_{13} \preceq c_{23} \tag{4.3}$$

$$c_{11} \preceq c_{21} \preceq c_{12} \preceq c_{13} \preceq c_{22} \preceq c_{23} \tag{4.4}$$

$$c_{11} \preceq c_{12} \preceq c_{21} \preceq c_{22} \preceq c_{13} \preceq c_{23} \tag{4.5}$$

$$c_{11} \preceq c_{12} \preceq c_{21} \preceq c_{13} \preceq c_{22} \preceq c_{23} \tag{4.6}$$

$$c_{11} \preceq c_{12} \preceq c_{13} \preceq c_{21} \preceq c_{22} \preceq c_{23} \tag{4.7}$$

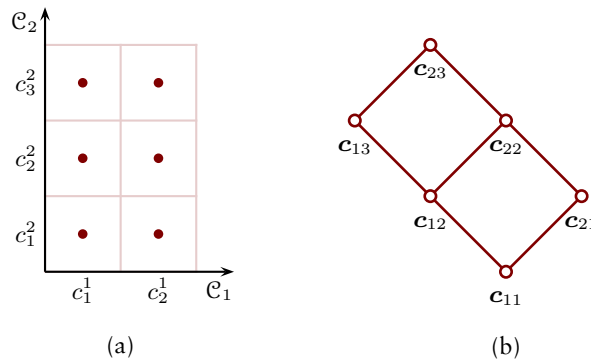


Figura 4.6: Exemplo de partição de um espaço de consequências com dois atributos. (a) Plano representando as consequências. (b) Diagrama de Hasse da ordem parcial induzida.

Note-se que as ordem lineares (4.3) e (4.7) são ordens lexicográficas. A primeira, ordem linear (4.3), mostra uma prioridade sobre o atributo \mathcal{C}_2 . Já a ordem (4.7), reflete uma prioridade sobre o atributo \mathcal{C}_1 . A Figura 4.8 apresenta as cadeias maximais de $\mathfrak{J}(\mathcal{C})$, note-se que as cadeias extremas (laterais mais escuras) estão relacionadas com as ordem lexicográficas. Com dois atributos existem apenas duas ordens lexicográficas.

Tome-se o caso $n = 2$, e $n_1 = n_2 = 3$. As ordens relacionadas a este problema são apresentadas nas Figuras 4.9 e 4.10. Neste caso são 42 cadeias maximais e portanto, 42 extensões lineares de \mathcal{C} . São duas ordens lexicográficas, cada uma delas com uma prioridade sobre um atributo, e 40 ordens lineares intermediárias para representar as preferências de um indivíduo. Cada uma delas representa um *tradeoff* entre as consequências.

Apresentou-se aqui apenas exemplos bidimensionais, *i.e.*, com dois atributos. Para dimensões maiores a diferença fundamental é o aumento do número de ordens lexicográficas. Se se trabalha com dimensão n , tem-se então $n!$ ordens lexicográficas. Pode-se imaginar as ordens lexicográficas como pontos extremos de um conjunto de ordens. Todas as outras ordens

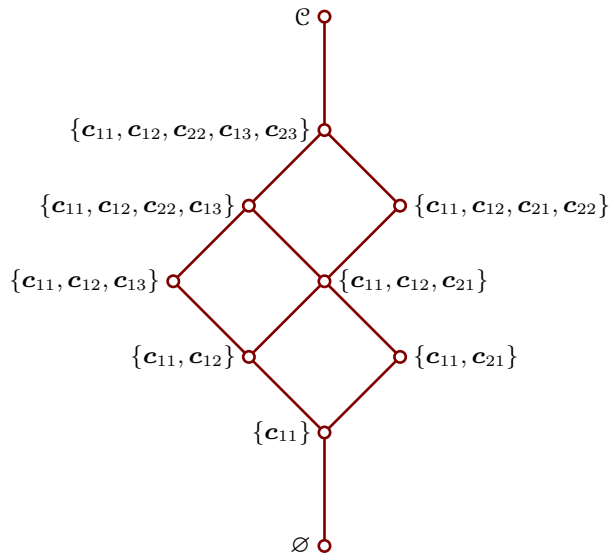


Figura 4.7: Diagrama de Hasse do conjunto parcialmente ordenado $\mathfrak{Z}(\mathcal{C})$.

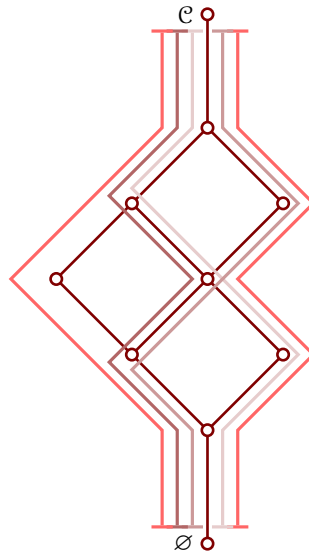


Figura 4.8: Cadeias maximais de $\mathfrak{Z}(\mathcal{C})$ ou extensões lineares de \mathcal{C} .

sendo intermediárias entre elas. Se existe uma ordenação sobre os atributos, pode-se reduzir o número de perguntas necessárias no questionário.

Uma questão de interesse é sobre o número máximo e o número mínimo de perguntas necessárias para estender \mathcal{C} . Considere ainda o caso bidimensional, *cf.* Figura 4.II.

Neste caso tem-se dois atributos, cada um deles particionado em três possibilidades (*cf.* Figura 4.II, $n_1 = n_2 = 3$).

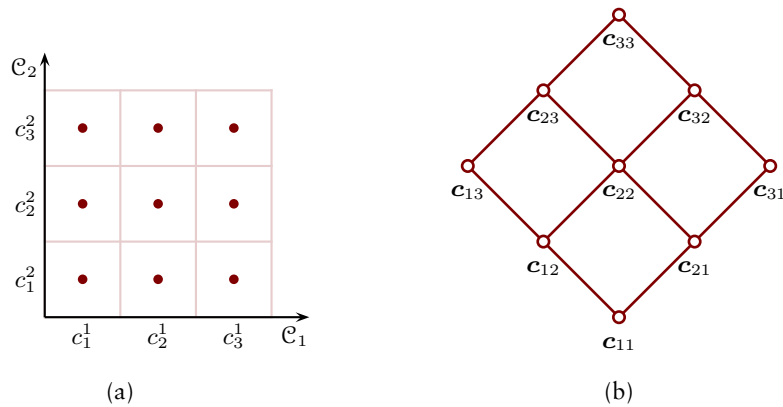


Figura 4.9: Exemplo de partição do plano de dois atributos mais simples possível. (a) Plano representando as consequências. (b) Diagrama de Hasse da ordem parcial induzida.

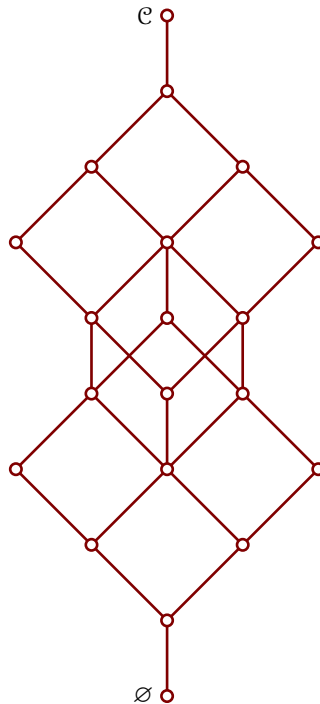


Figura 4.10: Exemplo de partição do plano de dois atributos mais simples possível.

Note-se que, no conjunto parcialmente ordenado \mathcal{C} , os elementos c_{13} e c_{31} não são comparáveis. Mais ainda, o elemento c_{13} não é comparável com nenhum dos elementos da área hachurada, *i.e.*, o fato das preferências sobre atributos serem monotônicas nada informa sobre estas relações. No entanto, note-se que, se o decisor ao responder um questionário afirmar que $c_{32} \preceq c_{13}$, então, da transitividade admitida nas relações de preferência, segue-se que

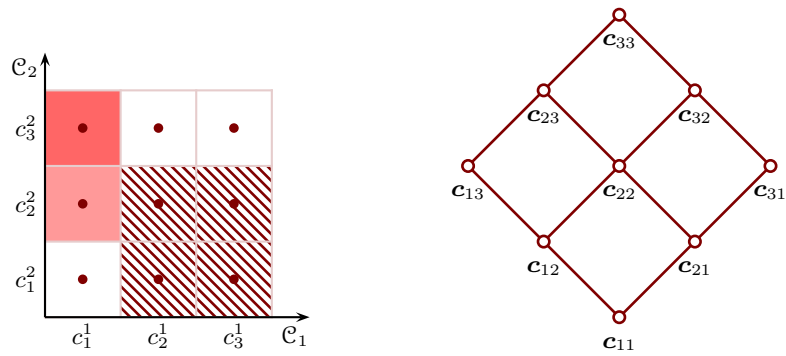


Figura 4.11: Ordem induzida para $n = 2$, $n_1 = 3$ e $n_2 = 3$.

ele também prefere c_{13} a qualquer outra consequência localizada na área hachurada. Caso contrário, ou seja, se o decisor responde que $c_{13} \preccurlyeq c_{32}$, isto nada nos informará sobre as outras comparações. Mas, se o decisor afirmar que $c_{32} \preccurlyeq c_{13}$ e depois que $c_{31} \preccurlyeq c_{12}$, então termina-se o processo de extensão e chega-se a uma ordem linear com apenas duas perguntas ao decisor.

Dito isto, fica clara a existência de uma maneira de se montar o questionário de modo que se aproveite ao máximo a transitividade e a monotonicidade das relações de preferência. Antes de entrar na construção do questionário, considera-se o caso com dois atributos e com uma divisão qualquer destes atributos.

Para o plano dividido de uma forma qualquer, com n_1 colunas e n_2 linhas, tem-se os seguinte limites para o número de perguntas, denotando-se por N_{\min} é o número mínimo de perguntas e por N_{\max} o número máximo de perguntas e $\|\mathcal{C}\| = n_1 \times n_2$.

$$N_{\min} = \text{Min}\{n_1 - 1, n_2 - 1\} \tag{4.8}$$

$$N_{\max} = \sum_i \sum_j (n_1 - i)(i - 1) = \frac{\|\mathcal{C}\|^2 - \|\mathcal{C}\|(n_1 + n_2 - 1)}{4}. \tag{4.9}$$

Para o caso $n_1 = n_2 = \eta$, a Tabela 4.1 apresenta alguns valores para estes limites.

Tabela 4.1: Exemplos de limites para o número de perguntas em função de η .

η	N_{\min}	N_{\max}
2	1	1
3	2	9
4	3	36
5	4	100
6	5	225
7	6	441
8	7	784
9	8	1296
10	9	2025

Esta área também inclui um dos problemas abertos mais famosos da teoria das ordens, a conjectura $1/3$ – $2/3$ devida a Fredman em 1975, cf. Brightwell [64]. Esta afirma que em um conjunto qualquer parcialmente ordenado \mathcal{C} existe um par de elementos para o qual a fração de extensões lineares de \mathcal{C} em que $c \preceq c'$, em relação ao número total de extensões lineares de \mathcal{C} , está entre $1/3$ e $2/3$.

Outra conjectura importante, devida a Pecarski [65], é conhecida como “*The Gold Partition Conjecture*.” Esta implica imediatamente na conjectura de Fredman. Além disso, implica num resultado interessante, apresentado a seguir. Denota-se por $\delta(\mathcal{C})$ o número de comparações necessárias e sempre suficientes para estender linearmente \mathcal{C} . A partir da conjectura de Pecarski, chega-se as seguintes cotas para $\delta(\mathcal{C})$ [65]:

$$\log_2 \mathfrak{M}(\mathcal{C}) \leq \delta(\mathcal{C}) \leq \frac{\log_2 \mathfrak{M}(\mathcal{C})}{\log_2 \phi}$$

em que ϕ é o número áureo:

$$\phi \equiv \frac{1 + \sqrt{5}}{2}.$$

O número $\mathfrak{M}(\mathcal{C})$ chamado aqui como “número de extensões lineares de uma ordem parcial \mathcal{C} ” pode receber diversos nomes como “número de extensões de \mathcal{C} para uma ordem total,” “número de linearizações de \mathcal{C} compatíveis em ordem,” “número de maneiras de ordenar \mathcal{C} topologicamente,” *etc.* dependendo do contexto.

Apresenta-se a seguir uma versão modificada de um algoritmo clássico de ordenação de ordens parciais, que servirá como questionário de educação de ordens lineares.

4.5 Ordenação topológica

Ordenação topológica consiste na construção de uma ordem linear ou total a partir de uma ordem parcial. A aplicação canônica da ordenação topológica (em inglês, *topological sorting* ou *topological ordering*) é a programação de uma sequência de trabalhos ou tarefas. O interesse em algoritmos de ordenação topológica surgiu no início da década de 1960 no estudo de técnicas PERT (*Program Evaluation and Review Technique*) para programação de tarefas em gerência de projetos [66–68].

Introduz-se o algoritmo de ordenação topológica para resolver o seguinte problema:

PROBLEMA 4.2 Encontrar uma extensão linear para ordem parcial \mathcal{C} que satisfaça às preferências do decisor.

Dado um conjunto parcialmente ordenado $\mathcal{C} = (\mathcal{C}, \preceq)$ e um conjunto completamente ordenado⁶ $\mathcal{L} = (\mathcal{L}, \preceq)$, sendo o conjunto \mathcal{L} inicialmente vazio, o algoritmo para ordenação topológica consiste, *grosso modo*, nos seguintes passos:

1. **Encontrar um elemento minimal de \mathcal{C}** , denotado por m .
2. **Remover m de \mathcal{C}** . Tecnicamente, isto significa substituir \mathcal{C} por $\mathcal{C} - m$ e remover todos os pares ordenados que contém m da relação binária \preceq . Intuitivamente, consiste em apagar m do diagrama de Hasse de \mathcal{C} , assim como suas ligações.
3. **Adicione m ao conjunto \mathcal{L} como “menor” (ou menos preferível) elemento**. Intuitivamente, isto significa colocar m abaixo do elemento mínimo de \mathcal{L} no diagrama de Hasse e ligá-los. Tecnicamente isto significa adicionar m em \mathcal{L} e o par ordenado (m, c) em \preceq , para todo c em \mathcal{L} .
4. Se \mathcal{C} está vazio, pare. Caso contrário, volta-se para etapa 1.

Pode-se modificar este algoritmo para que funcione como questionário de educação para se chegar a uma ordem linear que represente as preferências do decisor. Note-se que, na primeira etapa, a escolha de um elemento minimal \mathcal{C} pode ser feita de diversas maneiras, pois elementos minimais não são únicos (*cf.* Definição 4.3).

Considere o conjunto parcialmente ordenado \mathcal{C} , induzido por um espaço de consequências com dois atributos, onde um dos atributos (\mathcal{C}_1) possui dois níveis ($n_1 = 2$) e o outro (\mathcal{C}_2) possui três níveis ($n_2 = 3$). A Tabela 4.2 e a Figura 4.12 apresentam a execução do algoritmo.

⁶Note-se que se usa o mesmo símbolo (\preceq) para denotar relações binárias distintas.

Tabela 4.2: Aplicação do algoritmo de ordenação topológica de \mathcal{C} com $n = 2$, $n_1 = 2$ e $n_2 = 3$.

Etapa	\mathcal{C}	M	\mathcal{L}	Escolha
1	$\{c_{11}, c_{12}, c_{13}, c_{21}, c_{22}, c_{23}\}$	$\{c_{11}\}$	\emptyset	—
2	$\{c_{12}, c_{13}, c_{21}, c_{22}, c_{23}\}$	$\{c_{12}, c_{21}\}$	c_{11}	c_{12}
3	$\{c_{13}, c_{21}, c_{22}, c_{23}\}$	$\{c_{13}, c_{21}\}$	$c_{11} \preceq c_{12}$	c_{13}
4	$\{c_{21}, c_{22}, c_{23}\}$	$\{c_{21}\}$	$c_{11} \preceq c_{12} \preceq c_{13}$	—
5	$\{c_{22}, c_{23}\}$	$\{c_{22}\}$	$c_{11} \preceq c_{12} \preceq c_{13} \preceq c_{21}$	—
6	$\{c_{23}\}$	$\{c_{23}\}$	$c_{11} \preceq c_{12} \preceq c_{13} \preceq c_{21} \preceq c_{22}$	—
7	\emptyset	\emptyset	$c_{11} \preceq c_{12} \preceq c_{13} \preceq c_{21} \preceq c_{22} \preceq c_{23}$	—

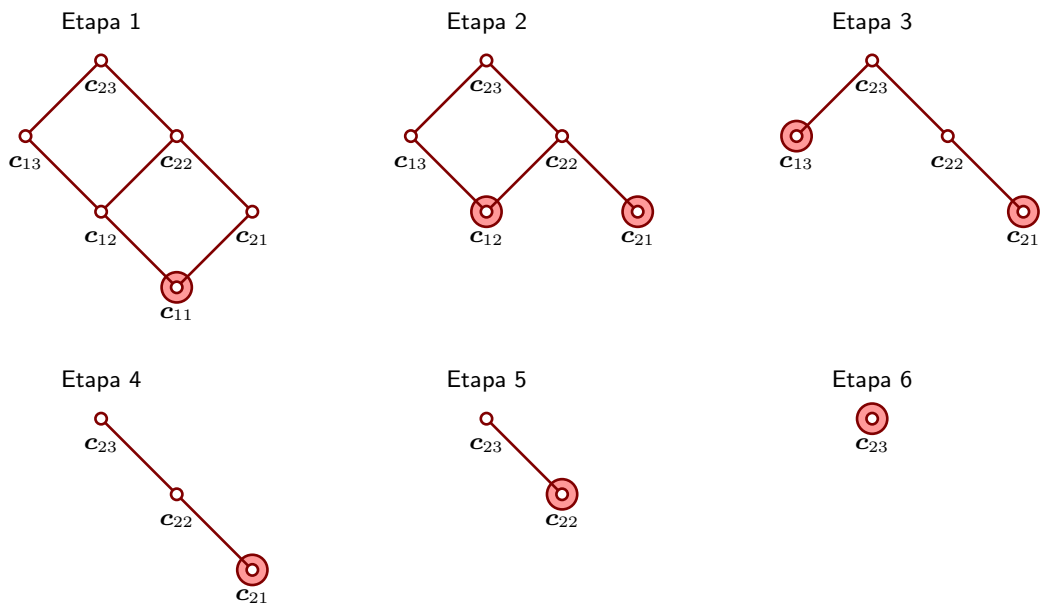


Figura 4.12: Aplicação do algoritmo de ordenação topológica de \mathcal{C} com $n = 2$, $n_1 = 2$ e $n_2 = 3$. Visualização a partir do diagrama de Hasse. Os vértices em destaque são elementos minimais.

Note-se que neste problema o decisor teve que responder a apenas duas perguntas, na segunda e terceira etapa, cf. Tabela 4.2 e a Figura 4.12.

Para implementação é conveniente outra abordagem que não a gráfica. Ao invés disso, pode-se usar a matriz de incidência para descrever o diagrama de Hasse de \mathcal{C} . Em tal matriz, uma linha i representa os sucessores (na ordem \preceq) do elemento i e uma coluna j representa os antecessores do elemento j . Usar-se-á células com cor para denotar a existência de ligação e células brancas em caso contrário. O número de antecessores imediatos para cada elemento de \mathcal{C} será denotado por $\alpha(\cdot)$ e a posição do elemento na ordem final por $\ell(\cdot)$.

O questionário consiste em usar o algoritmo de ordenação topológica e sempre que existir mais de uma possibilidade, isto é, sempre que houver mais de um elemento minimal, solicitar ao decisor que estabeleça sua preferência entre estes elementos.

A seguinte sequência de matrizes de incidência e diagramas de Hasse será usada para ilustrar o procedimento. Duas linhas serão acrescentadas na matriz de incidência para representar α e ℓ . Considerando o caso com dois atributos e cada um deles com três níveis, tem-se a condição inicial apresentada na Figura 4.13.

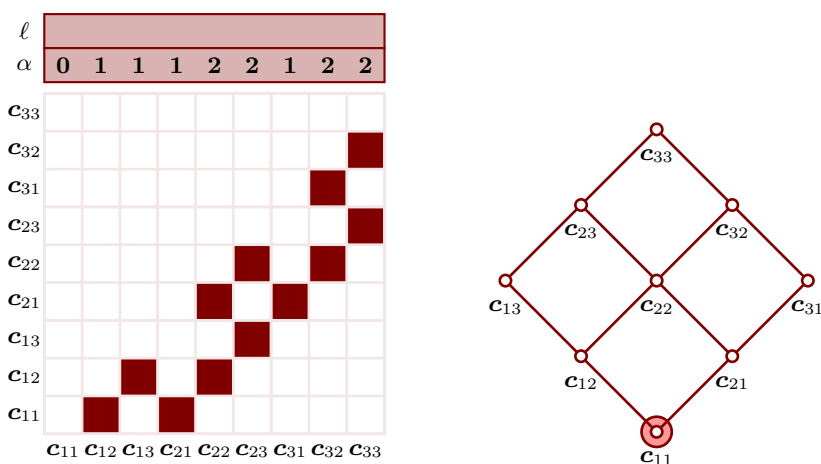


Figura 4.13: Matriz de incidência e diagrama de Hasse para ordem induzida. com $n = 2$, $n_1 = 3$ e $n_2 = 3$. Os vértices em destaque representam elementos minimais. Na matriz de incidência, isto significa que $\alpha(j) = 0$.

Existe uma coluna j tal que $\alpha(j) = 0$ e ainda não possui rótulo (ou seja, posição na ordem linear final). Esta é a coluna de c_{11} . Faz-se então $\ell(c_{11}) = 1$ (etapa 3 do algoritmo de ordenação topológica apresentado acima) e subtrai-se 1 da linha corresponde ao elemento c_{11} . Desse modo elimina-se o elemento c_{11} da ordem parcial (esta corresponde a etapa 2 do algoritmo), cf. Figura 4.14.

Tem-se agora duas colunas tais que $\alpha(j) = 0$ e j ainda não foi rotulada. Neste momento, deve-se questionar o decisor para que este decida entre $c_{12} \preceq c_{21}$ ou o contrário, não se permitindo empate. Suponha que este tenha afirmado o contrário, i.e., que $c_{21} \preceq c_{12}$. Faz-se então $\ell(c_{12}) = 2$ e subtrai-se 1 da linha corresponde ao elemento c_{12} , cf. Figura 4.15.

Continuando-se deste mesmo modo, tem-se as matrizes apresentadas nas Figuras 4.16 e 4.17 a seguir.

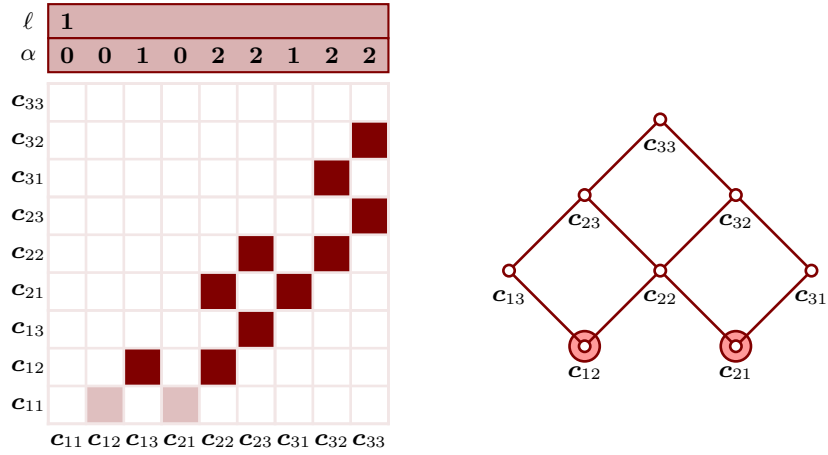


Figura 4.14: Matriz de incidência e diagrama de Hasse para etapa 2 do processo de educação da ordem. Note-se que o decisor precisa escolher entre dois elementos minimais.

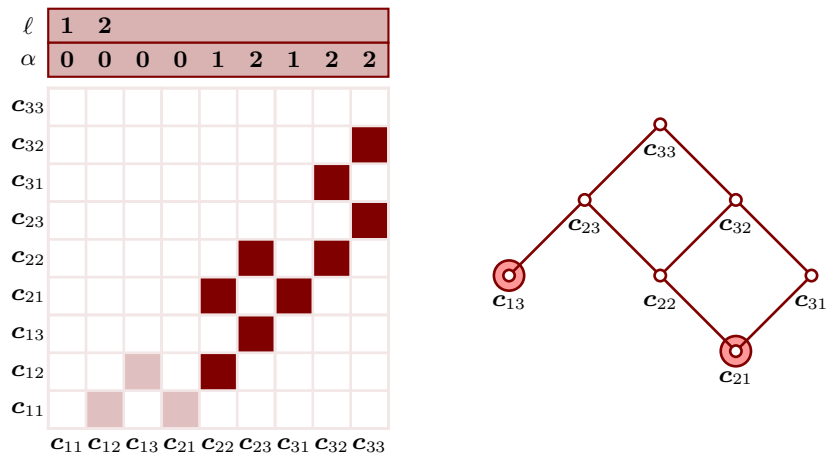


Figura 4.15: Matriz de incidência e diagrama de Hasse para etapa 3 do processo de educação da ordem. Novamente, o decisor precisa escolher entre dois elementos minimais.

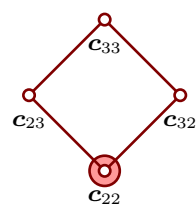
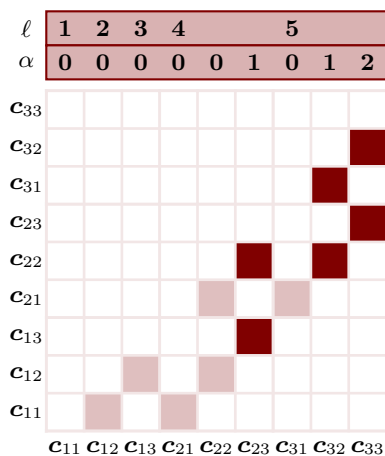
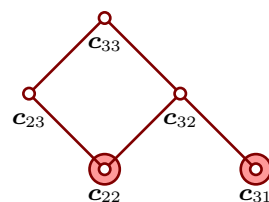
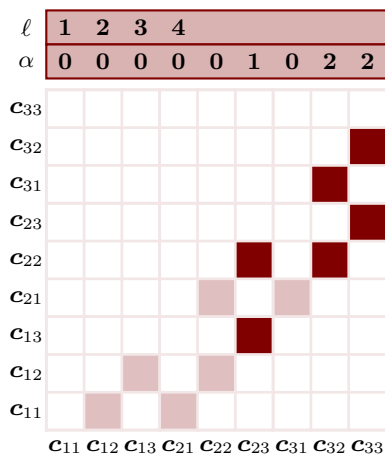
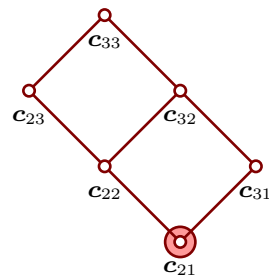
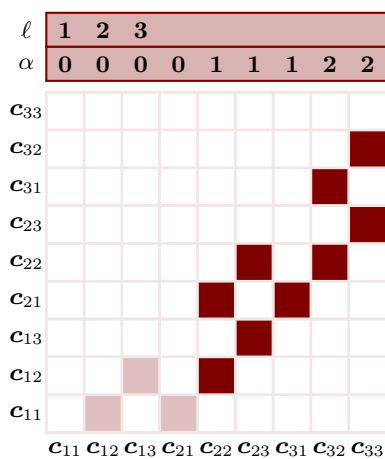


Figura 4.16: Matriz de incidência e diagrama de Hasse para ordem induzida. Etapas 4, 5 e 6.

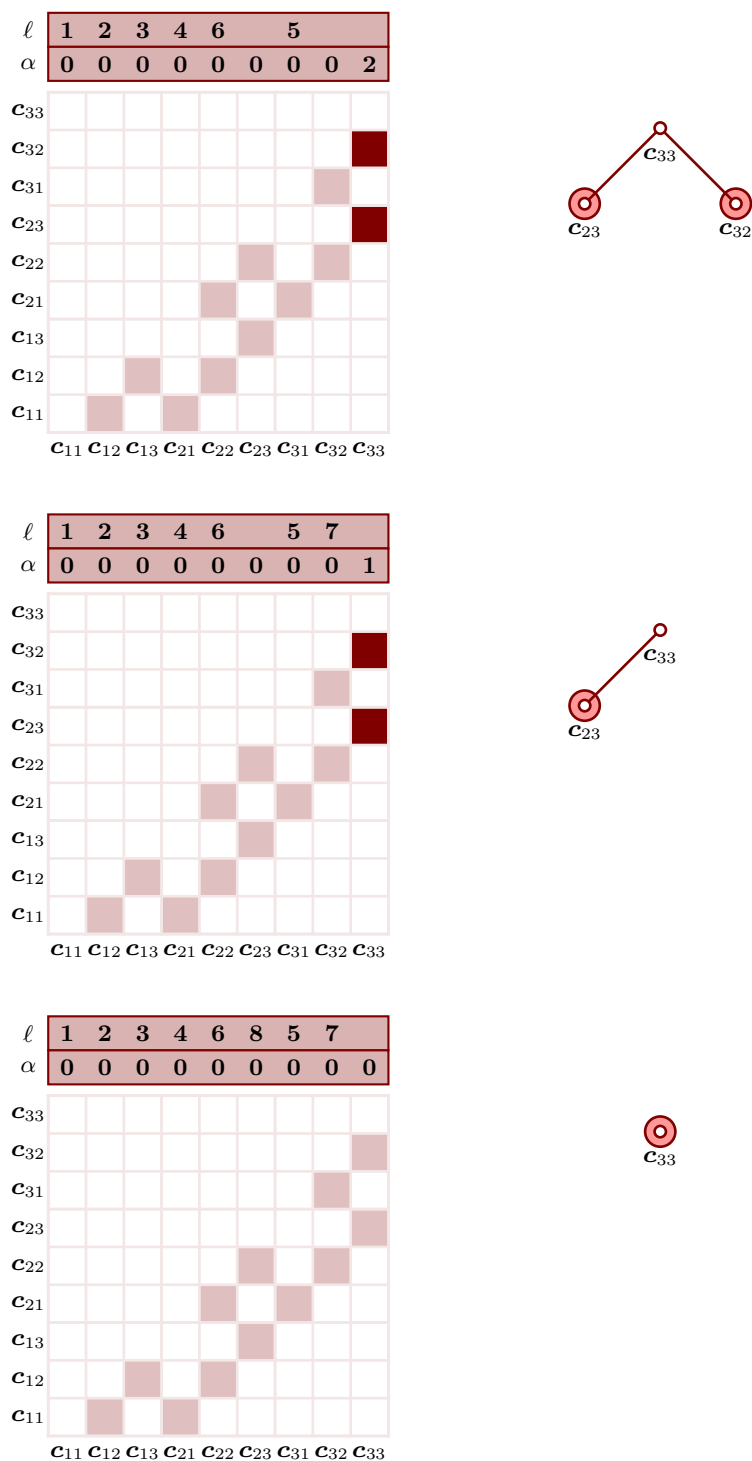


Figura 4.17: Matriz de incidência e diagrama de Hasse para ordem induzida. Etapas 7, 8 e 9.

Retirando-se o último elemento minimal do quadro mais abaixo da Figura 4.17, tem-se as seguintes posições, a partir da permutação definida por ℓ :

$$c_{11} \preceq c_{12} \preceq c_{13} \preceq c_{21} \preceq c_{31} \preceq c_{22} \preceq c_{32} \preceq c_{23} \preceq c_{33}$$

Note-se que quatro perguntas, nos quadros onde os elementos minimais não são únicos, foram necessárias para ordenação do conjunto de consequências.

4.6 Exemplo de educação: tempo livre *versus* salário

Considere o caso de um problema de decisão em que os atributos são *tempo livre* (lazer) e *salário*. Deseja-se eduzir a utilidade de decisores para estes dois atributos. Note-se que a hipótese de monotonicidade é claramente satisfeita, prefere-se maiores salários à menores salários. O mesmo acontece para tempo livre ou lazer.

Para facilitar o entendimento, utiliza-se o atributo *horas trabalhadas por dia*, ao invés de tempo livre. Neste caso, pode-se considerar uma faixa para horas de trabalho que vai de 6 horas por dia até 10 horas por dia. Supõe-se que outras possibilidades não são verossímeis. Divide-se ainda esta faixa da seguinte maneira:

1. 10h por dia;
2. 9h por dia;
3. 8h por dia;
4. 7h por dia;
5. 6h por dia;

Para o atributo trabalho, considera-se uma faixa que vai de dois mil reais até dez mil reais. Considera-se esta faixa de salários adequada para indivíduos que estão entrando no mercado de trabalho. Usou-se as seguintes possibilidades para o salário:

1. R\$ 2.000,00 por mês;
2. R\$ 4.000,00 por mês;
3. R\$ 6.000,00 por mês;
4. R\$ 8.000,00 por mês;
5. R\$ 10.000,00 por mês;

Considera-se que o atributo \mathcal{C}_1 seja o salários, com $n_1 = 5$ e que o atributo \mathcal{C}_2 seja o tempo livre, com $n_2 = 5$. Neste caso, tem-se que ordenar 25 elementos.

Este experimento foi conduzido com indivíduos em início de carreira (com exceção de um), com formações em medicina, engenharia, economia e matemática. Note-se que esta é uma amostra de conveniência, apenas para ilustrar o uso do protocolo. O algoritmo apresentado anteriormente foi implementado e utilizado para educação das utilidades. A Tabela 4.3 apresenta o resultado para a educação das ordens lineares. Note-se que a ordem ℓ está em função da ordem lexicográfica dos elementos de \mathcal{C} . Da Tabela 4.3, tem-se que a ordem eduzida do decisor número 3 é idêntica à ordem lexicográfica que tem como prioridade o salário, independente das horas trabalhadas.

Tabela 4.3: *Experimento trabalho versus tempo livre.*

decisor	1	2	3	4	5	6
c	ℓ_1	ℓ_2	ℓ_3	ℓ_4	ℓ_5	ℓ_6
c_{11}	1	1	1	1	1	1
c_{12}	2	2	2	2	2	2
c_{13}	4	3	3	3	3	3
c_{14}	5	4	4	4	4	4
c_{15}	6	7	5	6	5	5
c_{21}	3	5	6	5	6	6
c_{22}	7	6	7	7	7	7
c_{23}	9	9	8	8	8	8
c_{24}	11	10	9	10	9	9
c_{25}	12	14	10	12	12	10
c_{31}	8	8	11	9	10	11
c_{32}	10	11	12	11	11	12
c_{33}	14	15	13	14	13	13
c_{34}	16	17	14	15	14	14
c_{35}	20	18	15	17	17	21
c_{41}	13	12	16	13	15	15
c_{42}	15	16	17	16	16	16
c_{43}	19	19	18	18	18	18
c_{44}	21	21	19	20	20	20
c_{45}	22	22	20	22	22	24
c_{51}	17	13	21	19	19	17
c_{52}	18	20	22	21	21	19
c_{53}	23	23	23	23	23	22
c_{54}	24	24	24	24	24	23
c_{55}	25	25	25	25	25	25

Após a educação da ordem linear foi feita a educação da utilidade. Utilizou-se dois dos métodos apresentados no Capítulo 2, todos os dois da classe de métodos de equivalência em probabilidade. Para o decisor número 6 foi utilizado o método de valores extremos, apenas com o intuito de mostrar os mecanismos de erros inerentes a este método, apresentados no Capítulo 2. Para os demais usou-se o método de valores adjacentes também apresentado no Capítulo 2.

A vantagem de se utilizar o método de valores adjacentes é que este serve como teste de consistência da estimação de ordem, feita na etapa anterior. Note-se que as perguntas são da seguinte maneira, pergunta-se ao decisor qual a probabilidade (λ) que o deixa indiferente entre as loterias A e B :

$$A : \text{ganhar o par } c_{ij} \text{ com certeza.} \quad \text{e} \quad B : \begin{cases} \lambda \text{ de chance de ganhar o par logo acima de } c_{ij}. \\ 1 - \lambda \text{ de chance de ganhar o par logo abaixo de } c_{ij}. \end{cases}$$

em que c_{ij} representa um par salário e tempo livre, e os termos “logo abaixo” e “logo acima” definem os pares adjacentes segundo a ordem eduzida. Portanto, se a ordem linear eduzida for consistente, existe tal valor de probabilidade λ . Se houver alguma inconsistência na ordem eduzida, o decisor irá preferir ficar com uma das loterias.

Esta tipo de consistência não acontece quando se utiliza apenas os valores extremos da ordem para educação da utilidade. As utilidades eduzidas são apresentadas nas figuras a seguir. As Figuras 4.24 e 4.25 apresentam as utilidades para os decisores 1 e 2, em funções dos dois atributos. Note-se que apenas a educação do decisor 6 apresenta inconsistências (*cf.* Figura 4.23), pois a utilidade deveria ser crescente quando considerada sua ordem linear. Este fato segue diretamente do método para educação da utilidade, devido a dificuldade de ser considerar loterias com valores muito “distantes,” *cf.* Capítulo 2.

Em relação a dificuldade do processo de educação, é clara a vantagem deste protocolo. Os métodos clássicos [11] para educação da função utilidade multiatributo consistem basicamente das seguintes etapas: educação da função utilidade para cada atributo e educação de fatores de escala para combinação das utilidades. Ou seja, para o mesmo problema, o protocolo apresentado neste trabalho conduz apenas uma educação de ordem e uma função utilidade. Isto além do fato da combinação de função utilidade gerar controvérsia, pois a combinação de escalas de medida (ou seja, utilidades) não ser bem definida [69].

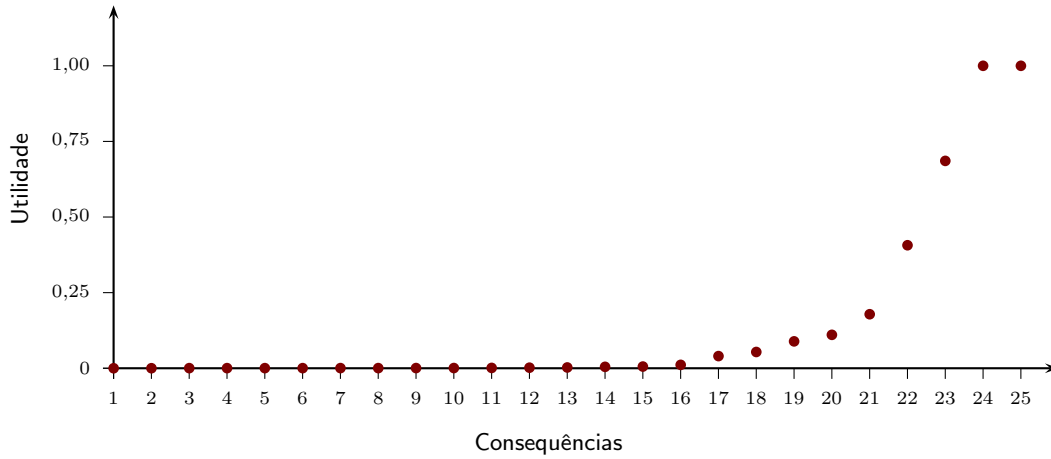


Figura 4.18: Utilidade eduzida para o Decisor 1. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.

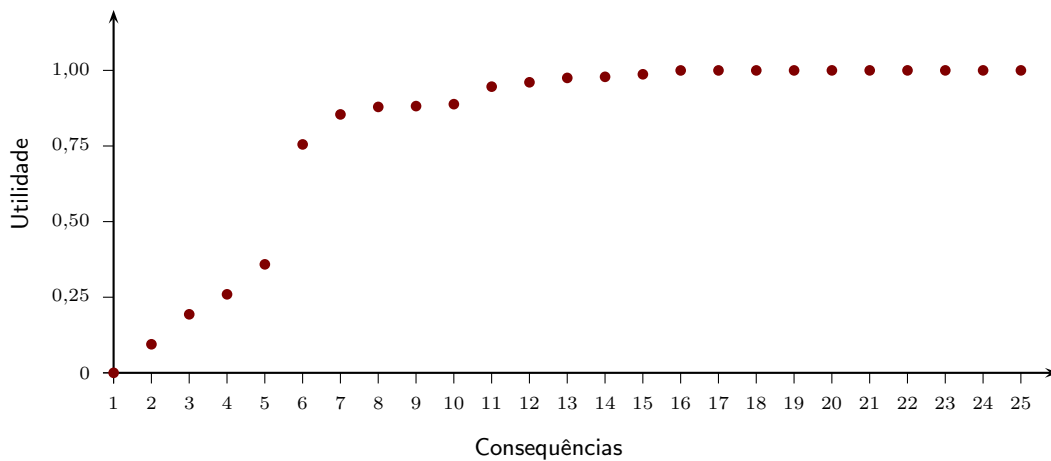


Figura 4.19: Utilidade eduzida para o Decisor 2. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.

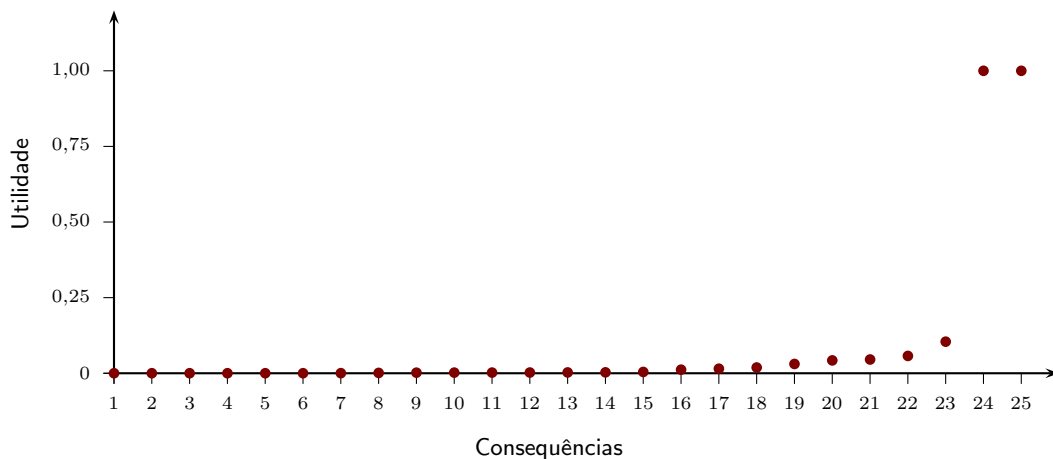


Figura 4.20: Utilidade eduzida para o Decisor 3. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.

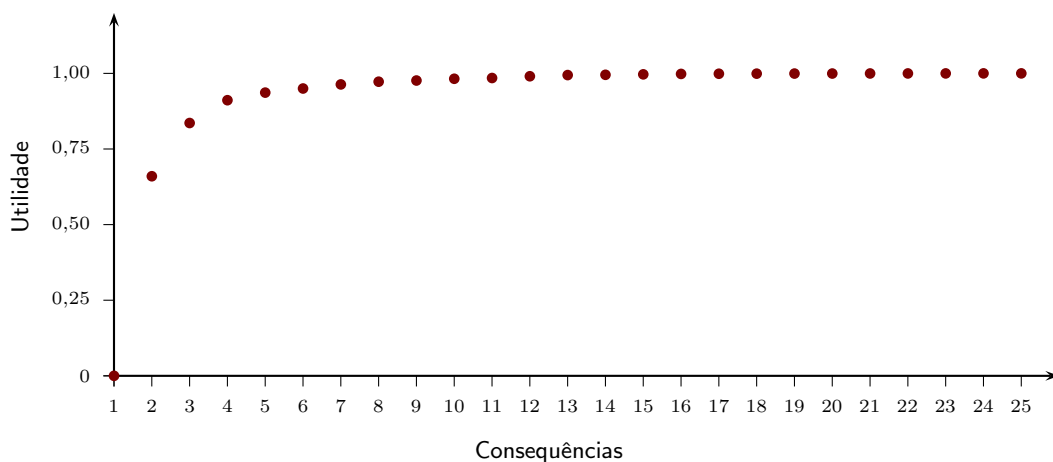


Figura 4.21: Utilidade eduzida para o Decisor 4. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.

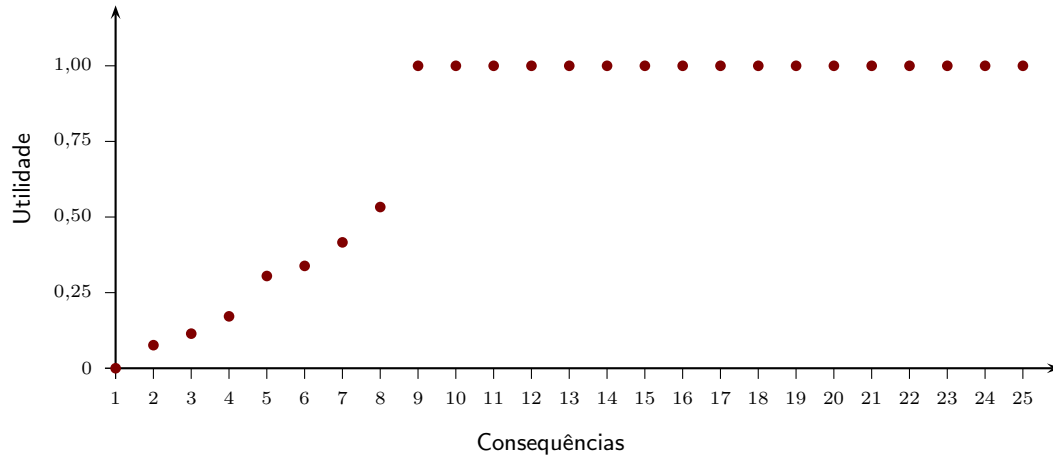


Figura 4.22: Utilidade eduzida para o Decisor 5. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.

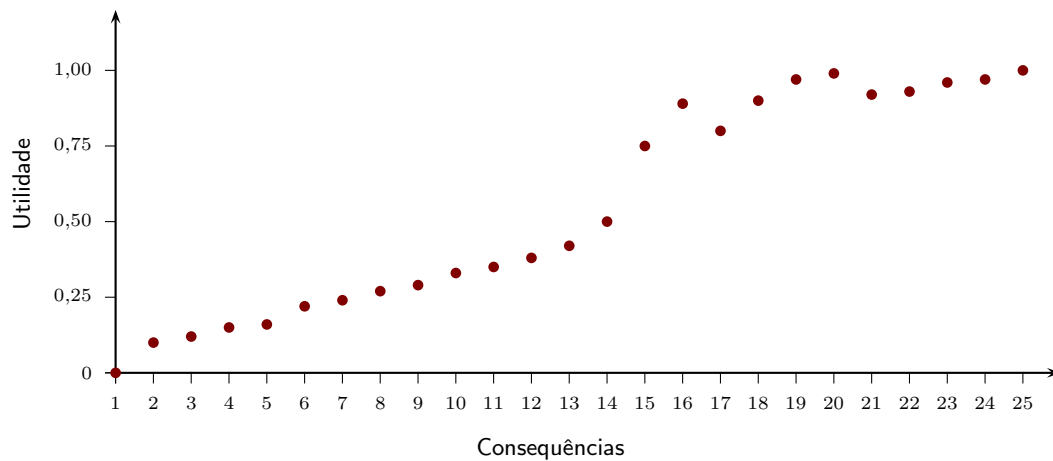


Figura 4.23: Utilidade eduzida para o Decisor 6. Note-se que o eixo horizontal está ordenado de acordo com a ordem linear eduzida.

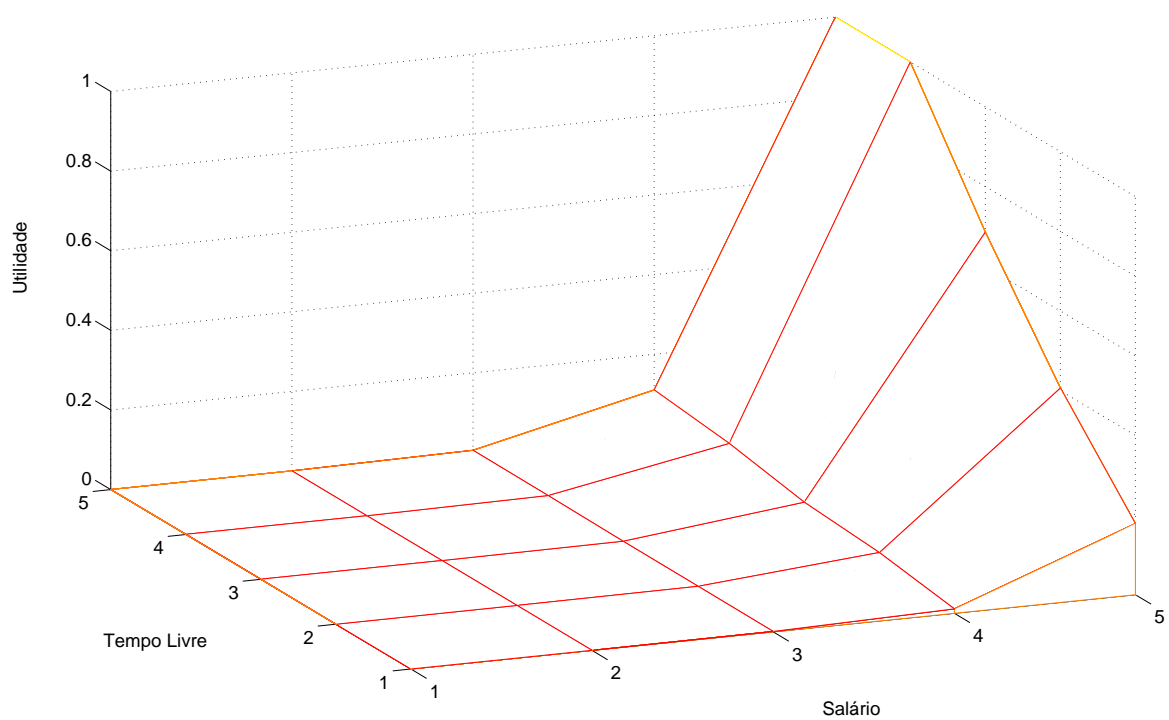


Figura 4.24: Utilidade biatributo do decisor 1, com as dimensões separadas.

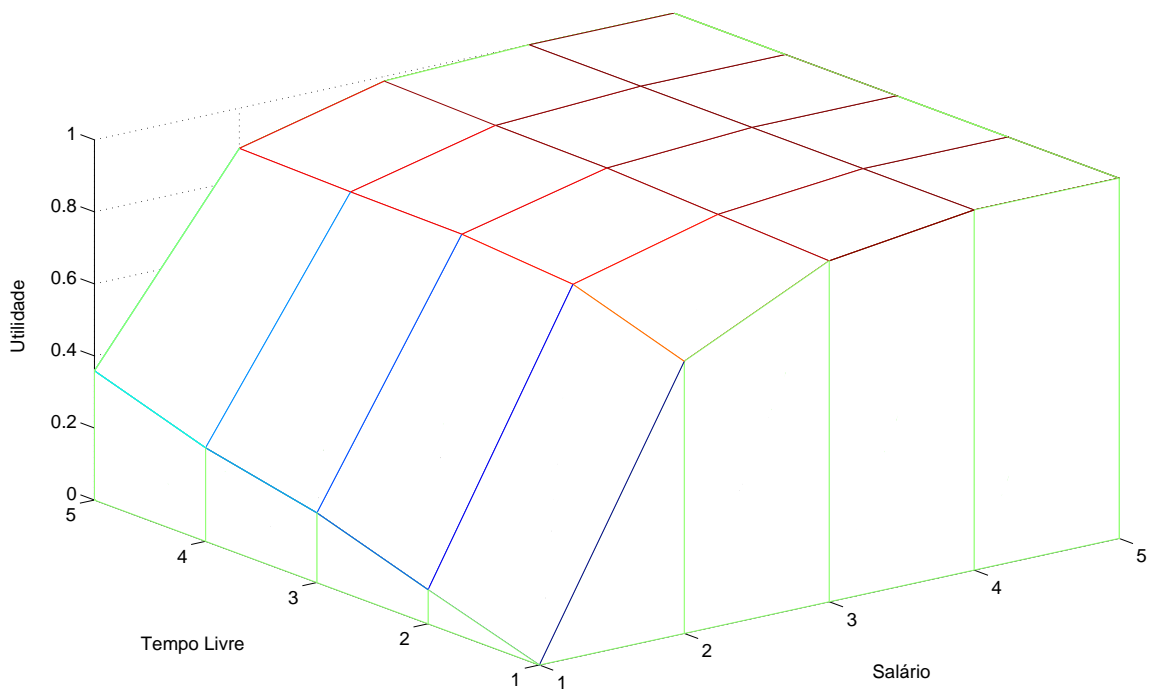


Figura 4.25: *Utilidade biatributo do decisor 2, com as dimensões separadas.*

4.7 Considerações finais

Apresentou-se neste capítulo uma nova maneira de eduzir funções utilidade multiatributo. A aplicação deste novo protocolo consiste nas seguintes etapas:

1. Escolha dos atributos pertinentes;
2. Escolha de escalas de medida adequadas a cada atributo e ordenação, quando necessária, dos atributos;
3. Educação da ordem linear, a partir do algoritmo proposto neste capítulo;
4. De posse de uma ordem linear, eduz-se a utilidade a partir de um dos métodos apresentados no Capítulo 2. Sugere-se o uso do método de valores adjacentes.

Este procedimento não requer nenhum tipo de independência entre os atributos, o que facilita o processo a construção do problema de decisão. O decisor é primeiro questionado com perguntas sobre preferência entre consequências determinísticas, com o intuito de se encontrar uma ordem linear sobre o espaço das consequências, mas também funciona como treinamento para perguntas sobre loterias, sabidamente com maior dificuldade cognitiva.

O processo de educação da utilidade, última etapa deste protocolo, serve também como verificação da ordem linear eduzida, pois na montagem das questões usa-se a ordem e qualquer impropriedade na ordem acarretará em perguntas cuja utilidade eduzida será 0 ou 1. Estes valores extremos de utilidade são reservados para os valores extremos da ordem e qualquer ponto intermediário com tais valores representa um equívoco na educação da ordem.

CONCLUSÕES, COMENTÁRIOS E SUGESTÕES

“The luxury of true-to-life performance tests [...] can rarely be afforded — the engineer must be right the first time. Thus no fruitful analytical avenue of approach can be overlooked.”

— A. E. Fitzgerald & C. Kingsley

O PROBLEMA DA EDUCAÇÃO da função utilidade tem sido tema central de pesquisas a partir da década de 1960. Ainda hoje as referências básicas continuam a ser os trabalhos de Keeney e Raiffa, principalmente o livro *Decision with multiple objectives: preference and values tradeoffs*, publicado em 1976 [11]. Reimpresso no início da década de 2000, os autores concluem:

In the intervening sixteen years, the theory most relevant to applications has been augmented a little and the application of the theory has greatly expanded.

O protocolo proposto aqui exige um esforço menor para o decisor, além de não exigir hipóteses de independência. A limitação está no fato de exigir a construção de uma escala finita e discreta para todos os atributos. Entretanto, acredita-se que esta não seja uma grande limitação, pois, de fato, a engenharia da tomada de decisão exige este procedimento. A construção de expressões analíticas pode ser desnecessária nos problemas de decisão. A educação de uma matriz de utilidades é suficiente para resolução do problema.

Outra vantagem do protocolo introduzido nesta tese é o fato de impedir que o decisor cometa intransitividades durante a construção da ordem. Além disso, combinado com a educação da utilidade utilizando o método de valores adjacentes, faz-se naturalmente um teste da consistência da ordem linear educada.

A educação de ordem como proposta aqui, além de ser intuitiva para o decisor e de reduzir os mecanismos de erro, tem o papel importante de preparar o indivíduo para a educação da utilidade, que sabidamente demanda um empenho cognitivo maior do decisor. Outra característica importante é que o algoritmo proposto é de implementação simples.

O estudo de conceitos de aversão ao risco, além de servir como um guia dos diversos conceitos existentes na literatura e suas ligações, cria um banco de dados de possíveis funções utilidades para uma posterior análise de regressão com os valores de utilidade eduzida.

Note-se que se sugere aqui utilizar os conceitos de aversão ao risco e independência de modo completamente diferente da literatura. A literatura recomenda realizar primeiro testes e hipóteses de independências, com isso chegando-se a uma estrutura para a função utilidade (que na maioria das vezes é aditiva). O proposto aqui consiste em eduzir primeiramente a utilidade, usando-se apenas as hipóteses de von Neumann e Morgenstern, depois disso, pode-se tentar, se for o caso, encontrar uma função utilidade que se ajuste a tais valores.

Fez-se no Capítulo 3 a revisão de diversos conceitos de aversão ao risco e estudou-se suas relações, assim como as implicações de comportamentos frente ao risco e estruturas da função utilidade. Apresentou-se a matriz de aversão ao risco de Duncan e estudou-se suas relações com a estrutura da função utilidade.

Como sugestões para trabalhos futuros, pode-se citar:

1. Realizar outros experimentos;
2. Construir medidas de aversão ao risco discretas para serem usadas diretamente com a matriz utilidade eduzida;
3. Considerar extensões do protocolo para decisões em grupo.
4. Estudar medidas de distâncias de conjuntos ordenados para comparações de ordens eduzidas e comportamentos de decisores.



REFERÊNCIAS

- [1] J. ELIASHBERG & J. R. HAUSER, Measurement error theories for von Neumann-Morgenstern utility functions, Northwestern University, Center for Mathematical Studies in Economics and Management Science, Discussion Paper N° 498, 1981.
- [2] A. WALD, **Statistical Decision Functions**. New York: John Wiley & Sons, 1950.
- [3] J. VON NEUMANN & O. MORGENSTERN, **Theory of Games and Economic Behavior**. Princeton University Press, 1944.
- [4] R. D. LUCE, **Individual choice behavior: a theoretical analysis**. New York: Dover Publications, 1959.
- [5] F. M. CAMPELLO DE SOUZA, **Decisões racionais em situações de incerteza**, 2ª ed. Recife: *Vade Mecum*, 2007.
- [6] P. WAKKER & A. M. STIGGELBOUT, Explaining distortions in utility elicitation through the rank-dependent model for risky choices. *Medical Decision Making*, v. 15, n. 2, p. 180–186, 1995.
- [7] P. WAKKER, S. J. T. JANSEN, & A. M. STIGGELBOUT, Anchor levels as a new tool for the theory and measurement of multiattribute utility, *Decision Analysis*, v. 1, n. 4, p. 217 – 234, 2004.
- [8] F. M. CAMPELLO DE SOUZA, Two-component random utilities, *Theory and Decision*, v. 21, n. 2, p. 129 –153, Sep. 1986.
- [9] ———, Mixed models, random utilities, and the triangle inequality, *Journal of Mathematical Psychology*, v. 27, n. 2, p. 183 – 200, 1983.
- [10] G. T. DUNCAN, A matrix measure of multivariate local risk aversion, *Econometrica*, v. 45, n. 4, p. 895 – 903, 1977.

- [11] R. L. KEENEY & H. RAIFFA, **Decision with multiple objectives: preference and values tradeoffs**. New York: John Wiley & Sons, 1976.
- [12] L. J. SAVAGE, **The Foundations of Statistics**, 2^a ed. New York: Dover Publications, 1954.
- [13] T. S. FERGUSON, **Mathematical statistics: a decision theoretical approach**. New York: Academic Press, 1967.
- [14] R. DORFMAN, P. A. SAMUELSON, & R. M. SOLOW, **Linear programming and economic analysis**. New York: McGraw-Hill, 1958.
- [15] H. R. VARIAN, **Microeconomic Analysis**, 3^a ed. London: W. W. Norton, 1992.
- [16] D. D. C. BEZERRA, Carteira de investimento usando a teoria da decisão, Dissertação (PPGEP), Universidade Federal de Pernambuco, Recife, 2003.
- [17] A. A. SILVA, Teoria da decisão em cardiologia, Dissertação (PPGEP), Universidade Federal de Pernambuco, Recife, 2002.
- [18] M. J. B. CUNHA, À tomada de decisão no planejamento estratégico: o caso do aeroporto industrial, Dissertação (PPGEP), Universidade Federal de Pernambuco, Recife, 2004.
- [19] A. WALD, Contributions to the theory of statistical estimation and testing hypotheses, *The Annals of Mathematical Statistics*, v. 10, n. 4, p. 299–326, Dec. 1939.
- [20] L. J. SAVAGE, The theory of statistical decision, *Journal of the American Statistical Association*, v. 46, n. 253, p. 55–67, Mar. 1951.
- [21] J. O. BERGER, **Statistical decision theory and Bayesian analysis**. New York: Springer, 1980.
- [22] A. WALD, Statistical decision functions which minimize the maximum risk, *The Annals of Mathematical Statistics*, v. 46, n. 2, p. 265–280, Apr. 1945.
- [23] M. ALLAIS, Le comportement de l’homme rationel devant le risque: critique des postulats et axiomes de l’école americaine, *Econometrica*, v. 21, n. 4, p. 503 – 546, Oct. 1953.
- [24] F. J. ANSCOMBE & R. J. AUMANN, A definition of subjective probabilities, *Annals of Mathematical Statistics*, v. 34, p. 199–205, 1963.
- [25] D. ELLSBERG, Risk, ambiguity and the Savage axioms, *Quarterly journal of Economics*, n. 75, p. 643 – 669, Nov. 1961.

- [26] M. J. MACHINA & W. S. NEILSON, The Ross characterization of risk aversion: strengthening and extension, *Econometrica*, v. 55, n. 5, p. 1139 – 1149, Sep. 1987.
- [27] A. TVERSKY & D. KAHNEMAN, Judgment under uncertainty: heuristics and biases, *Science*, v. 185, p. 1124 – 1131, 1974.
- [28] P. H. FARQUHAR, Utility assessment methods, *Management Science*, v. 30, n. 11, p. 1283 – 1300, Nov. 1984.
- [29] M. R. NOVICK & D. V. LINDLEY, Fixed-state assessment of utility functions, *Journal of the American Statistical Association*, v. 74, p. 306–311, 1979.
- [30] J. C. HERSHEY, H. C. KUNREUTHER, & P. J. H. SHOEMAKER, Sources of bias in assessment procedures for utility functions, *Management Science*, v. 28, p. 936 – 954, 1982.
- [31] M. A. ZAGORSKI, Risky decisions: attention and path dependency as a function of response mode, *Acta Psychologica*, v. 49, n. 2, p. 171 – 183, 1981.
- [32] R. KRZYSZTOFOWICZ & L. DUCKSTEIN, Assessment errors in multiattribute utility functions, *Organizational Behavior and Human Performance*, v. 26, p. 326 – 348, 1980.
- [33] I. FISHER, Is “utility” the most suitable term for the concept it is used to denote? *American Economic Review*, v. 8, p. 335–337, 1918.
- [34] B. RUSSELL, **Unpopular essays**. New York: Simon and Schuster, 1950.
- [35] J. W. PRATT, Risk aversion in the small and in the large, *Econometrica*, v. 32, n. 1–2, p. 122 –136, Jan. – Apr. 1964.
- [36] K. J. ARROW, The theory of risk aversion, In: **Essays in the theory of risk bearing**. Chicago: Markham, 1971.
- [37] E. KARNI, On multivariate risk aversion, *Econometrica*, v. 47, n. 6, p. 1391 – 1401, Nov. 1979.
- [38] B. DE FINETTI, Sulla preferibilità, *Giornale degli Economisti e Annali di Economia*, v. 11, p. 685 – 709, 1952.
- [39] S. F. RICHARD, Multivariate risk aversion, utility independence and separable utility functions, *Management Science*, v. 22, n. 1, p. 12 –21, Sep. 1975.

- [40] J. E. STIGLITZ, Behaviour towards risk with many commodities, *Econometrica*, v. 31, n. 4, p. 660 – 667, Oct. 1969.
- [41] R. L. KEENEY, Risk independence and multiattributed utility functions, *Econometrica*, v. 41, n. 1, p. 27 – 34, Jan. 1973.
- [42] R. E. KIHLSSTROM & L. J. MIRMAN, Risk aversion with many commodities, *Journal of Economic Theory*, v. 8, p. 361 – 388, 1974.
- [43] ———, Constant, increasing and decreasing risk aversion with many commodities, *Review of Economic Studies*, v. 48, n. 2, p. 271 – 280, Apr. 1981.
- [44] J. PAROUSH, Risk premium with many commodities, *Journal of Economic Theory*, v. 11, p. 283 – 286, 1975.
- [45] H. LEVY & A. LEVY, Arrow-Pratt measures of risk aversion: the multivariate case, *International Economic Review*, v. 32, n. 4, p. 891 – 898, Nov. 1991.
- [46] M. G. ERCOLANI, Risk aversion and risk loving in the small: a decomposition of the multivariate risk premium, *Bulletin of Economic Research*, v. 56, n. 1, p. 81 – 106, Jan. 2004.
- [47] S. A. SHAH, Comparative risk aversion when the outcomes are vectors, Centre for Development Economics. Delhi School of Economics. University of Delhi, Delhi, Working Paper N° 149, Sep. 2006.
- [48] R. F. MEYER, Some notes on discrete multivariate utility, Harvard Business School, Manuscript, Mar. 1972.
- [49] R. L. KEENEY, Multiplicative utility functions, *Operations Research*, v. 22, n. 1, p. 22 – 34, Jan. – Feb. 1974.
- [50] D. FUDENBERG & J. TIROLE, **Game theory**. Cambridge: MIT Press, 1991.
- [51] B. ROY, **Méthodologie multicritère d'aide à la décision**. Paris: Economica, 1985.
- [52] A. TVERSKY, Intransitivity of preferences, *Psychological Review*, v. 76, n. 1, p. 31 – 48, 1969.
- [53] V. CHANKONG & Y. Y. HAIMES, **Multiobjective decision making: theory and methodology**. Amsterdam: North-Holland, 1983.
- [54] S. S. STEVENS, On the theory of scales of measurement, *Science*, v. 103, n. 2684, p. 677 – 680, 1946.

- [55] P. A. SAMUELSON, **Economics**, 9^a ed. New York: McGraw-Hill, 1973.
- [56] E. ALENCAR FILHO, **Relações binárias**. São Paulo: Nobel, 1984.
- [57] B. A. DAVEY & H. A. PRIESTLEY, **Introduction to lattices and order**. New York: Cambridge, 1990.
- [58] G. BIRKHOFF, **Lattice theory**. New York: American Mathematical Society, 1948.
- [59] E. SZPILRAJN, Sur l'extension de l'ordre partiel, *Fundamenta Mathematicæ*, v. 16, p. 386 – 389, 1930.
- [60] A. F. SIDORENKO, Number of admissible linear orderings of a partially ordered set as a function of its incomparability graph, *Mathematical Notes*, v. 29, n. 1, p. 40 – 44, 1981, traduzido de *Mathematicheskije Zametki*, vol. 29, n. 1, pp. 75–82, Jan., 1981.
- [61] P. EDELMAN, T. HIBI, & R. P. STANLEY, A recurrence for linear extensions, *Order*, v. 6, p. 15–18, 1989.
- [62] R. P. STANLEY, **Ordered structures and partitions**, ser. Memoirs of the American Mathematical Society. New York: American Mathematical Society, 1972.
- [63] M. MORVAN & P. BALDY, Extending a partially ordered set: links with its lattice of ideals, *Order*, v. 16, p. 305 – 312, 1999.
- [64] G. R. BRIGHTWELL, Semiororders and the $1/3$ — $2/3$ conjecture, *Order*, v. 5, p. 369 – 380, 1989.
- [65] M. PECZARSKI, The gold partition conjecture, *Order*, v. 23, p. 89 – 95, 2006.
- [66] M. P. JARNAGIN, Automatic machine methods of testing PERT networks for consistency, U. S. Naval Weapons Laboratory, Dahlgren, Virginia, Technical Memorandum K-24/60, 1960.
- [67] A. B. KAHN, Topological sorting of large networks, *Communications of the ACM*, v. 5, n. 11, p. 558 – 562, 1962.
- [68] D. J. LASSER, Topological ordering of a list of randomly-numbered elements of a network, *Communications of the ACM*, v. 4, n. 4, p. 167 – 168, 1961.
- [69] J. BARZILAI, Measurement and preference function modelling, *International Transactions in Operational Research*, v. 12, p. 173 – 183, 2005.

SOBRE O AUTOR



O autor nasceu em Olinda, Pernambuco, no dia 15 de abril de 1979. Formado em Engenharia Elétrica, modalidade Eletrônica, pela Universidade Federal de Pernambuco (UFPE). Participou, durante a graduação, do programa de formação integrado de estudantes de engenharia brasileiros na França, na *Ecole Nationale Supérieure d'Ingénieurs Electriciens de Grenoble* (ENSIEG), com ênfase em *Signaux et Systèmes*.

Entre suas áreas de interesse estão os sistemas dinâmicos, controle e otimização de processos, fundamentos da teoria da probabilidade, probabilidades imprecisas, teoria da decisão, teoria da informação, filosofia, historia das ciências.

Endereço: Rua Aluizio de Azevedo, 236

Jardim Atlântico

Olinda – PE, Brasil

C.E.P: 53.140 – 100

e-mail: leite.andre@gmail.com

Esta tese foi diagramada usando $\text{\LaTeX} 2_{\epsilon}$ ¹ pelo autor.

¹ $\text{\LaTeX} 2_{\epsilon}$ é uma extensão do \LaTeX . \LaTeX é uma coleção de macros criadas por Leslie Lamport para o sistema \TeX , que foi desenvolvido por Donald E. Knuth. \TeX é uma marca registrada da Sociedade Americana de Matemática (\mathcal{AMS}). O estilo usado na formatação desta tese foi escrito por Dinesh Das, Universidade do Texas. Modificado em 2001 por Renato José de Sobral Cintra, Universidade Federal de Pernambuco, e em 2005 e 2008 por André Leite Wanderley.