UNIVERSIDADE ESTADUAL DE CAMPINAS
Instituto de Matemática, Estatística e Computação Científica
Departamento de Matemática Aplicada

Análise de Decisão e Valor da Informação

Relatório Final PIBIC - Iniciação Científica
Fevereiro de 2019 a julho de 2019

Nome: Raul Augusto Teixeira
RA: 205177
Orientador: Prof. Dr. Lúcio Tunes dos Santos
Local de Execução: Campinas - SP
Período de vigência da Bolsa: Agosto de 2018 a julho de 2019

1 Introdução

A habilidade de analisar matematicamente um processo complexo de decisão é uma ferramenta cada vez mais útil, dadas as inúmeras variáveis e probabilidades que nos cercam neste mundo caótico e cheio de informações. Uma área particularmente importante da Análise Estatística de Decisão é a Teoria de Utilidade, que trata de atribuir valores pessoais a cada possível resultado de uma situação.

Muitas áreas podem fazer uso da Análise de Decisão para resolver problemas, mas uma que é tratada em específico neste trabalho é a Geofísica.

Este relatório apresenta o que foi desenvolvido por mim durante a segunda metade de minha Iniciação Científica em Análise de Decisão e Valor da Informação.

2 Resumo das Atividades Desenvolvidas

Nesta última parte da iniciação científica, estudei o capítulo 4 do livro "Análise Estatística de Decisão" de Otto R. Bekman (ver referência 1), que trata de Teoria da Utilidade, e posteriormente utilizei os conhecimentos adquiridos ao longo do projeto na elaboração e resolução de um problema de Análise de Decisão na área de Geofísica.

2.1 Teoria da Utilidade

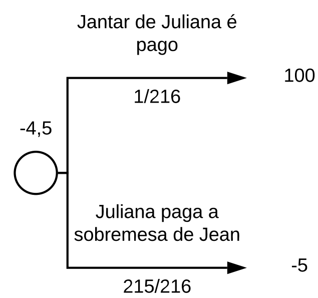
Em Análise de Decisão, algumas decisões podem envolver valores de retorno monetários, porém isso não é uma condição necessária para a aplicação dos métodos de resolução de problemas. Há casos que podem envolver preferências do decisor, que atribui valores de retorno para estados de natureza de maneira arbitrária. A isso é dado o nome de *Teoria da Utilidade*.

Uma boa maneira de ilustrar a atribuição de valores é o uso de *loterias*, que podem ser usadas para designar uma situação envolvendo prêmios. Considere o seguinte exemplo:

Exemplo 2.1.1 (Loteria): vamos supor que Juliana quer decidir se participará de uma loteria entre amigos jogando três vezes um dado não-viciado. Nesse jogo, se o dado resultar no número 1 três vezes seguidas, Juliana terá seu jantar de $R\$100,00$ pago inteiramente por seus companheiros. Entretanto, se o resultado não for esse, a garota deverá pagar a sobremesa de Jean, que custa $R\$5,00$.

A situação de Juliana pode ser modelada de acordo com a seguinte árvore de probabilidade:

Figura 2.1.1: Árvore de decisão de Juliana no exemplo 2.1.1



Nesta Figura, os valores na extremidade direita representam os dois possíveis valores de retorno da loteria. Acima das setas temos os possíveis resultados da loteria e, em baixo, suas probabilidades, que são calculadas a partir dos resultados de se lançar um dado honesto três vezes. A probabilidade de resultar em 1 três vezes vezes seguidas é de $(1/6)^3$, e a probabilidade de não resultar nisso é $1-(1/6)^3$. Além disso, o resultado $-4,5$ acima do nó de probabilidade é calculado multiplicando os possíveis resultados da loteria pelas respectivas probabilidades e somando-os. Assim, $-4,5 = (100 * \frac{1}{216}) + (-5 * \frac{215}{216})$.

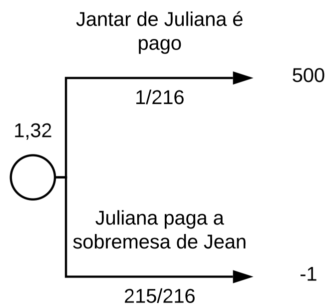
Levando o resultado ao pé da letra, não vale a pena para Juliana participar da loteria, pois o retorno esperado é de $R\$ - 4,5$ .Agora, de acordo com as preferências de Juliana, pode-se atribuir valores aos possíveis resultados do experimento.

Vamos supor que Juliana possui uma quantidade razoável de dinheiro, e portanto $R\$5,00$ de perda não fará muita diferença a ela. Por outro lado, ganhar $R\$100,00$ não é uma má ideia. Seja $u(x)$ uma função que associa ao valor monetário x um valor arbitrário escolhido por Juliana. De acordo com Juliana, considere uma utilidade de $u(-5) = -1$ e $u(100) = 500$. Considere também que não participar da loteria tem utilidade zero, ou seja, $u(0) = 0$. Assim, a árvore de decisão fica igual à Figura 2.1.2.

Logo, de acordo com as prefefências de Juliana e suas utilidades, vale a pena participar da loteria. Poderíamos continuar oferecendo vários valores possíveis de resultados monetários para Juliana e pedir que a garota fornecesse valores de utilidade para cada resultado de acordo com suas preferências, e montar uma tabela, como por exemplo a Tabela 2.1.1.

Outro conceito em Teoria de Utilidade é o *Equivalente Certo* de uma loteria. O equivalente certo da loteria em questão é definido como o valor pelo qual o decisor pagaria para não precisar participar da loteria.

Figura 2.1.2: Árvore de decisão de Juliana no exemplo 2.1.1 considerando os valores da Utilidade



Nesta árvore de decisão, os valores são calculados da mesma maneira aos da figura 2.1.1, com a exceção de que os valores de retorno dos Estados da Natureza são os valores das utilidades $u(100)$ e $u(-5)$.

Tabela 2.1.1: Valores de utilidade de Juliana

Valor	Utilidade
-500	-3500
-150	-1000
-100	-600
-50	-250
-5	-1
0	0
100	500
500	2400

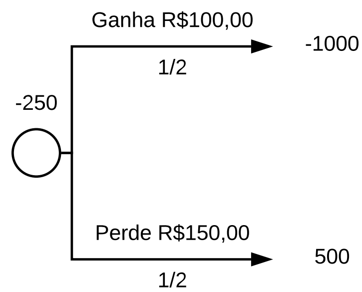
Tabela relacionando retornos monetários com os valores de utilidade relativos a eles escolhidos arbitrariamente por Juliana.

Exemplo 2.1.2 (Equivalente Certo): Vamos supor que Juliana participa de uma loteria com 50% de chance de perder $R\$150,00$ e 50% de chance de ganhar $R\$100,00$. Utilizando os valores de utilidade da Tabela 2.1.1, representamos o problema de acordo com a Figura 2.1.3.

O valor do Equivalente Certo dessa loteria é dado por X_{eq} tal que $u(X_{eq}) = -250$. Analisando a tabela 1, temos que o equivalente certo do problema é $X_{eq} = R\$ - 50,00$, ou seja, Juliana pagaria até $R\$50,00$ para não participar da loteria em questão.

Após ensinar esses conceitos, o Capítulo 4 do livro segue definindo os 6 Axiomas da Teoria da Utilidade, e depois ensina sobre como aproximar a *curva de utilidade* de um indivíduo. O processo consiste em fixar a utilidade de dois

Figura 2.1.3: Árvore de decisão de Juliana no exemplo 2.2.2



Árvore ilustrando uma loteria apresentada a Juliana. A representação segue o mesmo padrão das árvores anteriores.

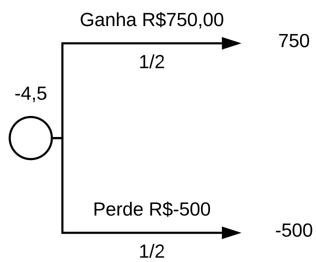
valores extremos, apresentar loterias ao indíviduo e perguntar quais são seus equivalentes certos em relação a elas, ou seja, perguntar quanto a pessoa gostaria de receber ou pagar para não ter que participar da loteria. Posteriormente, vamos construindo loterias similares com os valores de utilidade que obtemos e plotamos tudo num gráfico $u(x)$ versus x, em que x é o valor dos equivalentes certos do indivíduo. Num geral, o gráfico resultante é uma linha caso o indivíduo tenha indiferença em relação ao risco, curva com concavidade para baixo caso tenha aversão ao risco e curva com concavidade para cima caso tenha propensão ao risco.

Exemplo 2.2.3 (Função de Utilidade): vamos escolher arbitrariamente os valores -500 e 750 como extremos do gráfico, e colocar $u(-500) = 0$ e $u(750) = 1$. Depois, perguntamos a Juliana qual o seu equivalente certo para a loteria da figura 2.1.4.

Vamos supor que Juliana declara que pagaria $R\$50,00$ para não participar da loteria, ou seja, que seu equivalente certo é $Y_{eq} := -50$. Mas $u(Y_{eq}) = u(loteria) = 0,5 * u(750) + 0,5 * u(-500) = 0,5 \Rightarrow u(Y_{eq}) = u(-50) = 0,5$. Depois, vamos criar loterias com mesma probabilidade mas valores $\{1000, -50\}$ e $\{-50, -500\}$, e com os pontos que obtermos vamos plotar o gráfico. Uma aproximação pro gráfico obtido é o gráfico 2.1.1:

Note que Juliana apresenta aversão ao risco, o que é mais comum de se verificar na realidade.

Figura 2.1.4: Árvore de decisão de Juliana no exemplo 2.2.3



Árvore ilustrando uma loteria apresentada a Juliana. A representação segue o mesmo padrão das árvores anteriores.

Gráfico 2.1.1: Gráfico da Função de Utilidade de Juliana

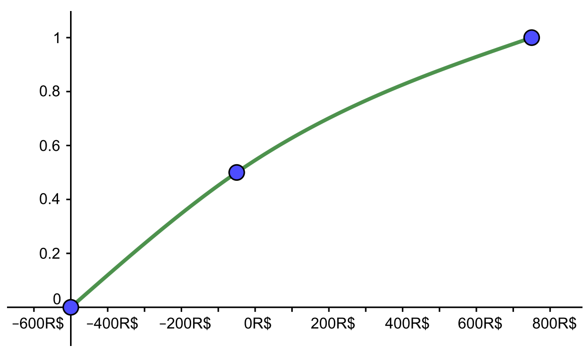


Gráfico apresentando a Função de Utilidade de Juliana de acordo com o Exemplo 2.2.3. O eixo das abscissas possui os valores dos equivalentes certos apresentados por Juliana e o das ordenadas, as respectivas utilidades.

2.2 Resolução de um problema geofísico a partir de Análise de Decisão

Para a última etapa de minha iniciação científica, construí um problema baseado no exemplo "*Cactus Petroleum*"do Subcapítulo 5.7 do livro *Operations Research: Principles and Practice*, de *Ravindran, Phillips* e *Solberg* (ver Referência 5).

Consideremos que uma determinada empresa chamada "Petrocamp"precisa decidir se deve ou não perfurar um poço de petróleo em um lugar específico, considerando que há probabilidades de o poço estar seco ou possuir petróleo pesado, médio ou leve, que são seus possíveis estados naturais. Além disso, a empresa leva em consideração a possibilidade de pagar por algumas medições sísmicas do solo que vão modificar as probabilidades do estado natural do poço, indicando propensão à existência de petróleo ou não. O objetivo é saber qual o Valor da Informação (V_I) do experimento e saber se vale a pena para a Petrocamp comprar as medições. O Valor da Informação é definido como o maior valor pelo qual vale a pena pagar pelas novas informações, e é dado pela média de lucro do problema com informação menos a média do problema sem a informação, ou seja, $V_I = E(decisao_{CI}) - E(decisao_{SI})$.

Assumamos que a empresa tem informações iniciais sobre o local (Tabela 2.2.1) e que o experimento é relativamente preciso (Tabela 2.2.2). Criando lucros baseando-se em valores reais de gastos da Petrobras no Pré-Sal em 2013 (Tabela 2.2.3) (ver Referência 4), podemos montar o problema. Nas Tabelas 2.2.1 e 2.2.2, O_k, $k = 1, 2$, denota os possíveis resultados do experimento, enquanto θ_j, $j = 1, 2, 3, 4$, denota os possíveis estados de natureza do problema. O problema de decisão sem experimento pode ser representado pela árvore de decisão da Figura 2.2.1.

Tabela 2.2.1: Probabilidade *a priori* dos estados da natureza

Estado Natural θ_j	Probabilidade P(θ_j)
Poço Seco	0,35
Petróleo Pesado	0,35
Petróleo Médio	0,20
Petróleo Leve	0,15

Tabela 2.2.2: Probabilidade do resultado do experimento dado o estado da natureza ($P(O_k \mid \theta_j)$)

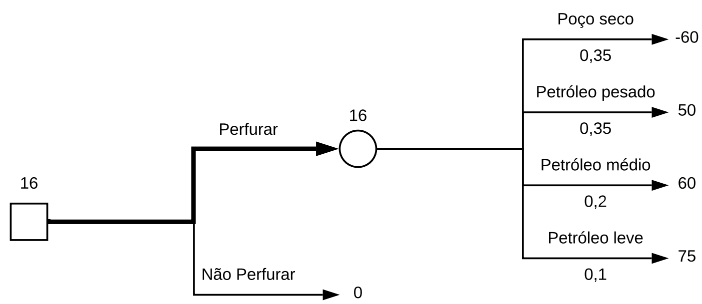
Estado Natural θ_j	Probabilidade de acordo com o experimento O_k	
	Propenso	Não Propenso
Petróleo Pesado	0,1	0,9
Petróleo Pesado	0,7	0,3
Petróleo Médio	0,8	0,2
Petróleo Leve	0,9	0,1

Analisando a figura 2.2.1, vemos que a média de lucro sem informação $E(decisao_{SI})$ ao se perfurar o poço é de $US\$16$ milhões de dólares. Portanto, sem levar em consideração as medições sísmicas, a empresa deve escolher perfurar o poço.

Tabela 2.2.3: Lucro total de cada estado natural do problema

Estado Natural	Lucro (milhões de US$)
Poço Seco	-60
Petróleo Pesado	50
Petróleo Médio	60
Petróleo Leve	75

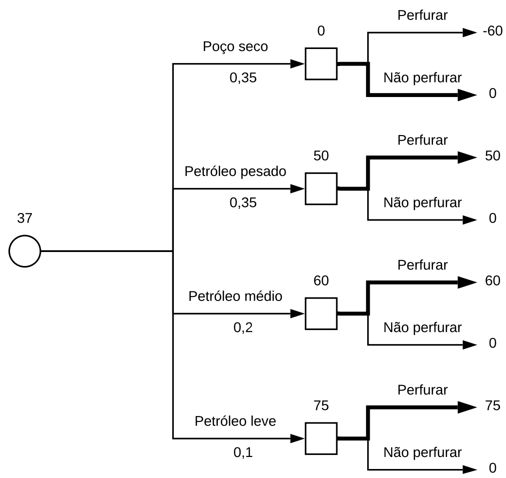
Figura 2.2.1: Árvore de decisão do problema da Petrocamp sem experimentação



Nesta árvore de decisão, o conceito é o mesmo das árvores anteriores. Porém, aqui adicionamos um nó de escolha (quadrado), responsável por representar a escolha do decisor em relação a duas possibilidades: perfurar ou não perfurar o poço. O valor acima do nó é o valor no maior nó relacionado à decisão.

Vamos passar agora para a análise do problema considerando a experimentação. O primeiro passo é analisar o Valor Esperado da Informação Perfeita $VEIP$ do problema: o valor máximo a se pagar por qualquer informação nova sobre os estados naturais. Para isso, devemos construir uma árvore de decisão que considere que em primeiro lugar é fornecida a informação real do estado da natureza, e depois é tomada a decisão em relação à perfuração, obtendo assim o valor da média do lucro com clarividência perfeita $E(decisao_{CP})$. O $VEIP$ então é dado por $VEIP = E(decisao_{CP}) - E(decisao_{SI})$. A árvore de decisão que modela o problema com clarividência perfeita é a Figura 2.2.2. De acordo com a análise dela, o valor máximo a se pagar por uma informação nova sobre o problema é de $VEIP = E(decisao_{CP}) - E(decisao_{SI}) = 37 - 16 = US\$21$ milhões de dólares.

Figura 2.2.1: Árvore de decisão do problema da Petrocamp com Clarividência Perfeita



Agora, vamos analisar a situação levando em consideração o experimento citado. A partir da Tabela 2.2.2, vamos utilizar o Teorema de Bayes para calcular os valores que usaremos para preencher a árvore de decisão do problema envolvendo o experimento.

Tabela 2.2.4: $P(O_k \cap \theta_j)$

Estado Natural θ_j	Probabilidade de acordo com o experimento O_k		**Total**
	Propenso	Não propenso	
Poço seco	0,03	0,32	0,35
Petróleo pesado	0,25	0,10	0,35
Petróleo médio	0,16	0,04	0,20
Petróleo leve	0,09	0,01	0,10
Total	0,53	0,47	1

Para construir esta Tabela, partimos da Tabela 2.2.2 e multiplicamos cada linha de probabilidades pelas respectivas linhas da Tabela 2.2.1. Além disso, adicionamos mais uma coluna e uma linha com os valores das somas das probabilidades.

Tabela 2.2.5: Probabilidade dos resultados do experimento $P(O_k)$

Resultados do experimento (O_k)	Probabilidades dos resultados ($P(O_k)$)
Propenso	0,53
Não propenso	0,47

Esta Tabela apresenta as probabilidades de cada possível resultado do experimento de medições sísmicas do solo, retiradas da Tabela 2.2.4.

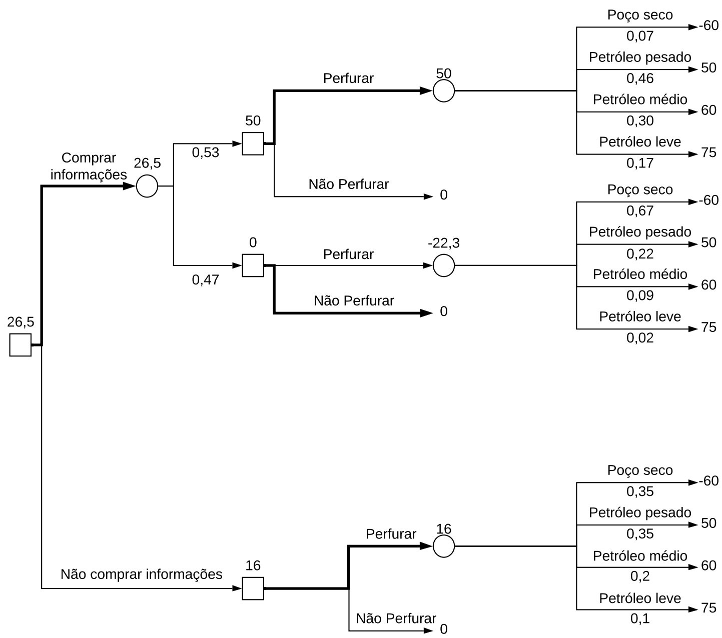
Tabela 2.2.6: Probabilidade dos estados da natureza dado o resultado do experimento $P(\theta_j \mid O_k)$

Estado Natural (θ_j)	Resultados das Medições (O_k)	
	Propenso	Não Propenso
Petróleo Pesado	0,07	0,67
Petróleo Pesado	0,46	0,22
Petróleo Médio	0,30	0,09
Petróleo Leve	0,17	0,02

Para montar esta última Tabela, partimos da Tabela 2.2.4, e dividimos cada coluna de probabilidades pela respectiva soma de todos os valores do coluna. Ou seja, dividimos cada valor da primeira coluna por $0,53$ e dividimos cada valor da segunda coluna por $0,47$

Temos que, de acordo com a figura 2.2.2 que modela o problema com a informação das medições sísmicas, o lucro esperado da decisão ao se perfurar o solo após comprar as medições sísmicas $E(decisao_{MS})$ é de $U\$26,5$ milhões de dólares, o que implica que o Valor da Informação do experimento é de $V_I = E(decisao_{CI}) - E(decisao_{SI}) = 26,5 - 16 = 10,5$ milhões de dólares. Como $VEIP = 21 > 10,5 = V_I$, vale a pena para a empresa Petrocamp comprar as medições sísmicas do solo se as medições custarem menos que $US\$10,5$ milhões.

Figura 2.2.2: Árvore de decisão do problema da Petrocamp com experimentação



A análise desta árvore de decisão é feita como nas árvores anteriores partindo da direita para a esquerda adicionando os valores obtidos nas tabelas citadas e escolhendo as melhores ações a serem tomadas pelo decisor.

3 Considerações

Houve alguns contratempos durante a segunda parte do meu projeto de Iniciação Científica. No primeiro semestre de 2019, me matriculei em uma disciplina complicadíssima que exigiu bastante tempo de estudo de minha parte, e portanto não consegui cumprir o cronograma original do projeto. A ideia após esse problema foi terminar de estudar Teoria da Utilidade e utilizá-la para analisar os *papers* da bibliografia. Porém, os trabalhos em questão mostraram-se demasiadamente complicados para o nível de conhecimento que adquiri durante o projeto, então busquei um exemplo mais simples no livro *Operations Research: Principles and Practice*, de *Ravindran, Phillips* e *Solberg* e o utilizei para criar o meu próprio exemplo de Análise de Decisão em Geofísica tentando seguir valores numéricos e conceitos um pouco mais realistas.

4 Conclusões

A Análise Estatística de Decisão é uma área da matemática aplicada que não está tão em voga no momento quanto outros assuntos, o que não significa que não seja uma área de estudos importante. No meu trabalho, resolvi um exemplo bem simplificado em Geofísica. Porém, como podemos ver nos *papers* da bibliografia, é possível aprimorar os problemas até chegar em questões mais palpáveis, mais perto da realidade. Portanto, apesar de não ter estudado um caso em que a Análise de Decisão possa ser utilizada na realidade, aprendi a base e as ferramentas necessárias para um dia poder aprimorar e aplicar meus conhecimentos nessa área de estudos fascinante.

5 Bibliografia

1. *O.R. Bekman & P.L.O. Costa Neto*, **Análise Estatística da Decisão**, Editora Blücher, 2009.

2. *J. Eidsvik, D. Bhattacharjya & T. Mukerji*, Value of Information of Seismic Amplitude and CSEM Resistivity, **Geophysics**, 73, R59-R69, 2008.

3. *J. Eidsvik, T. Mukerji, D. Bhattacharjya & G. Dutta*, Value of Information Analysis of Geophysical Data for Drilling Decisions, **Petroleum Geostatistics Conference**-EAGE, 2015.

4. Reduzimos em 55% o tempo de perfuração de poços no pré-sal. **Petrobras**, 01 de jul. de 2014. Disponível em: ¡http://www.petrobras.com.br/fatos-e-dados/reduzimos-em-55-o-tempo-de-perfuracao-de-pocos-no-pre-sal.htm¿. Acesso em: 08 de jul. de 2019.

5. *A. Ravindran, D. T. Phillips & J.J. Solberg*, **Operations Research: Principles and Practice.** Ed. 2. Canadá: John Wiley & Sons, 1987.

6. Árvores de decisão feitas utilizando as ferramentas do site www.lucidchart.com

7. Gráficos feitos no programa Geogebra

6 Agradecimentos

- Ao Conselho Nacional de Desenvolvimento Cientifico e Tecnológico - CNPq;
- Ao meu orientador Prof. Dr. Lúcio Tunes dos Santos, pela dedicação e apoio durante o projeto.