

Inspere
Graduação em Ciências Econômicas

Eduardo Carrijo Pires

Modelos de Escolha Estocástica

São Paulo

2021

Eduardo Carrijo Pires

Modelos de Escolha Estocástica

Monografia apresentada ao programa de Graduação em Ciências Econômicas como requisito parcial para obtenção do título de Bacharel em Economia.

Orientador: Professor Doutor José Heleno Faro.

São Paulo

2021

Carrijo Pires, Eduardo

Modelos de Escolha Estocástica

Eduardo Carrijo Pires – São Paulo, 2020

36.f

Monografia – Insper, 2021

Orientador: José Heleno Faro

1. Probabilidade. 2. Escolha 3. Violações 4.
Suposições

Eduardo Carrijo Pires

Modelos de Escolha Estocástica

Monografia apresentada ao programa de Graduação em Ciências Econômicas como requisito parcial para obtenção do título de Bacharel em Economia.

Orientador: Professor Doutor José Heleno Faro.

Banca Examinadora

Professor Doutor José Heleno Faro
Insper

Professor Antônio Bruno de Carvalho Morales
Insper

Agradecimento

É extremamente difícil expressar o quão grato sou pelas pessoas que me ajudaram e me incentivaram durante toda a graduação. São elas, minha família, meus professores, em especial, Hugo Valesini, Sérgio Martins, Gino Olivares, José Faro, Vínicius Müller, Louis de Oliveira, além de todos os outros profissionais do Insper; por fim, meus amigos.

Em mais detalhes, quero agradecer o meu orientador professor José Heleno Faro por sua disposição, dedicação e paciência em ter me ajudado sempre nas dificuldades enfrentadas com o projeto e em ter disponibilizado as melhores referências possíveis.

Em segundo, agradecer minha família que sempre me apoiou e me incentivou nos momentos difíceis que passei. Por último, agradecer o professor Antônio Bruno de Carvalho Morales por ter dedicado seu tempo para ler meu TCC e fazer as melhores críticas e sugestões, tornando o trabalho melhor.

Resumo

O objetivo do presente trabalho é apresentar a literatura de escolha estocástica, discutindo suas fundamentações normativas assim como alguns de seus problemas descritivos. Algumas das principais funções de escolhas estocásticas são introduzidas e os axiomas que fundamentam suas representações são apresentados. Especial atenção é dada ao modelo padrão de Luce, também conhecido como *multinomial logit*, levando à discussão do axioma conhecido como independência das alternativas irrelevantes (IIA, do inglês) e de sua representação como um modelo de utilidade randômica (“random utility model”). As limitações descritivas dados pelos fenômenos conhecidos como *efeito similaridade* (que surgiu do conhecido “duplicates problem”) e *efeito de atração* (“decoy effect”) são discutidos. Em seguida, são apresentados alguns modelos alternativos que enfraquecem o IIA e que são consistentes com tais fenômenos. Alguns exemplos são discutidos ao longo do texto, ilustrando os diferentes modelos.

Palavras-chave: 1. Probabilidade. 2. Escolha 3. Violações. 4. Suposições

Abstract

The aim of the present work is to present a literature of stochastic choice, discussing its normative foundations as well as some of its descriptive problems. Some of the main functions of stochastic choices are introduced and the axioms that underlie their representations are necessary. Special attention is given to Luce's standard model, also known as multinomial logit, leading to the discussion of the axiom known as independence from irrelevant alternatives (IIA) and its representation as a random utility model ("random utility model"). The descriptive limitations of data due to the phenomena known as the similarity effect (which arose from the well-known "duplicate problem") and the attraction effect ("decoy effect") are discussed. Next, there are some alternative models that weaken the IIA and are consistent with such phenomena. Some examples are discussed throughout the text, illustrating the different models.

Keywords: 1. Probability. 2. Choose 3. Violations. 4. Assumptions

Sumário

1.	INTRODUÇÃO	9
2.	REVISÃO DA LITERATURA	13
3.	METODOLOGIA	17
3.1	Fundamentação Axiomática Modelo de Luce.....	17
3.2	Problemas Descritivos	21
3.3	Fundamentação Axiomática Nested Logit	23
3.4	Fundamentação Axiomática Eliminação por Aspectos.....	24
3.5	Fundamentação Axiomática Modelo de Luce com Associação.....	29
4.	CONCLUSÕES	36
5.	REFERÊNCIAS	37

1. Introdução

Nos estudos desenvolvidos pelos economistas, uma questão central é como os indivíduos tomam suas decisões. Vários modelos se consolidaram ao longo do tempo, mas sempre apresentando inovações devido a observações que mostravam certos problemas descritivos. Com isso, tais modelos estão em constante aperfeiçoamento, dado que passam a ser adaptados com as observações empíricas e com a incorporação de características não consideradas anteriormente. Notadamente, observa-se na literatura o desenvolvimento de diferentes modelos para o problema da escolha, utilizando conceitos conhecidos como funções de utilidade, preferências e regras de escolha determinísticas. Posteriormente, estas últimas, foram adaptadas para o contexto de escolhas estocásticas, levando ao conceito de funções de escolha estocástica.

Os modelos de escolha padrões encontrados nos livros textos usualmente supõem agentes racionais dotados de preferências que satisfazem os axiomas de completude, continuidade, transitividade, dentre outros, assegurando a maximização de utilidade. Entretanto, isso não é verificado em muitas situações, como ocorre quando há violações da transitividade. Por exemplo, se um indivíduo prefere beber Coca-Cola à Pepsi e prefere beber Pepsi à Guaraná, a transitividade impõe que ele escolherá sempre o primeiro ao terceiro. No entanto, tal comportamento não é verificado em muitos casos já que o tomador de decisão poderá escolher Guaraná quando se depara com o menu de opções que contenha somente Guaraná e Coca-Cola. Numa perspectiva padrão, esse comportamento é visto como inconsistente por violar transitividade. A abordagem de maximização de utilidade não incorpora a possibilidade de um indivíduo escolher 60% das vezes um determinado produto e em 40% das vezes uma outra opção. Em consequência disso, o modelo padrão ignora várias possibilidades observáveis, já que os indivíduos diante de situações muito semelhantes podem vir a escolher opções diferentes. A área de escolha discreta surge com desenvolvimentos teóricos que incorporam a possibilidade de um indivíduo fazer escolhas distintas em situação idênticas ou semelhantes.

Nesse sentido, com desenvolvimentos alternativos à abordagem clássica de funções utilidade, surgiram teorias que incorporam um fator estocástico à tomada de decisão, em que se permite que os agentes econômicos escolham

de forma randômica as opções. Nessa classe de modelos, o mais conhecido foi o modelo de Luce (1959) em que cada alternativa x , dentro de um determinado menu de opções A , possui um valor, chamado valor de Luce e a probabilidade de se escolher tal alternativa será o seu valor dividido pelo valor dado pela soma dos valores de todas as opções pertencentes ao menu. O modelo de Luce assume a hipótese de que todas as opções possuem probabilidades positivas de serem escolhidas. A hipótese mais importante que caracteriza este modelo, conhecida como Independência das Alternativas Irrelevantes, é que dada duas opções arbitrárias x e y , a razão de suas probabilidades será constante independente do menu de opções em que estejam contidas.

Nesse sentido, se estabelecem fórmulas para o cálculo da probabilidade de se escolher a opção x dentre o menu de opções A , como por exemplo a fórmula $p_x(A)$. Esta expressão pode ser interpretada de duas formas: (i) este valor pode representar a probabilidade de um indivíduo escolher a opção x no menu A ; ou (ii) a proporção dos indivíduos numa dada população que escolhem a opção x no menu A , supondo que cada um integrante desta população escolhe somente uma das opções no menu A .

Macfadden (1973) fez importantes contribuições ao provar que, sob certas condições razoáveis, o modelo de Luce coincide com a versão de utilidade randômica do modelo Multinomial Logit que é uma formulação muito usada por estatísticos quando se quer prever probabilidades ou proporções de uma certa variável dependente explicada por variáveis independentes. Sendo assim, tal interpretação é muito aplicada na literatura de organização industrial quando se calcula a participação de mercado das empresas. Por exemplo, suponha um menu de alternativas A em que há apenas três marcas de refrigerante, desse modo, qual será a possível fração da população que escolherá a marca x de refrigerante? Sendo assim, pode-se concluir que o conhecimento das participações de mercado de cada marca é uma informação muito importante para as escolhas de produção, o que envolve as decisões sobre a demanda de insumos e a oferta do produto final. Nesse sentido, o modelo Multinomial Logit se consolidou como uma importante ferramenta para a estimação de demanda.

Entretanto, foram constatados alguns problemas descritivos do modelo no qual ferem algumas suposições, como o caso do efeito similaridade que viola a hipótese de Independência das Alternativas Irrelevantes (IAI), quando se observa a inclusão de produtos similares. Por exemplo, considere um restaurante com um menu oferecendo apenas as opções de peixe e de peito de frango, onde, segundo a demanda observada, cada opção tem probabilidade de 50% de ser escolhida. Entretanto quando se vai em um outro restaurante onde há um cardápio que contenha coxa de frango além das duas opções anteriores, faz sentido considerar a possibilidade de que a probabilidade de escolher peixe não muda, mas por outro lado, a probabilidade de se escolher peito de frango diminui – visto que há opção similar com a do peito de frango – fazendo com que a razão entre a probabilidade de se escolher peixe e a probabilidade de se escolher frango no primeiro menu de opções seja diferente da mesma razão observada no menu do segundo restaurante, violando assim o IAI.

Ademais, vale ressaltar que uma hipótese importante para o modelo de Luce é a regularidade. Considere a probabilidade de se escolher uma dada opção x num menu A . Ao se retirar qualquer opção do menu a probabilidade de se escolher x não pode diminuir. Analogamente, ao se introduzir uma nova opção a probabilidade de se escolher x não pode aumentar. Entretanto há casos em que tal hipótese não é verificada, ou seja, quando a adição de um elemento novo no menu de opções faz com que aumente a possibilidade de alguma opção que já pertencia ao cardápio. Tal efeito de uma alternativa sobre a outra é chamada de efeito *Decoy*. Por exemplo, imagine que o indivíduo se depare com duas opções de celulares, um de 280 gigas e outro de 560 gigas, o primeiro custando 800 reais e o segundo 1100 reais, com isso, se adicionarmos um outro celular que tenha 380 gigas custando 990 reais, faz com que a probabilidade de se comprar o celular de 560 gigas aumente visto que houve uma valorização deste quando a nova opção foi adicionada.

Com isso, tendo em vista tais fenômenos é possível constatar que o modelo de Luce apresenta algumas limitações, desse modo, o modelo pode se afastar da realidade, encontrando resultados menos consistentes. Em consequência disso, surgem novos modelos como o de Eliminação por Aspectos que lida com a falta da Independência das Alternativas Irrelevantes, o modelo de

Luce com Associação desenvolvido para resolver cenários onde há o efeito *Decoy*. Sendo assim tem-se como objetivo analisar e discutir uma série de restrições do modelo de Luce em mais detalhes e apresentar alguns desenvolvimentos mais recentes que levam em consideração tais possibilidades comportamentais.

Por fim, a discussão a seguir é desenvolvida em 4 partes. Seção 1, a revisão de literatura em que será mencionado estudos anteriores e posteriores ao de Luce e suas comparações. Seção 2, discussão da metodologia axiomática dos seguintes modelos: Modelo de Luce, Nested Logit, Eliminação Por Aspectos, Modelo de Luce com Associação, além disso, apresentar os problemas descritivos de tais trabalhos. Seção 3, a conclusão do presente trabalho.

2. Revisão da Literatura

A literatura de escolha estocástica começou com Thurstone (1927) onde teve como principal objetivo formular um modelo racional e quantitativo que incorpora o conceito de dispersão discriminatória na lei do julgamento comparativo que é o fato de que o indivíduo sob estímulos muito semelhantes desperta processos discriminatórios diferentes, isto é, toma decisões diferentes em situações parecidas. Para isso, o autor propôs uma escala psicológica ou a chamada *continuum* psicológico que é a escala existente nos pensamentos do consumidor, sendo esta, diferente de uma escala padrão com números de medidas prefixado onde o indivíduo julga suas preferências de acordo com tais números. Sendo assim, tal escala é obtida por uma escala ordinal e esta, por sua vez criada a partir de pares de estímulos no qual incorporará o conceito de julgamentos comparativos. Sendo assim, ele modela tal característica psicológica em função das diferenças entre dois estímulos e na medição da dispersão discriminatória (diferenciação perceptíveis dos estímulos) e a correlação entre tais desvios discriminatórios. Ademais, o modelo tinha como principal suposição de que haveria uma constante dispersão discriminatória.

Posteriormente, surgiu o modelo de Luce (1959), dando um destaque maior a essa área, onde criou-se uma função de escolha estocástica ρ no qual o indivíduo escolherá de forma randômica as opções em um dado menu de alternativas; sendo $\rho_x(A)$ a probabilidade de o indivíduo escolher a opção x dentro menu de alternativas A . Dessa maneira, para o modelo ser condizente com as situações de escolha, deve-se satisfazer as suposições Suporte Cheio aonde todas as opções de um menu de alternativas possuem probabilidade entre zero e um, o axioma de escolha e a principal, e a independência das alternativas irrelevantes (IAI) em que a inserção de uma opção gera uma queda proporcional das probabilidades das outras opções. Ademais, Luce prova que os dois últimos são equivalentes. Por fim, neste modelo cada probabilidade de uma opção x ser escolhida está descrita no valor de Luce v_x .

Sob uma outra perspectiva, Macfadden (1973) propôs uma generalização para a formulação de modelos econométricos que capture o comportamento da

escolha da população a partir das distribuições de regras de decisão do indivíduo (qualitativamente e quantitativamente), chamado de Multinomial Logit, além disso, provou que tal modelo equivale ao modelo de Luce, mas esta prova só se verifica quando existe a função ρ satisfazendo o Suporte Cheio e se esta seguir um modelo de utilidade randômica \tilde{U}_x que por sua vez, dependa de uma função de utilidade determinística $w(x)$ advinda da microeconomia mais um choque de utilidade randômica não observada $\tilde{\varepsilon}_x$, este último se encontra no conjunto de funções de utilidades aleatórias Ω em que abrange todas as possíveis mudanças de preferências do tomador de decisão. Ademais, tais choques precisam ser independentes e identicamente distribuídos e por fim, possuírem distribuição valor extremo com média zero. Ademais, tal resultado serviu para fazer aplicações em diversas áreas como escolha de número de filhos, migração, participação do mercado de trabalho, escolha de marcas de carro etc.

Dessa forma, surgiram alguns problemas descritivos referente ao Modelo de Luce, como o efeito de Similaridade ou *Duplications Problem* analisado por Debreu. São cenários em que a Independência das Alternativas Irrelevantes é violada, ou seja, quando surge um novo elemento no menu de alternativas e a probabilidade deste não diminui a probabilidade dos outros elementos, proporcionalmente. Tal irregularidade pode ser encontrada no paradoxo de *Red-bus/Blue-bus*, onde há duas opções para viajar, a de ônibus – três concorrentes, mas com serviços similares – e a de trem – duas concorrentes com o mesmo conforto. Entretanto, nem sempre as 5 alternativas estão disponíveis, sendo assim, supondo em um cenário que só há uma empresa de ônibus e uma de trem e que as probabilidades do indivíduo escolher tais opções são de 60% e 40% respectivamente, no entanto, surge a segunda empresa de ônibus e com isso, é razoável concluir que essa nova opção em nada impactará a probabilidade do indivíduo escolher a empresa de trem, somente a empresa de ônibus, desse modo, a empresa de trem continua com probabilidade de 40% de ser escolhida enquanto a outra terá uma redução de chance. Por fim, é razoável supor que o resultado seria análogo se surgisse a segunda empresa de trem. Sendo assim, para que existisse o IIA a empresa de trem e a primeira, de ônibus deveria diminuir proporcionalmente, para manter a razão de antes (1,5).

Tendo em vista tal irregularidade, surge o modelo Amos Tversky (1972) em que basicamente deixava de analisar as probabilidades pela escalabilidade simples proposto por Luce onde cada opção x possuía o respectivo valor $u(x)$. Sendo assim, o autor começa a analisar as alternativas em termos de características e cada uma delas possuía uma respectiva proporção ao valor da probabilidade de se escolher determinada opção. Desse modo, a cada estágio se selecionava uma característica fazendo com que certas opções fossem eliminadas do conjunto de alternativas até permanecer apenas uma opção. Com isso chegava à conclusão de que a probabilidade de se escolher x era a probabilidade de escolher os aspectos de x .

Entretanto, surge um novo problema descritivo, chamado de *Decoy Effect*, em que existem cenários onde a regularidade não é verificada, ou seja, o incremento de um novo elemento no menu de alternativas faz com que aumente a probabilidade de que uma outra opção seja escolhida. Tal anormalidade pode ser encontrada em algumas estratégias de marketing onde existem dois produtos (A e B) no mercado, e a empresa resolve lançar um novo produto parecido com o A, entretanto, inferior a este em todas as características, mas superior em algumas características do produto B, desse modo, a probabilidade do produto A aumenta com a chegada do novo produto, fazendo com que o produto B tenha uma menor chance de ser escolhido. Um exemplo claro foi o caso das assinaturas da revista The Economist aonde a assinatura digital era \$59,00 e assinatura completa (Digital+Impressa) era de \$125,00 e com isso muitos escolheram a primeira opção, entretanto, quando se coloca a terceira opção que é a assinatura impressa a \$125,00 também, faz com que muitos optem por fazer assinatura híbrida, ou seja, aumentando a chance de esta ser escolhida. Sendo assim, para que existisse a regularidade, a chance de os indivíduos escolherem a assinatura híbrida teria que no máximo permanecer constante.

Em consequência disso, surge o Modelo de Li & Tang (2016) que lida com o *Decoy Effect*, no qual é chamado de Modelo Luce com Associação, ou seja, uma extensão de Luce em que cada alternativa é caracterizada pelo valor de Luce atribuída àquela opção mais um valor de associação em que denota a capacidade de tal opção atrair a atenção do indivíduo. Ele se baseia na ideia de

que mesmo que o indivíduo não seja atraído pela alternativa x na primeira vez, ele pode ser atraído por uma alternativa que possua associação com ela e dessa forma, passará a considerá-la em sua tomada de decisão. Por exemplo, o indivíduo vai em uma livraria na busca de um livro que traria um alta utilidade para ele, mas quando chega no local se depara com outros livros também e passa a achar um deles mais desejável do que o livro que o fez ir na livraria. Além disso, tal modelo é fundamentado em duas fórmulas a de saliência e desejabilidade (serão discutidas na seção de metodologia). Desse modo, traz resultados consistentes em cenários onde a regularidade e a transitividade estocástica não são verificadas.

Por fim, Koppelman em seu paper (The Generalized Nested Logit Model) estuda 4 comportamentos de escolha de viagens intermunicipais de cidades que estão entre Montreal e Toronto. Para isso, utiliza um modelo criado por eles, o *Generalized Nested Logit* (GNL) em que incorpora que a elasticidade cruzada de pares de opções pode ser diferente através da alocação fracionária de cada alternativa para um menu de *nests*, cada um dos quais tem um parâmetro de dissimilaridade. Além disso, conclui que os modelos anteriores – Nested Logit (NL), Paired Combinatorial Logit (PCL), Cross-Nested Logit (CNL) etc. – são todos casos particulares do GNL, sendo assim, podendo ser utilizados apenas quando forem satisfeitas as devidas suposições. Ademais, verifica que o modelo possui uma metodologia para melhor explorar as elasticidades cruzadas das opções sem precisar estimar uma grande quantidade de modelos distintos, como é necessário quando se estima pelo modelo NL.

3. Metodologia

3.1 Fundamentação Axiomática Modelo de Luce.

Os modelos de escolha estocástica no contexto discreto assumem a existência de um conjunto finito e não vazio X , o que determina a coleção de todos os menus de alternativas, que é dado pelo conjunto das partes de X :

$$2^X := \{A: A \subseteq X\}.$$

Assim, tem-se que o número de possíveis menus não vazios é dado por:

$$\text{card}(2^X) - 1 = 2^{\text{card}(X)} - 1.$$

A principal classe de funções considerada em modelos de escolha estocástica toma as aplicações

$$\rho: 2^X \times (2^X \setminus \{\emptyset\}) \rightarrow [0,1].$$

Uma aplicação ρ é chamada de uma função de escolha estocástica se para todo $A \subseteq X$, $\rho(A, A) = 1$ e a função

$$\rho_x(A): X \rightarrow [0,1]$$

satisfazer, para todo $x \notin A$,

$$\rho_x(A) \equiv \rho(x; A) := \rho(\{x\}, A) = 0$$

por último,

$$\sum_{x \in A} \rho_x(A) = 1$$

Tendo em vista a estrutura moldada acima, pode-se introduzir a primeira propriedade de Luce que é a **Positividade** aonde $\rho(x; \{x, y, z\}) \equiv \rho(x; y, z) \in (0,1)$ para todo x e y pertencendo ao menu de alternativas X . Isto é, a probabilidade de se escolher qualquer x em um menu binário que contenha a opção x e y estará sempre entre zero ou um. Desse modo, ela cai na suposição de que por mais que uma alternativa seja muito ruim, o consumidor ainda tem chance de escolhê-la por erro.

Ainda, uma condição mais forte a descrita acima é o **Suporte Cheio**, aonde $\rho_x(A) \in (0,1)$ para todo x pertencente ao menu de alternativas “A” e que tal menu seja diferente de vazio. Isto é, a probabilidade de se escolher qualquer opção x em um menu genérico A sempre estará no intervalo de zero a um. Desse modo, é uma condição mais forte do que a de positividade.

Com isso, tendo em vista as suposições, é possível introduzir o **axioma da escolha** que é quando para todos os conjuntos não vazios (menus de alternativas com pelo menos uma opção) tal que $B \subseteq A \subseteq X$, $B' \subseteq A \subseteq X$, temos que:

$$\rho_x(A) = \rho_x(B)\rho(A; B)$$

$$\rho_x(A) = \rho_x(B')\rho(A; B')$$

Isto é, a probabilidade de escolher a opção x do menu de alternativas A é igual a probabilidade de escolher primeiro o sub-menu de alternativas B contido em A e depois escolher o elemento x de B . É análogo para o B' .

Por exemplo, tal axioma afirma que a probabilidade de um indivíduo comprar o livro “Irracionalidade dos Agentes Econômicos” na Amazon é a mesma probabilidade de o tomador de decisão ir primeiro na subseção da Amazon de livros de psicologia (B) e depois escolher tal livro e que, por sua vez é a mesma probabilidade de ele ir à subseção da Amazon de livros de economia (B') e em seguida, escolher tal livro.

Além disso, vale ressaltar que se ρ satisfizer o Axioma Escolha, temos que então a positividade valerá se somente se o Suporte Cheio valerá, ou seja, $\rho_x(A) > 0$ para todo menu de opções dado que $A \subseteq X$ e $x \in A$.

Ademais, a condição de Luce mais relevante do comportamento dos indivíduos é a chamada Independência das Alternativas Irrelevantes. A equação é escrita da seguinte forma:

$$\beta(x, y|x \cup y) := \frac{\rho(x; y)}{\rho(y; x)} = \frac{\rho_x(A)}{\rho_y(A)}$$

Isto é, a razão entre a probabilidade de escolher x quando se tem um menu binário de alternativas (x e y) e a probabilidade de escolher y quando se

tem esse mesmo menu de alternativas tem que ser igual a razão entre a probabilidade de escolher x quando se tem um outro menu de alternativa que contenha x e y , e a probabilidade de escolher y quando se tem esse mesmo outro menu de alternativas.

Por exemplo, se o indivíduo encontra uma barraca de bebidas e lhe oferecem suco e água com probabilidade de escolha de 60% e 40% respectivamente, a razão entre o primeiro e o segundo produto será de 1,5. Entretanto, na barraca ao lado há refrigerante além das outras duas bebidas, desse modo a inserção do terceiro elemento no menu de opções provocará uma queda proporcional na probabilidade das duas outras bebidas, dado que exista o IAI. Desse modo, se consideramos 50% a probabilidade de eles escolher refrigerante, a probabilidade de ele escolher suco e água será de 30% e 20% respectivamente, de forma a manter o valor para as duas razões.

Além disso, vale ressaltar que se ρ satisfizer a positividade, o Axioma da Escolha é válido se e somente se o IAI é válido.

Com tais propriedades se chega no **Teorema 1 (Luce 1959)**: A função de escolha estocástica ρ satisfaz positividade e o IIA se e somente se existir $v: X \rightarrow (0, +\infty)$ (ou seja, cada opção escolhida x tem um valor positivo de Luce v_x) de modo que para todo menu de opções A não vazios e $x \in A$. Temos que:

$$\rho_x(A) = \frac{v_x}{\sum_{y \in A} v_y}$$

Uma outra condição necessária ao modelo é que satisfaça a **Regularidade (Monotonicidade)**, isto é, quando para todo $x \in X$ e A, B sendo menus de opções não vazios tal que $B \subseteq A$:

$$\rho_x(B) \geq \rho_x(A)$$

Por exemplo, se um indivíduo se encontra com um menu com apenas duas cervejas Heineken e Colorado e que tem probabilidade de 60% e 40% respectivamente e de repente lhe é apresentado uma terceira cerveja sendo ela a Skol, desse modo, como o menu cresceu para três cervejas é esperado que a

probabilidade de se escolher a Hineken não aumente, isto é, possa cair ou no máximo manter o valor de que era antes com apenas duas opções.

Desse modo, incorporando todas as características descritas acima, pode-se chegar no modelo **Multinomial Logit**:

$$\frac{\exp(u_x)}{\sum_{y \in A} \exp(u_y)}$$

(Substituindo os valores de Luce v_x por $\exp(u_x)$, aonde u_x é o nível de utilidade gerado pela escolha da opção x e a função utilidade $u: X \rightarrow \mathbb{R}$ é único, sendo uma constante).

Por fim, estudos de autores posteriores nessa área (Block e Marschack), (1960) desenvolveram o modelo de função de utilidade randômica em um dado menu de alternativas X :

$$\tilde{U} = w(x) + \tilde{\varepsilon}_x$$

aonde $w: X \rightarrow \mathbb{R}$ é a função de utilidade determinística e $\tilde{\varepsilon}: \Omega \rightarrow \mathbb{R}^x$ é o choque aleatório e não observado de utilidade (assumindo que a distribuição de $\tilde{\varepsilon}$ tem densidade e satisfaz o Suporte Cheio). Ainda, se assumirmos que $\tilde{\varepsilon}$ for independente e identicamente distribuído a partir de $x \in X$ o modelo acima será equivalente ao modelo Probit se $\tilde{\varepsilon}$ tiver distribuição normal com média zero.

Ainda, tal modelo pode ser interpretado na maneira em que o indivíduo possui diversas funções de utilidade contidas no conjunto Ω , sendo assim:

$$\Omega_{(x;A)} := \{w \in \Omega: U_w(x) \geq U_w(y) \forall y \in A\}$$

pode ser entendido como a situação em que o indivíduo, dado o menu de alternativas A , maximizará sua função de utilidade quando escolhe a opção x . Entretanto, como é uma função randômica, o indivíduo em uma outra ocasião, poderá maximizar sua utilidade escolhendo uma outra opção y , dado que possui diferentes funções de utilidades.

Por exemplo, quando o indivíduo se encontra em um menu de opções em que só há Coca-Cola e Guaraná, possa ser que em um dia ele prefira o primeiro ao segundo, portanto, maximiza sua utilidade escolhendo o primeiro. Entretanto,

possa ser que em outro dia ele prefira o segundo ao primeiro, desse modo, maximizará a função escolhendo o segundo.

Vale ressaltar que Mcfadden (1973) provou em seu teorema de que se ρ tiver uma representação Logit com utilidade w , isto é, ρ sendo equivalente ao modelo de utilidade randômica aonde $\tilde{U} = w(x) + \tilde{\varepsilon}_x$, sendo que $\tilde{\varepsilon}_x$ é independente e identicamente distribuído e tendo uma distribuição de valor extremo com média zero. Desse modo, ρ será equivalente ao modelo de Luce tal que para todo $A \in 2^X \setminus \{\emptyset\}$ e $x \in A$:

$$\rho_x(A) = \frac{\exp(w(x))}{\sum_{y \in A} \exp(w(y))}$$

Note que de $\tilde{\varepsilon}_x$ pode assumir as seguintes funções de probabilidade acumulada:

$$F(x; \mu; \sigma; 0) = e^{-e^{-(x-\mu)/\sigma}}; \xi = 0, \mu = 0, x \in \mathbb{R}$$

$$F(x; \mu; \sigma; \xi) = \begin{cases} e^{-y^{-\alpha}}, & y > 0 \\ 0 & , y \leq 0 \end{cases}; \xi > 0, \mu = 0, x \in \mathbb{R}$$

$$F(x; \mu; \sigma; \xi) = \begin{cases} e^{-(-y)^{\alpha}}, & y < 0 \\ 1 & , y \geq 0 \end{cases}; \xi < 0, \mu = 0, x \in \mathbb{R}$$

Além disso, tais autores provaram o teorema de que se a função de escolha randômica ρ for equivalente a o modelo de utilidade randômica, logo, é satisfeito a regularidade, entretanto, a volta só é verdadeira para menu de alternativas com no máximo três opções.

3.2 Problemas Descritivos

Tendo em vista, a condição de Independência das Alternativas Irrelevantes do modelo de Luce, Debreu conclui que tal comportamento não é verificado em muitos dos casos (efeito similaridade), para isso, comprovou tal violação examinando um exemplo de três gravações, duas de Bethoven parecidas (B_1, B_2) e a de Debussy (D) , em seguida, supôs que o indivíduo é indiferente entre as três opções, ou seja, $\rho(B_1; B_2) = \rho(D; B_1) = \rho(D; B_2) = 1/2$,

desse modo, a razão entre as probabilidades de se escolher a gravação de Debussy e a primeira gravação de Bethoven $\frac{\rho(D;B_1)}{\rho(B_1;D)} = \frac{1/2}{1/2} = 1$, sendo assim, quando adiciona-se a segunda gravação de Bethoven no menu de opções – satisfazendo a condição do IAI – a razão das probabilidades deve-se permanecer a mesma $\frac{\rho(D;B_1,B_2)}{\rho(B_1;D,B_2)} = \frac{1/3}{1/3} = 1$, entretanto, espera-se que essa probabilidade $\rho(D;B_1,B_2) = 1/2$ pelo simples fato de que adicionar uma mesma gravação de Bethoven não mudará a indecisão do indivíduo de escolher Debussy tendo as três gravações como o menu de alternativas.

Além disso, uma outra hipótese fundamental ao modelo padrão é a Regularidade, entretanto, em muitos cenários tal hipótese não é verificada. Por exemplo, suponha que o indivíduo vá ao cinema e se depara com um cardápio que contenha pipoca pequena (p) por R\$10,00 e a grande (g) por R\$ 20,00, e as probabilidades de se escolher a pequena e a grande são respectivamente, $\rho(p;g) = 2/3$ e $\rho(g;p) = 1/3$. Por outro lado, tal indivíduo vai outra vez ao cinema e agora se depara com um cardápio novo que contém além do tamanho pequeno e grande, o médio (m) cujo preço é de R\$18,00, desse modo, seguindo a base normativa dos modelos anteriores espera-se que essa adição faça com que no máximo $\rho(g;p,m) = 1/3$, entretanto, isso não é verificado em muitos cenários, visto que, a inclusão dessa nova opção provoca uma valorização sobre o maior tamanho por parte do indivíduo já que a pipoca grande é apenas R\$2,00 mais cara que a média. Em consequência disso, a probabilidade de se escolher a opção grande aumenta de forma que $\rho(g;p,m) > 1/3$, configurando assim, o efeito *Decoy*.

Em consequência disso, surgem modelos como Nested Logit e o de Eliminação por Aspectos que lidam com casos em que o IAI não é validado e por outro lado, há o modelo de Luce com Associação que lida com a falta de regularidade. Tais modelos serão apresentados a seguir.

3.3 Fundamentação Axiomática Nested Logit

A principal característica do modelo logit aninhado é que as alternativas podem ser subdivididas em subconjuntos chamados nests, onde Independência das Alternativas Irrelevantes se mantém dentro de cada subconjunto (nest), entretanto, isso não se verifica para opções em subconjuntos diferentes.

Este modelo assume que o conjunto de alternativas X com m opções podem ser subdivididas em n subconjuntos sobrepostos denotados por B_1, \dots, B_n ninhos.

O Modelo usa a formulação de utilidade Randômica descrita anteriormente, aonde o tomador de decisão obtém da alternativa x encontrada no subconjunto B_k , em que $1 \leq k \leq n$

$$\tilde{U} = v_x + \tilde{\varepsilon}_x$$

Aonde, $w(x)$ é a utilidade observada e $\tilde{\varepsilon}_x$ a utilidade aleatória tal que o vetor $\varepsilon := (\varepsilon_1, \dots, \varepsilon_n)$ possui distribuição valor extremo, onde a distribuição marginal de ε_x é univariada.

Ainda, note que a função de distribuição acumulativa do vetor aleatório é:

$$\exp \left(- \sum_{k=1}^n \left(\sum_{x \in B_k} \exp(-\varepsilon_x / \lambda_k) \right)^{\lambda_k} \right)$$

onde, o parâmetro λ_k é o grau de independência do choque aleatório entre as opções no subconjunto (nest) k , note que quanto maior for o parâmetro, maior o grau de independência.

Ademais, vale ressaltar que diferentemente do modelo Logit, os choques podem ser correlacionados entre alternativas dos mesmos subconjuntos (o que garante analisar e estudar cenários em que o IIA não é satisfeito). Entretanto, é assumido que há independência entre alternativas que não se encontram no mesmo subconjunto (ninho).

Sendo assim, com tais resultados e suposições de comportamento dos choques aleatórios Daly and Zachary (1978), McFadden (1978) and Williams (1977) conseguiram modelar a função de Escolha Randômica Nested Logit ρ^{NL} tal que para todo $x \in B_k$:

$$\frac{\exp(v_x / \lambda_k) \left(\sum_{y \in B_k} \exp(v_y / \lambda_k) \right)^{\lambda_k - 1}}{\sum_{l=1}^n \left(\sum_{y \in B_l} \exp(v_y / \lambda_l) \right)^{\lambda_l}}$$

Vale ressaltar que quando $\lambda_k = 1$ (para todo $k \in \{1, \dots, n\}$) cada choque será independente um do outro, coincidindo assim, com o modelo Multinomial Logit aonde nesses cenários o IAI será válido. Por outro lado, se $\lambda_k \in (0,1)$, valerá a função Nested Logit e este será consistente com a maximização de utilidade.

Por fim, Train (2009) provou que o modelo de Eliminação de Aspectos é um caso particular do Nested Logit quando tais parâmetros que capturam a correlação são positivos e próximos a zero.

3.4 Fundamentação Axiomática Eliminação por Aspectos

O Modelo de Eliminação por Aspectos formado por Tversky se baseia na escolha de alternativas – em que cada uma possui um conjunto de características, algumas peculiares e outras comuns às alternativas em um certo conjunto de opções – por meio de estágios de eliminação de características, sendo estas tendo probabilidades proporcionais ao seu peso em cada alternativa.

Por exemplo, o indivíduo se depara com um menu de três opções, moto, carro e bicicleta. O primeiro estágio é a característica de transporte, sendo assim, nenhum é eliminado dado que tal característica é comum às três opções, no segundo estágio é o transporte que apenas duas rodas, desse modo, a opção do carro é eliminada, em seguida, a característica selecionada é que tenha motor, desse modo, a opção de bicicleta é eliminada e sobra apenas a opção de moto. Com isso, o modelo captura a probabilidade de tais características serem

selecionadas para assim obter a probabilidade de se escolher uma moto dentre as outras duas opções.

Por uma perspectiva formal, considere uma opção x que pertence ao conjunto não vazio T de alternativas e seja $x' = \{\alpha, \beta, \dots\}$ o conjunto de características pertencentes a opção x . Voltando ao exemplo descrito acima, o menu de opções seria $T = \{moto, carro, bicicleta\}$, aonde o conjunto de características de cada opção é:

$$x' = moto' = \{meio de transporte, possui motor, duas rodas, adrenalina\}$$

$$y' = carro'$$

$$= \{meio de transporte, possui motor, quatro rodas, confortável\}$$

$$z' = bicicleta'$$

$$= \{meio de transporte, gasta calorias, duas rodas, confortável\}$$

Sendo ω , as características em comum com todas (meio de transporte), tal aspecto pode ser removido dos cálculos dado que ele não elimina nenhuma opção; α, β, γ são as peculiaridades de cada opção (adrenalina, quatro rodas, gasto de calorias); θ, σ, p são respectivamente as características compartilhadas somente pela moto e o carro (possui motor), pela moto e a bicicleta (duas rodas), pelo carro e a bicicleta (confortável).

Seja u a função de escala que associa cada característica a um número positivo que represente a utilidade, e K a soma de todos os valores associados aos aspectos que estão sendo analisados, ou seja, $K = \sum_{\alpha} u(\alpha)$. Sendo assim, pode-se encontrar a probabilidade de se escolher a moto dadas as três opções $P(x; T)$.

Note que x pode ser escolhido de três formas diferentes, a primeira é se for selecionado suas peculiaridades ($u(\alpha) / K$), a segunda forma seria a seleção dos aspectos compartilhados somente por x, y , nesse caso, z seria eliminado ($u(\theta) \cdot \rho(x; y) / K$), por fim, a terceira seria a seleção dos aspectos compartilhados somente por x, z , com isso, y seria eliminado ($u(p) \cdot \rho(x; z) / K$). Desse modo, resulta-se na seguinte equação:

$$\rho(x; T) = (1 / K)[u(\alpha) + u(\theta) \cdot P(x; y) + u(p) \cdot \rho(x; z)]$$

onde:

$$\rho(x; z) = \frac{u(\alpha) + u(\theta)}{u(\alpha) + u(p) + u(\beta) + u(\sigma)}$$

Agora, considere Q como um conjunto qualquer finitos de opções sendo $A \subseteq Q$ de modo que $A' = \{\alpha | \alpha \in x' \text{ para algum } x \in A\}$, ou seja, são características de pelo menos uma opção que pertence ao menu de opções A e $A^0 = \{\alpha | \alpha \in x' \text{ para todo } x \in A\}$, ou seja, são as características em comum de todas as opções pertencentes ao conjunto A . Em consequência disso, o modelo aponta que para toda opção x pertencente a um conjunto $A \subseteq Q$ segue:

$$\rho(x; A) = \frac{\sum_{\alpha \in x' - A^0} u(\alpha) \cdot \rho(x, A_\alpha)}{\sum_{\beta \in A' - A^0} u(\beta)}$$

Onde $\rho(x; A_\alpha)$, é a probabilidade de se escolher x dentre as opções que contenham também a característica α , vale ressaltar que se for escolhida uma peculiaridade de x essa probabilidade será 1. Ainda, se todos os elementos de A compartilham das mesmas características $\rho(x; A) = 1/$ (número de elementos de A).

Por outro lado, considere um menu de opções em que todas as alternativas compartilhadas pelos pares de x, y e z, w são as mesmas, nesse sentido, tais características não afetam as probabilidades de escolha, logo, podem ser removidas do modelo, sendo assim, $x' \cap y' = \emptyset$ para todos x, y pertencentes a T , desse modo é gerado a seguinte equação:

$$\rho(x; A) = \frac{\sum_{\alpha \in x'} u(\alpha)}{\sum_{\beta \in A'} u(\beta)}$$

já que $\alpha \in x'$ implica que $A_\alpha = \{x\}$ e $\rho(x; x) = 1$. Fazendo com que:

$$u(x) = \sum_{\alpha \in x'} u(\alpha)$$

Assim, a equação se reduz ao modelo de Luce, quando as opções analisadas não possuem características em comum:

$$\rho(x; A) = \frac{u(x)}{\sum_{y \in A} u(y)}$$

Além disso, considerando um caso de escolha de probabilidade binária, obtemos o modelo de Restle (1961):

$$\rho(x; y) = \frac{\sum_{\alpha \in x' - y'} u(\alpha)}{\sum_{\alpha \in x' - y} u(\alpha) + \sum_{\beta \in y' - x'} u(\beta)}$$

Onde, $x' - y' = \{\alpha | \alpha \in x' \text{ \& } \alpha \notin y'\}$ é o conjunto de características que pertence a x mas não à y . De modo análogo, $y' - x' = \{\alpha | \alpha \notin x' \text{ \& } \alpha \in y'\}$ é o conjunto de características que pertence à y mas não a x .

Até agora, tal modelo foi formulado como a função de escala u , para isso, é preciso conhecer todas as características de cada alternativa, entretanto, muitas vezes não é possível conhecer os principais aspectos. Apesar disso, é possível trabalhar com o EPA pois podemos formular em termos dos subconjuntos de Q , ou seja, em termos das alternativas. Desse modo, podemos usar a notação $U(\bar{x})$ para denotar o valor do conjunto de todas as peculiaridades da opção x e não mais diferenciar as peculiaridades de x , dado que levam ao mesmo resultado. De forma análoga podemos denotar $U(\bar{x}, \bar{y})$, $U(\bar{x}, \bar{z})$, $U(\bar{z}, \bar{y})$, respectivamente, como o valor do conjunto de todas as características em comum de x e y , em comum de x, z e em comum de z, y . Sendo assim, a função U se comporta como a soma dos valores de escala u . Dessa forma, podemos expressar:

$$\rho_x(T) = \frac{U(\bar{x}) + U(\bar{x}, \bar{y})\rho(x; y) + U(\bar{x}, \bar{z})\rho(x; z)}{U(\bar{x}) + U(\bar{y}) + U(\bar{z}) + U(\bar{x}, \bar{y}) + U(\bar{x}, \bar{z}) + U(\bar{y}, \bar{z})}$$

onde:

$$\rho(x; z) = \frac{U(\bar{x}) + U(\bar{x}, \bar{z})}{U(\bar{x}) + U(\bar{z}) + U(\bar{x}, \bar{z}) + U(\bar{y}, \bar{z})}$$

Desse modo, pode-se concluir que a diferença entre u e U é de que a primeira função é definida em características individuais enquanto a segunda, em conjuntos de características.

Agora considere o conjunto $\bar{A} = \{\alpha \in Q' | \alpha \in x' \text{ para todo } x \in A \text{ \& } \alpha \notin y' \text{ para qualquer } y \notin A\}$, ou seja, tal conjunto engloba as características comuns de todas as opções de A e que não pertencem a nenhuma opção que não esteja em A . Por construção, $U(\bar{A}) = \sum_{\alpha \in \bar{A}} u(\alpha)$.

Vale ressaltar, que tal modelo é válido, se somente se, existe a escala U definida em $\{\bar{A}_i | A_i \subset Q\}$ tal que para todo $x \in A \subset Q$:

$$\rho_x(A) = \frac{\sum_{B_i \ni A} U(\bar{B}_i) P(x, A \cap B_i)}{\sum_{A_j \in W} U(\bar{A}_j)}$$

onde $w = \{A_j | A_j \cap A \neq A, \emptyset\}$, faz com que o denominador seja diferente de zero. Com isso, prova-se que não é necessário possuir os valores das características específicas, podendo assim, ser formulada a partir dos subconjuntos de Q . Portanto, podemos interpretar $U(\bar{A})$ como o benefício de todas as opções pertencentes ao menu A .

Agora, podemos examinar o exemplo de Debreu pelo modelo EPA, desse modo, suponha que todas as características compartilhadas da gravação de Debussy (D) com Bethoven (B1) seja igual as compartilhadas de D com a outra de Bethoven (B2), sendo assim, D é tratado como um conjunto disjunto no sentido das características. Além disso, tome $U(\bar{B}_1) = U(\bar{B}_2) = a$, $U(\bar{B}_1 \bar{B}_2) = b$, por fim, $U(\bar{D}) = a + b$. Com isso, dado que o indivíduo é indiferente quando se depara com tais escolhas binárias, podemos gerar a seguinte equação:

$$\rho(B_1; B_2) = \frac{a}{2a} = \frac{a + b}{2(a + b)} = \rho(D; B_1) = \rho(D; B_2)$$

No entanto, como previsto intuitivamente haverá uma desigualdade entre as probabilidades analisando, em conjunto, as três opções:

$$\rho(D; B_1, B_2) = \frac{a+b}{3a+2b} > \frac{a+b(a,2a)}{3a+2b} = \rho(B_1; B_2, D) = \rho(B_2; B_1, D)$$

Com isso, pode-se concluir que tal modelo lida com o problema levantado por Debreu em que o IAI é violado.

Ademais, segue abaixo uma descrição das suposições e consequências do modelo EPA:

Em primeiro lugar, é assumido no modelo a Regularidade: Para todo $x \in A \subseteq B$, $\rho(x; A) \geq \rho(x; B)$, ou seja, a probabilidade de se escolher o elemento x em um menu de opções com menos alternativas é no mínimo igual a probabilidade de se escolher tal opção em um menu de mais alternativas.

A primeira consequência do modelo é de que há transitividade estocástica aonde $\rho(x; z) \geq 1/2$ e $\rho(y; z) \geq 1/2$ implica que $\rho(x; z) \geq \min[\rho(x; y), \rho(y; z)]$

A segunda consequência do modelo é relacionar as escolhas binárias e trinárias pela propriedade chamada Desigualdade Multiplicativa aonde $\rho(x; y, z) \geq \rho(x; y) \rho(x; z)$, isto é, a probabilidade de se escolher x dentro do conjunto de opções x,y,z é no mínimo igual a multiplicação das probabilidade de se escolher x de um menu x,y e o menu x,z, assumindo que tais escolhas são independentes, generalizando o caso, seria $\rho(x; A \cup B) \geq \rho(x; A)\rho(x; B)$ para todo A, B contido em T.

Por fim, relacionando as equações da Regularidade e da Desigualdade chega no seguinte resultado: $\min[\rho(x; y), \rho(y; z)] \geq \rho(x; y, z) \geq \rho(x; y) \rho(x; z)$. Sendo assim, pode-se concluir que a escolha de probabilidade trinária possui um limite inferior pela consequência da multiplicativa e um superior pela regularidade.

3.5 Fundamentação Axiomática Modelo de Luce com Associação:

Tendo em vista, a condição de regularidade no modelo de Luce, Nested Logit, Eliminação por Aspectos, ou seja, a adição de um elemento novo em um conjunto de alternativas faz com que no máximo se mantenha as probabilidades dos que estavam antes da chegada de tal opção, entretanto, não é verificado em alguns casos quando ocorre o efeito *Decoy* em que a presença de uma nova alternativa faz com que aumente a probabilidade de algum dos elementos. Em consequência disso, os modelos anteriores podem trazer resultados que se afastam da realidade. Por conta de tal problema descritivo surge uma extensão ao modelo de Luce chamado de Luce com Associação no qual lida com essa peculiaridade.

Por uma perspectiva mais formal assumamos um menu de opções $X = \{x, y, z\}$. As alternativas x y z são elementos iguais, entretanto com qualidades diferentes, x e z possuem a mesma qualidade apesar de z ser mais caro e y possui uma qualidade inferior apesar de ser barata. Desse modo, espera-se que $\rho(x; y, z) > \rho(x; y)$, dado que z possui um efeito atração para se escolher x, ou seja, como z e y possui a mesma qualidade do produto eles geram uma

associação entre eles, com isso, se o indivíduo se deparar primeiramente com a opção z ele consideraria escolher a opção x também, por outro lado, isso não acontece com a opção y dado que é de uma qualidade inferior e não terá associação com z .

Seja X um conjunto finito não vazio que engloba todas as alternativas possíveis. Seja D os subconjuntos de X . Assuma que o domínio satisfaça a seguinte condição: $A \in D$ para todo A que contenha 4 opções e $B \in D$ sempre quando $B \subset A$ e $A \in D$. Seja $D_+ = D \setminus \emptyset$. Os elementos de D_+ são vistos como conjuntos factíveis que o indivíduo se depara.

Seja $\rho(A', A)$ a probabilidade de o indivíduo escolher uma alternativa de A' quando se depara com o menu de alternativas A . Além disso, Aditividade é o requisito para que a função $\rho(\cdot, A)$ seja de probabilidade. $\rho(x, A) \in (0,1)$ implica que a função de escolha randômica é não degenerativa, ou seja, cada alternativa x no menu A é escolhida com probabilidade positiva.

Em um menu C que contenha os conjuntos de opções A, B , nós denotamos a probabilidade relativa de escolher as opções em A e B dado o menu C como:

$$\beta(A, B|C) = \frac{\rho(A; C)}{\rho(B; C)}$$

Para todos os menus disjuntos em pares $A, B, E \in D_+$, define-se $\beta(A, B|A \cup B \cup E) - \beta(A, B|A \cup B)$ como o impacto probabilidade relativa de se escolher A e B quando se adiciona o menu E .

Vale lembrar que a função de escolha ρ é uma função de Luce, se somente se, ela satisfaça a Independência das Alternativas Irrelevantes (IAI), com isso no caso mostrado acima:

$$\beta(A, B|A \cup B \cup E) - \beta(A, B|A \cup B) = 0$$

Axioma APR – Aditividade das Probabilidades Relativas: para todos os menus disjuntos em pares $A, B, C \in D_+$:

$$\beta(A \cup B, C|A \cup B \cup C) = \beta(A, C|A \cup C) + \beta(B, C|B \cup C)$$

Tal axioma não implica que A não tem impacto na probabilidade relativa de escolher B e C , ou que B não tem impacto na probabilidade relativa de se escolher A e C , entretanto, implicam que tais impactos se cancelam. Nesse sentido, o teorema mostra que o APR é uma outra formulação do IAI de Luce.

Teorema 1: A função de escolha é uma função de Luce, se e somente se, ela satisfaz o Axioma APR.

Tendo em vista as definições, pode-se chegar no modelo de Luce com Associação em que cada alternativa é caracterizada por um valor de saliência, valor de Luce e o valor de associação das alternativas. Sendo assim, o tomador de decisão é primeiramente atraído por uma opção x de acordo com a fórmula de Luce baseada na saliência e depois escolhe entre as alternativas que se associam com opção x de acordo com a fórmula de Luce baseada na desejabilidade.

Retomando, o Valor de Saliência é uma função $\gamma: D_+ \rightarrow R_{++}$ que captura a habilidade de uma alternativa atrair a o indivíduo na tomada de decisão. É importante notar que $\gamma_x := \gamma(x)$. Tal função deverá ser aditiva como o valor de Luce. Com isso, para todo $A \in D_+$, temos que:

$$\gamma(A) := \sum_{x \in A} \gamma_x$$

Associação das Alternativas: Duas alternativas x, y possuem associação quando mesmo que o indivíduo não seja atraído primeiramente por x ele considerará x em sua tomada de decisão por ter sido atraído por y . Tal característica implica que a probabilidade relativa de x e y dado qualquer conjunto será constante, isto é, seguem o IAI.

Definição 1: duas alternativas x e y são associadas se em qualquer menu de opções A que contenha x e y , $\beta(x, y|A) = \beta(x, y|x \cup y)$.

Denotamos $x \sim y$ se estes são associados. A relação \sim é simétrica e reflexiva. Ademias, denota-se $A \approx B$ se para algum $x \in A$ e $y \in B$, temos que $x \sim y$. Além disso, chama-se que um conjunto de alternativas A é uma coleção de opções com associação se $x, y \in A$ e $x \sim y$, portanto, $A \in \mathcal{L}$. Por fim, se A e B possuem elementos em comum estão conectados, mas mesmo que não tenham

podem estar conectados se existirem associação entre pelo menos um de seus respectivos elementos.

Definição 2: A e B são desconectados quando $x \in A$ e $y \in B$ e x e y não possuem associação. A notação será $A \perp B$. O Axioma seguinte define que \sim possui uma relação de transitividade, ou seja, se $x \sim y$ e $y \sim z$ implica que $x \sim z$.

Axioma T – Transitividade: A relação \sim é transitiva; desse modo, implica que se pode fazer partições sobre o conjunto que engloba todas as alternativas X da seguinte forma: $\{P_i\}_{i=1}^n$, tal que $P_i \in \mathcal{L}$, para qualquer $i = 1, 2, \dots, n$ e $P_i \perp P_j$ para qualquer $i, j = 1, 2, \dots, n$ e $i \neq j$.

Suponha que $A \cup B \perp C$. Desse modo se o indivíduo for atraído pelas opções do conjunto $A \cup B$ ele não considerará as alternativas do menu C quando fizer a escolha e vice-versa. Em outros termos, a probabilidade relativa de se escolher $A \cup B$ e C condicionado a $A \cup B \cup C$ depende apenas da atração de $A \cup B$ relativo à atração de C , e atração $A \cup B$ pode ser separada em duas partes, a atração de A e de B . Em consequência disso, podemos afirmar que tal probabilidade pode ser decomposta em duas, que é a probabilidade relativa de se escolher A e C dado o menu $A \cup C$ mais a probabilidade relativa de se escolher B e C dado o menu $B \cup C$. Por último, segue o axioma:

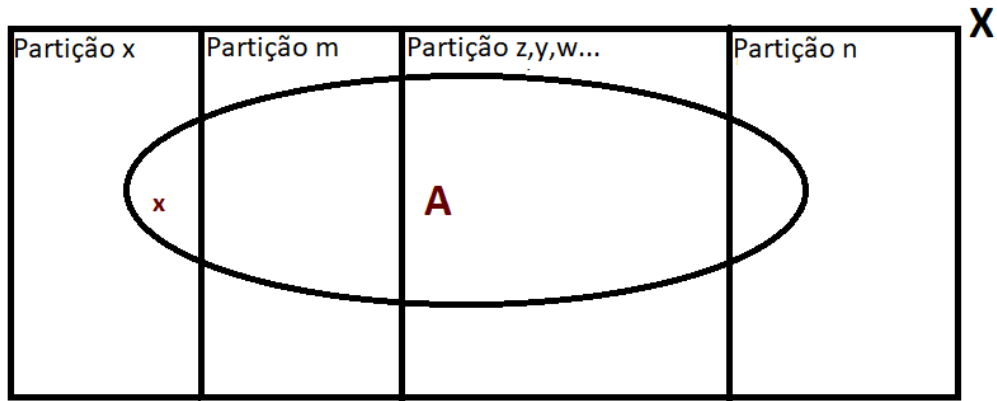
Axioma Fraco APR – Aditividade Fraca de Probabilidade Relativa: se $AB \perp C$ e $A \cup B \perp C$, então:

$$\beta(A \cup B, C | A \cup B \cup C) = \beta(A, C | A \cup C) + \beta(B, C | B \cup C)$$

Modelo de Luce com Associação: para uma partição de X e $x \in X$, denota-se P_x a área que contenha x . Assim, ρ será uma função de associação de Luce se existe uma partição $(P_i)_{i=1}^n$ de X , um valor de saliência γ e o valor de Luce v tal que:

$$\rho_x(A) = \frac{\gamma(A \cap P_x)}{\gamma(A)} \frac{v(x)}{v(A \cap P_x)}$$

desde que $x \in A \in D_+$. Além disso, $((P_i)_{i=1}^n, \gamma, v)$ representa ρ se a equação acima assegura para qualquer $x \in A$. Segue abaixo uma representação das partições:



Axioma Fraco APR*: Se $A \cap B = \emptyset$ e $A \cup B \perp C$ onde $C \in \mathcal{L}$, então $\beta(A \cup B, C | A \cup B \cup C) = \beta(A, C | A \cup C) + \beta(B, C | B \cup C)$.

É importante notar que o Axioma APR implica no Axioma APR*. O Teorema a seguir mostra que o modelo de Luce com Associação é a única função que satisfaz o Axioma T e o Axioma Fraco APR*.

Teorema 2: O modelo de escolha é o de Luce com Associação, se somente se, satisfaz o Axioma T e o Axioma Fraco APR*

Por exemplo, seja $X = \{x, y, z, w\}$, suponha que $x \sim y$ $z \sim w$ e que $(\{\{x, y\}, \{z, w\}\}, \gamma, v)$ represente a função ρ , onde $\gamma_x = \gamma_z = \gamma_w = 1$, $\gamma_y = 2$, $v_x = v_w = 1$, $v_y = v_z = 2$. Com isso é simples de verificar que $\beta(x, z | x \cup y \cup z) = \beta(x, z | x \cup z)$ e $\beta(x, w | x \cup y \cup w) = \beta(x, w | x \cup w)$. Isto é, y não tem nenhum impacto nas probabilidades relativas apesar de ter associação com x .

Sendo assim, concluímos que $x \in X$ é uma opção isolada se para qualquer $y \in X$, não tem associação com x ou que $\beta(x, z | x \cup y \cup z) = \beta(x, z | x \cup z)$. Dessa forma, denota-se o conjunto de alternativas isoladas sob a função ρ de S_ρ .

Teorema 2: se $((P_i)_{i=1}^n, \gamma, v)$ representa a função de escolha ρ , então $P_i \cap S_\rho = \emptyset$ ou $P_i \subset S_\rho$ para todo $i \in \{1, 2, \dots, n\}$. Além disso, se $((P_i)_{i=1}^n, \gamma, v)$ e $((P_j)_{j=1}^m, \gamma', v')$ representa ρ , $P'_i \subset X \setminus S_\rho$ implica que P'_i é um bloco de $(P_i)_{i=1}^n$.

Por último, vale ressaltar que o modelo de Luce com Associação se reduz ao Modelo de Luce tradicional quando todas as alternativas possuem

associações com elas mesmos, visto que, as razões dos gamas (γ) serão iguais a 1 e apenas restará os valores de Luce (v).

Tendo em vista a anomalia da irregularidade, segue um exemplo em que o modelo de Luce com Associação lida:

Seja $\gamma_x = 1$, $\gamma_y = \gamma_z = 2$, $v_x = v_y = 2$, $v_z = 1$, onde x e z estão no mesmo bloco e y em um diferente. Sendo assim, temos o efeito atração:

$$\rho(x; y, z) = \frac{\gamma_x + \gamma_z}{\gamma_x + \gamma_y + \gamma_z} \frac{v_x}{v_x + v_z} = \frac{2}{5}$$

$$\rho(x; y) = \frac{\gamma_x}{\gamma_x + \gamma_y} = \frac{1}{3}$$

Proposição 1: As duas sentenças são equivalentes

1. $\rho(x; y, z) > \rho(x; y)$;
2. x e z estão no mesmo bloco e y em um diferente, isto é,

$$\frac{v_x}{v_z} > \frac{\gamma_x}{\gamma_z} \frac{\gamma_x + \gamma_y + \gamma_z}{\gamma_y}$$

Com isso, pode-se concluir que se x e z possuem associação, mas x e y não, a adição da alternativa z diminui a probabilidade de o indivíduo ser atraído primeiramente por x e y , mas por outro lado, quando o tomador de decisão é atraído por z ele ainda considera x mas y ele descarta, desse modo, a probabilidade de se escolher y sempre diminuirá, já para a opção x vai depender do valor de Luce, se x tem um valor relativamente maior que z então a probabilidade de x ser escolhido aumenta quando se adiciona a opção z . Mostrando assim, o efeito atração.

Por fim, se voltarmos ao primeiro exemplo de que há três elementos iguais x, y, z , entretanto x e z possuem a mesma qualidade só que z sendo mais caro e y de uma qualidade mais barata e concluímos que $\rho(x; y, z) > \rho(x; y)$, agora se adicionarmos w que seja da mesma qualidade de x e z , entretanto, mais caro do que z . Com isso pelo mesmo raciocínio conclui-se que $\rho(z; y, w) > \rho(z; y)$.

Definição 3: Transitividade por Atração. Para qualquer $x, y, z, w \in X$, $\rho(x; y, z) > \rho(x; y)$ e $\rho(z; y, w) > \rho(z; y)$ implica que $\rho(x; y, w) > \rho(x; y)$.

Proposição 2: o Modelo de Luce com Associação satisfaz a Transitividade por Atração.

Em outras palavras, pode-se concluir que se adicionarmos uma opção w que possua um efeito atração sobre a opção z – considerando que a adição desta aumente a probabilidade de se escolher x contra y –. Pode-se afirmar, pelo efeito da transitividade, que a adição w aumentará também a probabilidade de se escolher x contra y .

4. Conclusão

O presente trabalho discutiu a literatura de escolha estocástica desde o modelo Luce, apresentando suas características fundamentais como a Positividade, a Independência das Alternativas Irrelevantes e por fim, a Regularidade. No entanto, foi mencionado alguns problemas descritivos ao modelo padrão, justamente por apresentar cenários em que o IAI não é satisfeito, causado pelo efeito similaridade, casos em que a regularidade não se verifica acarretado pelo efeito *Decoy*. Com isso, tais problemas suscitaram o desenvolvimento de novas aplicações que capturassem esses efeitos, como o modelo Eliminação por Aspectos, o Modelo de Luce com Associação e entre outros.

Em consequência disso, surgiram novas opções para se estimar resultados em áreas de tomadas de decisão dos indivíduos, organização industrial etc. Vale ressaltar que o analista se depara com *trade-offs* nas escolhas de qual modelo usar dado que se utilizar o modelo mais simples, no caso o Multinomial Logit, ignora-se o efeito da similaridade e o efeito *Decoy*, com isso, dependendo da área que se estuda, o modelo poderá trazer conclusões distantes da realidade. Por outro lado, é um modelo simples com poucas hipóteses e que não tem muitos parâmetros para se estimar. Em contrapartida, usar modelos que capturem tais anormalidades, gerará resultados mais próximos da realidade, entretanto, haverá um custo maior dado que há mais características a serem estimadas, terá que incorrer em vários testes e outras dificuldades.

Por fim, tendo em vista a discussão, cabe ao leitor verificar os fenômenos existentes nos cenários em que está estudando, se elas aparecem de forma recorrente. Pois caso se confirme essas circunstâncias deve-se tender a usar modelos mais completos, caso contrário, é interessante utilizar o modelo de Luce por sua simplicidade.

5. Referências

1. Debreu, G. (1960): "Review of 'Individual Choice Behavior' by R. D. Luce," *American Economic Review* 50, 186-188.
2. Huber, J., J. W. Payne, e C. Puto. (1982): "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research* 9, 90-98.
3. MacFadden, D. (1973): "Conditional Logit Analysis of Qualitative Choice Behaviour", *Frontiers in Econometrics*, Academic Press.
4. Williams, H.C.W.L (1977): "On the formulation of travel demand models and economic evaluation measures of user benefit", *Environment and Planning A*, 9 (3) (1977), pp. 285-344
5. Wen, C. H., e Koppelman, F. S. (2001): "The Generalized Nested Logit Model", *Transportation Research Part B: Methodological* Volume 35, Issue 7, Pages 627-641
6. Thurstone, L.L. (1927): "A law of comparative judgment," *Psychological Review* 34, 273-286.
7. Marschak, J. (1960): "Binary choice constraints and random utility indicators," paper 155, Cowles Foundation.
8. Li, J., e Tang, R. (2016): "Associationistic Luce Rule," National University of Singapore.
9. Luce, R. D. (2005). *Individual choice behavior: A theoretical analysis*. Dover Publications.