

2

Decisão sob incerteza

Problema de decisão sob incerteza, como o próprio nome sugere, são problemas que contemplam variáveis aleatórias em algum dos dados de sua instância, de forma a caracterizar a incerteza para a qual o agente tomador de decisão está sujeito. Além disso, deve existir um subconjunto de suas variáveis de decisão (denominadas decisões de primeiro estágio) que deverão ser selecionadas ex ante sem a informação prévia da realização das incertezas (que se realizarão no segundo estágio, ex post), mas somente com base na distribuição de probabilidade destas.

Durante todo este trabalho, o termo incerteza e risco terão significados iguais. Em um contexto mais geral, a incerteza se caracteriza por uma ambigüidade na definição da própria distribuição de probabilidade das variáveis aleatórias, onde as distribuições são desconhecidas e desta maneira, subjetivas e individuais de cada agente. Contudo, o termo incerteza utilizado em todo o escopo desta tese deverá ser compreendido como o risco devido a uma distribuição de probabilidade objetiva, que assumiremos que possa ser estimada pelos agentes. Esta simplificação pode ser justificada pela existência de modelos probabilísticos que modelam de forma bastante razoável as distribuições das v.a.s que serão tratadas.

Recomenda-se a leitura completa da Parte II de [16] para uma excelente descrição de incerteza, subjetividade e outros aspectos relacionados a este trabalho.

Grande parte dos problemas de decisão da vida real são problemas de decisão sob incerteza, contudo, é bastante usual substituir os parâmetros incertos por suas previsões (por exemplo, o valor esperado). Em alguns casos tal procedimento simplifica bastante a solução e torna os modelos muito mais amigáveis e intuitivos, entretanto em casos onde as decisões sejam bastante sensíveis às variáveis de incerteza e que estas apresentem um considerável grau de variabilidade, torna-se importante que o modelo considere em suas decisões de

primeiro estágio a “dispersão” ou o risco das possíveis realizações das incertezas e os seus respectivos impactos.

A palavra dispersão foi utilizada entre aspas por não “expressar” todas as possíveis preocupações que se possa eventualmente exigir nestes modelos, mas já introduz o conceito de que, de alguma forma, as decisões de primeiro estágio deverão considerar toda uma gama de possíveis cenários que as incertezas podem proporcionar (ver [27]). Além disso, a importância que cada agente atribui aos possíveis resultados caracteriza o seu perfil de preferências. Este perfil também deve ser levado em consideração na seleção das variáveis de segundo estágio (este tema é justamente o foco desta tese e será amplamente discutido e referenciado ao longo dos próximos tópicos).

É importante ressaltar de que as decisões de primeiro e segundo estágios não estão necessariamente associadas às decisões do primeiro e segundo período, no sentido de unidade temporal. O primeiro e o segundo estágios se distinguem exclusivamente pela posse da informação das realizações dos parâmetros de incerteza.

Por exemplo, se um “dado produtor industrial” deve contratar hoje os montantes de suprimento firme de matéria prima para a sua produção para os próximos três meses a preços pré-fixados, essas podem ser encaradas como decisões de primeiro estágio. Desta maneira, se no decorrer destes três meses subsequentes de produção a demanda pelos seus produtos superar a quantidade respaldada pelos montantes de matéria prima já contratados, o industrial deverá buscar tais insumos no mercado de curto prazo ou à vista (*spot*), caso haja disponibilidade para os montantes requeridos e o preço seja interessante. Assim, as compras no mercado à vista são decisões do segundo estágio que dependem de três fatores de incerteza: (i) demanda pelos produtos em cada período, (ii) preço dos insumos no mercado à vista e (iii) disponibilidade da quantidade requisitada no mercado à vista. Como estas decisões são tomadas após a verificação da realização das incertezas de cada período, tais decisões são funções das incertezas e das decisões de primeiro estágio.

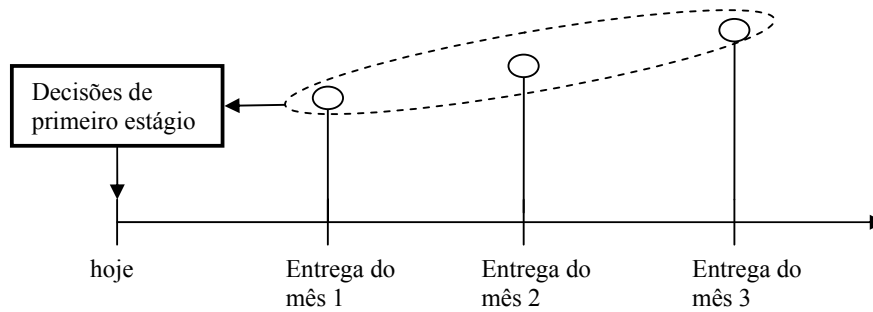


Figura 2-1 – Decisões de primeiro estágio em problemas de decisão sob incerteza.

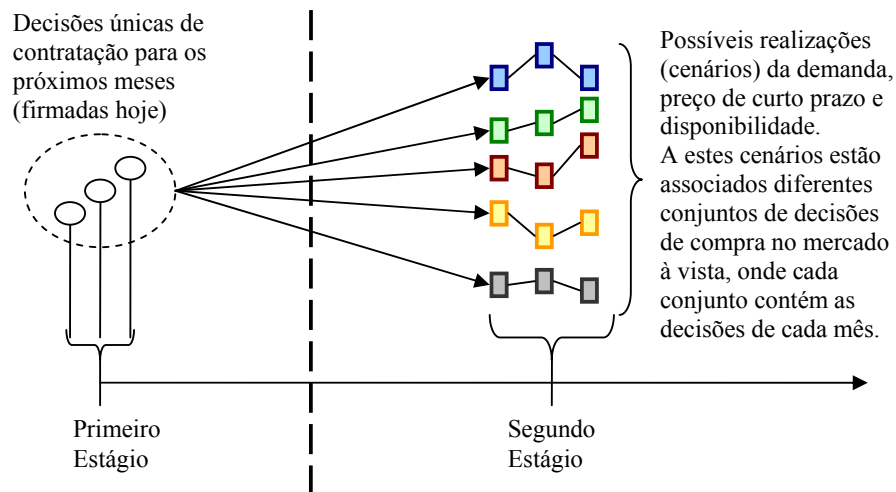


Figura 2-2 – Diagrama de decisões (primeiro vs segundo estágio) em problemas de decisão sob incerteza.

É importante notar que existe um contexto mais geral que o descrito anteriormente, onde as decisões sob incerteza são tomadas ao longo do tempo, ou seja, à medida que as informações parciais das incertezas são reveladas. Este caso é denominado multi-estágio e pode ser visto como uma sucessão de problemas de dois estágios acoplados entre si (ver referência [4] para exemplos na indústria onde as técnicas de opções reais são bastante ilustradas). Tais problemas sofrem do “mau da dimensionalidade” por serem constituídos de uma árvore de decisão que cresce exponencialmente com o número de períodos, porém, em alguns casos estes podem ser decompostos e resolvidos como uma seqüência de problemas menores de dois estágios, aplicando, por exemplo, a técnica de Programação

Dinâmica Dual Estocástica (PDDE) foi introduzida em [37] para o problema de planejamento da operação de sistemas hidrotérmicos de grande porte. Tal classe de problemas, multi-estágio, não será abordada nesta tese, que se restringirá aos problemas de dois estágios com conseqüências multi-período, ou seja, com mais de um período temporal.

2.1

Modelagem das incertezas

Nos modelos de decisão sob incerteza que serão desenvolvidos em todo este trabalho as incertezas serão caracterizadas através de Simulações de Monte Carlo que produzirão amostras das variáveis aleatórias e processos estocásticos. Tais amostras aproximaram as distribuições de probabilidade reais e proporcionam propriedades desejáveis pelos modelos de otimização e de precificação de ativos (ver [39] para aplicação).

Devido à alta complexidade de se caracterizar a distribuição conjunta de variáveis aleatórias multivariadas, a aproximação por cenários tornou-se bastante comum com o crescimento do poder de processamento dos computadores. Neste sentido, uma descrição completa das incertezas, dada por complexas funções de probabilidade de natureza não linear, é trocada por um conjunto de dados (amostras de dados) suficientemente grande para caracterizar as dependências e momentos das variáveis aleatórias de forma aproximada. Desta forma, uma vez que a aproximação dada pelos cenários de ocorrência e suas respectivas probabilidades seja aceita, a variável aleatória pode ser tratada como uma variável de distribuição discreta e, portanto, incorporada de forma eficiente em modelos de programação linear para se resolver os problemas de decisão sob incerteza.

Além disso, existem processos estocásticos para os quais não se tem acesso às distribuições analíticas, pois são originários de procedimentos computacionais complexos envolvendo, inclusive, processos de otimização dentro da formação de suas realizações. Assim sendo, estes só podem ser caracterizados através dos seus cenários simulados.

No Setor Elétrico Brasileiro os processos que regem os preços de curto prazo e geração das usinas são formados por procedimentos de otimização que visão minimizar o custo de atendimento à demanda (para mais detalhes ver

Anexo A: Formação de preços no mercado de curto prazo). Desta maneira, as incertezas presentes nos fluxos financeiros que serão analisados nesta tese, serão todas oriundas de simulações do despacho do sistema Brasileiro e das respectivas aplicações das regras comerciais, conforme tratado no mesmo anexo. Um detalhamento completo das regras de comercialização de energia pode ser obtido no site da Câmara de Comercialização de Energia Elétrica (CCEE – [38]).

Como notação, todas as vezes que estivermos mencionando uma variável aleatória ou um processo estocástico, este aparecerá em itálico. Por exemplo, um fluxo estocástico de renda aparecerá na seguinte forma: $\{R_t\}_t$ (em itálico). Entretanto, ao fazermos menção ao conjunto de cenários deste processo, a seguinte notação será adotada: $\{R_{ts}\}_{ts}$ (em formato normal acompanhado dos respectivos sub-índices de cada cenários).

2.2

Teoria de Utilidade

A teoria de utilidade de von Neumann-Morgenstern (1944) [56] apresenta uma forma analítica de comparação e de representação de preferências entre variáveis aleatórias (v.a.). Neste trabalho, as v.a.'s são representadas por um conjunto de resultados de uma “loteria”, ou jogo, e comparados segundo o valor esperado da utilidade. Desta maneira, se um agente tomador de decisão diz que uma determinada v.a. é preferível frente a outra, então, o valor esperado da utilidade desta primeira deverá ser maior que o da segunda.

A função de utilidade (FU) leva em consideração toda a gama de cenários, atribuindo uma utilidade ao consumo de cada cenário de receita financeira. Neste caso, o agente tomador de decisão buscará tomar suas decisões de forma a maximizar a utilidade esperada das rendas futuras, onde a função utilidade do agente descreve sua atitude frente ao risco (seu “perfil de risco”), que pode ser de aversão, neutralidade ou propensão a risco.

Por exemplo, um investidor avesso a risco apresentaria uma FU côncava, como se vê na Figura 2-3a. Neste caso, a perda devida a um “mau” resultado não é “compensada” pelo ganho advindo de um “bom” resultado de mesma magnitude. Já um investidor indiferente a riscos apresentaria uma FU linear, como

na Figura 2-3b. Isto significa que um aumento de receita tem o mesmo impacto (em módulo) que uma redução. Finalmente, um investidor que arrisca (propenso a risco) teria uma função de utilidade convexa, conforme se vê na Figura 2-3c.

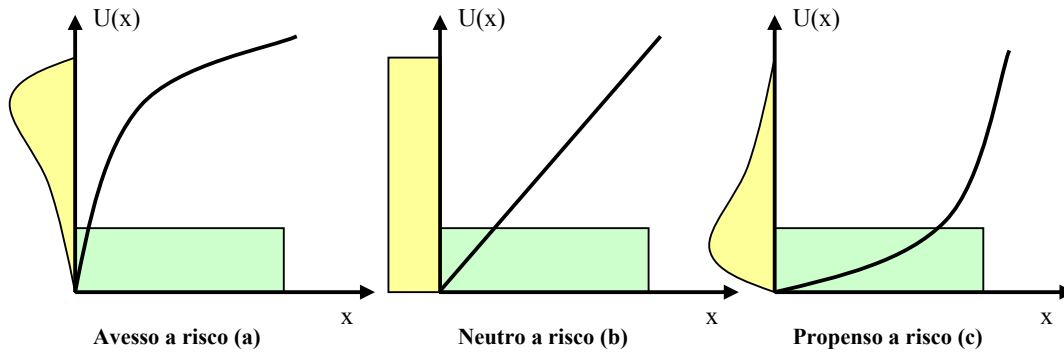


Figura 2-3 – Função de Utilidade e perfil de risco

Na seqüência abaixo (Figura 2-4, Figura 2-5 e Figura 2-6) são ilustrados os conceitos de aversão, neutralidade e propensão a risco apresentados para três indivíduos (indivíduos “k”, “y” e “z” – respectivamente avesso, neutro e propenso a risco) cujas funções utilidades são as mesmas da figura anterior:

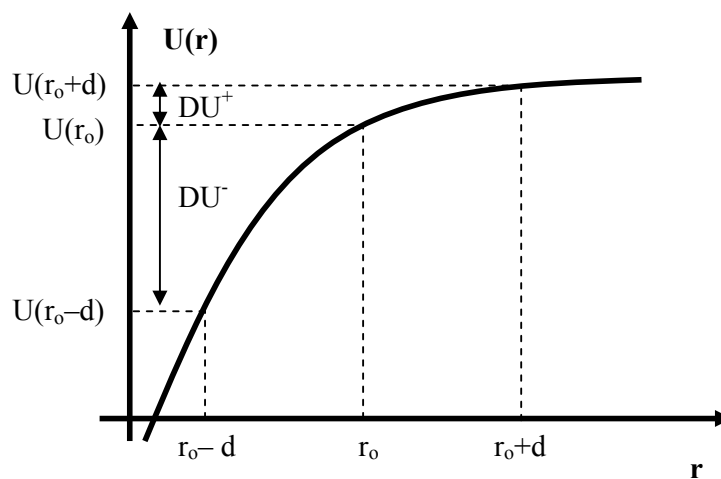


Figura 2-4 – Características do perfil de aversão a risco.

A principal característica do agente “k” (avesso a risco ou dito “conservador”) é que este é muito mais sensível a perdas do que a lucros. Por

exemplo, para uma mesma variação de $+d$ em torno do ponto r_0 (representando um aumento da renda), o “ganho” da utilidade do indivíduo é menor, em módulo, que o decréscimo de utilidade resultante da mesma variação negativa ($-d$) em torno de r_0 , ou seja, $DU^+ \leq DU^-$. Matematicamente isso pode ser explicado através de uma característica de funções côncavas diferenciáveis, na qual a segunda derivada é negativa, ou seja, decresce a primeira derivada ao longo do domínio da função. Assim, à medida que se avança no sentido positivo de r (aumento da renda), o benefício marginal da utilidade (primeira derivada) é decrescida de forma monótona.

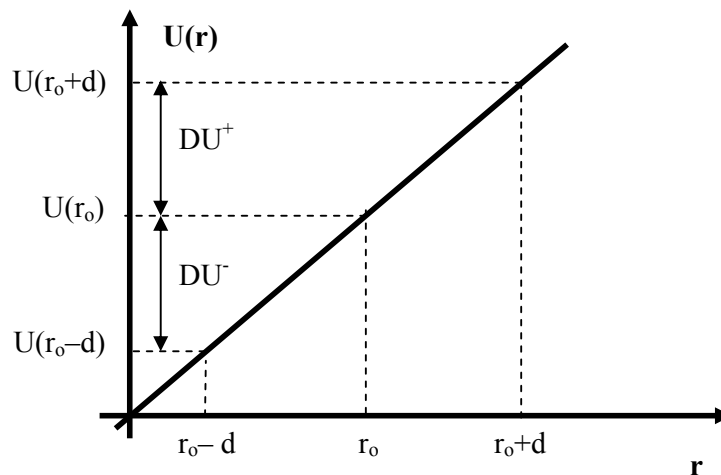


Figura 2-5 – Características do perfil de neutralidade a risco.

Para o perfil do agente “y” (neutro a risco), como a utilidade apresenta primeira derivada constante ao longo de todo o domínio, variações positivas e negativas de mesmo módulo, em torno de um mesmo ponto, proporcionam variações idênticas (em módulo) de utilidade, ou seja, $DU^+ = DU^-$.

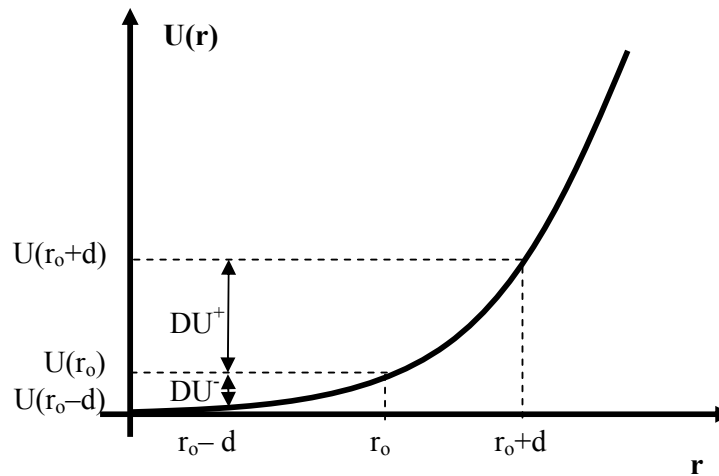


Figura 2-6 – Características do perfil de propensão ao risco.

No perfil do agente “z” (propenso ao risco), ocorre o oposto do caso de aversão. O agente dá muito mais “valor” a variações positivas do que às variações negativas de mesma magnitude, em torno de um mesmo ponto. Por isso a utilidade deste tipo de perfil é convexa, ou seja, a primeira derivada aumenta à medida que a renda cresce.

2.3

O coeficiente de aversão a risco (CAR)

O coeficiente de aversão a risco, ou grau de aversão a risco, tem por objetivo caracterizar o comportamento frente ao risco, que o agente expressa para cada ponto do espaço das possíveis realizações da variável aleatória R . Esse comportamento é modificado à medida que a concavidade da função utilidade varia em seu domínio.

Assim, seja uma v.a. R e uma função utilidade $U(\cdot)$ contínua e diferenciável, o coeficiente de aversão ao risco (CAR), também conhecido como Coeficiente de Arrow-Pratt, pode ser definido por:

$$CAR_U(r) = -U''(r)/U'(r) \quad (2-1)$$

Onde,

$U'(r)$ é a primeira derivada de $U(r)$,

$U''(r)$ é a segunda derivada de $U(r)$.

A primeira derivada aparece no denominador para padronizar o CAR de utilidades equivalentes, ou seja, se $U(r)$ e $V(r)$ são equivalentes, ou seja, $V(r)=aU(r)+b$, para $a > 0$ e $b \geq 0$ terão o mesmo $CAR(r)$.

2.4

Função Utilidade Linear por Partes (FULP)

A partir da função de utilidade o perfil de risco de um agente pode ser caracterizado. Foi visto também, que para expressar um perfil de aversão a risco, é necessário que esta função seja côncava, por tanto não linear. Contudo, a dificuldade de se “levantar” (estimar) a aversão a risco de um agente para todos os pontos de renda faz, com que a utilidade linear por parte torne-se bastante interessante do ponto de vista de sua especificação. Aliado a este fato, a forma linear por partes ainda apresenta a vantagem de permitir o uso de algoritmos de programação linear.

Esse tipo de função apresenta a propriedade do agente ser localmente neutro ao risco, porém, globalmente avesso⁴. A grande virtude desta forma é que se pode escolher os pontos de “quebra” onde os coeficientes angulares mudarão, utilizando como base parâmetros financeiros da empresa. Em outras palavras, o agente tem a “liberdade” de selecionar a taxa de crescimento da sua “satisfação” (utilidade) para alguns intervalos conhecidos de sua renda, e assim não ficar “preso” a uma só forma, definida por um único parâmetro. Além disso, esta forma permite uma interpretação financeira direta dos seus parâmetros com base nos pontos relevantes do balanço da empresa. Esta interpretação será exibida no próximo item.

Desta forma, podemos definir uma Função de Utilidade Linear por Partes (FULP) através de um problema de programação linear (PL) de maximização, para o caso de um agente avesso a risco (função côncava) e um PL de minimização para um agente propenso ao risco (função de utilidade convexa). Assim sendo, para o caso de aversão a risco, a FULP pode ser expressa por:

⁴ Para pequenas variações “locais” da renda que sejam restritas a um único segmento da FULP, o funcional baseado nesta função exibirá um comportamento de neutralidade ao risco por se tratar de uma utilidade “localmente” linear. Contudo, para v.a.s que apresentem cenários em mais de um segmento o funcional resultaria em um comportamento de aversão ao risco, uma vez que mais de um cenário estaria exposto a diferentes incrementos marginais de utilidade.

$$U(r) = \text{Maximizar}_{(u)} u \quad (2-2)$$

Sujeito a:

$$u \leq a_k r + b_k \quad k = 1, \dots, K \quad (2-3)$$

Onde

K Número de segmentos lineares

a_k Coeficiente angular k-ésimo segmento.

b_k Coeficiente linear do k-ésimo segmento.

u Variável que é sempre menor que todos os segmentos.

Como $U(\cdot)$ é crescente, $a_k \geq 0$, $k = 1, \dots, K$

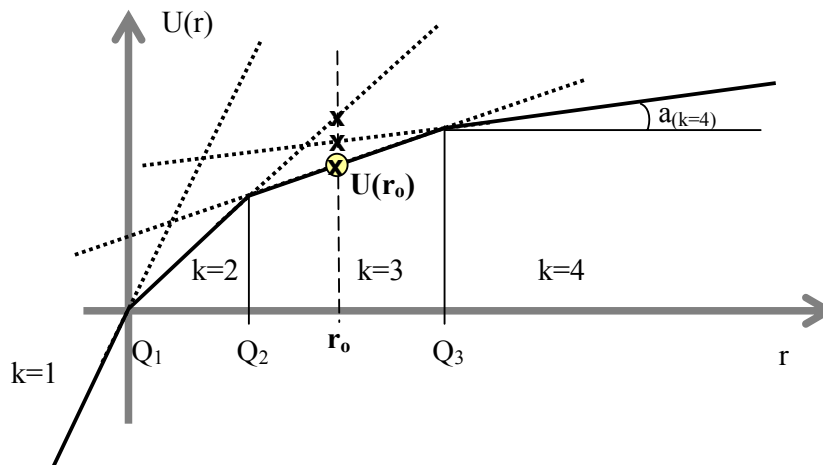


Figura 2-7 – Função de utilidade linear por partes com 4 segmentos.

A figura acima ilustra uma FULP com quatro segmentos, onde cada ponto Q_k representa uma “quebra”, ou seja, uma mudança de inclinação e, portanto, uma mudança na taxa de crescimento de “satisfação” com a renda do agente. Pode-se notar que o PL, que define esta curva, seleciona para um dado ponto r_0 o segmento que proporciona o menor valor no contradomínio. A figura anterior ilustra este fato, para um caso onde o ponto está localizado no terceiro segmento.

Esta formulação resulta na envoltória inferior definida entre todos os segmentos dada por:

$$U(r) = \inf_{\{1 \leq k \leq K\}} \{ a_k r + b_k \} \quad (2-4)$$

Outro aspecto importante para este tipo de função é que o CAR “convencional” não pode ser definido, pois existem pontos onde a função é não diferenciável e na verdade, o grau de aversão a risco da FULP se dá na mudança de coeficiente angular, assim pode-se definir o coeficiente de aversão a risco por partes (CARP) onde o que se expressa não é a aversão a risco local, mas entre segmentos.

$$\text{CARP}_k = -(a_k - a_{k-1})/a_{k-1} \quad (2-5)$$

Este coeficiente, através da razão entre a diferença das primeiras derivadas de dois segmentos adjacentes e a derivada do primeiro deles, expressa uma medida da taxa de variação da primeira derivada, que é a própria definição de segunda derivada, no caso de variações infinitesimais.

Assim, com base neste resultado, pode-se especificar a função utilidade através de um perfil de risco desejado, por exemplo, encontrando os coeficientes a_k que satisfaçam um conjunto de $\text{CARP}_{k=2,\dots,K}$, segundo a expressão abaixo.

$$a_k = (1 - \text{CARP}_k)a_{k-1} \quad (2-6)$$

Essa será a forma de função de utilidade adotada neste trabalho, onde a maior parte da metodologia será baseada em programação linear.

2.4.1

Interpretação Financeira e Especificação da Função

Conforme mencionado anteriormente, um aspecto bastante relevante no uso de modelos baseados em utilidade esperada para aplicações práticas é a especificação da função utilidade. Alguns experimentos podem ser realizados no sentido de interpolar a função utilidade de um agente. Para isso é realizado um questionário onde diversas perguntas relativas a preferências binárias são impostas ao agente e com base nestas respostas, alguns pontos da utilidade são encontrados [17].

Porém, ao utilizarmos a FULP, uma abordagem de interessante interpretação associando aos pontos de “quebra” da função linear por partes os pontos de interesse do balanço da empresa. Desta forma, as inclinações dos segmentos poderiam traduzir, dentre outros aspectos de interesse para o investidor, o benefício marginal da renda em função da taxa de juros que uma

instituição financeira cobraria por emprestar dinheiro à empresa. Ou seja, os coeficientes angulares da FULP podem ser associados aos benefícios marginais em taxa de juros com o acréscimo da renda.

Neste sentido, a empresa deve definir os seus pontos de interesse no eixo das rendas, onde, por exemplo:

- R₁ Renda a partir da qual a empresa “não sobrevive” um ano.
- R₂ Renda a partir da qual a empresa obtém seu “*breakeven*” anual.
- R₃ Renda anual projetada (meta).

Já os coeficientes angulares (derivadas em cada segmento) podem ser especificados de diversas formas com base nos parâmetros que expressem da melhor forma possível a diferença entre as preferências e benefícios marginais em cada segmento. Um exemplo pode ser dado com base nos *spreads* de juros cobrados para se obter fundos no mercado. Desta forma os coeficientes poderiam ser obtidos conforme as classificações de *ranking* de crédito:

- a₁ 1 + taxa de juros cobrada por empréstimo a uma empresa que oferece um alto risco de inadimplência (15% a.a. por exemplo).
- a₂ 1 + Taxa de cobrada por um empréstimo a uma empresa que oferece um risco médio de inadimplência (por exemplo 10% a.a.).
- a₃ 1 + Taxa de cobrada por uma empresa classificada como boa pagadora (por exemplo 8% a.a.).

Esses pontos devem ser definidos pelo investidor ou tomador de decisão, segundo a sua percepção do mercado e de distinção de taxas que serão cobradas ou obtidas em investimentos.

O intuito desta interpretação e objetivo da metodologia de especificação da utilidade, acima apresentada, é criar um funcional de preferência que possa refletir o comportamento desejado pelo investidor ou tomador de decisão com base em parâmetros financeiros que este tenha acesso e utilize no seu dia-a-dia para verificar a qualidade e direção de suas decisões, tornando este processo mais natural e intuitivo. Neste sentido, o funcional especificado se torna uma ferramenta de tomada de decisão de fácil utilização e alta transparência, podendo ser utilizado por outros agentes dentro da empresa que tomarão decisões com base no perfil estimado do investidor.

A referência [54] discute outros métodos para determinar a função utilidade de um agente. Entretanto, ao passo que funções utilidades contínuas definidas por expressões analíticas proporcionam uma caracterização de aversão a risco para qualquer variação de renda, em termos práticos um investidor pode perfeitamente assumir um comportamento localmente neutro a risco, porém globalmente avesso. Neste sentido, o argumento baseia-se no fato de que um investidor teria dificuldades de especificar o seu perfil de aversão a risco para pequenas variações de renda em torno de um dado resultado, contudo, através dos seus balanços patrimoniais ou de metas traçadas pelo conselho, pode-se mapear os pontos de “quebra” nos resultados financeiros a partir dos quais a preferência entre estar acima ou abaixo do mesmo é notoriamente diferente. Por exemplo, variações que façam o resultado da empresa sair de uma faixa onde o *breakeven* operacional é obtido para uma faixa onde é necessário obter capital externo devem ser diferentemente penalizadas relativamente a variações que proporcionem o cumprimento das metas financeiras estipuladas para um período.

Neste caso, o que se pode perceber é que o caso da função utilidade linear por partes proporciona uma alta flexibilidade para o agente especificar as “penalidades” ou “benefícios” marginais em um determinado segmento ou faixa de renda, sem exigir que este defina um grau de aversão a risco para um conjunto infinito de pontos do domínio, como é necessário no caso contínuo.

Como será visto, o objetivo do modelo de decisão proposto nesta tese é a maximização do VPL dos equivalentes certos de cada período. O capítulo a seguir apresenta a visão geral da metodologia adotada para o caso linear por partes. No capítulo consecutivo será realizada a conexão entre o “mundo” das utilidades e o das medidas de risco para o caso específico do *Conditional Value-at-Risk* (CVaR).